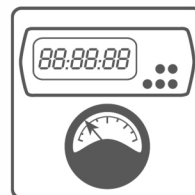
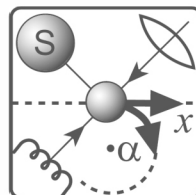


**Sergey Y. Yurish**

**Editor**

**Sensors, Measurements and Networks**



## Sensors, Measurements and Networks



Sergey Y. Yurish  
Editor

# **Sensors, Measurements and Networks**

**Book Series: Advances in Sensors, Vol. 8**



International Frequency Sensor Association Publishing

Sergey Y. Yurish  
*Editor*

Sensors, Measurements and Networks  
Advances in Sensors, Vol. 8

Copyright © 2022 by International Frequency Sensor Association Publishing, S. L.

E-mail (for orders and customer service enquires): [ifsa.books@sensorsportal.com](mailto:ifsa.books@sensorsportal.com)

Visit our Home Page on <http://www.sensorsportal.com>

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (IFSA Publishing, S. L., Barcelona, Spain).

Neither the authors nor International Frequency Sensor Association Publishing accept any responsibility or liability for loss or damage occasioned to any person or property through using the material, instructions, methods or ideas contained herein, or acting or refraining from acting as a result of such use.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identifies as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

ISBN: 978-84-09-44953-8  
e-ISBN: 978-84-09-44954-5  
BN-20221015-XX  
BIC: TJFC

## Acknowledgments

As Editor I would like to express my undying gratitude to all authors, editorial staff, reviewers and others who actively participated in this book. We want also to express our gratitude to all their families, friends and colleagues for their help and understanding.



# Contents

<b>Contributors.....</b>	<b>13</b>
<b>Preface .....</b>	<b>17</b>

## Chapter 1

<b>Single-Coil Eddy Current Sensors .....</b>	<b>19</b>
1.1. <i>Introduction.....</i>	19
1.2. <i>Design Features and Operating of Single-Coil Eddy Current Sensors.....</i>	22
1.3. <i>Measuring Circuits with SCECS.....</i>	26
1.3.1. <i>Differential Measuring Circuit with Blumlein Bridge .....</i>	26
1.3.2. <i>Measuring Circuits with a Differentiating Device based on an Operational Amplifier .....</i>	28
1.4. <i>Temperature Effect on the Informative Parameters of Single-Coil Eddy Current Sensors and the Solutions to Remove it.....</i>	31
1.5. <i>Varieties of Single-Coil Eddy Current Sensors .....</i>	33
1.5.1. <i>Single-Coil Eddy Current Gap Sensors .....</i>	34
1.5.1.1. <i>Single-Coil Eddy Current Gap Sensor with a Sensing Element as a Segment of a Linear Conductor .....</i>	34
1.5.1.2. <i>Cluster Compositions of Single-Coil Eddy Current Sensors with a Sensing Element as a Segment of a Linear Conductor .....</i>	37
1.5.1.3. <i>Screen-type Single-Coil Eddy Current Sensors with a Sensing Element as a Single Current-carrying Coil.....</i>	39
1.5.2. <i>Single-Coil Eddy Current Angular-Movement Sensors.....</i>	40
1.5.3. <i>Single-Coil Torque-Measuring Devices .....</i>	40
1.5.4. <i>Cluster Single-Coil Eddy Current Sensor for Debris Monitoring .....</i>	43
1.6. <i>Conclusions.....</i>	44
<i>References .....</i>	45

## Chapter2

### Optimum Design, Measurement Accuracy and Calibration of Optical Fiber

<b>Sensors Based on Strain Transfer Theory .....</b>	<b>49</b>
2.1. <i>Introduction.....</i>	49
2.2. <i>Problems of Optical Fiber-based Sensors for the Monitoring of Civil Structures .....</i>	49
2.2.1. <i>Measurement Limitations of Optical Fibers .....</i>	51
2.2.2. <i>Sensors Based on Strain Transfer Theory.....</i>	52
2.3. <i>Review on Strain Transfer Theory of Optical Fiber-based Sensors.....</i>	53
2.3.1. <i>Embedded and Surface-attached Sensors .....</i>	54
2.3.2. <i>Multi-layered Sensing Model under Different Loading Types .....</i>	55
2.3.3. <i>Multi-layered Sensing Model under Different Damage Modes.....</i>	59
2.4. <i>Design and Calibration of Optical Fiber-based Sensors .....</i>	61
2.4.1. <i>Priority Design Parameters .....</i>	61
2.4.2. <i>Improved Calibration Test for End-expanded Sensors .....</i>	62
2.5. <i>Conclusions.....</i>	65
<i>Acknowledgements.....</i>	66
<i>References .....</i>	66



**Chapter 3**

<b>Fine Measurements of Field Torsion and Its Spectra .....</b>	<b>69</b>
3.1. Introduction .....	69
3.2. Magnetic Nature of the Torsion .....	71
3.3. Torsion Energy .....	76
3.4. Electronics Gaugings and Experiments on the Measurement of Torsion .....	77
3.5. Conclusions .....	81
Acknowledgements .....	81
References .....	82

**Chapter 4**

<b>Optical Waveguides for Integrated Photonics: Mainstream Technologies Versus Low-cost Silica-titania Sol-gel Approach .....</b>	<b>85</b>
4.1. Introduction .....	85
4.2. Waveguide Structures .....	87
4.2.1. Conventional Waveguide Structures .....	87
4.2.1.1. Ridge Waveguide .....	87
4.2.1.2. Rib Waveguide .....	88
4.2.1.3. Diffused Waveguide .....	88
4.2.2. Unconventional Waveguide Structures .....	89
4.2.2.1. Slot Waveguide .....	89
4.2.2.2. Hybrid Plasmonic Waveguide .....	90
4.2.2.3. SWG Waveguide .....	91
4.2.2.4. Plasmonic Waveguides .....	91
4.3. Mainstream Photonic Integrated Circuits (PIC) Technologies .....	92
4.3.1. SOI WGs .....	92
4.3.2. InP/InGaAsP WGs .....	95
4.3.3. SiN WGs .....	96
4.4. Silica-titania Waveguides Developed Through the Sol-gel Dip-coating Approach .....	97
4.5. Future of Silica-titania Waveguide Optical System .....	100
4.6. Examples .....	101
4.6.1. Numerical Modelling of Bandstop Filters Based on Silica-titania Platform .....	101
4.6.2. Numerical Modelling of 1×2 optical power splitter based on silica-titania platform .....	101
4.6.3. Mode Sensitivity Analysis of a Rib Waveguide Based on Silica-titania Platform .....	103
4.7. Final Remarks .....	104
Acknowledgement .....	105
References .....	105

**Chapter 5**

<b>Detection, Classification and Re-identification of Road Users Across Heterogeneous Video Surveillance Networks Using Artificial Intelligence and Calibration Techniques .....</b>	<b>109</b>
5.1. Introduction .....	109
5.2. Object Detection .....	110
5.3. Metric Calibration .....	115
5.4. Height Estimation .....	118
5.5. Calibration with Known Extrinsic Parameters .....	121
5.6. Vehicle Re-identification across a Heterogeneous Camera Network .....	125
5.7. Conclusions .....	129
Acknowledgements .....	130
References .....	130

<b>Chapter 6</b>	
<b>Measurement of Microclimate as a Comfort Indicator .....</b>	<b>135</b>
6.1. <i>Introduction</i> .....	135
6.2. <i>Microclimate</i> .....	137
6.2.1. Sensor Technology .....	141
6.2.2. Sweating Source .....	144
6.3. <i>Applications</i> .....	147
6.3.1. SWEATOR and SWEATLOG .....	147
6.3.2. Validation of Proband and SWEATOR Test .....	148
6.3.3. Skin Integrity on Mattresses .....	149
6.4. <i>Conclusion</i> .....	150
<i>Acknowledgements</i> .....	151
<i>References</i> .....	151
<b>Chapter 7</b>	
<b>From Food to Mining: Dual Energy XRT Applications .....</b>	<b>153</b>
7.1. <i>Introduction</i> .....	153
7.2. <i>Dual Energy X-ray Transmission: The Method</i> .....	154
7.2.1. Dual Energy-X-ray Transmission Measurements .....	154
7.2.2. Basis Material Decomposition .....	155
7.2.3. Limits of BMD .....	157
7.3. <i>Mining</i> .....	159
7.4. <i>Food Safety</i> .....	160
7.4.1. Calibrated Test Bodies .....	160
7.4.2. Realistic Foreign Bodies .....	165
7.5. <i>Laundry</i> .....	166
7.6. <i>Metal Discrimination</i> .....	168
7.7. <i>Summary and Outlook</i> .....	169
<i>Acknowledgment</i> .....	169
<i>References</i> .....	169
<b>Chapter 8</b>	
<b>Industrial Measurements, Communications and Protocols .....</b>	<b>173</b>
8.1. <i>Introduction</i> .....	173
8.2. <i>Industrial Measurements</i> .....	174
8.3. <i>Industrial Communication Protocols</i> .....	176
8.3.1. Current Loops .....	176
8.3.2. HART .....	178
8.3.3. Fieldbus (FF) .....	181
8.4. <i>Case Study: FF Prototype</i> .....	184
8.4.1. Experimental Results .....	185
8.5. <i>Conclusions</i> .....	187
<i>References</i> .....	188
<b>Chapter 9</b>	
<b>Industrial Ethernet and Time Sensitive Networking for Industry 4.0 .....</b>	<b>191</b>
9.1. <i>Introduction</i> .....	191
9.2. <i>IIoT and Industry 4.0</i> .....	192
9.3. <i>Industrial Ethernet</i> .....	193
9.3.1. Real Time .....	193
9.3.2. Collision Avoidance and Synchronization .....	195
9.3.3. Industrial Internet .....	196

9.4. Time Sensitive Networking.....	198
9.5. Conclusions.....	203
References.....	203

## Chapter 10

<b>Towards a Digital Electronic Conditioning Approach Dedicated to High Sensitivity Giant Magnetoimpedance Sensor .....</b>	<b>205</b>
10.1. Introduction .....	205
10.2. Origin and Context of the Giant Magneto Impedance Effect.....	206
10.3. From Ferromagnetic Wire to Making a Sensor.....	209
10.3.1. GMI Sensor Conditioning Strategy.....	210
10.3.2. Noise Performance of a GMI Sensor .....	213
10.4. Hardware Architecture of a Digital off Diagonal GMI Sensor .....	214
10.4.1. The FPGA Zynq SOC and Design Process.....	216
10.4.2. FPGA Implementation of the Direct Digital Synthesizer (DDS) .....	217
10.4.3. Implementation of the Digital-Down-Converter (DDC).....	218
10.4.4. Results and Discussions.....	222
10.5. Conclusions.....	228
References.....	228

## Chapter 11

<b>R-testbench – How to Build a Reliable and Open-source Software Solution for the Remote Control of Measurements Instruments and Experimental Test Benches? .....</b>	<b>231</b>
11.1. Introduction .....	231
11.2. State of the Art: an Overview of the Existing Solutions .....	232
11.2.1. Proprietary Solutions .....	232
11.2.2. Open-source Alternatives.....	233
11.2.3. Comparison and Discussion.....	233
11.3. System-level Specifications .....	234
11.3.1. Reliability .....	234
11.3.1.1. Reliability of Control-instrument Communications .....	234
11.3.1.2. Reliability of Source Code (Core Development).....	235
11.3.2. Adaptability .....	236
11.3.2.1. Portability.....	236
11.3.2.2. No Brands or Models Restrictions.....	236
11.3.2.3. Communication Bus Multi-compatibility.....	236
11.3.3. User Interface.....	236
11.3.4. Open-source Access.....	237
11.3.5. Summary of the Specifications .....	237
11.4. Software Architecture and Implementation .....	237
11.4.1. Selection of the Programming Language .....	238
11.4.2. Design of the Software Architecture.....	238
11.4.2.1. A Front End to Call the VISA Libraries.....	238
11.4.2.2. Automated Instrument Recognition.....	239
11.5. Data Management Considerations .....	240
11.5.1. From the Instruments to the Computer: the Fetch Operation.....	241
11.5.2. From Acquisition to Temporary Saving: the Local Storage.....	241
11.5.3. From Temporary to Permanent Saving: the Dump Action.....	242
11.5.4. From Raw Data to Refined Information: the Data (Post-) Processing .....	242
11.6. Validation: Tests, Results, and Discussion .....	243
11.6.1. Continuous Integration and Alpha Tests.....	243
11.6.2. Performance Characterization.....	244

11.6.3. Public Release .....	246
11.6.4. Results Discussion.....	246
11.7. Conclusions.....	246
Acknowledgements .....	248
References .....	248
<b>Chapter 12</b>	
<b>Gas Sensing Technologies Application in Combustion .....</b>	<b>251</b>
12. Introduction.....	251
12.1. Combustion Process.....	252
12.2. Gas Sensing Technologies for Combustion Process Control.....	254
12.3. Oxygen Gas Sensing Technologies in Combustion .....	254
12.3.1. Zirconia Oxygen Potentiometric Technology .....	255
12.3.2. Tunable Diode Laser Spectroscopy (TDLS).....	257
12.3.3. Paramagnetic Oxygen Gas Sensing Technology .....	260
12.4. Carbon Monoxide (CO) Gas Sensing Technologies in Combustion .....	262
12.4.1. Calorimetric Catalytic CO Gas Sensing Technology.....	262
12.4.2. Mixed Potential Electrochemical CO Gas Sensing Technology .....	264
12.4.3. Tunable Diode Laser (TDL) and Quantum Cascade Laser (QCL) Spectroscopies .....	269
12.5. Methane (CH <sub>4</sub> ) Gas Sensing Technologies in Combustion .....	270
12.5.1. Mixed Potential Electrochemical CH <sub>4</sub> Gas Sensing Technology.....	270
12.5.2. Calorimetric Catalytic CH <sub>4</sub> Gas Sensing Technology .....	271
12.5.3. Tunable Diode Laser (TDL) and Quantum Cascade Laser (QCL) Spectroscopies .....	273
Acknowledgements .....	274
References .....	274
<b>Index .....</b>	<b>279</b>



# Contributors

**Christine Bauer** Fraunhofer IIS, Fraunhofer Institute for Integrated Circuits IIS, Division Development Center X-Ray Technology, Flugplatzstr. 75, 90768 Fürth, Germany

**Sergey Borovik** Samara Federal Research Scientific Center RAS, Institute for the Control of Complex Systems RAS, Samara, Russian Federation

**Francisco Bulnes** Research Department in Mathematics and Engineering, TESCHA, Chalco, Mexico  
Electronics Engineering Division, TESCHA, Chalco, Mexico  
E-mail: francisco.bulnes@tesch.edu.mx

**Muhammad Ali Butt** Warsaw University of Technology, Warsaw, Poland

**José Miguel Costa Dias Pereira** ESTSetúbal/IPS, Instituto Politécnico de Setúbal, Setúbal, 2914-508, Portugal  
Instituto de Telecomunicações, Lisboa, 1049 – 001, Portugal

**Evelyne Daubie** University of Mons, Nuclear and Subnuclear Physics Unit, 6 Avenue du Champ de Mars, 7000 Mons, Belgium

**Fortunato Dualibe** University of Mons, Electronics & Microelectronics Unit, 31 Boulevard Dolez, 7000 Mons, Belgium

**Markus Firsching** Fraunhofer IIS, Fraunhofer Institute for Integrated Circuits IIS, Division Development Center X-Ray Technology, Flugplatzstr. 75, 90768 Fürth, Germany  
E-mail: markus.firsching@iis.fraunhofer.de

**M. Y. Flores-García** Research Department in Mathematics and Engineering, TESCHA, Chalco, Mexico  
Electronics Engineering Division, TESCHA, Chalco, Mexico

**A. U. García Galicia** Research Department in Mathematics and Engineering, TESCHA, Chalco, Mexico  
Electronics Engineering Division, TESCHA, Chalco, Mexico

**J. C. García-Limón** Research Department in Mathematics and Engineering, TESCHA, Chalco, Mexico  
Electronics Engineering Division, TESCHA, Chalco, Mexico  
E-mail: jcgarcia@tese.edu.mx

**Anna Jusza** Warsaw University of Technology, Warsaw, Poland

**Paweł Karasiński** Warsaw University of Technology, Warsaw, Poland

**Andrzej Kaźmierczak** Warsaw University of Technology, Warsaw, Poland

**Bernhard Kurz** Institut für angewandte Ergonomie GmbH, Unterschleissheim, Germany

**Johannes Leisner** Fraunhofer IIS, Fraunhofer Institute for Integrated Circuits IIS,  
Division Development Center  
X-Ray Technology, Flugplatzstr. 75, 90768 Fürth, Germany

**Aleksej Makarov** Vlatacom Institute of High Technologies, 11070 Belgrade, Serbia

**Véronique Moeyaert** University of Mons, Telecommunications & Electromagnetism  
Unit, 31 Boulevard Dolez, 7000 Mons, Belgium

**L. A. Ortíz-Dumas** Research Department in Mathematics and Engineering, TESCHA,  
Chalco, Mexico  
Electronics Engineering Division, TESCHA, Chalco, Mexico

**Ryszard Piramidowicz** Warsaw University of Technology, Warsaw, Poland

**Alexandre Quenon** University of Mons, Electronics & Microelectronics Unit, 31  
Boulevard Dolez, 7000 Mons, Belgium

**Christoph Russ** Institut für angewandte Ergonomie GmbH, Unterschleissheim, Germany

**Victor A. Sánchez-Suárez** Research Department in Mathematics and Engineering,  
TESCHA, Chalco, Mexico  
Electronics Engineering Division, TESCHA, Chalco, Mexico

**Yuriy Sekisov** Samara Federal Research Scientific Center RAS, Institute for the Control  
of Complex Systems RAS, Samara, Russian Federation

**Pavel Shuk** Saint-Gobain Norpro, 3840 Fishcreek Road, Stow, 44224 Ohio, USA  
E-mail: Pavel.Shuk@saint-gobain.com, <https://www.norpro.saint-gobain.com/>

**Jakub Sochor** Brno University of Technology, 62700 Brno, Czech Republic

**Jakub Špaňhel** Brno University of Technology, 62700 Brno, Czech Republic

**Stanisław Stopiński** Warsaw University of Technology, Warsaw, Poland

**Papa Silly Traore** Ecole Supérieure Polytechnique of Dakar, Senegal

**Miljan Vučetić** Vlatacom Institute of High Technologies, 11070 Belgrade, Serbia

**Rebecca Wagner** Fraunhofer IIS, Fraunhofer Institute for Integrated Circuits IIS,  
Division Development Center  
X-Ray Technology, Flugplatzstr. 75, 90768 Fürth, Germany

**Huaping Wang** School of Civil Engineering and Mechanics, Lanzhou University,  
Department of Civil and Environmental Engineering, The Hong Kong Polytechnic  
University

School of Civil Engineering, Dalian University of Technology, China  
E-mail: wanghuaping1128@sina.cn; hpwang@lzu.edu.cn

**Pavel Zemčik** Vlatacom Institute of High Technologies, 11070 Belgrade, Serbia





# Preface

It is my great pleasure to present the 8<sup>th</sup> volume from our popular Book Series ‘Advances in Sensors’ started by the IFSA Publishing in 2012. This Book Series was accepted by all sensor community with a great enthusiasm.

According to *Next Move Strategy Consulting*, the global sensor market is projected to more than double in size between 2019 and 2030. While the market was sized at some 163.84 billion U.S. dollars in 2019, it is expected to reach the size of around 426.2 billion U.S. dollars in 2030. Emerging trends, which have a direct impact on the dynamics of the industry, include development of new product technologies, increasing usage of sensors in various applications, and advances in gas sensor materials and manufacturing.

Since the Vol. 4 of this book series it is published as an Open Access Book in order to significantly increase the reach and impact of this volume, which also published in two formats: electronic (pdf) with full-color illustrations and print (paperback).

The 8<sup>th</sup> volume entitled ‘Sensors, Measurements and Networks’ contains twelve chapters with the descriptions of latest advances in sensor related area written by 33 authors from academia and industry from 11 countries: Belgium, China, Czech Republic, Germany, México, Poland, Portugal, Russia, Senegal, Serbia, and the USA.

Like the first seven volumes of this Book Series, the 8<sup>th</sup> volume also has been organized by topics of high interest. In order to offer a fast and easy reading of each topic, every chapter in this book is independent and self-contained. All chapters have the same structure: first an introduction to specific topic under study; second particular field description including sensing or/and measuring applications. Each of chapter is ending by well selected list of references with books, journals, conference proceedings and web sites.

This book ensures that our readers will stay at the cutting edge of the field and get the right and effective start point and road map for the further researches and developments. By this way, they will be able to save more time for productive research activity and eliminate routine work.

Coverage includes current developments in physical sensors, chemical sensors, measurements and sensor networks. With this unique combination of information in each volume, the Advances in Sensors book Series will be of value for scientists and engineers in industry and at universities, to sensors developers, distributors, and users.

I hope that readers enjoy this book and that can be a valuable tool for those who involved in research and development of various physical sensors, chemical sensors, measuring systems and sensor networks.

I shall gratefully receive any advices, comments, suggestions and notes from readers to make the next volumes of Advances in Sensors: Reviews book Series very interesting and useful.

*Dr. Sergey Y. Yurish*

*Editor  
IFSA Publishing*

*Barcelona, Spain*

# Chapter 1

## Single-Coil Eddy Current Sensors

**Sergey Borovik and Yuriy Sekisov**

### 1.1. Introduction

Eddy current probes (ECP) [1, 2] are widely used for non-contact measurement of the displacements of electrically conductive objects, their mutual position, geometric dimensions, vibrations, technical diagnostics, and defect detection [3–7]. As its name implies, the operation of the probe is based on the eddy current measurement principle, which entails the determination of the monitored object's properties by analyzing the results of the electromagnetic interaction between the ECP's sensing element (SE) and the target. When the SE is supplied with a current, it causes a primary electromagnetic field to form in the surrounding area. If an electrically conducting object is placed in this field, the secondary electromagnetic field, distorting the field created by the current in SE, is induced according to the Faraday's induction law. The intensity of the secondary electromagnetic field, and therefore the degree of its influence on the primary field created by the current in SE, are determined by the mutual location of the SE and the monitored object, the defects (cavities) in the object, etc.

Of particular note is the applicability of the eddy current measurement principle only for electrically conductive objects (ferromagnetic or non-ferromagnetic). Therefore, ECP are not sensitive to non-conductive non-ferromagnetic materials and media such as dirt, dust, oil, etc. This circumstance predestines the ECP for applications in harsh industrial environment.

The coil-based ECP [1] are the most widely used among the variety of existing probes based on the eddy current principle. A typical design of the coil-based ECP usually contains one or two inductance coils, which are connected to a measuring circuit intended to convert their natural output parameters (inductances) [7, 8]. The inductance coils can both envelope the monitored object, for example, a pipeline (annular/ring ECP), and be located above the target (pancake-type probes, horseshoe-shaped coil probes, spiral coil

probes, coil probe arrays) [1]. Depending on the purpose, the ECP may also contain magnetic circuits, screens and other additional elements which make it possible to convert the desired characteristics of the monitored object into the probe's magnetic circuit parameters change.

Traditionally, the inductance coils used in ECP (at least in annular/ring ECP and in pancake-type probes) are multi-component. They contain a frame, turns of wire in insulation and other structural elements, whose material is often limited to the maximum possible operating temperature. Therefore, the temperature range of such probes is usually limited to a value of about 180...200 °C [6, 9, 10]. This significantly limits, and sometimes makes impossible to use ECP in aerospace engine building, nuclear power engineering, etc. Additionally, the ECP with orthogonal orientation of the multiturn inductance coil (pancake-type probes) have another special feature which reduces the potential sensitivity of the probe. This feature is associated with the decrease in the «efficiency» of turns of the inductance coil as they are removed from the monitored surface [11].

There are known some solutions that extend the operation temperature range of the coil-type ECP. The solutions are based on the forced water cooling of the probes [1, 12] or making the ECP working coils of platinum with placing the coils on a ceramic frame [13-15]. Unfortunately, these technologies are expensive and are not always applicable. For example, the water cooling is extremely problematic in research and testing of aircraft and rocket engines. An alternative without these disadvantages is single-coil eddy current sensors (SCECS) which constitute a separate and independent branch among the existing ECP [16].

The first developments of SCECS date back to the late 1980s, early 1990s [17, 18]. Theoretical and experimental studies in this field were initiated in USSR in the city of Kuibyshev (now – Samara, Russia) by the Chief Designer of aircraft engines of the NK brand, Academician N.D. Kuznetsov. Their goal was to create the tools for the collecting and processing of the measuring information about radial clearances between the blade tips and the stator of the compressor and turbine of the aircraft gas-turbine engine (GTE). As it is noted in [19] the main energy, strength, economic and environmental indicators of the GTE (thrust, specific fuel consumption, exhaust toxicity, the boundaries of stable operation, engine operation safety, etc.) are largely determined by the value of the radial clearances. Therefore, the SCECS were initially oriented to very harsh and even extreme conditions in the gas-air path of the aircraft engine, which are associated with high temperatures, reaching 1200 °C or more in GTE turbines, high linear speeds of the monitored blades tips (up to 400 m/s), contamination of the gas-air path by fuel combustion products, oil particles and dust, high vibration level of structural elements on which the sensors are fixed, etc. It was also necessary to ensure the sensors performance and the repeatability of their metrological characteristics during the whole engine test cycle.

To consider all these requirements, the developers came up with the idea to use the simplest geometrical shape for the SE in a form of a single current loop or its part (a segment of a linear conductor). It is obvious that such design of the SE is highly reliable,

is high-tech, and does not require the use of the high temperature winding insulation. Furthermore, the use of the materials with a high melting point (for example, the same steel alloys that are used in turbine blades) for the manufacture of the SE and other sensor's structural elements, provides the required mechanical and thermal resistance of the sensor and its operability under extreme temperatures in GTE gas-air path without additional cooling of the SE.

It should also be noted that the extremely low SE's self-inductance (of about few units-tens of nH) is the main disadvantage of the SCECS. At the same time the useful changes in the SE's inductance from the influence of the measured parameters are only about 1...2 % of its self-inductance. Therefore, the direct connection of the SE in the measuring circuit (MC) of the secondary converter is difficult and the matching transformer (MT) is used for these purposes. MT is located in a zone with mild temperature and is connected to the SE through the current leads with a minimum intrinsic inductance. In addition, the SCECS are designed to be used in the industrial conditions which are characterized by the impact of intense electromagnetic interference. Then, an acceptable signal-to-noise ratio should be ensured especially considering the fact that the changes in the SE's inductance from the influence of the measured parameters are small.

The effective method to increase the level of the useful ECP signal is to use a pulsed sensor power supply. On the opposition to the traditionally used in ECP high-frequency harmonic power supply [1, 8, 13, 20], the pulsed mode makes it possible to increase the amplitude of the information signal at the output of the MC by increasing the voltage value of the sensor's supply pulse while maintaining the average power dissipated in the sensor. The desired thermal mode of the sensor can be saved by reducing the pulse duration [21]. With SCECS pulsed power supply, the signal at the output of the MC also has a pulsed character. It greatly simplifies further switching and analog-to-digital conversion. The effect of suppressing undesirable influences of non-informative parameters (for example, the active resistance of the ECP, varying widely from the influence of temperature in the measuring area) can be also obtained in measuring devices [22] by choosing the ratio between the pulse duration and the MC's constant time.

It was precisely those decisions that were applied in the SCECS with SE in the form of a conductor segment, designed to measure the displacements of structural elements of power plants of various classes [16–18, 23, 24]. The sensors were used as a part of measurement systems during bench tests of full-size GTE, their components, and assemblies. The tests confirmed the high performance and reliability of the sensors.

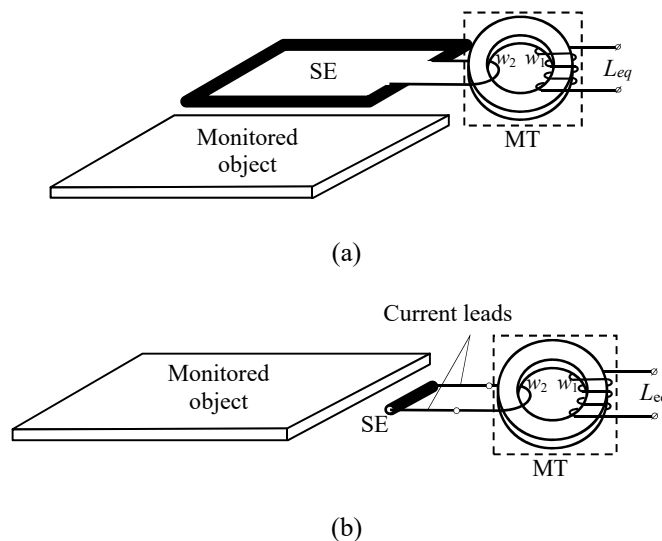
SCECS are not only used for measuring of the displacements of the GTE's constructive elements under the extreme temperatures. There are also other applications of the SCECS. For example, the SCECS with full-hole SE are used in tribodiagnostics for the detection of metal wear particles in the lubrication systems of the bearing units of the power plants [25]. In turn, the SCECS with the parallel orientation of the SE contours have found their application for measuring of the kinematic and dynamic characteristics of the power plants (angular movements, shaft torques, etc.) [26].

The design features of the SCECS, the principles of converting the measured parameters in a single-coil SE and further in a MC, as well as the practical applications of the sensors are described in this book chapter.

## 1.2. Design Features and Operating of Single-Coil Eddy Current Sensors

SCECS can be divided into two main groups according to their own unique design features. The first group includes the SCECS with SE as the whole current-carrying coil (circuit) and the second one includes the SCECS with SE as a fragment of the coil (for example, a segment of a linear conductor) that is connected using "non-inductive" current leads (usually in the form of closely spaced and insulated conductive strips or coaxial cylinders) to the MT.

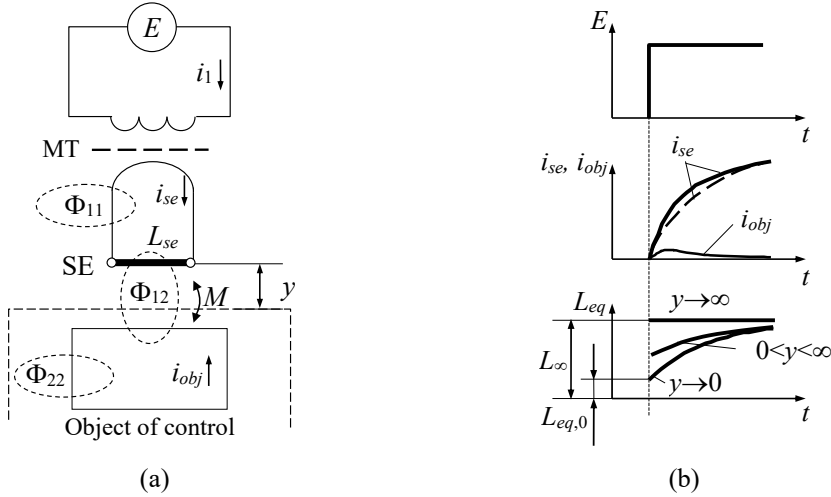
The schematic images of both types of SCECS are shown in Fig. 1.1. The simplest double-contour model of the electromagnetic interaction between the SE of the SCECS and the electrically conductive object is presented in Fig. 1.2a. The time diagrams of changes in the currents and equivalent inductance of the SE on an example of the measuring of radial clearances between the blade tips and the stator of a power plant are presented in Fig. 1.2b.



**Fig. 1.1.** Schematic images of the SCECS with SE as a single current-carrying coil (a), and with SE as a segment of a linear conductor (b).

The monitored object (the blade tip) is located at a distance  $y$  from the SE (Fig. 1.2a). The sensor is powered by the rectangular voltage pulses with amplitude  $E$ . Let's assume that the MT does not distort the front edge of the supply voltage, which excites the growing current  $i_{SE}$  in the SE circuit ( $i_1$  is a current in the primary MT winding). Current  $i_{SE}$  causes a primary electromagnetic field to form in the surrounding area. The eddy current  $i_{obj}$  appears under the action of the electromagnetic field in the contour that imitates the target

and forms the secondary electromagnetic field. Magnetic fluxes  $\Phi_{11}$  excited by the current  $i_{se}$ , and  $\Phi_{12}$  excited by the current  $i_{obj}$  pass through the surface bounded by the primary circuit. The flux linkage  $\Psi_1$  is determined by the total magnetic flux  $\Phi_1 = \Phi_{11} - \Phi_{12}$ .



**Fig. 1.2.** Double-contour model of the electromagnetic interaction between the SE of the SCECS and the electrically conductive object (a), and the time diagrams of changes in the currents and equivalent inductance of the SE (b).

The flux linkage  $\Psi_1$  can also be expressed through the SE's self-inductance  $L_{se}$  and mutual inductance  $M$  [27] as  $\Psi_1 = L_{se} \cdot i_{se} - M \cdot i_{obj}$ . Then the equivalent inductance of the SE (Fig. 1.2b) can be written as a function of the ratio of currents  $i_{se}$  and  $i_{obj}$ :

$$L_{eq} = \Psi_1 / i_{se} = L_{se} - M \cdot (i_{obj} / i_{se}). \quad (1.1)$$

When the monitored object is at a large distance from the SE ( $y \rightarrow \infty$ ), the electromagnetic interaction between the SE and the target is absent ( $M=0$  and  $i_{obj}=0$ ). If the influence of eddy currents in the SE and other sensor's structural elements associated with the magnetic field caused by the current  $i_{se}$  can be neglected, the current  $i_{se}$  is determined only by the self-inductance  $L_{se}$  at the beginning of the transition process. In this case the  $i_{se}$  changes in time will have an exponential growing character [28] (Fig. 1.2b, dashed curve) and in accordance with (1.1) the SE's equivalent inductance will equal its self-inductance ( $L_{eq} = L_{se}$ ).

With the approach of the monitored object to the SE ( $y \rightarrow 0$ ), the eddy currents appear in the target under the action of a magnetic field created by  $i_{se}$  and time-varying  $i_{obj}$  current appears in the contour that imitates the target. The current  $i_{obj}(t)$  affects the resulting magnetic field and this leads to changes in the shape of the current  $i_{se}(t)$  and its deviation from the exponential dependence (Fig. 1.2b, solid curve). Such deviation can be interpreted as the influence of a time-variable equivalent inductance  $L_{eq}(t)$ , whose instability in the transition mode is explained by the influence of eddy currents in

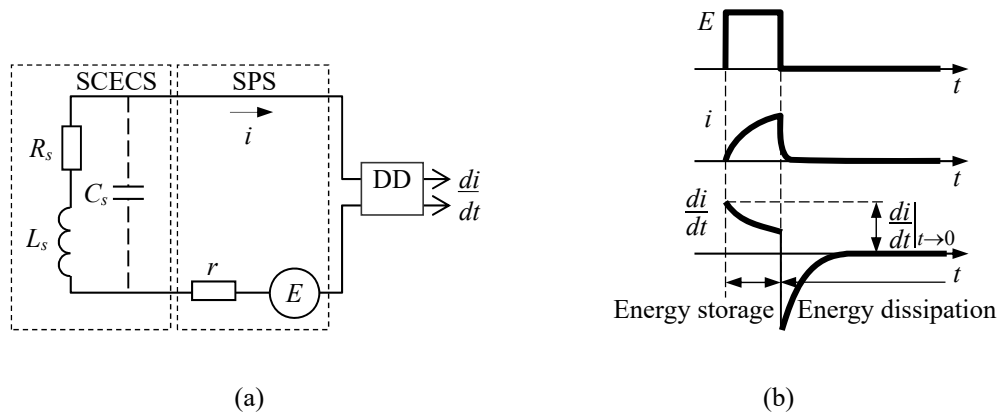


electroconductive object. It is easy to show that the equivalent inductance  $L_{eq}$  is minimal at the beginning of the transition process (at  $t \rightarrow 0$ ), and in accordance with (1) it is defined only by the SE's self-inductance ( $L_{se}$ ) and contours' mutual inductance ( $M$ ) which in turn depends on the transformation parameter – the distance between the SE and the target ( $y$ ):

$$L_{eq}(0) = L_{eq,0} = \lim_{t \rightarrow 0} (L_{se} - M(i_{obj}/i_{se})). \quad (1.2)$$

In steady state at  $t \rightarrow \infty$ , the eddy currents are attenuated ( $i_{obj} = 0$ ) and the equivalent inductance  $L_{eq}$  increases and tends to the self-inductance  $L_{se}$  at the infinite distance of the target from the SE (the set of the  $L_{eq}(t)$  characteristics for the different distances between SE and the monitored object are also given in Fig. 1.2b). Obviously, the beginning of the transition process is characterized by the greatest sensitivity of the inductance  $L_{eq}$  to the distance ( $y$ ) changes between SE and the target. In this connection the time point  $t \rightarrow 0$  is the most attractive for subsequent transformations.

Among the pulse techniques used for converting the parameters of the inductive and eddy-current sensors, the method of the first derivative [22] (one of the test transition methods) provides the minimum duration of the pulse of a power supply, and therefore the maximum permissible amplitude. The method allows to convert the instantaneous value of the equivalent inductance of the SE at the start of the transition process at the time of formation of the read pulse (at  $t \rightarrow 0$ ). The circuit and the time diagrams explaining the method of the first derivative are given in Fig. 1.3.



**Fig. 1.3.** Circuit (a), and time diagrams (b) explaining the method of the first derivative.

The circuit (Fig. 1.3a) contains a switching power supply (SPS) with electromotive force (emf)  $E$  and internal resistance  $r$ , a SCECS represented by its equivalent parameters ( $L_s$ ,  $R_s$ ,  $C_s$ ) and a current differentiation device (DD) with a zero-input resistance. The equivalent inductance of the sensor ( $L_s$ ) that is recalculated to the primary winding of the MT is associated with the equivalent inductance of the SE ( $L_{eq}$ ) by the expression

$L_s = n_t^2 \cdot L_{eq}$ , where  $n_t$  is the MT transformer ratio ( $n_t = w_1/w_2$ ,  $w_1$  and  $w_2$  – the number of turns, and  $w_2 = 1$ ) [23].

If the self-capacity of the MT primary winding ( $C_s$ ) is not taken into account, then the transition process in its circuit with pulsed supply  $E$  and the time-varying of  $L_{eq}(t)$  (and consequently  $L_s(t)$ ) is described by the equation:

$$L_s(t) \frac{di}{dt} + i \cdot \left[ \frac{dL_s(t)}{dt} + R_s + r \right] = E. \quad (1.3)$$

Since  $L_s(t)$  is the continuous function, the value of the current derivative can be found without solving the equation (1.3) for zero initial conditions at the moment of the generation of voltage  $E$  (at  $t \rightarrow 0$  current  $i=0$ ):

$$i \cdot \left[ \frac{dL_s(t)}{dt} + R_s + r \right] = 0, \quad \left. \frac{di}{dt} \right|_{t \rightarrow 0} = \frac{E}{L_s(0)}, \quad (1.4)$$

where  $L_s(0)$  is the inductance at  $t \rightarrow 0$  ( $L_s(0) = n_t^2 \cdot L_{eq}(0) = w_1^2 \cdot L_{eq,0}$ ).

This means that the current derivative does not depend on the resistance  $R_s$  and is determined by the instantaneous value of the inductance. The conversion process can be repeated after the derivative is fixed and the energy dissipation in the circuit is completed, (Fig. 1.3b).

Thus, the method of the first derivative provides the minimum conversion time, the maximum sensitivity to the distance to the target at the time of  $t \rightarrow 0$  and an increased accuracy by eliminating the influence of  $R_s$  changes.

Considering the self-capacitance ( $C_s$ ) of the SCECS that shunts  $L_s$  and  $R_s$ , the current  $i(t)$  jumps and is determined by the capacitive component at the beginning of the transient process. And  $i(t)$  is determined by the inductive component when the capacitance is charged. In this case the derivative reaches its maximum not at  $t \rightarrow 0$  as in the idealized circuit (when  $C_s=0$ ), but after some time  $t_m$ . The area of the transient process, where the current  $i$  is determined by the capacitive component has an irregular character and is not suitable for the converting of the inductance. Therefore, the maximum of the derivative at the beginning of the regular area of the transient process ( $t=t_m$ ), where the influence of the main parameter of the circuit becomes decisive, is taken as the output signal.

Obviously, self-capacitances in real circuits slow down the performance and worsen the resistance suppression ( $R_s$ ). However, the time  $t_m$  does not exceed tens of nanoseconds and the first derivative method provides a significant suppression of the influence of the active resistance of the sensor even under unfavorable parameter ratios [23]. And the non-linear character of the dependence of the DD's output signal on the SCECS's inductance should be also noted.

### 1.3. Measuring Circuits with SCECS

#### 1.3.1. Differential Measuring Circuit with Blumlein Bridge

The implementation of the first derivative method that ensures the transformation of the changes in the equivalent inductance of the SCECS during the process of the interaction of its SE with the monitored object, is carried out in differential MC [23]. Historically, the MC based on the Blumlein transformer measuring bridge (Fig. 1.4) were the first to be used to convert the parameters of the SCECS [29].

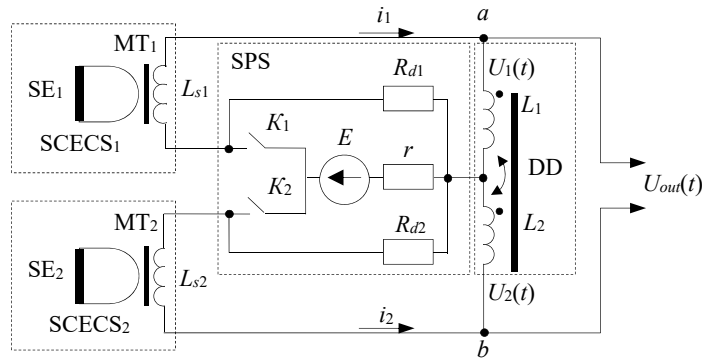


Fig. 1.4. Differential measuring circuit with Blumlein bridge.

The MC contains a bridge measuring circuit whose two adjacent arms include SCECS<sub>1</sub> and SCECS<sub>2</sub> with sensing elements SE<sub>1</sub>, SE<sub>2</sub> and matching transformers MT<sub>1</sub>, MT<sub>2</sub>. The sensors are powered by short rectangular voltage pulses with the frequency of up to several megahertz from a switching power supply (SPS). The feed pulses are formed by the contactless key elements ( $K_1$ ,  $K_2$ ) in the DC voltage source ( $E$ ). The energy stored in the differential circuit is dissipated through the  $R_{d1}$  and  $R_{d2}$  resistors during the time of the feed pulse (the key elements  $K_1$ ,  $K_2$  are open when the energy is dissipated). Magnetically coupled inductors (the coupling coefficient is equal to one) are used as a current differentiation device (DD). In this case, the DD's input resistance in relation to the clamps  $a$ ,  $b$  is determined only by the active resistance of the inductance coils. It is assumed that the monitored object is at an equal distance from the centers of SE<sub>1</sub> and SE<sub>2</sub>.

Moving an object relative to its initial position, for example from SE<sub>1</sub> to SE<sub>2</sub>, will lead to changes in the equivalent inductance of the sensors:

$$L_{s1} = L_0 + \Delta L, \quad L_{s2} = L_0 - \Delta L, \quad (1.5)$$

where  $L_0$  is the self-inductance of the SCECS (to simplify it is assumed that SCECS<sub>1</sub> and SCECS<sub>2</sub> are completely similar in their parameters and therefore  $L_{s1,0} = L_{s2,0} = L_0$ );  $\Delta L$  are the changes of the SCECS inductances caused by the movement of the target in the sensitivity zone of the sensors. Since the voltages on the bridge arms are determined at

$t \rightarrow 0$  only by the inductances and do not depend on the active resistances of the circuit, the voltage at the DD output can be defined as:

$$U_{out}(0) = U_1(0) - U_2(0) = L \cdot \left( \left. \frac{di_1}{dt} \right|_{t \rightarrow 0} - \left. \frac{di_2}{dt} \right|_{t \rightarrow 0} \right) = 2E \frac{\Delta L}{L_0} \cdot \frac{1}{1 + \frac{1}{2} \cdot \frac{L_0}{L} \left[ 1 - \left( \frac{\Delta L}{L_0} \right)^2 \right]}, \quad (1.6)$$

where  $U_1$  and  $U_2$  are the voltages on the inductances  $L_1$  and  $L_2$  ( $L_1=L_2=L$ ). The self-capacitances of the inductance coils of the sensors and DD are not considered.

Taking into account that  $L_0 \gg \Delta L$  and  $L \gg L_0$ , the equation (1.6) can be rewritten as:

$$U_{out}(0) \approx 2E \frac{\Delta L}{L_0}. \quad (1.7)$$

It can be seen from (1.7) that the Blumlein bridge provides an almost linear conversion and the instability of the inductance  $L$  has any effect on the result [30].

If it is impossible to obtain the opposite in sign changes of the equivalent inductances of the SCECS<sub>1</sub> and SCECS<sub>2</sub> (the monitored object moves relative to the SE only of a single SCECS, for example, as it happens when the radial clearances between the stator and the blades tips of the GTE rotor are measured), one of the sensors (for example, SCECS<sub>1</sub>) is used as a working sensor (its equivalent inductance is  $L_{s1}=L_0+\Delta L$ ), and the second one (SCECS<sub>2</sub>) is used as a compensation sensor (its equivalent inductance  $L_{s2}=L_0$  does not depend on the position of the target, but is influenced by the same interfering factors and primarily by high temperature). In this case, the nature of the dependence of  $U_{out}$  on inductance changes remains similar, but the sensitivity of the MC is reduced by half:

$$U_{out}(0) \approx E \frac{\Delta L}{L_0}. \quad (1.8)$$

In real circuits the increase of  $L$  leads to growing influence of the DD's winding stray capacitances, which turn out to be much larger than the self-capacitances of the sensors. The transient process in the measuring circuit becomes oscillatory and the first maximum of the half-wave of the output voltage is taken as the conversion result. When the sensor is removed from the DD, the equivalent capacitance includes the capacitance of the communication line and its influence increases.

The fixation time of the maximum of the derivative ( $t_m$ ) changes almost by order depending on the parameters of the SCECS, DD and the length of the communication line. For example, it is noted in [23] that the change of the length of the communication line from 1 m to 30 m in the MC with SCECS intended for radial clearances measuring in GTE, modifies  $t_m$  from 0.1  $\mu$ s to 1  $\mu$ s. The duration of the dissipation of the energy accumulated in the MC during the conversion also increases with the increase of the line

length. The MC must be returned to zero or close-to-zero initial conditions to be recovered. The recovery time is about 1  $\mu\text{s}$  for a line length of about 1 m and increases to 30  $\mu\text{s}$  for a line length of about 30 m [23]. This means that the power pulses repetition rate can vary from tens to thousands of kHz depending on the location of the signal converter of the SCECS.

### 1.3.2. Measuring Circuits with a Differentiating Device Based on an Operational Amplifier

Another variant of MC with SCECS that implements the first derivative method is the circuit in which the operational amplifiers (op-amps) are used as a differentiating device instead of inductance coils with close magnetic coupling [23, 31, 32]. It should be noted that the use of a microelectronic base creates a perspective for the MC's miniaturization and its integration in the SCECS or embedding into the communication line at a short distance from the sensor. Moreover, the integration of MC with the analog-to-digital converters (ADC) and microcontroller-based computing devices within a single design allows to create so-called "intelligent sensors" [33, 34] with automatically performed self-control and advanced information processing [34].

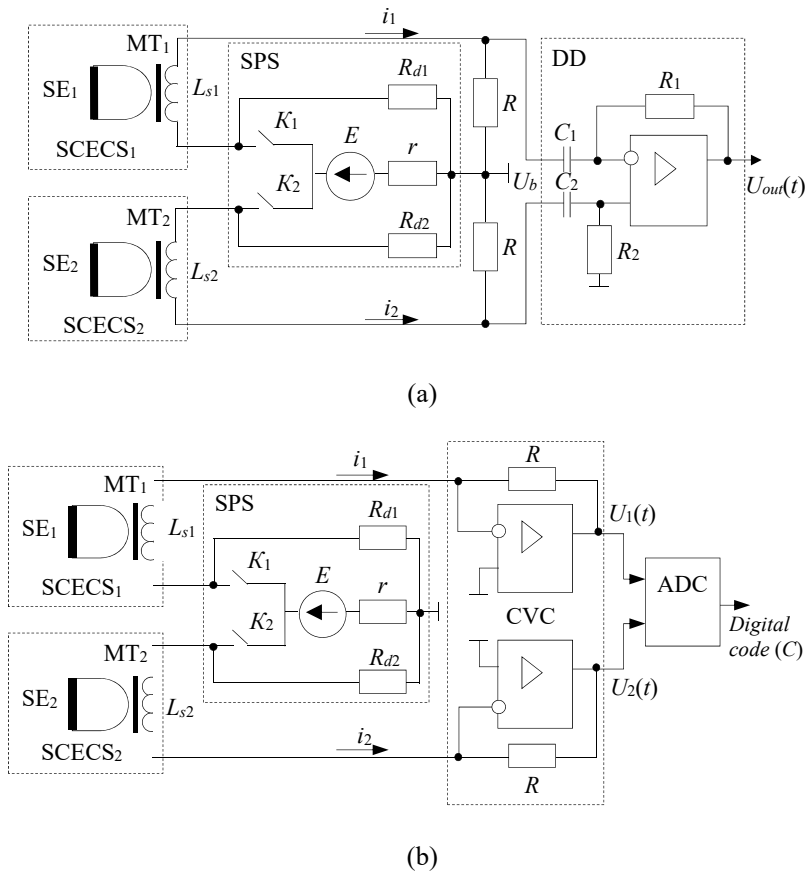
Currently there are two varieties of the existing MC built on electronic element base with op-amps. The first one (Fig. 1.5a) contains non-equilibrium  $LR$ -bridge at the input of the op-amp operating in a voltage differentiation mode. The adjacent arms of the bridge include the MT of both sensors (SCECS<sub>1</sub> and SCECS<sub>2</sub>), as well as resistors with equal resistance  $R$ . The stabilized SPS is included in the bridge's power supply diagonal and forms a brief pulsed voltage  $U_p$ . The voltage on the bridge output  $U_b$  is the potential difference between the resistors  $R$  formed by the exponentially increasing currents  $i_1$  and  $i_2$ . The voltage  $U_b$  is fed to the difference differentiator based on the op-amp [35] where it is differentiated.

Relevant to the scheme on Fig. 1.5a it should be noted that the resistance  $R$  is typically many times greater than the ohmic resistance and the equivalent resistance of the primary winding of the SCECS's MT. Considering the stray self-capacitances, the maximum of the derivative is fixed at the moment  $t_m$  that may be commensurate with the circuit's time constant which is determined by the inductance of the SCECS's MT and the resistor  $R$ .

This contradicts the fundamentals of the method of the first derivative under which the measurement should be carried out at the beginning of the transition process. That is why the new (improved) variant of the MC with op-amp as a differentiating device was developed. In [36] it was named as the «electronic analog of the Blumlein bridge» (Fig. 1.5b). In this variant of the MC the resistors  $R$  in the adjacent arms of the bridge are replaced by the current-to-voltage converters (CVC) on the base of op-amps with resistors  $R$  in the feedback circuits which convert the currents  $i_1(t)$  and  $i_2(t)$  into voltages  $U_1(t)$  and  $U_2(t)$  at their outputs [37].

The feature of the «electronic analog» scheme is that the currents in SCECS and the time constants of the MC are the same as in the Blumlein bridge with a similar power supply.

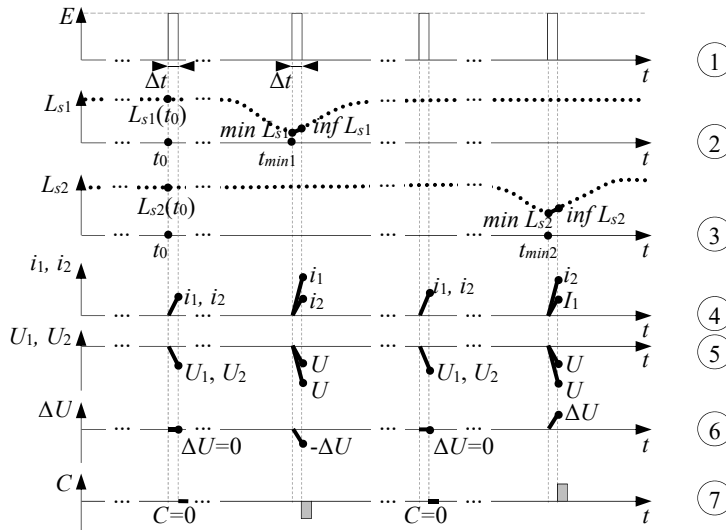
The difference voltage from the CVC outputs ( $\Delta U=U_1-U_2$ ) is either differentiated by the op-amp or the so-called operation of “approximate differentiation” is carried out. In the last case the difference voltage is fixed after a short time interval equal to the duration of the feed pulse and then the analogue-to-digital conversion is produced [38, 39]. The output voltages ( $U_1, U_2$ ) are delivered directly to the differential inputs of the ADC (Fig. 1.5b). This ensures that the digital code at the ADC output will be proportional to the voltage difference  $\Delta U$  at the end of the pulse of the power supply.



**Fig. 1.5.** Measuring circuits with the differentiating device on the base of op-amp with a non-equilibrium LR-bridge (a), and current-to-voltage converters (b).

Fig. 1.6 presents time diagrams that explain the functioning of the variant of the MC in the form of the «electronic analogue of the Blumlein bridge» with approximate differentiation. The measuring of the radial clearances between the blades tips and the GTE stator is considered as an example. The sensors are placed on a stator shell with a shift multiple of  $1/2$  blade pitch. In this case when the blade tip passes the SE of the first sensor (for example,  $SCECS_1$ ), the SE of the second one ( $SCECS_2$ ) is in an inter-blade gap, and vice versa.

The diagram 1 shows the power supply of the SCECS by voltage pulses with the amplitude equal to  $E$  (source voltage). The pulses rate is determined by the time interval during which a monitored object is in the sensor sensitivity zone. The time interval, in turn, depends on the maximum possible linear speed of the target, its size, the dimensions of the SCECS's SE, etc. For example, if the rotor speed of a turbocompressor is 10000 rpm and the blade tips' linear speed is 500 m/s, then each blade will pass the sensitivity zone of the SCECS with SE of 10 mm length during  $20 \cdot 10^{-6}$  s. At a pulse rate of 1 MHz (pulse period  $10^{-6}$  s) there will be 20 power pulses in the specified interval (only one pulse from the sequence of pulses is shown in diagram 1 both when the blade is in the SE's sensitivity zone and when it is out of the zone).



**Fig. 1.6.** Timing diagrams of the operation of the MC in the form of the «electronic analogue of the Blumlein bridge» with approximate differentiation.

The time intervals relevant to the passage of the monitored object through the sensitivity zone of the SCECS are shown in diagrams 2 and 3. The minimum values of the inductances ( $\min L_{s1}$ ,  $\min L_{s2}$ ) fixed at the moments ( $t_{\min1}$ ,  $t_{\min2}$ ) when the rising edge of the power supply pulses appear in accordance with the first derivative method, are indicated in each of the intervals. At this particular moment in time when the SE equivalent inductance has a minimum value, which then monotonically increases over the time [23]. However, the conversion of  $\min L_{s1}$ ,  $\min L_{s2}$  at the moments  $t_{\min1}$ ,  $t_{\min2}$  is possible only in the ideal MC. In the real MC the derivative is fixed after some time interval  $\Delta t$  during which the inductance changes (increases) from  $\min L_{s1}$ ,  $\min L_{s2}$  to some values corresponding to the moments of the end of the feed pulse. These values are taken to be the informative ( $\inf L_{s1}$ ,  $\inf L_{s2}$ ).

The variations with time in currents  $i_1$  and  $i_2$  at the input of both SCECS and voltages  $U_1$  and  $U_2$  at their outputs (Fig. 1.5b), as well as the difference voltage  $\Delta U$  are shown in the diagrams 4, 5, 6. A transient process begins in SCECS<sub>1</sub> and SCECS<sub>2</sub> as a response to the

rising edge of the feed pulse at the moment  $t_0$ . Since the  $MT_1$ ,  $MT_2$  have the same resistances and the equivalent inductances of their primary windings, as well as the time constants of both arms of the input circuit are equal, the currents  $i_1$  and  $i_2$  are also equal and we can assume that they are changing according to the linear law during a period of the power supply pulse  $\Delta t$  (diagram 4). Considering the inversion, the voltages  $U_1$  and  $U_2$  remain equal and change over the time in the same manner (diagram 5). Therefore, the difference voltage  $\Delta U = U_1 - U_2$  ( $\Delta U = -U_1 - (-U_2)$  considering the sign) at the CVC outputs remains zero as well as the digital code ( $C$ ) at the ADC output (diagrams 6, 7). The energy dissipation is intended to be instantaneous at the end of the power supply pulse (after  $\Delta t$  time from the moment  $t_0$ ) (diagrams 4, 5, 6). In addition, it is assumed that the times  $t_{min1}$ ,  $t_{min2}$  coincide with the rising edges of the feed pulses and correspond to the inductances' minimum values  $min L_{s1}$  and  $min L_{s2}$  (diagrams 2, 3). The dissipation of the energy that was accumulated during the interval  $\Delta t$  takes some time in reality. The mismatches between  $t_{min1}$ ,  $t_{min2}$  and the rising edge of the feed pulse are also possible and even more probable. These facts can cause the additional error.

The inductance  $min L_{s1}$  is less than  $L_{s1}(t_0)$  while  $L_{s1}(t_0)$  is equal to  $L_{s2}(t_0)$  and  $L_{s2}$  remains its value at  $t_{min1}$ . Then the time constant of the SCECS<sub>1</sub> will be less than the time constant of the SCECS<sub>2</sub> and the slope of the curve  $i_1(t)$  will be greater than the slope of the curve  $i_2(t)$  at this moment in time. The slope of the curve  $U_1(t)$  will also be larger compared to the slope of the curve  $U_2(t)$  (diagrams 4, 5). This means that the difference voltage  $\Delta U$  at the end of the time interval  $\Delta t$  will become negative ( $\Delta U = -U_1 - (-U_2) = -U_1 + U_2 < 0$ ) (diagram 6). The digital code  $C$  at the ADC output will also be negative.

In turn, the inductance  $min L_{s2}$  is less than  $L_{s2}(t_0)$  while  $L_{s1}(t_0)$  is equal to  $L_{s2}(t_0)$  and  $L_{s1}$  remains its value at  $t_{min2}$ . Then the time constant of the SCECS<sub>2</sub> will be less than the time constant of the SCECS<sub>1</sub> and the slope of the curve  $i_2(t)$  will be greater than the slope of the curve  $i_1(t)$  at this moment in time. As a result, the difference voltage  $\Delta U$  will be positive as well as the digital code  $C$  at the ADC output (diagrams 4-7).

Considering that  $\Delta U$  can be both positive and negative polarity, the implementation of this variant of the MC requires a high-resolution ADC with differential inputs [40, 41]. If it is not appropriate or possible for any reason to use a bipolar power supply, then an asymmetric MC is used with a scaling amplifier at the CVC output and a single-supply ADC [32]. The inductance of one of the arms of the bridge circuit is increased by adding the extra inductance  $L_a$  to ensure a single-supply voltage at the output of the scaling amplifier and the nominal operation mode of the ADC. The  $L_a$  value is about 1 % of the inductance  $L_0$  and it cannot be lower than the maximum possible changes in the inductance of the MT's primary winding caused by the displacements of the monitored object.

#### 1.4. Temperature Effect on the Informative Parameters of Single-Coil Eddy Current Sensors and the Solutions to Remove it

It is noted in the introduction that despite the great potential of SCECS for diagnostics and testing of complex technical facilities their appearance and main purpose are mainly related to the tasks of experimental research of GTE and, in particular, for the



measurement of radial clearances in the flow path of a compressor or turbine. Meanwhile the SE and other structural elements of the SCECS operate in the gas-air path under extreme temperatures up to 650 °C in the compressor and 1500 °C in the turbine.

It was shown in the experiments that the changes of the informative parameter of SCECS (equivalent inductance of the MT primary winding) at a temperature of 1000 °C reach 5 % and it is many times higher than the changes of the same parameter associated with radial clearances (about 1 per cent by 1 mm of the clearance) [32, 42]. Therefore, extreme temperature in the measuring area is the main interfering factor affecting the result of the conversion of measured parameters in the sensors of this class [42, 43].

The differential MC with two identical SCECS (working SCECS and witness (compensation) SCECS) placed in the same environmental conditions makes it possible to reduce the temperature effect on the sensor, but not to remove it. The main reason why the complete elimination of the temperature effect is not feasible is the technological impossibility to manufacture two absolutely identical SCECS. In addition, in cases when the monitored object moves relative to the SE of only one sensor (measuring of the radial clearances in GTE), the result of the electromagnetic interaction between the SE of the working SCECS and the monitored object (blade tip) depends not only on the temperature of the sensor's elements, but also on the temperature (more precisely, temperature changes in electrophysical parameters) of the target [44].

To further reduce of the temperature effect on the SCECS and the monitored object an algorithmic method for the correcting the temperature effect is additionally used [23]. Correction involves the experimental obtaining of the calibration characteristics of the measuring channels with SCECS in the form of dependences of voltages (or codes) at the MC output on the measured parameters, considering the temperature changes  $\theta$  in the SE locating area. The families of calibration characteristics are preliminary obtained on the special calibration devices with a thermal chamber and a fragment of the target before the sensors are installed on a product or test bench [23]. In this case the SCECS is completed with at least one thermocouple which hot junction is placed next to the sensor's SE and provides the temperature measurement in the working area.

The high labour intensity of the experimental obtaining of the calibration characteristics is evident. It can be reduced by minimizing the number of points (increasing the step between the adjacent samples) at which the calibration characteristic is measured. Considering the fundamentally non-linear nature of the dependence of the equivalent inductance of the SCECS (and, accordingly, the MC output signal) on the measured parameter, the possibilities of this method of the reducing the labour intensity of the calibration are limited by potentially large errors in defining the measured parameters. In addition, the significant costs for the calibration device, which ensures the high accuracy of the target positioning relative to the SE of SCECS, as well as for the thermal chamber in which the sensor and the target must be in the conditions identical to the working ones, should be noted.

The alternative approach to the problem of reducing the temperature effect on the measured by SCECS parameters is proposed in [42, 43]. It assumes the rejection from the

experimental acquisition of the families of calibration characteristics and their replacement by families of the conversion functions in the form of the similar dependences of voltages (codes) at the output of the MC with SCECS on measured parameters at different temperatures. The computer models of electromagnetic interaction between the SCECS's SE and the target considering further conversion of the equivalent parameters of the sensor in the MC are used to calculate the conversion functions [45-49]. Of course, the validity of the families of conversion functions, and therefore the value of the final temperature error of the measuring channel with SCECS, is determined by the adequacy of the used models of the sensor and the MC.

### 1.5. Varieties of Single-Coil Eddy Current Sensors

The classifications of SCECS varieties by purpose and by design are presented in Fig. 1.7 a, b. The intersections of these classifications for current variants of the sensors are shown in Table 1.1 in a matrix form. The options that are possible but not implemented are indicated in grey.

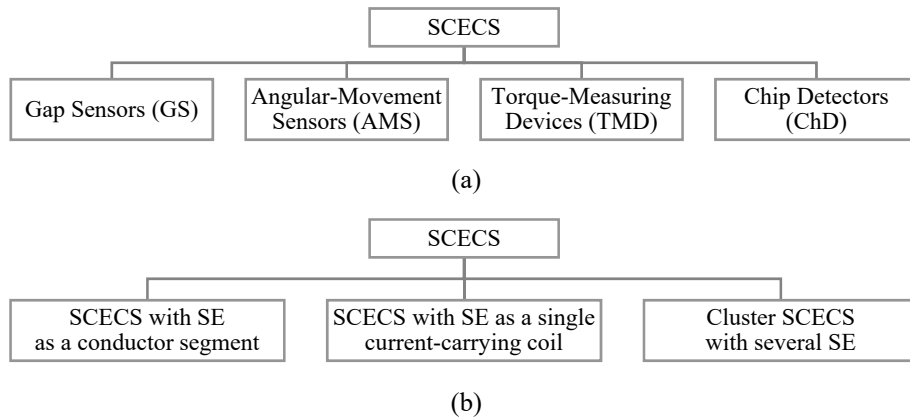


Fig. 1.7. Classification of SCECS varieties by purpose (a), and by design (b).

Table 1.1. Current variants of the SCECS.

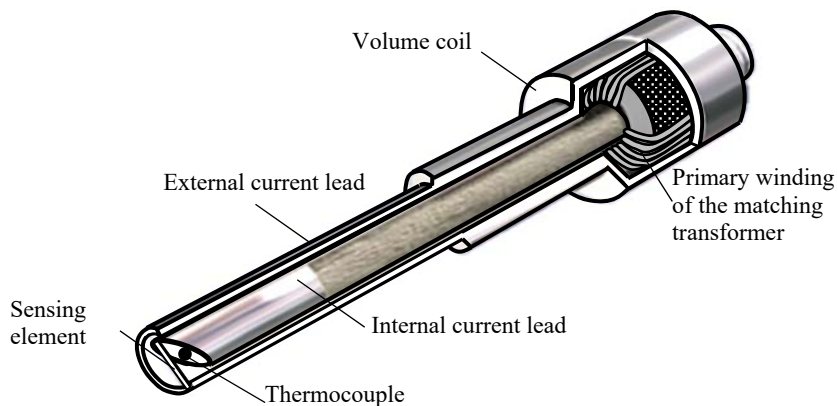
<b>Purpose</b> \ <b>Design</b>	SE as a segment of a linear conductor	SE as a single current-carrying coil	Cluster SCECS with several SE
Gap Sensors (GS)			
Angular-Movement Sensors (AMS)			
Torque-Measuring Devices (TMD)			
Chip Detectors (ChD)			

### 1.5.1. Single-Coil Eddy Current Gap Sensors

#### 1.5.1.1. Single-Coil Eddy Current Gap Sensor with a Sensing Element as a Segment of a Linear Conductor

Historically, the SCECS with SE in the form of a conductor segment were the first among the entire set of the existing single-coil sensors [17]. The main purpose of the sensors is the non-contact measurements of the radial clearances between the inner surface of the stator and the tips of the working blades of the rotor of the compressor or turbine of GTE [16, 23, 32]. Sensors are also used for non-contact monitoring of the displacements of GTE structural elements with a smooth surface such as the shafts in the seals or bearing assemblies [50, 51].

Although there currently exists many varieties of SCECS with SE in the form of a conductor segment, they are all structurally identical. The schematic image of the sensor is shown in Fig. 1.8 [16].



**Fig. 1.8.** Schematic image of a typical SCECS with SE in the form of a conductor segment.

The sensor consists of three main elements: SE in the form of a conductor segment usually with square or rectangular cross-section, "non-inductive" (to a first approximation) current leads and MT. The SE of the sensor is made from heat-resistant steel alloy (for example, BZh98, EI868, XH60BT) [52]) similar to the one used in turbine blades and does not require an insulating coating. Due to the simplest design the SE retains its structural strength at temperatures above 1000 °C and ensures the sensor's efficiency under extreme conditions in the GTE gas-air path.

Directly into the measurement area the SE is placed through the mounting hole by means of elongated (up to 500 mm) current leads, which are usually made as coaxial cylinders. The current leads connect the SE of the sensor with the secondary winding of the MT in the form of a "volume coil". The current leads also provide the MT removal from the zone

with high temperatures to the area with acceptable external conditions, for example, to the outside of the engine's stator shell. If the temperature conditions in the measurement area are acceptable for the magnetic circuit of the MT, then it is no longer necessary to use the elongated current leads. In this case the current lead may be completely absent, and the SE is directly connected to the sensor's body [18, 23, 53].

To minimize the sensor's size its MT windings are placed on the core, which made from the high-frequency ferrite with a high magnetic permeability. Since the loss of the ferromagnetic properties of the ferrite core occurs when it is heated above the Curie point, the ferrite core is the most dependent on the temperature element of the sensor's design. Therefore, if the temperature in the MT's location area is high, then the ferrite magnetic circuits with a high Curie point value (up to +450 °C) or the forced air cooling of the MT are used. The construction of the SCECS with so-called «air» MT without ferrite core is described in [32]. The frame of the primary winding of the MT is made from high-temperature dielectric material. At the same time, the use of «air» MT is accompanied by a loss of the sensor's sensitivity. The compensation of the specified effect requires a greater MT and sensor dimensions.

The current leads and the “volume coil” of the MT as well as the SE are made from heat-resistant stainless steel. The external and internal current leads are separated by a thin layer of ceramic insulating coating. Near the SE in the end part of the internal current lead a hot junction of a thermocouple is placed. The thermocouple monitors the changes in the temperature of the SE and is intended for use in the thermal correction channel of the measuring system. All elements are connected by laser weld. The rigid construction of the SCECS ensures the high stability of the sensor's characteristics during repeated cyclic temperature changes.

The characteristics of some variants of SCECS with SE as a conductor segment are presented in the Table 1.2. The sensors were used in GTE bench tests for radial clearances measuring.

It should be noted that the geometry and the electrical parameters of the sensors are individual and depend on the specific practical task. In particular, the diameter of the end part of the SCECS is determined by the maximum value of the measured radial clearances and usually exceeds it by 3-4 times. The length of the current leads depends on the distance between the inner and outer surfaces of the engine stator. The SCECS is mounted on the GTE by means of a special sleeve, which is rigidly based on the inner surface of the stator and is the engine part. The same sleeve is used in the calibration device for the experimental obtaining of the calibration characteristics of the measuring channels with SCECS. This makes it possible to minimize the difference in conditions during the sensor's calibration and sensor's use in the operating mode on the power plant.

**Table 1.2.** Characteristics of SCECS with SE as a conductor segment for GTE bench tests.

Sensor type	Characteristics
 <p>SCECS-K-3(SS)</p>	<p>Measured displacements range 0...3 mm</p> <p>Operating temperature range -50...+600°C</p> <p>Nominal inductance value 30 uH</p> <p>Inductance deviation 1.8 uH</p> <p>Mounting hole diameter 10 mm</p> <p>Dimensions 10×25×100 mm</p> <p>Body material Stainless steel</p> <p>Application GTE compressor</p>
 <p>SCECS-T-3(SS)</p>	<p>Measured displacements range 0...3 mm</p> <p>Operating temperature range -50...+1100°C</p> <p>Nominal inductance value 30 uH</p> <p>Inductance deviation 2.0 uH</p> <p>Mounting hole diameter 14 mm</p> <p>Dimensions 14×34×115 mm</p> <p>Body material Stainless steel</p> <p>Application GTE turbine</p> <p>*Forced air cooling of the matching transformer is required</p>
 <p>SCECS-DP-15(SS)</p>	<p>Measured displacements range 0...15 mm</p> <p>Operating temperature range -50...+250°C</p> <p>Nominal inductance value 90 uH</p> <p>Inductance deviation 7.0 uH</p> <p>Mounting hole diameter 32 mm</p> <p>Dimensions 54×32 mm</p> <p>Body material Stainless steel</p> <p>Application Ducted propfan</p>
 <p>SCECS-S-3(SS)</p>	<p>Measured displacements range 0...2.5 mm</p> <p>Operating temperature range -50...+650°C</p> <p>Nominal inductance value 40 uH</p> <p>Inductance deviation 5.0 uH</p> <p>Mounting hole diameter 15 mm</p> <p>Dimensions <sup>1</sup>15×34×300 mm <sup>2</sup>15×34×241 mm</p> <p>Body material Stainless steel</p> <p>Application Seals</p> <p>*Forced air cooling of the matching transformer is required</p>
 <p>SCECS-PB-0.2(SS)</p>	<p>Measured displacements range 0...0.2 mm</p> <p>Operating temperature range -50...+240°C</p> <p>Nominal inductance value 7 uH</p> <p>Inductance deviation 0.6 uH</p> <p>Mounting hole diameter 14 mm</p> <p>Dimensions 14×18 mm</p> <p>Body material Stainless steel</p> <p>Application Plane bearings</p>

### 1.5.1.2. Cluster Compositions of Single-Coil Eddy Current Sensors with a Sensing Element as a Segment of a Linear Conductor

As most papers devoted to the problems of radial clearances measuring in power plants [16, 23, 31, 32, 50, 54-56] note, in real conditions the engine's constructive elements perform complex multi-coordinate displacements associated with the factors of different physical nature. For example, the temperature, elastic and plastic deformations of the rotor structure elements (including the deformations of the blades themselves) caused the radial displacements of the blade tips and are one of the main reasons of radial clearances' changing. Thermal elongation of the shaft and the displacement of the rotor wheel in radial thrust bearings due to the aerodynamical loads are the reasons of blade tips displacement in the direction of the GTE axis. Aerodynamic loads also lead to bending of the blade pen, etc.

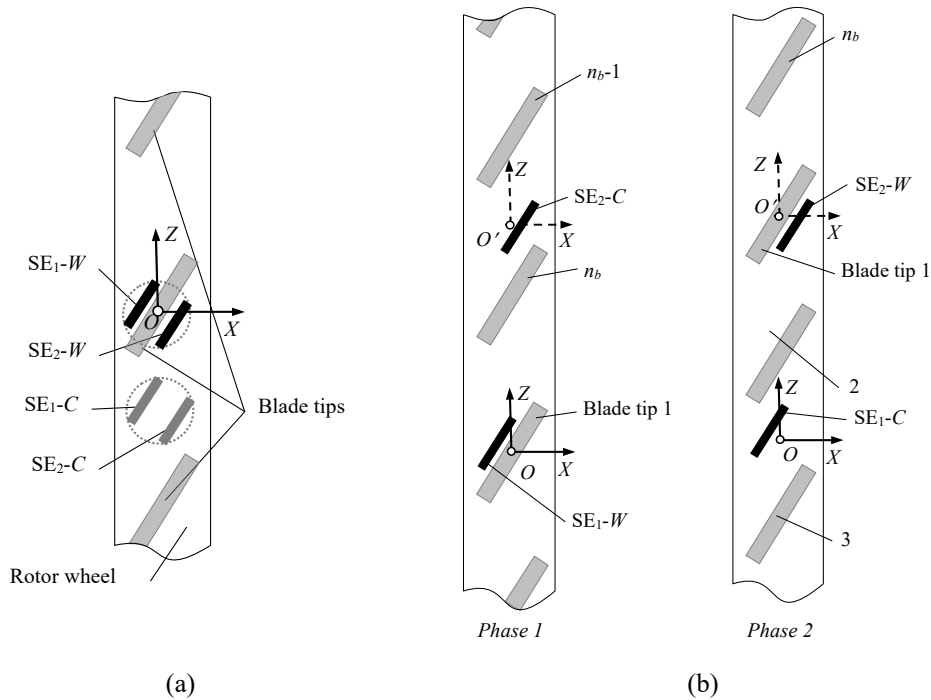
Therefore, it can be argued that in GTE operating mode the displacement of the material point selected at the monitored constructive element has a fundamentally vector character and is determined by several coordinates in the Cartesian reference system  $OXYZ$ , the center of which (point  $O$ ) is rigidly attached to the stator at the SCECS position (to the geometric center (g.c.) of the sensor's SE). These movements have a significant impact on the output signal of the SCECS, which integrally contains information about all coordinates of the displacements of the monitored object. So, on the one hand, the components of target movements in the direction of the axis  $X, Z$  are an interfering factor that makes it difficult to obtain information about the radial clearance (movement in the direction of the axis  $Y$ ). On the other hand, they are a source of important information about the processes occurring in the power plant. For example, the destruction of the radial thrust bearing can be determined on the basis of the monitoring of axial displacements of the blade tips [51]. The monitoring of the bending movements along the  $X$  and  $Z$  axes allows to detect vibrations of the blade tips and diagnose the compressor stall and the pre-surge state of the GTE [57-59].

Measuring the displacements of the engine's constructive elements along several coordinates is provided by using the methods on the basis of the clusters of SCECS (groups of identical sensors) whose SE are oriented in a certain way toward the monitored object and the number of sensors in the cluster corresponds to the number of monitored coordinates. These methods were called "cluster methods" in [31, 32].

Several variants of the cluster methods related to the task of radial clearance measuring in a compressor or turbine of GTE are considered in [16]. One version of the methods (Fig. 1.9a) provides the concentration of the SCECS in close proximity to the observation point (point  $O$  is the origin and the g.c. of the "concentrated cluster" of SCECS). The cluster contains two working SCECS (they are shown in Fig. 1.9a by their sensing elements  $SE_1-W$  and  $SE_2-W$ ) and is intended for measuring the axial ( $x$ ) and radial ( $y$ ) displacements of the blade tips.

An obvious limitation of the "concentrated cluster" of the SCECS is the need to make several mounting holes in the stator shell on a relatively small area. This can cause the loss of stator strength especially in the case of increasing the number of monitored

coordinates and the corresponding number of sensors in the cluster. The number of mounting holes on a stator is doubled due to the need of using additional witness SCECS (in Fig. 1.9a they are also represented by their sensing elements  $SE_1-C$  and  $SE_2-C$ ) to compensate the temperature effect on the sensors.



**Fig. 1.9.** Sensors' locations in “concentrated” (a), and “distributed” (b) clusters of SCECS.

The use of a “distributed cluster” of SCECS allows to eliminate the excessive concentration of SCECS and mounting holes while reducing the number of the sensors. The location of the distributed cluster with two SCECS intended for the measuring the same  $x,y$ -coordinates is presented in Fig. 1.9b (the sensors are shown by their sensing elements  $SE_1$  and  $SE_2$ ).  $SE_1$  is in the same position as in “concentrated cluster”.  $SE_2$  is equidistantly shifted at an angle  $1.5 \cdot \Delta\psi_b$ , where  $\Delta\psi_b$  is the angular pitch of the blades on the impeller (this corresponds to the shift of the g.c. and the reference system to the distance  $OO'$ ). The  $x,y$ -coordinates are measured at two phases. At phase 1 the root of blade 1 passes g.c. (point  $O$ ) and  $SE_1$  performs the working functions ( $SE_1-W$ ) and  $SE_2$  performs the compensation functions ( $SE_2-C$ ). At phase 2 the root of blade 1 passes the point  $O'$  and  $SE_1$  reverses its functions from working to compensation ( $SE_1-C$ ) and  $SE_2$  vice versa – from compensation to working functions ( $SE_2-W$ ).

And when the number of allowed mounting holes in the stator shell is less than the number of monitored coordinates and the corresponding number of SCECS in the “concentrated” or “distributed” cluster, so-called “incomplete cluster” of the SCECS is used. In the “incomplete cluster” “unmeasured” coordinates are calculated using specially developed

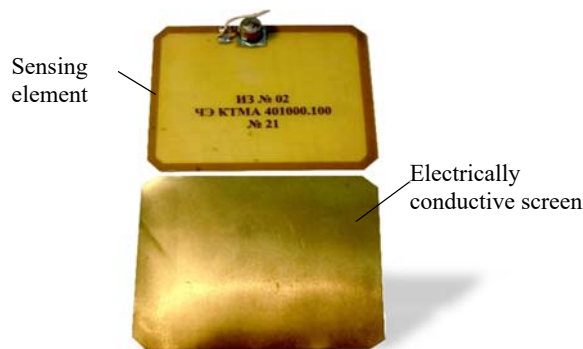
real-time models based on the current parameters of the engine regime and its environment [54, 60]. For example, the use of an “incomplete cluster” of SCECS in combination with on-line modeling of the blade bend in the system for the radial clearances measuring between the propeller blade tips and the stator shell of the ducted propfan made it possible to reduce the number of SCECS to one sensor at each control point on a stator shell [61].

The works [53, 62] contains the descriptions of the prototypes of so-called cluster sensors, which combine several SCECS in a single body and solve the problem of a large number of mounting holes in a limited space on the stator of a power plant. The experiments have confirmed the operability of the “cluster SCECS” and the identity of their characteristics to the existing varieties of SCECS (with a slight decrease in sensitivity) [53]. However, the “cluster SCECS” is significantly larger in size than traditional SCECS. Nevertheless, “cluster SCECS” can be used in power plants where the placement of oversized sensors is not critical.

### 1.5.1.3. Screen-type Single-Coil Eddy Current Sensors with a Sensing Element as a Single Current-carrying Coil

The screen-type SCECS with SE in the form of a single current loop [11] are another variation of SCECS used to measure the clearances. Such sensors are normally used to study the gaps between two adjacent deformable surfaces under bench conditions at normal or close to them temperatures. This circumstance makes it possible to minimize the weight of the SCECS by using foil materials on a rigid or flexible substrate during the sensors’ manufacture. The screen and the sensor’s SE are placed opposite each other on the adjacent surfaces (both electrically conductive and non-electrically conductive) the distance between which is monitored.

The example of such SCECS which is intended to measure the gaps between the photocells of the adjacent flaps of the solar arrays during their vibration tests is shown on Fig. 1.10. The SE of the sensor is made of a thin foil fiberglass (weight is no more than 5 g) and it works together with a copper foil screen (weight is no more than 0.5 g). The range of measured displacements of the screen relative to the SE is 0...20 mm.



**Fig. 1.10.** Screen-type SCECS for measuring the gaps between the photocells of the adjacent flaps of the solar arrays.



### **1.5.2. Single-Coil Eddy Current Angular-Movement Sensors**

The angular-movement sensors (AMS), as well as SCECS for measuring the gaps between adjacent deformable surfaces (Section 1.5.1.3), belong to the group of screen-type SCECS with SE in the form of a single current-carrying coil. The electromagnetic unit of a typical AMS includes a cylindrical frame made of electrically insulating material. The inner surface of the frame contains the cluster of the SE on a film substrate. The non-magnetic metal cylinder with a system of raster windows which overlap the surface of the SE cluster in a sector of 180 degrees is arranged inside the frame with the clearance 0.1...0.2 mm. The inductance value is the primary information parameter of the AMS. It depends on the degree of the overlapping of the raster windows when the cylinders are shifted relative to each other. While changing the number and the relative position of the SE in the cluster we can create various designs of AMS for different limits of the angular movements.

The sensors have one or two shafts which are fixed in the bearing assemblies. The shafts provide the mechanical linkage with a monitored object. The outer and inner cylinders of the sensors intended for measuring of the angular movements of two shafts are connected to the sensor's input and output shafts, respectively. If the sensor is intended for measuring the angular position of one shaft only, then the inner cylinder is connected to the shaft of the sensor and the outer cylinder is locked in the body of the monitored object. The construction of AMS with through axial hole in an internal perforated cylinder makes it possible to place the electromagnetic unit of the sensor directly on the shaft even inside the monitored object. The conversion of the information parameter of the sensor into an electrical signal (analogue or digital) is performed by an external or built-in electronic unit.

The AMS examples for different ranges of an angular displacement are given in Table 1.3.

### **1.5.3. Single-Coil Torque-Measuring Devices**

The single-coil torque-measuring devices (TMD) are based on the principle of the intermediate conversion of the torque into the angular movements of the resilient element that transmits this moment from an energy source to a consumer. Further the angular movements are converted into information signal using small angular motion sensors, e.g., AMS-5 (Table 1.3).

Two types of elastic elements are usually used in TMD. The first one is a torsion shaft, and the other is a disk-shaped elastic element with milled or forged "U"-shaped springs. The elastic element is connected to the input and output shafts of the sensor, which are locked in reinforced bearing assemblies capable to receive increased bending moments. This makes it possible to transmit the torque without installing the sensor's body on a support. The TMD fixation in the shaft's rotation direction is provided by a floating lever. If the task is to measure the torque of the shaft, which is a part of the engine's unit, the electromagnetic assembly of the sensor is installed directly on the monitored shaft. In this case the shaft performs the function of an elastic element. The conversion of the sensor's

information parameter into an electrical signal (analogue or digital) is performed by the remote electronic unit.

**Table 1.3.** Single-coil AMS.

Sensor type	Characteristics
 <p>AMS-5</p>	<p>Measured angular movements range <math>\pm 5</math> degrees</p> <p>Output voltage range 0...5 V</p> <p>Maximum fiducial error 0.5 %</p> <p>Dimensions 28×57 mm</p> <p>Mass 150 g</p> <p>Electronic unit External</p> <p>*Is used for embedding into the mechanisms for determining the relative angular movements of two rotating shafts</p>
 <p>AMS-60</p>	<p>Measured angular movements range 0...60 degrees</p> <p>Output voltage range 0...5 V</p> <p>Maximum fiducial error 0.5 %</p> <p>Dimensions 100×50 mm</p> <p>Mass 300 g</p> <p>Electronic unit Built-in</p>
 <p>AMS-90</p>	<p>Measured angular movements range 0...90 degrees</p> <p>Output voltage range 0...5 V</p> <p>Maximum fiducial error 1 %</p> <p>Dimensions 56×4×20 mm</p> <p>Mass 50 g</p> <p>Electronic unit Built-in</p> <p>*Is used for replacement of the potentiometric sensor of the angular position of the throttle valve in the internal combustion engine</p>
 <p>AMS-900</p>	<p>Measured angular movements range 0...900 degrees</p> <p>Output voltage range 0...5 V</p> <p>Maximum fiducial error 0.1 %</p> <p>Dimensions 30×60 mm</p> <p>Mass 250 g</p> <p>Electronic unit External</p> <p>* Is used for identification of the angular position of the multi-turn shaft. The sensor contains a wave gear from the input shaft to the internal cylinder with the raster windows</p>
 <p>AMS-120(AR)</p>	<p>Measured angular movements range 0...120 degrees</p> <p>Output voltage range 0...5 V</p> <p>Maximum fiducial error 0.1 %</p> <p>Dimensions 30×60 mm</p> <p>Mass 1200 g</p> <p>Electronic unit Built-in</p> <p>* High reliability sensor. Probability of survival is 0.997 (T=5 years).</p> <p>**Is used for embedding into the mechanisms for determining the relative angular movements of two rotating shafts</p>

The TMD (as the AMS) use the raster window system, which consists of the outer and the inner hollow cylindrical screens. Each screen is mechanically connected to one of the ends of the elastic element. The screens have one or two rows of perforated windows. The rows of the windows of the internal screen are in the same cross-sections as the rows of the windows of the external screen. Each row is covered by a corresponding SE on a stationary cylindrical framework. If the torque is absent, then the windows of the interior screen are covered by half of the windows of the external screen. The shaft's torsion changes the relative angular position of the screens. As the result, the area of the window overlapping and, therefore, the SE inductances are increased or decreased.

The examples of the most typical TMD are given in the Table 1.4.

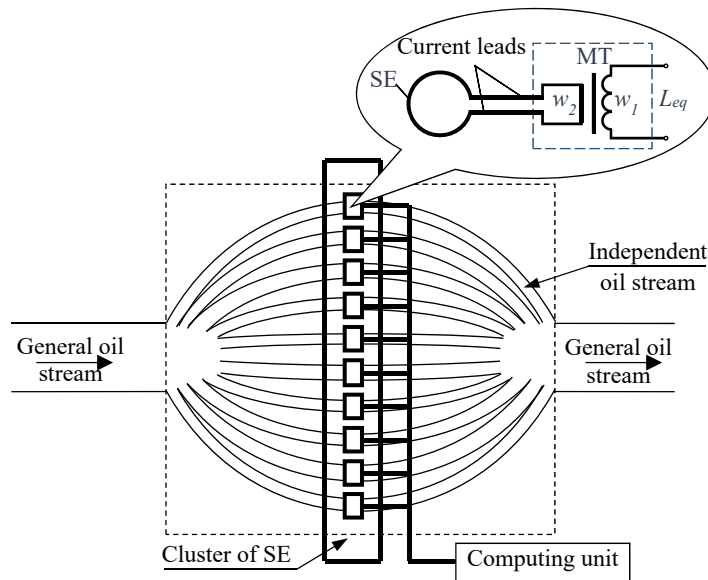
**Table 1.4.** Single-coil TMD.

Sensor type	Characteristics
 <p>TMD-500</p>	<p>Measured torque force range <math>\pm 500</math> Nm            Permitted rotation speed of the input shaft 5000 rpm            Output voltage range 0...5 V            Maximum fiducial error 0.5 %            Dimensions 120×120 mm            Mass 2.5 kg</p> <p>*Connection to the monitored object is done through the centering disks with 6 holes 10 mm in diameter.            **Is used for embedding into the mechanisms for determining the torque within one revolution of the shaft.</p>
 <p>TMD-200</p>	<p>Measured torque force range <math>\pm 200</math> Nm            Permitted rotation speed of the input shaft 5000 rpm            Output voltage range 0...5 V            Maximum fiducial error 0.5 %            Dimensions 220×160 mm            Mass 7.5 kg</p> <p>*Contains elastic element with 8 forged "U"-shaped springs.</p>
 <p>TMD-100</p>	<p>Measured torque force range <math>\pm 100</math> Nm            Permitted rotation speed of the input shaft 1000 rpm            Output voltage range 0...5 V            Maximum fiducial error 2 %            Dimensions 50×220 mm            Mass 2.6 kg</p> <p>*Contains disk-shaped elastic element with 8 forged "U"-shaped springs.            **Is used for embedding into the mechanisms for determining the torque within the limitations on the object's length.</p>
 <p>TMD-90</p>	<p>Measured torque force range <math>\pm 90</math> Nm            Permitted rotation speed of the input shaft 5000 rpm            Output voltage range 0...5 V            Maximum fiducial error 0.5 %            Dimensions 155×84 mm            Mass 1.5 kg</p> <p>* Connection to the monitored object is done through the centering disks with 6 holes 8 mm in diameter.</p>

### 1.5.4. Cluster Single-Coil Eddy Current Sensor for Debris Monitoring

It is noted in [16] that ECP are the most popular sensors that are used in on-board systems for the diagnosis of the state of the bearing mount assemblies of the power plants on the base of the detecting of metal wear particles in the engines' lubrication systems. The approach for construction of a cluster single-soil eddy current sensor for debris monitoring in the GTE oil system [63, 64] is also considered in [16]. The sensor is invariant to the diameter of the oil pipeline and enables to determine the type of the metal particles (ferromagnetic or non-ferromagnetic), to estimate their sizes, and to identify if several particles are running in the same cross section of the pipeline of the power plant's lubricant system.

The essence of the proposed solution is that the general oil stream on the input of the sensor is split into  $N$  independent streams with a smaller cross-section area (Fig. 1.11) [65]. Each independent stream is covered by a single current coil located in the pipeline of the engine oil system. The proposed stream separation method, in addition to increasing the sensitivity, also increases the probability of detecting the single metal particles running in the general stream of the pipeline of the power plant's oil system.



**Fig. 1.11.** Splitting of the general oil stream into  $N$  independent streams and placing the single-coil eddy current SE in them.

To convert the changes of each SE's inductance a differential MC is used. It provides a zero level of the output signal in the absence of a metal particle in the oil channel of the sensor. For this, the second similar SE additionally installed in the cross section of each independent oil channel. The additional SE ( $SE_2$ ) is shifted from the first one ( $SE_1$ ) with the respect to the axis of the channel by a specified distance  $h$ . The sequential movement of the metal particle first through the  $SE_1$  contour and then through the  $SE_2$  contour, leads

to the formation of two consecutive voltage pulses of the opposite polarity at the MC output. The pulse amplitude is determined by the size of the metal particle and its location in the plane of the SE contour, and the polarity is determined by the magnetic properties of the particle metal. For example, the voltage pulse of the positive polarity is formed at the MC output when a non-ferromagnetic particle passes through the SE<sub>1</sub> contour and, on the contrary, the polarity of the pulse at the MC output changes from positive to negative when a non-magnetic metal particle passes through the SE<sub>2</sub> contour. A ferromagnetic metal particle also changes the polarity of the voltage pulses at the MC output. But unlike the case with a non-ferromagnetic metal particle, the pulse polarity changes from negative to positive when the magnetic particle passes through the SE<sub>1</sub> and SE<sub>2</sub> contours. This allows to identify the material of a particle by its magnetic properties.

Setting the two SCECS in the same oil channel at a fixed distance from each other allows also to evaluate the oil flow rate in the oil system of the power plant according to the speed of the metal particle in the sensor's oil channel. For these purposes the time moments of the appearance of the amplitude values of the MC output voltages are additionally fixed. They correspond to the moments of the sequential passage of the metal particles through the SE<sub>1</sub> and SE<sub>2</sub> contours. The oil flow rate is determined by the calculated time interval  $\Delta t$  and the distance  $h$  between the SE<sub>1</sub> and SE<sub>2</sub> [66].

It should be also noted that the use of two SE and the differential MC allows for partial to compensate the influence of some interfering factors (primarily temperature) on the process of transforming the informative parameter. The existence of two consecutive pulses of the opposite polarity when the particles pass through the oil channel of the sensor is also a good diagnostic property that makes it possible to uniquely identify a signal at the SCECS output as a signal from a metal particle. It allows to reduce the number of false alarms (type I errors) of the measuring system with such sensors. In this case the metal particle is fixed only when two pulses of the opposite polarity occur consistently at the MC output.

As for the determination of particle sizes, their calculation is carried out on the basis of the calibration characteristic in the form of the dependence of the MC output signal on the position of the particle in the SE contour and the size of the particle. The position of a metal particle in the SE contour during the measurements can be determined through the average oil flow rate, the dimensions of the oil system pipeline and the distribution of elementary flow velocities over the pipeline cross section [67]. It should be noted that in most practical tasks it is enough to estimate the size of the wear particle by assigning it to one of the fixed dimension groups to determine the defect nature in the bearing unit.

## **1.6. Conclusions**

Single-oil eddy current sensors have almost 40-years history of their existence. The simplicity of the sensors' design ensures their operation in difficult and even extreme conditions at high temperatures, a significant level of vibrations, pollution in the measurement area, etc. The sensors find their main application primarily in the experimental research of high-power plants (first GTE), their components and assemblies.

SCECS and their cluster compositions were used for measuring the radial clearances in compressors and turbines of full-size GTE, angular movements and torques of the engines' shafts, for the diagnosis of the state of the power plants' friction pairs, etc. During such studies the high performance and reliability of the sensors were confirmed. The creating of new varieties of SCECS and measuring tools based on them continues at present time concerning modern tasks. Research is mainly aimed at improving the measurement accuracy, reducing the number of the sensors placed on the monitored object, expanding sensors' functionality, including intelligent processing of the measurement results, etc.

## References

- [1]. J. García-Martín, J. Gómez-Gil, E. Vázquez-Sánchez, Non-Destructive Techniques Based on Eddy Current Testing, *Sensors*, Vol. 11, Issue 3, 2011, pp. 2525-2565.
- [2]. M. Wright, Eddy Current Testing Technology, *Eclipse Scientific Incorporated*, 2015.
- [3]. Applications with eddy current probes (<https://www.eddylab.com/applications/Eddy-Current-Probes>).
- [4]. Using Eddy-Current Sensors for Thread Detection (<https://www.lionprecision.com/using-eddy-current-sensors-for-thread-detection/>)
- [5]. Eddy current sensor: Operating principle and applications (<https://www.bestech.com.au/blogs/eddy-current-sensor-principle/>).
- [6]. Eddy Current displacement, distance and position sensors (<https://www.micro-epsilon.co.uk/displacement-position-sensors/eddy-current-sensor/>).
- [7]. KAMAN Measuring and Memory Systems ([https://www.kamansensors.com/wp-content/uploads/2020/01/Kaman\\_Applications\\_Handbook\\_WEB.pdf](https://www.kamansensors.com/wp-content/uploads/2020/01/Kaman_Applications_Handbook_WEB.pdf)).
- [8]. E. Doebelin, Measurement Systems: Applications and Design – 4<sup>th</sup> ed., *McGraw-Hill Publishing Company*, 1990.
- [9]. Eddy Current Probes (<https://www.eddycurrentprobe.com/>).
- [10]. KAMAN Precision Products/Measuring. Sensor Family Data Sheet: Advanced family of high-precision position sensors using Eddy Current Technology ([https://www.kamansensors.com/pdf\\_files/Kaman\\_sensor\\_family\\_data\\_sheet\\_web.pdf](https://www.kamansensors.com/pdf_files/Kaman_sensor_family_data_sheet_web.pdf)).
- [11]. S. Borovik, Yu. Sekisov, Single-coil eddy-current transducers for measuring mechanical parameters, *Vestnik of Samara State Aerospace University*, Vol. 42, Issue 4, 2013, pp. 94-102.
- [12]. Puftechnik: NDT Accessories | Product Catalog ([https://www.puftechnik.com/fileadmin/Products-Services/Products/NDT/NDT-Accessories/Encoders-Paint-markers-Warning-units-Test-defect/Downloads/PRUEFTECHNIK\\_NDT\\_Accessories\\_Catalog\\_EN-DE.pdf](https://www.puftechnik.com/fileadmin/Products-Services/Products/NDT/NDT-Accessories/Encoders-Paint-markers-Warning-units-Test-defect/Downloads/PRUEFTECHNIK_NDT_Accessories_Catalog_EN-DE.pdf)).
- [13]. Z. Zhao, Z. Liu, Y. Lyu, Y. Gao, Experimental Investigation of High Temperature-Resistant Inductive Sensor for Blade Tip Clearance Measurement, *Sensors*, Vol. 19, Issue 1, 2019, 61.
- [14]. R. Przysowa, E. Rokicki, Inductive Sensors for Blade Tip-Timing in Gas Turbines, *Journal of Konbin*, Vol. 36, Issue 4, 2015, pp. 147-164.
- [15]. V. Sridhar, K. Chana, M. Pekris, High temperature eddy current sensor system for turbine blade tip clearance measurements, in *Proceedings of the 12<sup>th</sup> European Conference on Turbomachinery Fluid dynamics & Thermodynamics (ETC'12)*, Stockholm, Sweden, 3-7 April 2017, paper ETC2017-217.
- [16]. S. Borovik, Y. Sekisov, Single-Coil Eddy Current Sensors and Their Application for Monitoring the Dangerous States of Gas-Turbine Engines, *Sensors*, Vol. 20, Issue 7, 2020, p. 2107.

- [17]. O. P. Skobelev, Yu. N. Sekisov, A. A. Khritin, High-temperature conductor eddy-current converter, *Patent SU 1394912 A1*, 1986.
- [18]. B. K. Raykov, Yu. N. Sekisov, O. P. Skobelev, et al., Clearance eddy current probes with sensitive elements in the form of a conductor segment, *Devices and Control Systems*, Issue 8, 1996, pp. 27-30.
- [19]. N. D. Kuznetsov, V. P. Danilchenko, V. E. Reznik, Control of radial clearances in turbocompressors of aviation gas-turbine engines, *Samara Aviation Inst.*, 1991.
- [20]. V. G. Gerasimov, V. V. Klyuev, V. E. Shaternikov, Methods and devices for electromagnetic control, ed. V. E. Shaternikov, *Spectr*, 2010.
- [21]. G. I. Peredelskiy, Pulse-powered bridge circuits, *Energoatomizdat*, 1988.
- [22]. O. P. Skobelev, Conversion methods based on test transients, *Measurement, Control, Automation*, Issue 1, 1980, pp. 11-17.
- [23]. Yu. N. Sekisov, O. P. Skobelev, L. B. Belenki, et al., Displacements of Structural Components of Power Plants, *SamNTs RAN*, 2001.
- [24]. M. M. Kuteynikova, B. K. Raykov, O. P. Skobelev, Structural varieties of high-temperature single-coil eddy current sensors, in *Proceedings of the XIV International Conference 'Complex systems: control and modelling problems'*, Samara, Russian Federation, 19-22 June 2012, pp. 595-601.
- [25]. A. Blinov, S. Borovik, M. Luchsheva, F. Muhutdinov, Yu. Sekisov, Monitoring the state of power plants' friction pairs on the basis of single-coil eddy-current sensors, *J. Phys.: Conf. Ser.*, Vol. 1891, 2021, 012053.
- [26]. V. P. Eremin, N. V. Eremin, A. N. Kirilin, et al., Creation of a new generation of electromechanical drives for transformable spacecraft systems, *JSC SRC Progress*, 2011.
- [27]. V. A. Taft, Electrical circuits with variable parameters, *Energiya*, 1968.
- [28]. M. I. Kontorovich Operational calculus and processes in electrical circuits, *Sovetskoe Radio*, 1975.
- [29]. Electrical Engineering Dictionary, ed. Ph. A. Laplante, *CRC Press LLC*, 2000.
- [30]. G. P. Nubert, Measuring converters of non-electric quantities. Introduction to the theory calculation and design, Translation M.M. Fedorov, *Energiya*, 1970.
- [31]. L. B. Belenki, S. Y. Borovik, B. K. Raykov, et al., Cluster Methods and Tools for Measuring Stator Deformations and Displacement Coordinates of Blade Tips and Blades in Gas Turbine Engines, Ed. O.P. Skobelev, *Mashinostroenie*, 2011.
- [32]. V. N. Belopukhov, S. Y. Borovik, M. M. Kuteynikova, et al., Cluster Methods and Tools for Measuring Radial Clearances in Turbine Flow Section, *Innovatsionnoe Mashinostroenie*, 2018.
- [33]. A. S. Morris, R. Langari, Measurement and Instrumentation: Theory and Application, Third Edition, *Academic Press*, 2021.
- [34]. L. Finkelstein, Intelligent and knowledge-based instrumentation. An examination of basic concepts, *Measurement*, Vol. 14, Issue 1, 1994, pp. 23-29.
- [35]. A. J. Peyton, Y. Walsh, Analog electronics with Op Amps: a source book of practical circuits, *Cambridge University Press*, 1993.
- [36]. L. B. Belenki, O. P. Skobelev, An electronic analogue of the measuring circuit in the form of Blumlein bridge with single-coil eddy current sensors, in *Proceedings of the XIII International Conference 'Complex Systems: Control and Modelling Problems'*, Samara, Russian Federation, 15-17 June 2011, pp. 300-304.
- [37]. B. Carter, T.R. Brown, Handbook of operational amplifier applications, *Texas Instruments Incorporated*, 2001.
- [38]. L. B. Belenki, O. P. Skobelev, Measuring circuit with single-coil eddy current sensors and approximate differentiation, in *Proceedings of the XIV International Conference 'Complex Systems: Control and Modelling Problems'*, Samara, Russian Federation, 19-22 June 2012, pp. 602-606.

- [39]. L. B. Belenki, M. M. Kuteynikova, A. V. Logvinov, et al., Device for measuring of multi-coordinate displacements of the blades' tips, *Patent RF 2525614*, 2014.
- [40]. C. L. King, Differential ADC Biasing Techniques, Tips and Tricks, <https://www.microchip.com/content/dam/mchp/documents/MCU08/ApplicationNotes/ApplicationNotes/00842a.pdf>.
- [41]. ADS1210, ADS1211 - 24-Bit Analog-to-Digital Converter, <http://www.gaw.ru/pdf/TI/adc/ads1210-11.pdf>
- [42]. S. Y. Borovik, M. M. Kuteynikova, Y. N. Sekisov, et al., Analysis of temperature influence on the informative parameters of single-coil eddy current sensors, *Optoelectron. Instrument. Proc.*, Vol. 53, Issue 4, 2017, pp. 395–401.
- [43]. S. Y. Borovik, M. M. Kuteynikova, Y. N. Sekisov, et al., Influence of Turbine Wheelspace Temperature on Measurements of Radial and Axial Displacements of Blade Tips, *Optoelectron. Instrument. Proc.*, Vol. 54, Issue 1, 2018, pp. 105–111.
- [44]. P. E. Podlypnov, Temperature impact on the controlled and adjacent blades of the compressor impeller 550 when measuring the radial clearances with self-compensation of temperature effects on the sensor, *Vestnik SamGTU. Ser. 'Tekhnicheskie Nauki'*, Vol. 59, Issue 3, 2018, pp. 106-117.
- [45]. S. Y. Borovik, Y. V. Marinina, Y. N. Sekisov, Model of a cluster single-coil eddy current sensor based on the finite element method, *Vestnik SamGTU. Ser. 'Tekhnicheskie Nauki'*, Volume 19, Issue 1, 2007, pp. 76–83.
- [46]. M. M. Kuteynikova, Yu. N. Sekisov, Models of electromagnetic interaction of sensitive elements of single-coil eddy current sensors and blades. Status and development, in *Proceedings of the XIV International Conference 'Complex Systems: Control and Modelling Problems'*, Samara, Russian Federation, 19-22 June 2012, pp. 589-594.
- [47]. M. M. Kuteynikova, Yu. N. Sekisov, O. P. Skobelev, Models of electromagnetic interaction of sensitive elements of single-coil eddy current sensors and the blade with complex shape, in *Proceedings of the XV International Conference 'Complex Systems: Control and Modelling Problems'*, Samara, Russian Federation, 25-28 June 2013, pp. 627-635.
- [48]. S. Y. Borovik, M. M. Kuteynikova, Y. N. Sekisov, et al., Model of the measuring circuit with variable in time inductances of single-coil eddy current sensors, in *Proceedings of the XVI International Conference 'Complex Systems: Control and Modelling Problems'*, Samara, Russian Federation, 30 June – 03 July 2014, pp. 687-691.
- [49]. S. Yu. Borovik, M. M. Kuteynikova, P. E. Podlipnov, et al., Modeling the process of measuring radial and axial displacements of complex-shaped blade tips, *Optoelectron. Instrument. Proc.*, Vol. 51, Issue 5, 2015, pp. 512–522.
- [50]. S. Yu. Borovik, Measuring of power-plant structural units displacement in underdetermined conditions, in *Proceedings of the II IASTED International Multi-Conference on Automation, Control and Applications*, Novosibirsk, Russian Federation, 20-24 June 2005, pp. 7-12.
- [51]. V. Belosludtsev, S. Borovik, V. Danilchenko, Y. Sekisov, Wear Diagnostics of the Thrust Bearing of NK-33 Turbo-Pump Unit on the Basis of Single-Coil Eddy Current Sensors, *Sensors*, Vol. 21, Issue 10, 2021, p. 3463.
- [52]. A. S. Zubchenko, M. M. Koloskov, Yu. V. Kashirskiy, Steels and alloys grade guide, *Innovatsionnoe Mashinostroenie*, 2021.
- [53]. L. B. Belenki, B. K. Raykov, Yu. N. Sekisov, O. P. Skobelev, Single-coil eddy current sensors: from cluster compositions to cluster constructions, in *Proceedings of the VI International conference 'Complex Systems: Control and Modelling Problems'*, Samara, Russian Federation, 14-17 June 2004, pp. 437-443.
- [54]. S. Yu. Borovik, Yu. N. Sekisov, O. P. Skobelev, Generalized representation of the methods for obtaining measurement information about the coordinates of the displacements of the blades and propellers tips, *Mekhatronika, Avtomatizatsiya, Upravlenie*, Issue 3, 2007, pp. 19-24.



- [55]. L. B. Belenki, S. Yu. Borovik, A. V. Logvinov, Raykov B. K., et al., Methods for measuring of blade faces displacements in compressors and turbines on basis of distributed clusters of sensors. Part 1. Measurement techniques: reasoning and description, *Mekhatronika, Avtomatizatsiya, Upravlenie*, Issue 4, 2009, pp. 16-19.
- [56]. S. Y. Borovik, M. M. Kuteynikova, B. K. Raykov, et al., Method for measuring radial and axial displacements of complex-shaped blade tips, *Optoelectron. Instrument. Proc.*, Vol. 51, Issue 3, 2015, pp. 302–309.
- [57]. S. Yu Borovik, V. M. Belkin, Mescheryakov, A. A., et al., Method for radial clearances measuring and detecting the blade vibrations of a turbomachine rotor, *Patent RF 2258902*, 2005.
- [58]. S. Yu Borovik, B. K. Raykov, Yu. N. Sekisov, et al., Method for measuring the multidimensional displacements and detecting the vibrations of the blade tips of a turbomachine rotor, *Patent RF 2272990*, 2006.
- [59]. S. Yu Borovik, B. K. Raykov, Yu. N. Sekisov, et al., Method for surge detecting and estimating parameters of surge oscillations in compressors of gas turbine plants, *Patent RF 2273831*, 2006.
- [60]. S. Yu Borovik, B. K. Raykov, Yu. N. Sekisov, et al., Method for measuring the radial displacement of the fan blades using channels of physical and virtual correction, in *Proceedings of the V International conference 'Complex Systems: Control and Modelling Problems'*, Samara, Russian Federation, 17-21 June 2003, pp. 512-520.
- [61]. S. Yu. Borovik, S. M. Ignachkov, S. A. Il'inskii, et al., A system of radial clearance measurements in the ducted propfan installation, *Izvestiya Vysshikh Uchebnykh Zavedenij. Aviatsionnaya Tekhnika*, Issue 3, 2004, pp.77-79.
- [62]. S. Yu. Borovik, A. V. Logvinov, B. K. Raykov, O. P. Skobelev, Cluster single-coil eddy current sensors with built-in converter, in *Proceedings of the XIII International Conference 'Complex Systems: Control and Modelling Problems'*, Samara, Russian Federation, 15-17 June 2011, pp. 305-310.
- [63]. V. A. Beosludtsev, S. Yu. Borovik, Yu. N. Sekisov, et al., Wear diagnostics of aircraft bearings on the base of single-coil eddy-current sensors, *Gazoturbinnye Tekhnologii*, Issue 6, 2015, pp. 26-31.
- [64]. S. Yu. Borovik, Yu. N. Sekisov, A. V. Blinov, et al., Transformation of the information in the monitoring system of wear-and-tear particles of friction pairs on the basis of the group of single-coil eddy-current sensitive elements, *Turbiny i Dieseli*, Vol. 73, Issue 4, 2017, pp. 10-17.
- [65]. S. Yu Borovik, I. G. Korshykov, Yu. N. Sekisov, et al., Method for detecting metal wear particles in the oil stream of a running gas turbine engine, *Patent RF 2646520*, 2018.
- [66]. S. Yu. Borovik, I. G. Korshykov, Yu. N. Sekisov, et al., Method for detecting metal particles in the oil of the friction unit lubrication system and determining the oil flow, *Patent RF 2668513*, 2018.
- [67]. S. Yu. Borovik, I. G. Korshykov, V. A. Belosludtsev, et al., Method for detecting metal particles in the lubrication system of friction units of power plants with grouping of the particles by size, *Patent RF 2674577*, 2018.

# **Chapter 2**

# **Optimum Design, Measurement Accuracy and Calibration of Optical Fiber Sensors Based on Strain Transfer Theory**

**Hua-Ping Wang**

## **2.1. Introduction**

Optical fiber sensors are the most promising technique in monitoring physical and chemical variables of civil structures. For the brittle material characteristics, bare optical fiber is prone to breakage under the shear or torsional action existed in the construction and operation. To guarantee the survival and long-term service of the sensors, the packaging measure is particularly significant. This treatment generates an intermedium layer between the sensing fiber and the monitored structure, which leads to the strain of the host material not entirely transferred to the sensing fiber for a portion of strain loss in the transferring path. To correct the error and improve the measurement accuracy, strain transfer theory is developed to establish the quantitative strain relationship between the sensing fiber and the host material. Considerable attempts have contributed to exploring the strain transfer mechanism of optical fiber sensors. A state-of-the-art review on strain transfer theory of optical fiber-based sensors developed for civil structures is addressed [1]. It aims to demonstrate the advance, the application and the challenge of strain transfer theory and provide scientific guidance for the better understanding of the multi-layered sensing model. Besides, the study also provides theoretical instruction for the optimum design, measurement accuracy, calibration and durability improvement of optical fiber sensors.

## **2.2. Problems of Optical Fiber-based Sensors for the Monitoring of Civil Structures**

The sensing fiber used in civil engineering usually constitutes of fiber core, cladding and coating. The configuration and physical photo of optical fiber is displayed in Fig. 2.1. The

materials of fiber core and cladding are both silicon dioxides, which constitutes the effective sensing fiber. The coating used for protecting the sensing fiber can be made of acrylate polymer. The interfacial slippage between the fiber core, cladding and coating can lead to the fiber broken or measurement failure, which is not desired in testing. It should be noted that the coating may induce measurement error for the strain transfer effect. However, compared with the practical engineering scale, the thickness of the coating ( $62.5\ \mu\text{m}$ ) is extremely small and the correlated measurement error can be neglected. Therefore, the optical fiber is regarded as one layer to jointly bear the deformation induced by external action, with radius equal to  $125\ \mu\text{m}$ .

When bare optical fiber is adopted to measure physical parameters of a structure, the packaging measure is essential to configure the functional sensors and guarantee the effective monitoring. A general sensing system for structural health monitoring is presented in Fig. 2.2.

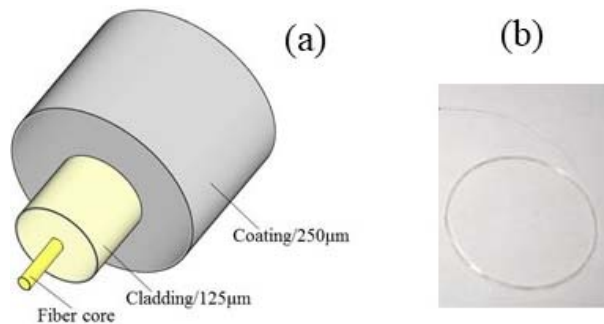


Fig. 2.1. Optical fiber: (a) Structure; (b) Physical photo.

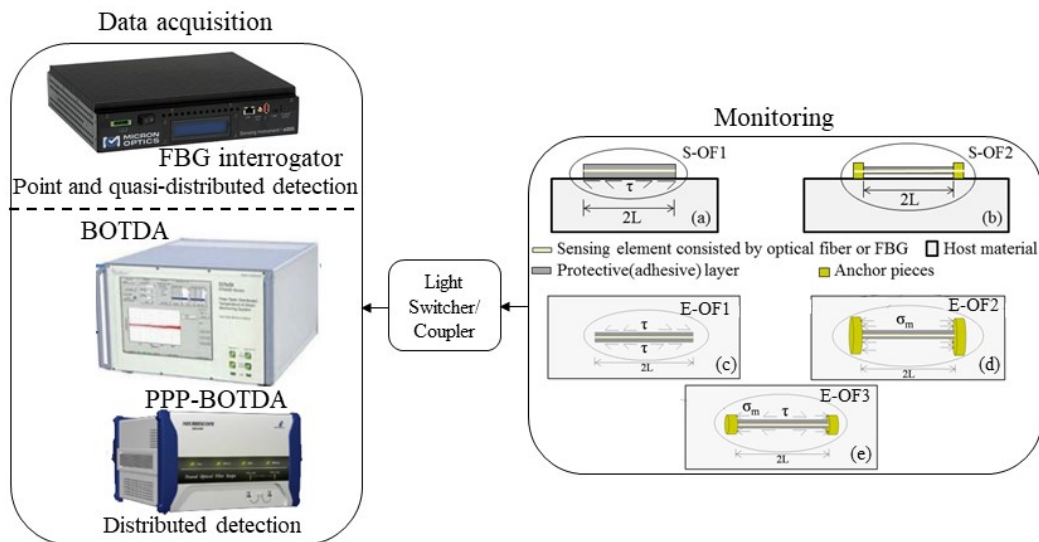


Fig. 2.2. The sensing system for structural health monitoring.

The sensing system mainly contains the monitoring part and the data acquisition part. The five types of optical fiber based-sensors most universally used in engineering are illustrated. MOI Si255 is used for collecting the dynamic and static information of fiber Bragg grating (FBG) sensors. Brillouin Optical Time Domain Analysis (BOTDA) and Pulse Pre-Pump (PPP) -BOTDA are adopted for interpreting the strain and temperature information of Brillouin optical fiber separately with long span and high spatial resolution. It should be noted that the monitoring part is very critical, which determines the effective and reliable information collection of the structures. In this process, the correct design, calibration and installation of sensors are the most significant procedures.

### **2.2.1. Measurement Limitations of Optical Fibers**

Fundamental tests indicate that the strain measurement range of a bare optical fiber-based sensor is about 4000  $\mu\epsilon$ . In most cases, this value cannot meet the requirements of practical projects. Therefore, the packaging measure is deployed to decrease the strain sensitivity of the sensors and then indirectly enlarge the measurement amplitude. Referring to the temperature measurement, special packaging materials can be coated on the bare sensing fiber to increase the temperature sensitivity and measurement stability.

For the point measurement of tensile and compressive strains, displacements, pressures and accelerators, the packaging measure can also be used to establish the quantitative relationship between the target parameters and the wavelength change of FBG (or Brillouin frequency shift of optical fiber). FBGs in series and distributed optical fiber can be packaged with proper materials to realize the local and global measurement of strain, displacement and settlement profiles. For the dynamic distributed strain monitoring of structures (i.e., subsea risers and flowlines, pipelines, bridges, railways and high-rise towers), Brillouin optical fiber sensor is limited to static condition for the essential long acquisition time due to the frequency sweep technique, and FBGs in series can be configured to establish the long-gauge sensors for quasi-distributed detection.

It should be noted that the sensors suitable for practical engineering should be robust, rugged and easy to installation, and cause no negative effect on the integrity, durability and reliability of the monitored structures. Besides, due to the large-scale application, the cost of the sensors should be economical. For the different measurement requirements of civil structures, the optical fiber-based sensors can be divided into three kinds: point, quasi-distributed and distributed sensors. Generally, the FBG sensors can be used for point measurement and FBGs in series can be integrated for quasi-distributed detection. Brillouin optical fiber sensors can be designed for quasi-distributed and distributed detection. For the different installation modes, optical fiber-based sensors can be classified into two categories: surface-attached sensors and embedded sensors. Normally, the material properties, constraints and loading types of a structure and the target parameters to be measured determine the installation mode of a sensor. For example, surface-attached sensors are required for steel structures, and embedded sensors are employed for pavements. According to the different paths for transferring the deformation of host material to the sensing fiber, optical fiber-based sensors can be majorly divided into three types: (1) the first type is that deformation of host material is transferred to the sensing

fiber by the interfacial shear force (i.e., S-OF1 and E-OF1 in Fig. 2.2); (2) the second type is that deformation of host material is transferred by the two anchor pieces at the two ends of the sensing fiber (i.e., S-OF2 and E-OF2); (3) the third type is that the transferring of deformation depends on both the interfaces and two anchor pieces (i.e., E-OF3).

In general, the limitations induced by sensing and material properties of optical fibers can be remedied by specific packaging measure and optical fiber-based sensors can be configured to satisfy the various measurement requirements in civil engineering. However, the packaging technique can introduce strain transfer error. For some kinds of sensors (i.e., S-OF2), this error can be corrected by traditional calibration tests. For some sensors, especially for embedded cases (i.e., E-OF2 and E-OF3), traditional calibration test cannot accurately describe the true relationship between the sensor and the monitored structure. Improved calibration test or strain transfer analysis has been proposed to improve the accuracy of the calibration coefficient [2]. Traditional calibration test just involves the sensor, while the improved calibration test covers the sensor and the host material (a portion of the monitored structure) [3, 4]. The major difference is that the latter considers the influence of the interaction between the host material and the sensor on the measurement accuracy.

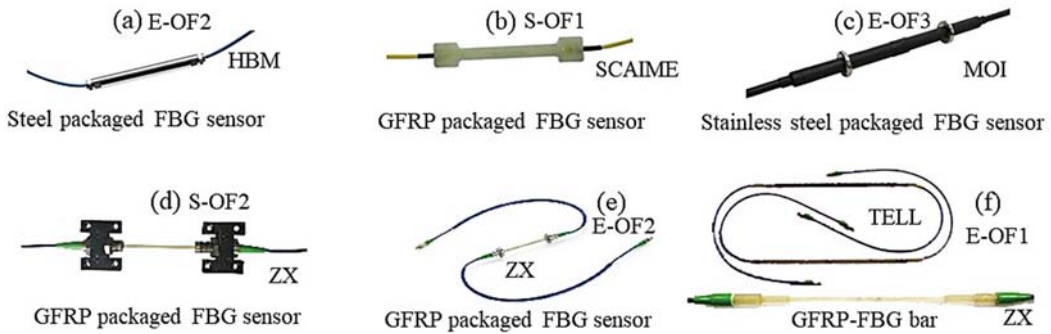
### 2.2.2. Sensors Based on Strain Transfer Theory

Given the analysis above, it can be concluded that the development of strain transfer theory is accompanied with the application of optical fiber-based sensors. The objective of the strain transfer analysis is to provide a scientific basis for accurately interpreting the relationship between the signal of the sensor and the measured parameters of the host material. The published reference [2] also points out that strain transfer theory plays an important role in the kind of sensors that depends on the interfacial shear forces to transfer deformation (i.e., S-OF1/E-OF1). That explains why the published strain transfer studies are based on multi-layered models.

However, current research has yet associated the analyzed strain transfer model and the actual applied sensors. The designers from the optical-electrical engineering cannot well understand the mechanical interaction between the sensor and the host material. If the designer doesn't know how the deformation of the host material is transferred to the sensing fiber, it is difficult to conduct the fine design or reinforcement of optical fiber-based sensors. On the other hand, the scholars in civil engineering can understand the strain transfer mechanism of multi-layered model but are poor at improving the measurement accuracy of the sensors. It should be emphasized that understanding the working principles of optical fiber-based sensors and the interaction between the sensor and the host material is particularly important for the design of sensors through strain transfer analysis.

Fig. 2.3 presents various packaged FBG sensors available in the market. It can be known that the relatively accurate strain transfer coefficients of the sensors in Figs. 2.3(a), (c), (d) and (e) can be obtained by traditional and improved calibration tests. For the sensors in Fig. 2.3(b) and (f), strain transfer theory is required to establish the strain relationship

between the sensor and the host material. The implementation of the optical fiber-based sensors on site addresses further challenges. The surface-attached sensors (S-OF1) are prone to debonding, particularly under the cyclic load [5-6]. Besides of selecting the adhesive with high bonding performance to decrease the occurrence of interfacial debonding, strain transfer analysis can be an effective approach to avoid this phenomenon by understanding the complex interactions between the sensor and the host material and identifying the material and geometrical parameters that influences the strain and bonding strength at the interface. The design of the packaging layer is desired to have minimum possibility of debonding from the host material and efficient bonding capacity to transfer strain from the host material to the sensor [7-11]. The strain transfer analysis of optical fiber-based sensors has also been used to predict the residual strain distribution and characterize the interfacial shear stress-strain state of composites with metal-coated optical fiber sensors bonded on the surface [12-14].



**Fig. 2.3.** Industrialized FBG sensors available in the market: (a) Steel packaged sensor designed by HBM; (b) GFRP packaged sensor developed by SCAIME; (c) Stainless steel packaged sensor provided by MOI; (d)-(e) GFRP packaged sensors produced by Zhixing S&T, Ltd.; (f) Smart GFRP-FBG bars offered by Tider and Zhixing S&T, Ltd.

In summary, the encapsulation and packaging of optical fiber and FBG for strain monitoring of civil structures in embeddable or surface-attached form can lead to the strain directly read from the sensor not entirely representing the actual strain of the host material. Strain transfer theory is thus developed to conduct the accurate interpretation of the collected data, analyze the effects of the installed sensors on the structural integrity and strength and improve the design and calibration of the sensors for the long-term, effective and accurate measurement.

### 2.3. Review on Strain Transfer Theory of Optical Fiber-based Sensors

Strain transfer analysis is intended to improve the measurement accuracy of optical fiber-based sensors encountered in engineering. When the installation forms of the sensors are different and the sensing models are under different loading types and damage modes, the theoretical models and the correlated assumptions can be distinguished.

### 2.3.1. Embedded and Surface-attached Sensors

For the embedded case (i.e., E-OF1), the general theoretical model is usually consisted of the sensing fiber, the protective coating and the host material, as shown in Fig. 2.4. The embedded model has been simplified as multi-layered cylinder under axial loading [15-21]. During the theoretical analysis, different assumptions have been made to obtain the strain transfer relationship. Ansari and Yuan [16] introduced the assumption that bonded length  $L$  was far larger than the radius of optical fiber and obtained the stress relationship ( $\tau_f$  and  $\tau_p$ ) through the static equilibrium function. Besides, the compatibility of strains of host material and optical fiber at the axis of symmetry ( $\sigma_f/E_f = \sigma_m/E_m$ ) was made to produce the differential governing equation. However, this assumption is not very rigorous, for the existence of protective layer (as shown in Fig. 2.5(a)), the strain of host material at the symmetric center cannot be totally transferred to the sensing fiber. It may lead to the derived strain transfer coefficient much higher than the actual case. Li et al. [17] proposed the improved assumption that the strain gradients can be approximately the same to obtain the differential equation.

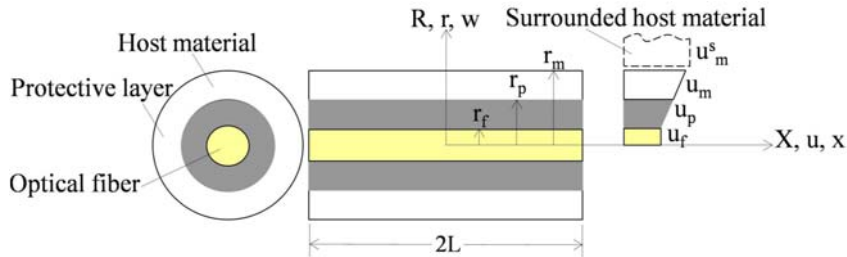


Fig. 2.4. Structure of a three-layered sensing model.

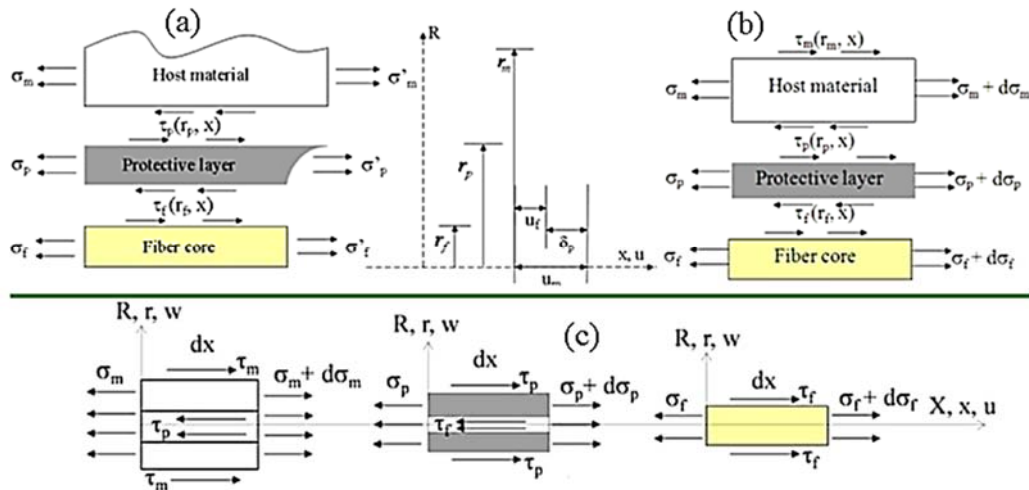
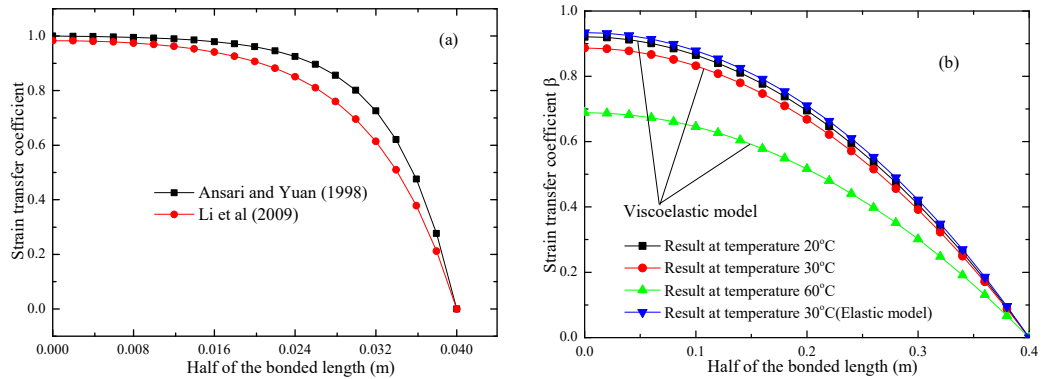


Fig. 2.5. Stress state of micro sections: (a) Ansari and Yuan [16]; (b) Li et al. [17]; (c) Wang and Xiang [4].

The comparison of the analyzed result is presented in Fig. 2.6(a). Besides, the influence of surrounded host material was considered by introducing  $\tau_m$ , as displayed in Fig. 2.5(b). However, the shear stress  $\tau_m$  was supposed to be 0 during the derivation. In other words, it ignored the influence of surrounded host material on the strain transfer efficiency of the three-layered sensing model. Wang and Xiang [4] considered the influence of surrounded host material by introducing Goodman's hypothesis to describe the shear relationship  $\tau_m$ . Besides, the viscoelastic property of host material was involved during the analysis. The distribution of strain transfer coefficient under different temperatures is plotted in Fig. 2.6(b).



**Fig. 2.6.** The distribution of strain transfer coefficient along the bonded length.

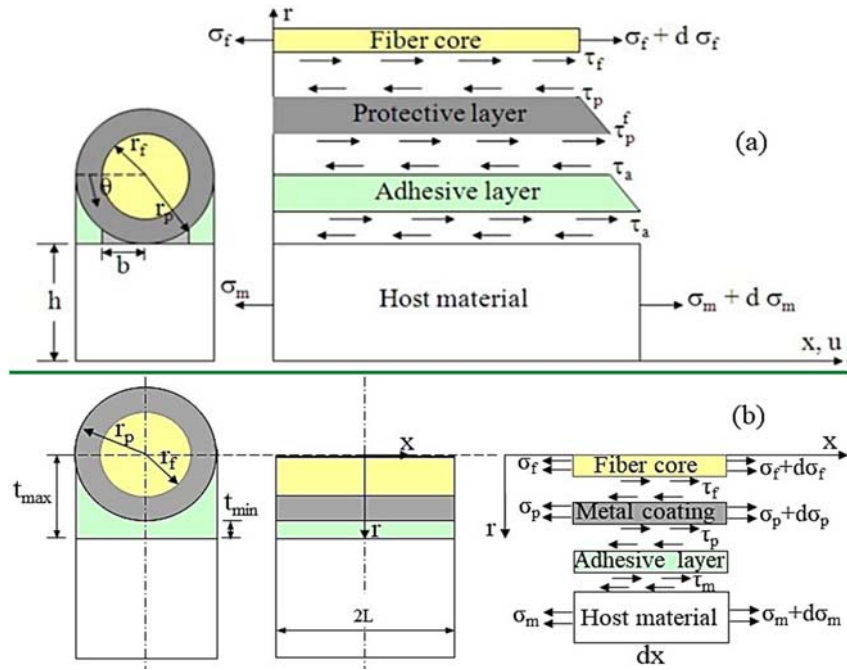
For the surface-attached case (i.e., S-OFF1), the general theoretical model is usually consisted of the sensing fiber, the protective coating, the adhesive layer and the host material, as shown in Fig. 2.7. Wan et al. [22] introduced the assumption that the surface-attached analytical model was simplified into an axisymmetric model and validated its feasibility by comparison with the results provided by Ansari and Yuan [16]. Her and Huang [23] proposed improved assumption that the protective and adhesive layers are just subjected to shear deformation, as displayed in Fig. 2.7(a). Besides, the equilibrium equations of the fiber core and protective layer were illustrated with integral forms to consider the non-uniform influence of the analytical model. Kim et al. [12] pointed out that for a metal coating with high stiffness, the normal stress ( $\sigma_p$ ) of the protective layer needed to be considered. Moreover, the bottom thickness ( $t_{min}$ ) of the adhesive layer was also involved, as shown in Fig. 2.7(b). The assumption that the strain gradient of the fiber and that of the protective coating was approximately equal was adopted during the theoretical analysis.

### 2.3.2. Multi-layered Sensing Model under Different Loading Types

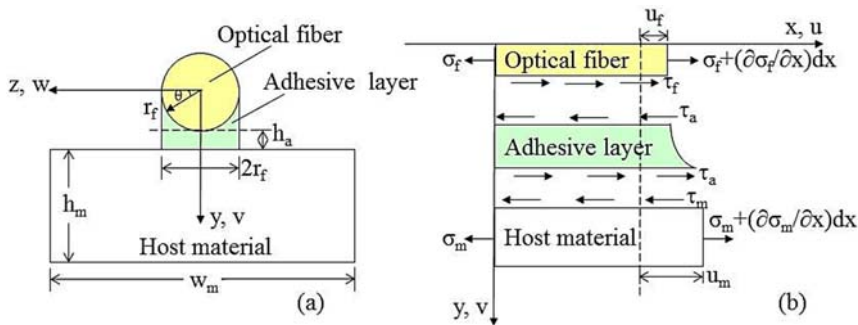
Analysis of the above-mentioned cases are under static and quasi-static actions. Along with the rapid development of advanced dynamic signal demodulation technology, FBG and optical fiber sensors have been used to measure the dynamic strain. It has been noted that the sampling frequency of Si255 for FBG signal acquisition can be up to 5 kHz, and



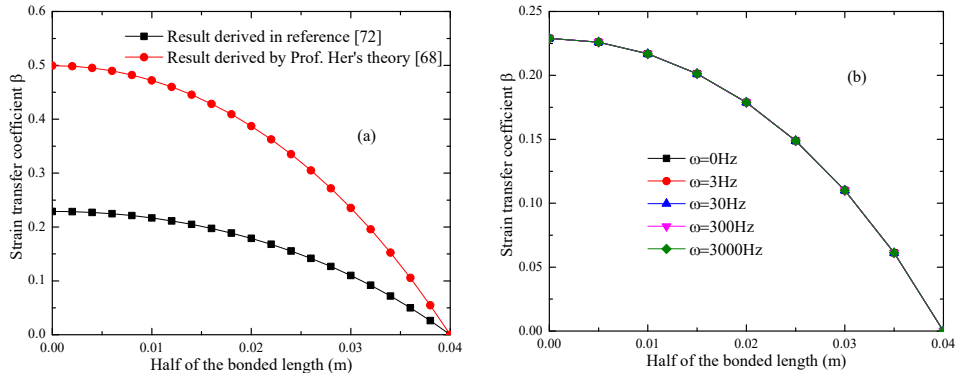
that of PPP-BOTDA for optical fiber signal sweep can be 500 Hz. Therefore, the theoretical investigation on strain transfer mechanism of optical fiber sensing model under fatigue load was conducted [24]. The mechanical model is presented in Fig. 2.8. In the theoretical analysis, inertial force was considered, and the damping effect induced by flexible adhesive layer was ignored. The loading mode was steady state response excited by stationary harmonic vibration. The strain transfer coefficients under static and dynamic loads and different vibration frequencies were separately described in Fig. 2.9. It can be observed that the static strain transfer coefficient is far larger than that under dynamic response, which indicates that the error modification is particularly important. Besides, the vibration frequency is insensitive to the strain transfer efficiency.



**Fig. 2.7.** Surface-attached model and stress state of micro sections: (a) Her [23]; (b) Kim et al. [12].



**Fig. 2.8.** Structure of the three-layered sensing model under dynamic load [24].

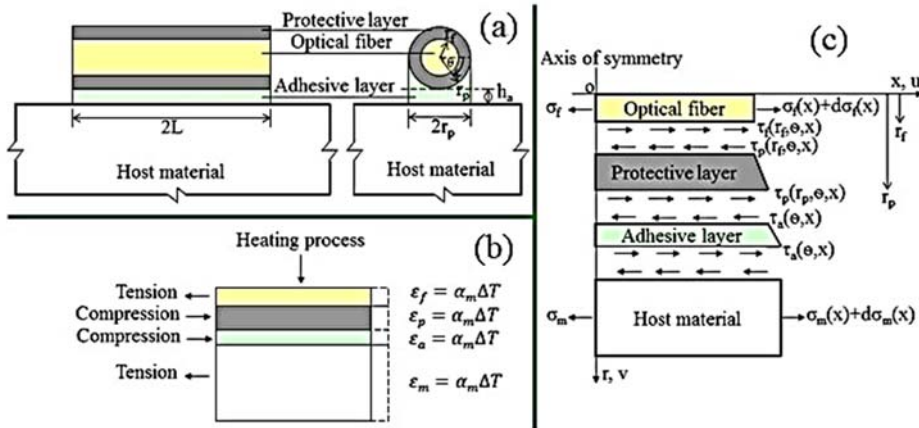


**Fig. 2.9.** Strain transfer coefficients: (a) Comparison with static state; (b) Influence of vibration frequency.

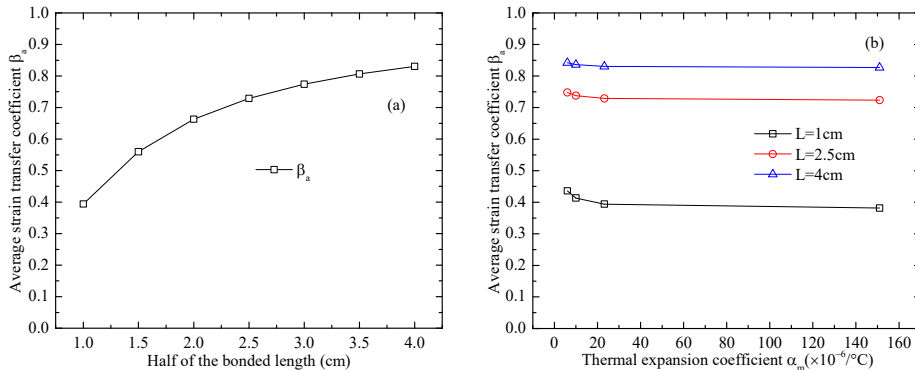
Due to environmental temperature changes during the service of structures, optical fiber-based sensors have been extensively used to monitor the thermal strains. Therefore, the strain transfer mechanism of the multi-layered sensing model subjected to thermal loading has been investigated. When the testing model is subjected to mechanical loading, the host material is subjected to applied loads, and deformation of the sensing fiber is subsequently induced by the strain transferred from the host material. However, when the testing model is subjected to thermal loading, the sensing fiber, the intermediate medium, and the host material are simultaneously exposed to the load. The shear lag effect is also influenced by the interfacial thermal stresses inside the continuous layers, which are caused by the mismatching thermal expansion coefficients of each layer. Due to the mutual constraints imposed by the contacted interfaces between the sensing fiber, the intermediate medium, and the host material, the axial extension (or shortening) of each layer cannot occur freely, which will induce interfacial thermal stresses in the joint layers. These interfacial thermal stresses bring about an additional strain in the optical fiber sensor and result in a measurement error. Taking the heating process as an example, the theoretical model is shown in Fig. 2.10. The theoretical derivation involves the thermal expansion coefficients [25]. The evolution of the average strain transfer coefficient along with the bonded length is presented in Fig. 2.11(a). For different bonded lengths, the variation of the average strain transfer coefficient with the thermal expansion coefficient is provided in Fig. 2.11(b). It can be concluded that the bonded length is the most significant factor influencing the strain transfer efficiency, and the strain transfer efficiency is relatively less sensitive to the thermal expansion coefficient of the host material.

Due to environmental temperature changes during the service of structures, optical fiber-based sensors have been extensively used to monitor the thermal strains. Therefore, the strain transfer mechanism of the multi-layered sensing model subjected to thermal loading has been investigated. When the testing model is subjected to mechanical loading, the host material is subjected to applied loads, and deformation of the sensing fiber is subsequently induced by the strain transferred from the host material. However, when the testing model is subjected to thermal loading, the sensing fiber, the intermediate medium, and the host material are simultaneously exposed to the load. The shear lag effect is also

influenced by the interfacial thermal stresses inside the continuous layers, which are caused by the mismatching thermal expansion coefficients of each layer. Due to the mutual constraints imposed by the contacted interfaces between the sensing fiber, the intermediate medium, and the host material, the axial extension (or shortening) of each layer cannot occur freely, which will induce interfacial thermal stresses in the joint layers. These interfacial thermal stresses bring about an additional strain in the optical fiber sensor and result in a measurement error. Taking the heating process as an example, the theoretical model is shown in Fig. 2.10. The theoretical derivation involves the thermal expansion coefficients [25]. The evolution of the average strain transfer coefficient along with the bonded length is presented in Fig. 2.11(a). For different bonded lengths, the variation of the average strain transfer coefficient with the thermal expansion coefficient is provided in Fig. 2.11(b). It can be concluded that the bonded length is the most significant factor influencing the strain transfer efficiency, and the strain transfer efficiency is relatively less sensitive to the thermal expansion coefficient of the host material.



**Fig. 2.10.** The four-layered sensing model under thermal load: (a) Structure; (b) Loading direction; (c) Stress state.



**Fig. 2.11.** The average strain transfer coefficient under thermal load: (a) Influence of bonded length; (b) Influence of thermal expansion coefficient of host material.

### 2.3.3. Multi-layered Sensing Model under Different Damage Modes

Local micro cracks are very common to many structures, such as concrete beam, steel component and asphalt pavement, and these host structures with cracks are still required to work for a long period of time before the failure. Distributed optical fiber sensors have been used to identify the possible random cracks. To decrease the interfacial debonding between the sensor and the host material, the packaging measure is essential, and thus strain transfer analysis with cracks in host material should be investigated. Feng et al. [26] established a four-layered model to explore the influence of crack in host material on the strain profile of the sensing fiber, as shown in Fig. 2.12(a). The model was simplified to symmetric cylinder and the interfacial constraints of host material on the transferring efficiency was ignored. Influence of the intermedium layers in elastic-plastic and plastic stages on the strain transfer coefficient was also discussed. For embedded model, Wang and Xiang [4] studied the influence of one crack in host material on the strain transfer efficiency, as shown in Fig. 2.12(b). The connection between strain transfer relationship of a cracked model and the spatial resolution of the BOTDA interrogator system was firstly established, as displayed in Fig. 2.13. The occurrence of crack induces a significant strain transfer loss to the sensors. To identify crack damage with high accuracy, the interrogator system with higher spatial resolution was preferred.

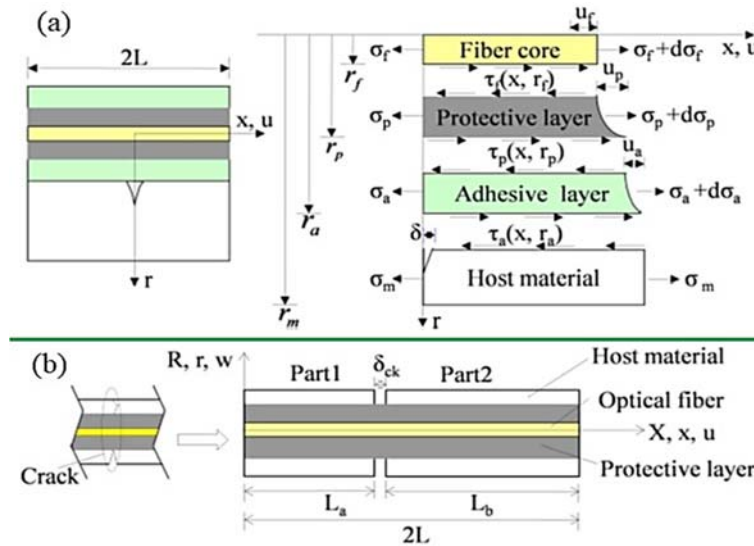
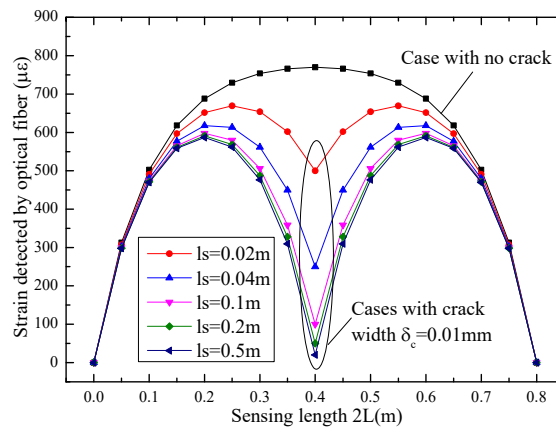


Fig. 2.12. The sensing model considering crack damage of host material: (a) Feng et al. [26]; (b) Wang and Xiang [4].

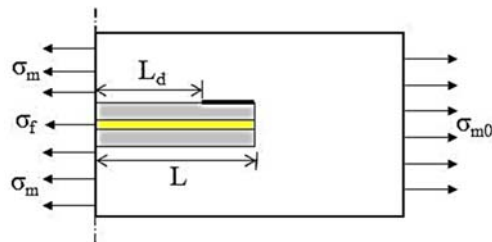
For both surface-attached and embedded sensors, the interfacial debonding between the sensor and the host material is always a problem to address further discussion. The local interfacial debonding brings about the decrease of strain transfer efficiency, and the strain transfer efficiency can drop significantly with the increase of the debonding length. It may affect the measurement accuracy and long-term effectiveness of the testing. This phenomenon can become severe for the embedded sensor, since it is difficult to conduct

the rehabilitation. Therefore, a general three-layered model (as shown in Fig. 2.14) was established to explore the occurrence principle of local interfacial debonding, the influence of local interfacial debonding on the strain transfer efficiency and the preventive design to avoid or decrease the probability of occurrence of interfacial debonding damage [27]. The distribution of interfacial shear stress is presented in Fig. 2.15(a), and the influence of local interfacial debonding on the strain transfer coefficient is displayed in Fig. 2.15(b). The shear stress shows a nonlinear increase by approaching the end of the packaged sensor, which means that the interfacial debonding damage can start from the end. Special measure can be made on the two ends of the sensor to increase the bonding strength. Generally, a larger packaging thickness and longer bonded length can be better. It also should be noted that the structural integrity should not be compromised during the packaging design of sensors.

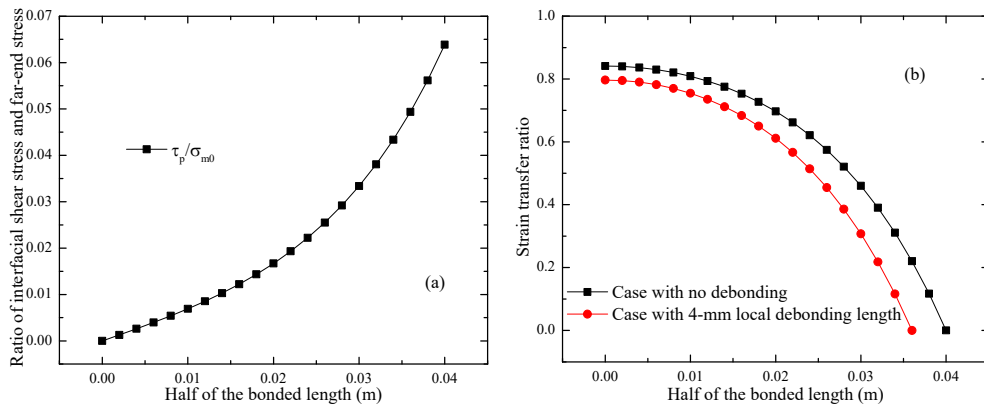
The influence of multi-layered sensing model suffered from elastic-plastic deformation, creep effect and fatigue damage on the strain transfer efficiency has also been explored based on the simplified theoretical derivation, experimental investigation and finite element validation [28-30]. However, these studies are closely correlated with the action time and the degradation process is relatively long. The weight of these damage factors on the strain transfer coefficient has been concluded to be much lighter than other damage modes.



**Fig. 2.13.** Strain transfer relationships with one crack in host material in different spatial resolutions.



**Fig. 2.14.** The sensing model considering the interfacial debonding [27].



**Fig. 2.15.** Distribution of the ratios along half of the bonded length: (a) Ratio of interfacial shear stress and remote stress; (b) Strain transfer ratio with and without local interfacial debonding.

## 2.4. Design and Calibration of Optical Fiber-based Sensors

### 2.4.1. Priority Design Parameters

According to the analysis on the application of strain transfer theory in the optimization design of optical fiber-based sensors, it can be known that the sensitive parameters influencing the strain transfer coefficient in the sensors dependent on the interfacial forces to transfer deformation and the sensors dependent on the distance change between the two anchor pieces are very different.

For the sensors (S-OF1 and E-OF1), the measurement accuracy is significantly impacted by the strain transfer efficiency. Therefore, the design of this kind of sensors should aim to obtain the maximum average strain transfer coefficient of the multi-layered sensing model. According to the parametric sensitive analysis and optimization analysis, when the two conditions that the protective layer guarantees a high survival ratio and causes no influence on the structural performance of host material, the priority design parameters of this kind of sensors can be in turn as: bonded (sensing) length, interfacial cohesion strength, elastic modulus and radius of protective layer, and thickness and shear modulus of the adhesive layer.

For the sensors (S-OF2 and E-OF2), the strain of host material is majorly transferred to the sensing fiber by the two anchor pieces located at the two ends of the sensor. Since the coefficient reflecting the strain transfer efficiency can be always acquired with relatively high precision by the calibrated test, the related error can then be removed and the modified value much accurately represent the average strain of host material in the gauge length. The priority design parameters of this kind of sensor are the material and geometrical parameters of the anchor pieces.

For the sensor (E-OF3), to much accurately obtain the strain transfer relationship between host material and optical fiber, the interfacial shear stress should be tested. Compared with

the sensor numbered E-OF2, the priority design parameters are the geometrical and material parameters of the anchor pieces and then the interfacial cohesion coefficient between the host material and the embedded sensor.

#### 2.4.2. Improved Calibration Test for End-expanded Sensors

To distinguish the difference in design and calibration of the sensors numbered E-OF2 and E-OF3, a case study on the end-expanded FBG sensors applied to the asphalt concrete is conducted. The FBG sensor with FRP blocks as the anchor pieces and armoring wire as the protective pipe is attached on the surface of the asphalt concrete, as shown in Fig. 2.6. The host material will suffer from deformation under the external loading  $\sigma_m A_m$ , with average strain  $\varepsilon_m$  in the gauge length. The corresponding elongation of the sensor is  $2\Delta L$ . According to the linear strain theory, the strain of the sensor,  $\varepsilon_s$ , can be expressed as

$$\varepsilon_s = \frac{2\Delta L}{2L} = \frac{\varepsilon_m \cdot 2L}{2L} = \varepsilon_m, \quad (2.1)$$

where it is equal to the actual deformation of the host material.

For FBG sensors, strain of the sensor,  $\varepsilon_s$ , is reflected by the central wavelength change of FBG,  $\Delta\lambda$ , and the relationship between  $\Delta\lambda$  and  $2\Delta L$  can be conveniently calibrated by experiments, as shown in Fig. 2.16. The following equation is obtained from the three loading and unloading cycle tests:

$$\Delta\lambda = 1.61989 \times 2\Delta L - 0.12552 \quad (2.2)$$

Uniting Eq. (2.1) and Eq. (2.2), the relationship between the strain of host material and the central wavelength change of FBG can be established

$$\varepsilon_m = \frac{\Delta\lambda + 0.12552}{1.61989 \times 2L} = \frac{\Delta\lambda + 0.12552}{3.23978L}, \quad (2.3)$$

where it illustrates that the relationship between the strain of host material and the variable of the sensor with two anchor pieces located at the two ends can be obtained by calibration test with relatively high precision. The sensor can be optimized by considering the other parameters (i.e., compression strength, stability and durability), but not dependent on the precision-control condition.

When the sensor is embedded in host material, as shown in Fig. 2.18, expect for the action of the stress of host material,  $\sigma_m$ , the sensor is also affected by the interfacial shear stress between protective pipe and host material. According to the mechanical action, the following equation can be established

$$\sigma_m \cdot \pi(r_{ep}^2 - r_p^2) + \tau_m \cdot 2\pi r_p L = E_p \varepsilon_p \cdot A_p, \quad (2.4)$$

where  $E_p$ ,  $\varepsilon_p$ ,  $A_p$  and  $r_p$  are the elastic modulus, strain, cross section and radius of protective layer.  $r_{ep}$  is the radius of anchor pieces, cross section  $A_p = \pi[r_p^2 - (r_p - t_p)^2]$ ,  $t_p$  is the thickness of the protective layer.

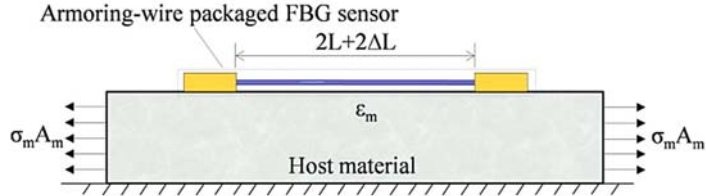


Fig. 2.16. Mechanical model of the FBG sensor bonded on the surface of structure.

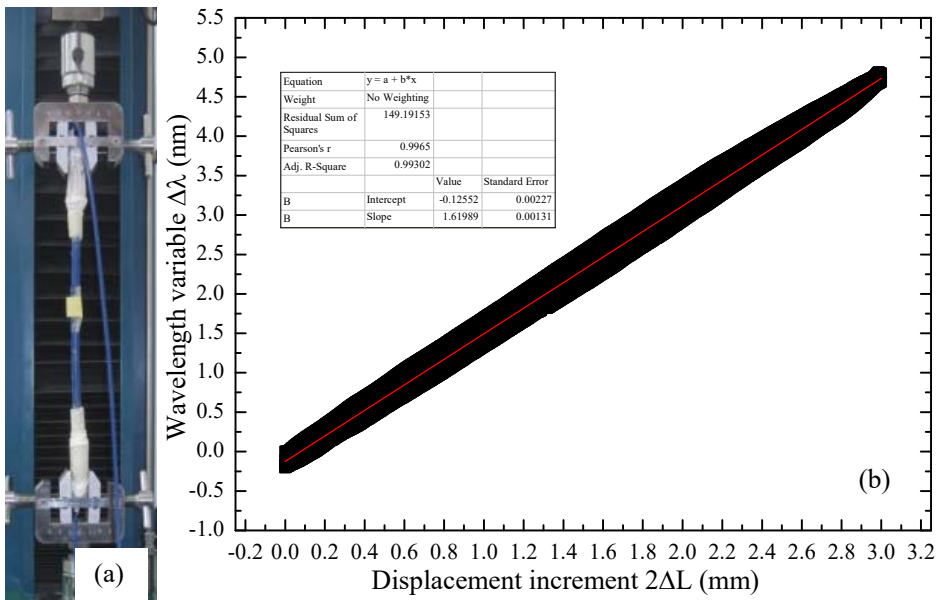


Fig. 2.17. Traditional tensile calibration test: (a) Set-up; (b) Experimental data.

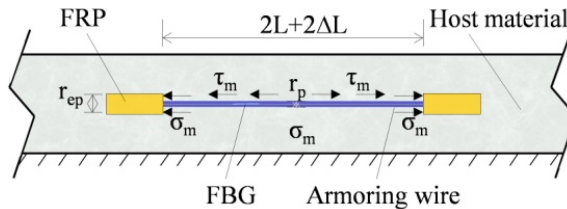


Fig. 2.18. Mechanical model of the FBG sensor embedded in the structure.



Introducing Goodman's hypotheses [31] and using the linear strain principle, the shear stress can be expressed as

$$\tau_m = k_m \cdot (\varepsilon_m - \varepsilon_p)L, \quad (2.5)$$

where  $k_m$  is the interfacial cohesion coefficient. Substituting Eq. (2.5) into Eq. (2.4), it yields

$$\varepsilon_p = \frac{E_m(r_{ep}^2 - r_p^2) + 2kr_p L^2}{2kr_p L^2 + E_p[r_p^2 - (r_p - t_p)^2]} \varepsilon_m \quad (2.6)$$

Substituting Eq. (2.2) into Eq. (2.6), the relationship between central wavelength change and strain of host material can be expressed as

$$\varepsilon_m = \frac{\Delta\lambda + 0.12552}{3.23978L} \cdot \frac{2kr_p L^2 + E_p[r_p^2 - (r_p - t_p)^2]}{E_m(r_{ep}^2 - r_p^2) + 2kr_p L^2} \quad (2.7)$$

Comparing Eq. (2.3) and Eq. (2.7), it can be noted that: the strain transfer coefficients of the same sensor for surface-attached case and embedded case are different. Compared with the surface-attached case, the mechanical action of embedded sensor is much complicated. The strain transfer coefficient involves the interfacial shear stress and the strain transfer coefficient cannot purely be calibrated by experiment. If ignoring the influence of the shear stress  $\tau_m$  on the measurement, Eq. (2.3) can be adopted in the embedded case to reflect the actual deformation.

Besides, for the embedded FBG sensor as displayed in Fig. 2.8, using the traditional calibration test aided with theoretical transformation by Eq. (2.4) to Eq. (2.7) to obtain the quantitative relationship between the strain of host material and central wavelength increment, improved calibration test is also proposed. That is embedding the FBG sensor into the standard host material sample, as shown in Fig. 2.8, and doing the tensile test to get the relationship of strain of host material and central wavelength shift. The core point is to introduce the host material sample in the calibration test, not just calibrating the sensor, but calibrating the host material with sensor embedded. The strain transfer analysis proposed here is equivalent to the second calibration of optical fiber sensors applied in practical engineering proposed by Habel [32-33].

Given the analysis above, it can be noted that: (1) For the optical fiber sensor dependent on the interfacial shear force to deliver deformation, the strain transfer relationship has more randomness, and the error modification analysis is adopted to optimize the sensor according to the priority index and control the measurement error of the sensor within the allowable scope of the engineering project. (2) For the optical fiber sensor with two anchor pieces, although the material and geometrical parameters of the protective layer and the anchor pieces affect the strain transfer coefficient, the relatively accurate strain transfer

coefficient can always be conveniently calibrated and no obvious priority design parameter is proposed [34-39].

It should be noted that the design of a sensor is closely correlated with the installation cases, constraint conditions, load types, damage modes, material and geometrical properties of the monitored structure and the intermedium layers. To improve the durability of optical fiber-based sensors for the long-term and effective monitoring, the proper selection of the packaging material for protecting sensing fiber and adhesive glue for fixing optical fiber sensors is the most significant. It determines the quality and the measurement accuracy of the sensor. It has been desired that the protective layer and the adhesive layer are always in the elastic stage during the working period of the sensor. According to the strain transfer analysis, it can be concluded that the intermedium layers may be damaged under large deformation. Therefore, for the different given host material, the special fine design of a sensor is required through the strain transfer theory. Besides, fundamental tests are in essential to check the measurement and structural performance of the sensor. Generally, to avoid the interfacial failure and guarantee the measurement accuracy, the sensing model for surface-attached case should be no more than four layers, and that for embedded case had better to be no more than three layers. The best choice for a sensor is that the adhesive layer and the protective layer are the same material. It can be concluded that the durability of the sensor can be finally attributed to the proper selection of a packaging material and adhesive glue, except for the measurement accuracy problem that can be solved by strain transfer theory and calibration test.

## **2.5. Conclusions**

Motivated by the outstanding advantages over electrical sensors, optical fiber-based sensors have extensive use in civil engineering. However, the measurement accuracy and durability of the sensors have been questioned with the increasing of service period and the worsening of service environment. Sensors with enhanced function are in high demand to satisfy the various measurement requirements of different structures, such as asphalt pavements, tunnels, bridges, high-rise buildings and high-speed railways. Improved schemes on the optimization design, calibration, robustness and durability of industrialized optical fiber-based sensors are imperative. Strain transfer theory of optical fiber-based sensors has thus been developed and investigated to provide theoretical guidance for the implementation of the improved schemes. A review on strain transfer theory of industrialized optical fiber-based sensors in civil engineering has been presented in this article. The following conclusions can be drawn from this study:

- (1) The measurement limitations induced by the sensing and material properties of optical fibers can be remedied by special packaging techniques. The configuration of optical fiber-based sensors to satisfy the various requirements in civil engineering brings about strain transfer error, which promotes the development of the strain transfer theory.
- (2) According to the different paths for transferring the deformation of host material to the sensing fiber, the optical fiber-based sensors extensively used in civil engineering can be generally divided into five types: S-OF1, E-OF1, S-OF2, E-OF2 and E-OF3. For some

kinds of sensors (i.e., S-OF2, E-OF2 and E-OF3), traditional and improved calibration tests can be adopted to correct the strain transfer error.

(3) The strain transfer theory is particularly important for the sensor dependent on the interfacial shear stress to transfer deformation (i.e., S-OF1 and E-OF1), for the error cannot be eliminated by calibration tests. It also explains why the published strain transfer studies are based on multi-layered models.

(4) Strain transfer theory can be adopted to conduct the accurate interpretation of the collected data, analyze the influence of the installed sensors on the structural integrity and strength and improve the design and calibration codes of the sensors for the long-term, effective and accurate monitoring.

(5) The strain transfer mechanism of embedded and surface-attached sensors under different loading types and damage modes has been presented to provide guidance for the correlated strain transfer error modification. The strain transfer analysis considers the characteristics of the monitored structures and the performance of the adopted interrogator system.

## Acknowledgements

The work described in this chapter was supported by the National Natural Science Foundation of China (Grant No. 51908263 and 11932008). The authors would also like to appreciate the funding support by Provincial Projects (2020-0624-RCC-0013 and JK2021-18) and National Foreign Expert Project of China (DL2021175003L and G2021175026L). Special thanks are due to Prof. Jinping Ou and Prof. Zhi Zhou of Dalian University of Technology, and Prof. Youhe Zhou and Prof. Ning Huang of Lanzhou University. The findings and opinions expressed in this article are only those of the authors and do not necessarily reflect the views of the sponsors.

## References

- [1]. H. Wang, P. Xiang, L. Jiang, Strain transfer theory of industrialized optical fiber-based sensors in civil engineering: A review on measurement accuracy, design and calibration, *Sensors and Actuators A: Physical*, Vol. 285, 2019, pp. 414-426.
- [2]. H. Wang, L. Jiang, P. Xiang, Priority design parameters of industrialized optical fiber sensors in civil engineering, *Optics and Laser Technology*, Vol. 100, 2018, pp. 119-128.
- [3]. H. P. Wang, Strain transfer of optical fiber under damage conditions and its application in multi-layered pavements, PhD Thesis, *Dalian University of Technology*, Dalian, 2015.
- [4]. H. Wang, P. Xiang, Strain transfer analysis of optical fiber-based sensors embedded in asphalt pavement structure, *Measurement Science and Technology*, Vol. 27, Issue 7, 2016, pp. 75106-75117.
- [5]. F. Ansari, Practical implementation of optical fiber sensors in civil structural health monitoring, *Journal of Intelligent Material Systems and Structures*, Vol. 18, Issue 8, 2007, pp. 879-889.
- [6]. K. Oman, B. V. Hoe, K. Aly, K. Peters, G. V. Steenberge, N. Stan, S. Schultz, Instrumentation of integrally stiffened composite panel with fiber Bragg grating sensors for vibration measurements, *Smart Materials and Structures*, Vol. 24, 2015, 085031.

- [7]. C. K. Y. Leung, X. Y. Wang, N. Olson, Debonding and calibration shift of optical fiber sensors in concrete, *Journal of Engineering Mechanics*, Vol. 126, Issue 3, 2000, pp. 300-307.
- [8]. P. Biswas, S. Bandyopadhyay, K. Kesavan, S. Parivallal, B. A. Sundaram, K. Ravisankar, K. Dasgupta, Investigation on packages of fiber Bragg grating for use as embeddable strain sensor in concrete structure, *Sensors and Actuators A: Physical*, Vol. 157, 2010, pp. 77-83.
- [9]. B. Glisic, Influence of the gauge length on the accuracy of long-gauge sensors employed in monitoring of prismatic beams, *Measurement Science and Technology*, Vol. 22, Issue 3, 2011, 5206.
- [10]. P. A. Calderon, B. Glisic, Influence of mechanical and geometrical properties of embedded long-gauge strain sensors on the accuracy of strain measurement, *Measurement Science and Technology*, Vol. 23, Issue 9, 2012, 065604.
- [11]. Z. Zhou, W. Q. Liu, Y. Huang, H. P. Wang, J. P. He, M. H. Huang, J. P. Ou, Optical fiber Bragg grating sensor assembly for 3D strain monitoring and its case study in highway pavement, *Mechanical Systems and Signal Processing*, Vol. 28, pp. 36-49.
- [12]. S. W. Kim, M. S. Jeong, I. Lee, E. H. Kim, I. B. Kwon, T. K. Hwang, Determination of the maximum strains experienced by composite structures using metal coated optical fiber sensors, *Composites Science and Technology*, Vol. 78, 2013, pp. 48-55.
- [13]. S. W. Kim, M. C. Cha, I. Lee, E. H. Kim, I. B. Kwon, T. K. Hwang, Damage evaluation and strain monitoring of composite plates using metal-coated FBG sensors under quasi-static indentation, *Composites: Part B*, Vol. 66, 2014, pp. 36-45.
- [14]. D. S. Xu, H. B. Liu, W. L. Luo, Evaluation of interface shear behavior of GFRP soil nails with a strain-transfer model and distributed fiber-optic sensors, *Computers and Geotechnics*, Vol. 95, 2018, pp. 180-190.
- [15]. M. LeBlanc, Interaction mechanics of embedded single-ended optical fibre sensors using novel in-situ measurement techniques, PhD Thesis, *University of Toronto*, Canada, 1999.
- [16]. F. Ansari, L. B. Yuan, Mechanics of bond and interface shear transfer in optical fiber sensors, *Journal of Engineering Mechanics*, Vol. 4, 1998, pp. 385-394.
- [17]. H. N. Li, G. D. Zhou, L. Ren, D. S. Li, Strain transfer coefficient analysis for embedded fiber Bragg grating sensors in different host materials, *Journal of Engineering Mechanics*, Vol. 135, Issue 12, 2009, pp. 1343-1353.
- [18]. Y. W. Yang, B. S. Divsholi, L. H. Tang, L. Zhang, Strain transfer models for macrofiber-composite strain actuators, *Materials and Manufacturing Processes*, Vol. 25, 2010, pp. 237-242.
- [19]. L. B. Yuan, L. M. Zhou, Sensitivity coefficient evaluation of an embedded fiber-optic strain sensor, *Sensors and Actuators A: Physical*, Vol. 69, 1998, pp. 5-11.
- [20]. D. S. Li, H. N. Li, L. Ren, G. B. Song, Strain transferring analysis of fiber Bragg grating sensors, *Optical Engineering*, Vol. 45, Issue 2, 2006, 024402.
- [21]. G. Duck, M. LeBlanc, Arbitrary strain transfer from a host to an embedded fiber-optic sensor, *Smart Materials and Structures*, Vol. 9, 2000, pp. 492-497.
- [22]. K. T. Wan, C. K. Y. Leung, N. G. Olson, Investigation of the strain transfer for surface-attached optical fiber strain sensors, *Smart Materials and Structures*, Vol. 17, 2008, 035037.
- [23]. S. C. Her, C. Y. Huang, Effect of coating on the strain transfer of optical fiber sensors, *Sensors*, Vol. 11, Issue 7, 2011, pp. 6926-6941.
- [24]. H. Wang, P. Xiang, X. Li, Theoretical analysis on strain transfer error of FBG sensors attached on steel structures subjected to fatigue load, *Strain*, 2016, Vol. 52, Issue 6, pp. 522-530.
- [25]. H. Wang, J.-G. Dai, Strain transfer analysis of fiber Bragg grating sensor assembled composite structures subjected to thermal loading, *Composites Part B: Engineering*, Vol. 162, 2019, pp. 303-313.

- [26]. X. Feng, J. Zhou, C. S. Sun, X. T. Zhang, F. Ansari, Theoretical and experimental investigation into crack detection with BOTDA distributed fiber optic sensors, *Journal of Engineering Mechanics*, Vol. 139, Issue 12, 2013, pp. 1797-1807.
- [27]. H. Wang, L. Jiang, P. Xiang, Improving the durability of the optical fiber sensor based on strain transfer analysis, *Optical Fiber Technology*, Vol. 42, 2018, pp. 97-104.
- [28]. Q. B. Li, G. Li, G. L. Wang, et al., Elasto-plastic bonding of embedded optical fiber sensors in concrete, *Journal of Engineering Mechanics*, Vol. 128, Issue 4, 2002, pp. 471-478.
- [29]. J. L. Li, Z. Zhou, J. P. Ou, Interface transferring mechanism and error modification of FRP-OFBG strain sensor based on standard linear viscoelastic model, *Proceedings of SPIE*, Vol. 6174, 2006, pp. 1124-1135.
- [30]. Z. Zhou, Q. Wang, X. W. Hao, H. Y. Yang, J. P. Ou, Strain transfer analysis of the FBG sensor considering the creep of the concrete host, *China Measurement & Test*, Vol. 42, Issue 5, 2016, pp. 1-6.
- [31]. J. Uzan, M. Livneh, Y. Eshed, Investigation of adhesion properties between asphaltic-concrete layers, *Association of Asphalt Paving Technologies Proceedings*, 1978, pp. 495-521.
- [32]. W. R. Habel, V. G. Schukar, N. Kusche, How do application-related issues influence the reliability of fiber optic strain measurements?, *IEEE Sensors*, Vol. 2011, 2011, pp. 947-950.
- [33]. W. R. Habel, V. G. Schukar, N. Kusche, Fibre-optic strain sensors are making the leap from lab to industrial use – Reliability and validation as a precondition for standards, *Measurement Science and Technology*, Vol. 24, 2013, 094006.
- [34]. H. Wang, J.-G. Dai, X. Wang, Improved temperature compensation of fiber Bragg grating-based sensors applied to structures under different loading conditions, *Optical Fiber Technology*, Vol. 63, 2021, 102506.
- [35]. H. Wang, W. Liu, J. He, X. Xing, D. Cao, X. Gao, X. Hao, H. Cheng, Z. Zhou, Functionality enhancement of industrialized optical fiber sensors and system developed for full-scale pavement monitoring, *Sensors*, Vol. 14, 2014, pp. 8829-8850.
- [36]. H. Wang, Z. Zhou, Y. Huang, P. Xiang, J. Ou, Strain transfer mechanism of quadrate-encapsulated FBG sensors embedded in rectangular structures, *Journal of Civil Structural Health Monitoring*, Vol. 5, Issue 4, 2015, pp. 469-480.
- [37]. H. Wang, P. Xiang, L. Jiang, Optical fiber sensor based in-field structural performance monitoring of multi-layered asphalt pavement, *Journal of Lightwave Technology*, Vol. 36, Issue 17, 2018, pp. 3624-3632.
- [38]. H. Wang, P. Xiang, L. Jiang, Optical fiber sensing technology for full-scale condition monitoring of pavement layers, *Road Materials and Pavement Design*, Vol. 21, Issue 5, 2020, pp. 1258-1273.
- [39]. P. Xiang, H. Wang, Optical fibre-based sensors for distributed strain monitoring of asphalt pavements, *International Journal of Pavement Engineering*, Vol. 19, Issue 9, 2018, pp. 842-850.

## **Chapter 3**

# **Fine Measurements of Field Torsion and Its Spectra**

**Dr. Francisco Bulnes, J. C. García-Limón,  
Victor A. Sánchez-Suárez, L. A. Ortiz-Dumas,  
M. Y. Flores-García and A. U. García Galicia**

### **3.1. Introduction**

The torsion is a field observable, which in geometry is a second curvature. In field theory, the torsion is the better evidence of the birth of gravity and their different effects until our days with the gravitational waves that are echoes detected from big bang in astronomical observatories.

Through of electronics is designed an analogue of torsion measurement, considering the electronic signals detected by a Hall sensor gauge to detect the magnetic field of a dilation (gauge particle of magnetic nature). This moves along a spiral trajectory with certain torsion (see the Fig. 3.1), that in this case and due to the instrumental limitations was considered with constant torsion [1]. An important result obtained was that spectral torsion its spectral curvature and thus its torsion energy [1].

A consideration important raised of some investigations using Hall sensors gauged to different ranges of magnetic field is the obtaining of the torsion as energy spectra whose energy density is a spectral torsion determined through the energy curvature developed in other spectral studies of curvature [2-4], where curvature energy can be obtained through electromagnetic fields. However, in the case of the torsion, is required a trajectory whose torsion is detectable and measured by fine instruments.

The application of a Hall sensor to detect torsion using magnetic field [1], helps to conjecture the spin nature of the torsion, which interacts with a movement, where this movement depends of microscopic spin oscillations. However, we need a field that gives orientation to the matter without the parasite microscopic currents that carries an electric

field. Then this carry us to conjecture that the torsion energy can be obtained purely from magnetic field. To this, we consider in first instance the lemma corresponding to the Effect Hall sensor used to detect torsion through magnetic field [1], and we remember that the Effect Hall sensor deflects the current to create a potential through magnetic field  $H$ .

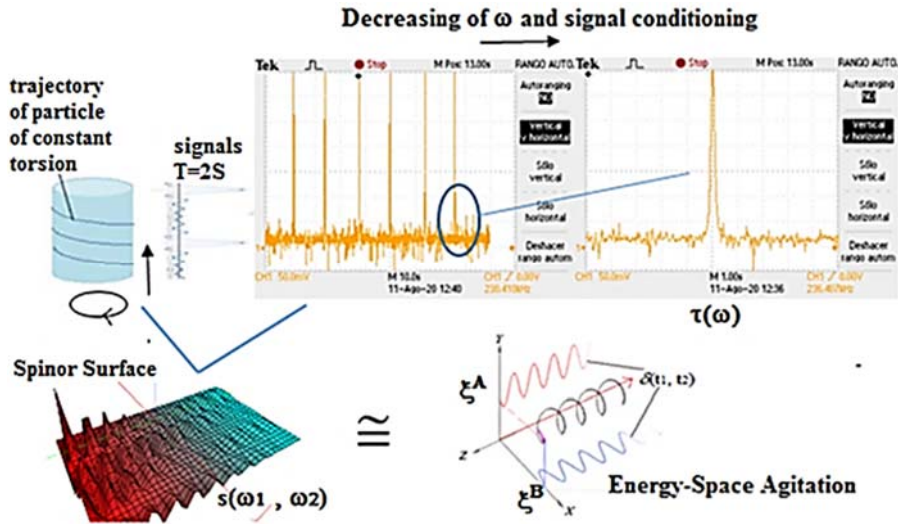


Fig. 3.1. Complete method by Hall Effect sensor for detecting of field torsion [1].

Then considering the Ohm's law to relate potential with current and characterizing this deflection (or deviation) as curvature, further of relate the curvature energy with torsion, we have that:

$$\tau = H \frac{Rb}{a^2}, \quad (3.1)$$

which tell us that the torsion is generated from the magnetic field intensity.

This meets with the fact that in sidereal objects formation, the existence of intersidereal magnetic field is fundamental to the ordering of spins of matter particles and the consequent evidence of gravitational waves existence that start from a galaxy or star centre, conforming the gravity when the evolution of particles in baryogenesis proliferates as hadrons and other particles.

Then also verifies the conjecture<sup>1</sup>, and some theorems established in other studies in theoretical physics and mathematical physics. Likewise, using a Hall type sensor is detected the variation of the magnetic field produced from a magnetic dilation [5], which

<sup>1</sup> **Conjecture (F. Bulnes)**. The torsion is the geometrical invariant of the interaction between energy and space. This interaction between energy and space happens when a magnetic field acts as "energy screw in the space", agitating the space.

is moved along a cylindrical spiral where the matter agitates the space and from which emanates gravity (*from the kinematic tensor theory in duality with spinor waves*), also interacts with the magnetic field of the proper dilation (Fig. 3.1).

### 3.2. Magnetic Nature of the Torsion

**Conjecture 3. 2. 1.** The torsion is of magnetic nature.

To prove this conjecture, we need the fact that to a magnetic state potential  $\phi_m$ , of the magnetic field intensity  $H$ , is satisfied that:

$$\operatorname{div} \operatorname{grad} \phi_m = \operatorname{div} H = 0, \quad (3.2)$$

which establish that the magnetic intensity lines follow being closed to any exterior point to the electric field current.

**Theorem 3. 2. 1 (F. Bulnes).** The torsion as geometrical invariant of the space is of field magnetic nature in constant case as well as variable case.

*Proof.* We analyse two cases. We consider the movement that generates constant torsion. For this movement where  $I$  is the electric current,  $a$  is the radius of a spire,  $l$  is a distance, which could correspond to a cycle of the cylindrical spiral, when is re-walked a distance  $b$  on its axis; the total magnetic field is:

$$H = \int_s dH = \frac{I}{4\pi} \int_s \frac{ads}{l^3} = \frac{I}{2\pi} \frac{a^2}{l^3} \quad (3.3)$$

However, we remember that the Effect Hall sensor deflects the current to create a potential through magnetic field  $H$ , then considering the Ohm's law to relate potential with current and characterizing this deflection (or deviation) as curvature, further to use the Corollary 3. 2. 1, [1] we have (3.1). Then substituting (3.1) in (3.3) we have:

$$\tau = \frac{I}{2\pi} \frac{a^2}{l^3}, \quad (3.4)$$

which being considered a cylindrical helix trajectory for the sensing, the torsion (3.4) has its corresponding geometrical term of constant torsion  $\frac{b}{a^2 + b^2}$

$$\tau = \frac{V}{2\pi} \frac{b}{a^2 + b^2} \frac{1}{l} \quad (3.5)$$



Now, we consider a movement that generates variable torsion. We consider of a vortex trajectory, which are defined for the conic helix<sup>1</sup>. Likewise, analogous to the cylindrical helix case, we have  $l$ , as the distance, which could correspond to a cycle of the conical spiral, although for this case, is the function  $l(t) = bt$ , where  $b$  is the angle, then we write the torsion as:

$$\tau(t) = \frac{\nu(6+l^2)}{4+l^2+r^2(2+l^2)^2} \quad (3.6)$$

We define the space of the vorticity  $\Omega_H$

$$C(p) = \{\mathbf{v} \in T_p(\Sigma) \mid \Omega_H = \text{rot } \mathbf{v}\}, \quad (3.7)$$

with solid angle of vortex  $\Sigma$ , where current circulate on circular ring which can increase until the equator of the sphere of radius  $r$ , see the Fig. 3.2.

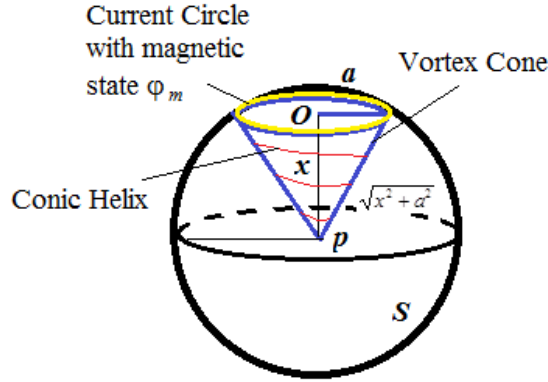


Fig. 3.2. Vorticity and conic spiral or helix.

We consider a little sphere  $S$  of radius  $r$ . In the movement on cone, the variation of the solid angle is given by

$$d\Sigma = f'(r)dA = -\frac{2h}{r^3}dA \quad (3.8)$$

<sup>1</sup> We consider the conic helix  $\gamma(t) = (tr \cos \alpha t, tr \sin \alpha t, t)$ . Then the variable torsion is getting by

$$\tau(t) = \frac{\det(\gamma', \gamma'', \gamma''')}{\|\gamma' \times \gamma''\|^2} = \frac{\alpha(6 + \alpha^2 t^2)}{4 + \alpha^2 t^2 + r^2(2 + \alpha^2 t^2)^2}.$$

<sup>2</sup>  $\nu$  is the velocity of cycles persecond (a frequency).

Then the current circle can increase until the sphere equator, having that  $\cos \theta = \frac{\sqrt{x^2 + a^2}}{\sqrt{x^2 + a^2}} = 1$ . Then for a cone whose variations are given by (3.8) from its base until vertex in  $p$  we have that the solid angle is:

$$\begin{aligned} \Sigma &= \iint_S d\Sigma = -2 \iint_S \frac{1}{r^2} r dr d\theta = \frac{2\pi h}{r} = 2\pi (\cos \theta_2 - \cos \theta_1) = \\ &= 2\pi (1 - \cos \theta_1) = 2\pi \left( 1 - \frac{x}{\sqrt{x^2 + a^2}} \right) \end{aligned} \quad (3.9)$$

Working a little part, the expression  $1 - \frac{x}{\sqrt{x^2 + a^2}}$ , we have its equivalent expression<sup>1</sup>:

$$\Sigma = 2\pi \left[ \frac{a^2}{x^2 + a^2 + x\sqrt{x^2 + a^2}} \right], \quad (3.10)$$

which involves a torsion of the corresponding conic helix<sup>2</sup> in vorticity. Indeed, we can apply the complex transformation:

$$T : D_{\Sigma} \rightarrow \mathbb{C}, \quad (3.11)$$

with correspondence rule

$$(x^2, a) \mapsto \frac{a^2}{a^2 + z\bar{z} + r^2(a + z\bar{z})^2}, \quad (3.12)$$

with  $z = 2i$ ,  $\bar{z} = -2i$ ,  $x^2 = b^2 t^2 = z\bar{z}$ ,  $a = 2$ , and  $r = \sqrt{x} = bt$ . Then we have the torsion given in the footnote 2, considering the frequency  $b$ , depending of other parameters.

Likewise, we consider the annular magnetic fields defined by the magnetic contender that come given by the Maxwell equation  $\text{div} H = 0$ , which establish along the current inside  $\mathcal{R}$ , current circles of length  $2\pi a$ ,  $a > 0$ . Then its magnetic field is the gradient of its magnetic potential:

---

<sup>1</sup>  $1 - \frac{x}{\sqrt{x^2 + a^2}} = \frac{\sqrt{x^2 + a^2} - x}{\sqrt{x^2 + a^2}} \left[ \frac{\sqrt{x^2 + a^2} + x}{\sqrt{x^2 + a^2} + x} \right] = \frac{a^2}{x^2 + a^2 + x\sqrt{x^2 + a^2}}$ .

<sup>2</sup> The torsion of the conic helix of vector function  $\gamma(t) = (x(t), y(t), z(t)) = (tr \cos bt, tr \sin bt, t)$  with the scalar function  $\tau(t) = \frac{b(6 + b^2 t^2)}{4 + b^2 t^2 + r^2(2 + b^2 t^2)^2}$ .

$$H = -\text{grad}\phi_m = -\text{grad}\left(\frac{I}{4\pi\gamma}\Sigma\right) = -\frac{I}{4\pi\gamma}\frac{d\Sigma}{dx} = -\frac{I}{4\pi\gamma}2\pi\left[\frac{a^2}{\left(\sqrt{a^2+x^2}\right)^3}\right] =$$

$$= \frac{I}{2\gamma}\frac{a^2}{\left(\sqrt{a^2+x^2}\right)^3}$$

The magnetic field involves the torsion given in (3.12). ♦

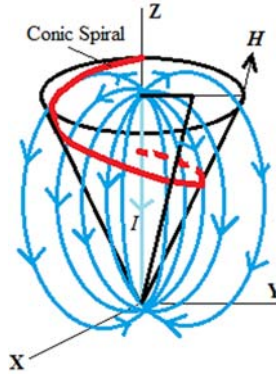


Fig. 3.3. Representation of electromagnetic interacting

**Corollary 3. 2. 2 (F. Bulnes).** The corresponding dimensional analysis [8]<sup>1</sup> to constant and variable torsion energy is

$$\tau = H \frac{Rb}{a^2} \left( = \frac{\text{Volts}}{(\text{meter})^3} \right), \quad (3.13)$$

*Proof.* By the Theorem 3. 2. 1, we have that the magnetic intensity for any dilation varies accord to

$$H = -\frac{I}{4\pi\gamma}\frac{d\Sigma}{dx} = \frac{I}{2\gamma}\frac{a^2}{\left(\sqrt{a^2+x^2}\right)^3}, \quad (3.14)$$

where the sign – is by convention of the gradient. Likewise using the magnetic field dilation (3.14) in the torsion formula (3.13) we have (considering left-handed-spin<sup>2</sup>):

<sup>1</sup> Remember that the torsion is a second curvature.

<sup>2</sup> Of fact in the experiment of variable torsión of Hall magnetic sensor we realize the two spins in the experiment: the right-handed-spin and the left-handed-spin.

$$\tau = H \frac{Rb}{a^2} = \frac{d\Sigma}{dx} \frac{I}{4\pi\gamma} \left[ \frac{Rb}{a^2} \right], \quad (3.15)$$

Remember that the Effect Hall sensor deflects the current to create a potential through magnetic field  $H$ , that is to say, creates a deviation that is curvature. We need detect this deviation energy considering an electric component given by the electrical resistance  $R(=\Omega)$ , that reflects in the Ohm's law. In this case, also the Ohm's law to relate potential with current, characterizing this deflection (or deviation) as curvature, further to use the Corollary 3. 2. 1, published in [1]. Then (3.15) takes the form:

$$\tau = H \frac{Rb}{a^2} = \frac{d\Sigma}{dx} \frac{V}{4\pi\gamma R} \left[ \frac{Rb}{a^2} \right] = 2\pi \frac{a^2}{(\sqrt{a^2 + x^2})^3} \frac{V}{4\pi\gamma} \frac{b}{a^2} = 2\pi \frac{a^2}{(a^2 + x^2)r} \frac{V}{4\pi\gamma} \frac{b}{a^2},$$

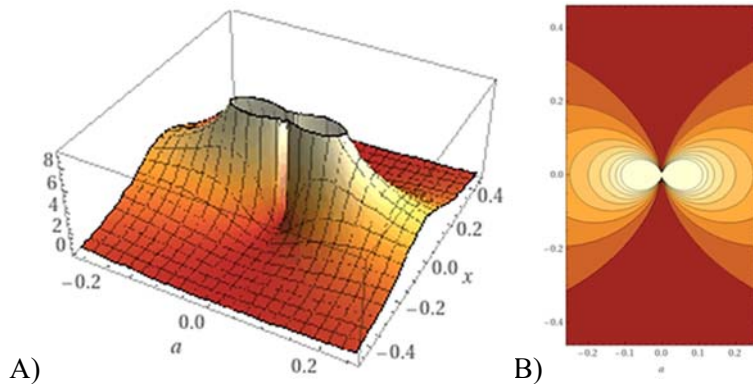
which we have the final expression of torsion

$$\tau = \frac{1}{2\gamma} \frac{V}{(a^2 + x^2)} \frac{b}{r} \left( = \frac{\text{Volts}}{(\text{meter})^3} \right), \quad (3.16)$$

which proves the corollary. ♦

Likewise, the curvature energy finally will be  $(\text{meter})^{-1}$ , that is the torsion unit from a point of geometrical purely view [2].

The rational or cremonian function  $\Sigma'(x) = 2\pi \frac{a^2}{(a^2 + x^2)^{3/2}}$ , express and determine the magnetic field Hall sensing through the lobules that determines in the space (Fig. 3. 4).



**Fig. 3.4.** A). Lobules of magnetic field Hall sensing. Each lobule represents the sensing of a dilation pole. These are interspersed along the dilation movement and its detection by the Hall magnetic sensor.

### 3.3. Torsion Energy

We have demonstrated (*theorem (F. Bulnes)* [1]) that the torsion energy comes given by the Fourier spectra:

$$\begin{aligned}
 \tau(t_1, t_2) &= \frac{c}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 2\pi F\{K(t_1, t_2)\} \delta(t_1, t_2) d\omega_1 d\omega_2 = \\
 &= \frac{2\pi c}{2\pi} 2\pi \kappa(-\omega_1, -\omega_2) \delta(t_1, t_2) = \\
 &= \pm 2\pi \cot \Psi \kappa(\omega_1, \omega_2) \delta(t_1, t_2),
 \end{aligned} \tag{3.17}$$

which involves a Dirac function as force of impulse of the screw effect in the space. Then we need to know the spectral density of the magnetic field intensity. This spectral density corresponds to the magnetic signal responses obtained in the Hall sensor. Likewise, we have the following lemma.

**Lemma 3. 3. 1.** The spectral torsion (torsion energy) on the conic trajectory of dilaton obtained for the Hall sensor is:

$$\tau(\omega_1, \omega_2) = \frac{1}{2\gamma} \frac{1}{(a^2 + x^2)} \frac{b}{r} T L S a^2(\omega_1, \omega_2), \tag{3.18}$$

*Proof.* We consider the two displacements of the dynamic device of Hall sensor (see the figure), where to vertical displacement we have the signal condition of initial voltage  $\Pi_1\left(\frac{t}{T}\right)$ , and the signal condition of initial voltage to horizontal displacement given by  $\Pi_2\left(\frac{t-\tau}{L}\right)$ , then their joint realization is their convolution in the system. Likewise the output variation of the spectral torsion will comes as:

$$\tau(\omega_1, \omega_2) = \frac{1}{2\gamma} \frac{1}{(a^2 + x^2)} \frac{b}{r} \int_{-L/2}^{L/2} \int_{-T/2}^{T/2} \Pi_1\left(\frac{t}{T}\right) \Pi_2\left(\frac{t-\tau}{L}\right) e^{-j(\omega_1 + \omega_2)t} d\tau_1 d\tau_2,$$

where applying the properties of the Fourier transform on functions convolution, we have:

$$\begin{aligned}
 \tau(\omega_1, \omega_2) &= \frac{1}{2\gamma} \frac{1}{(a^2 + x^2)} \frac{b}{r} \int_{-L/2}^{L/2} \frac{e^{j(\omega_1 + \omega_2)T} - e^{-j(\omega_1 + \omega_2)T}}{j(\omega_1 + \omega_2)} \Pi_2\left(\frac{t}{L}\right) \times e^{-j(\omega_1 + \omega_2)t} d\tau_2 = \\
 &= \frac{1}{2\gamma} \frac{1}{(a^2 + x^2)} \frac{b}{r} T \left\{ \frac{e^{j(\omega_1 + \omega_2)T} - e^{-j(\omega_1 + \omega_2)T}}{j(\omega_1 + \omega_2)} \right\} \left\{ \frac{e^{j(\omega_1 + \omega_2)L} - e^{-j(\omega_1 + \omega_2)L}}{j(\omega_1 + \omega_2)} \right\} = \\
 &= \frac{1}{2\gamma} \frac{1}{(a^2 + x^2)} \frac{b}{r} T L S a^2(\omega_1, \omega_2),
 \end{aligned}$$

where the function  $Sa(\omega) = \frac{\sin \omega}{\omega}$ , which evidences the torsion existing in the space where happens the dilaton movement and the interaction with electromagnetic field of the Hall sensor. The rotation energy is given by the implicit torsion. ♦

### 3.4. Electronics Gaugings and Experiments on the Measurement of Torsion

For the sensing of magnetic field is used a module of Hall Effect KY-024 (see the Fig. 3.5), which has a sensor of linear Hall Effect SS49E. This linear Hall Effect component will obtain an output voltage proportional to the magnetic field magnitude detected with a sensibility of the order of  $1.4 \text{ mV/Gauss.}$ , permitting also the orientation magnetic field detection [6-8].

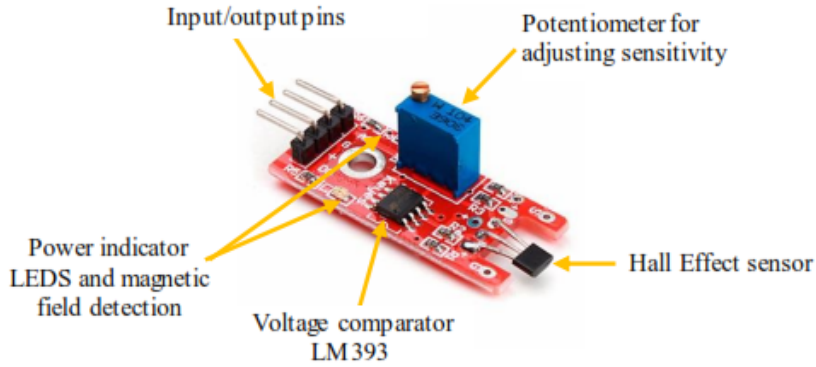


Fig. 3.5. Module for magnetic field detection.

**Corollary 3. 4. 1.** In the Hall sensor  $\mathcal{L}_{\text{Hall}}^H$ ,<sup>1</sup> [1] the torsion force is obtained only in the vertical Z-axis.

*Proof.* We consider the Hall sensor electronics device, which realizes translations and rotations. Due to that, the rotation movement is realized in the Z-axis; by the Theorem 2.1, [1], we have that the curvature energy has kernel given as  $e^{-j(\omega_1 t_1 + \omega_2 t_2)}$ , which involves the waves that interact in a space  $\mathbb{R}^3$ , in YZ and XZ – plane having terms  $A \cos n(\omega_1 + \omega_2)t + jB \sin n(\omega_1 + \omega_2)t$ , in both direction in the plane. To the conic case (Fig. 3.6B), the translation movements, vertical as well as horizontal only detect the scale or magnitude of the signal (proximity to the sensor). Considering (3.17) we have

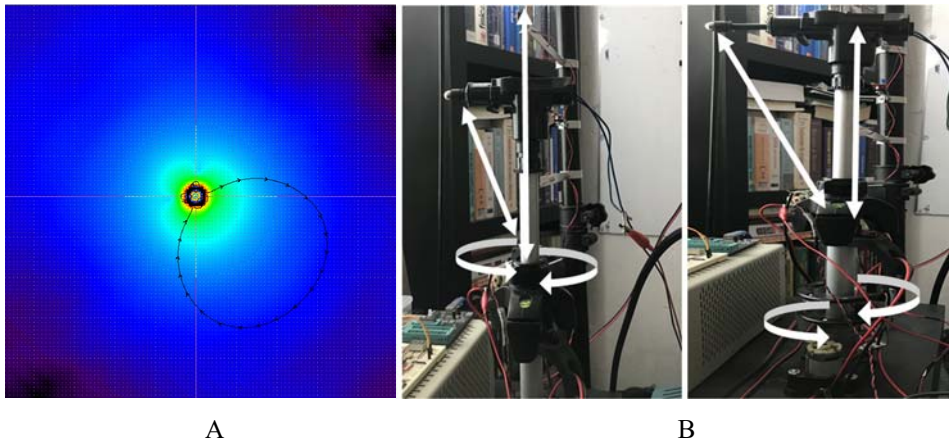
<sup>1</sup> It is defined as the topological space

$$\mathcal{L}_{\text{Hall}}^H = \left\{ \delta H \in C(\mathbb{R}^3 \setminus \{0\}) \mid \delta H = \frac{1}{4\pi\gamma} \frac{\delta s}{a^2 + b^2}, a > 0, b \in \mathbb{R} \right\}.$$

$$z = \tau(t_1, t_2) = k\tau(\omega_1, \omega_2)\delta(t_1, t_2), \quad (3.19)$$

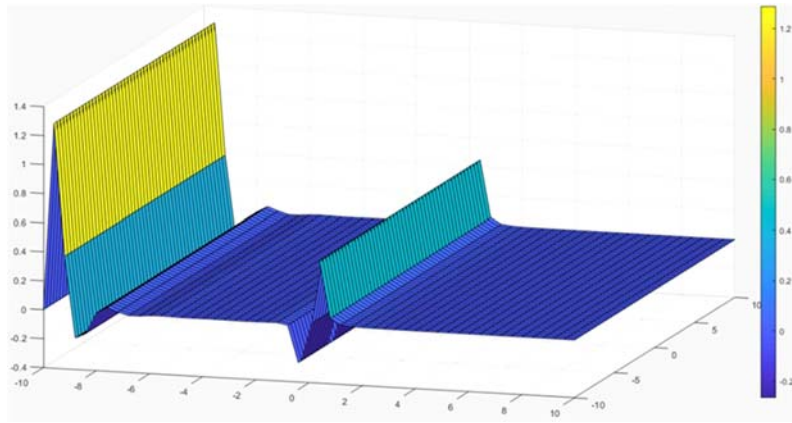
where by the Corollary 2.1, [1], the spectral torsion is its curvature energy. Here  $k = \frac{2\gamma(a^2 + x^2)r}{bTL}$ , where involves the lengths  $a$ , and  $x$ . ♦

The corollary conclusions meet with the fact of the displacement of the Hall sensor device, which requires a force in the two senses of the displacement (see the Fig. 3.6B and Fig. 3.7) and the sensing action is realized on the magnetic dilation producing signals of the form  $\frac{\sin L\omega}{L\omega}$ , (Fig. 3.8) increasing or decreasing; according to the proximity of dilation with the Hall magnetic sensor (Fig. 3.6B and Fig. 3.9A). After through the three movements we obtain the signals torsion evidence (Fig. 3.8 and Fig. 3.9B).

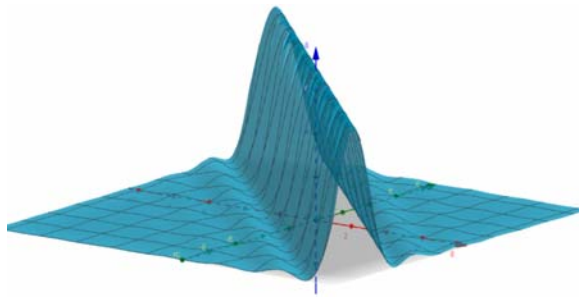


**Fig. 3.6.** Magnetic sensing. A). The magnetic field lobules in green-clear blue represent the magnetic sensing along of two cycles of the conic spiral trajectory of the dilation in upside (also see the Fig. 3.4B). Device of Hall sensor chip using the three synchronize movements to permits an inverted cone as well as right cone.

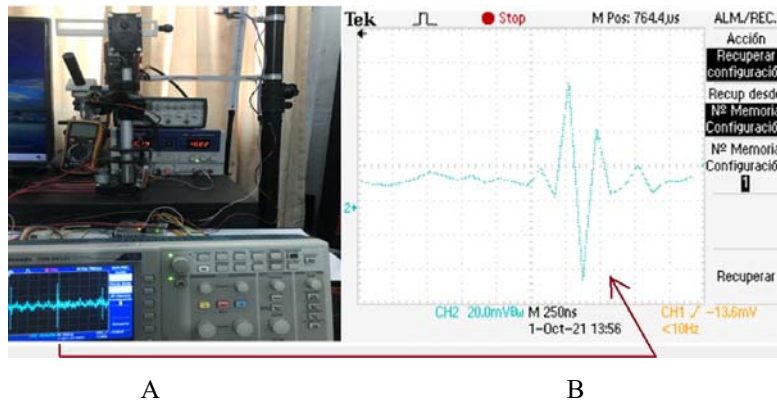
In the control system we need establish synchronization between revelation motors and geared motor considering pass to pass in microseconds of the electronic signal to establish the three movements, accord to the conic spiral trajectory. Likewise, we consider in the fine measurement of torsion a monolithic circuit L298 integrated in two pockets multi-watt and PowerSO20 in fifteen derivations. This is a controller of complete H-Bridge of high voltage and high current, designed to control revelators, solenoids, CC and pass to pass motors. For this relegation functions are required rectifications of voltage in each pass or step. Then for whole functioning are considered in each step the voltage rectangle pulses with certain delay  $\Pi_1\left(\frac{t-\tau}{L}\right)$  [9]. This realizes a voltage rectification to the synchronicity of the three movements in the magnetic sensing (see Table 1).



**Fig. 3.7.** Real magnetic sensing surface. This represents the magnetic sensing from top to bottom to a conditioning of voltage signal. It is not 2-dimensional spectral density surface of torsion. In the top of the first undulation, the intensity of magnetic sensing is high. This goes decreasing conform goes downing (second ondulation). This variation is due to the proximity of dilation to the Hall sensor.








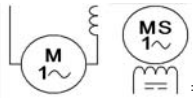
**Fig. 3.8.** 2-dimensional Spectral torsion signal given by (3.18). The spectral torsion force is given in the vertical direction, Z-axis. The energy agitation can be viewed in lateral undulations to the top in the surface.



**Fig. 3.9.** A. Hall sensor devise to detect torsion to conic trajectory. B. Evidence torsion signal. Here the output signal show in the display is enlarged in period.



**Table 1.** Electronic gauges and components to detect the torsion in variable regime.

Regime	Electronics component and Dynamical System	Associated Functions to the Torsion
Variable		$\pm \delta(t_1, t_2)$ <p>Impulse</p>
Variable		$H = -\frac{I}{4\pi\gamma} \frac{d\Sigma}{dx} = \frac{I}{2\gamma} \frac{a^2}{\left(\sqrt{a^2 + x^2}\right)^3}$ <p>Total Magnetic Field</p>
Variable		<p>Vertical/Horizontal Displacements:</p> $\Sigma = 2\pi \left[ \frac{a^2}{x^2 + a^2 + x\sqrt{x^2 + a^2}} \right]$
Variable		$\Pi_1 \left( \frac{t - \tau}{L} \right)$ <p>Rectifications</p>
Constant/ Variable		<p> <math>V = 12\text{Volts},</math>  <math>\nu = 55\text{RPM},</math>  <math>I = 0.2\text{A},</math>            Motor Length = 38mm   = 12 Kgcm  <math>P = 2.4\text{watts},</math>            Motor Diameter: 27.7 mm            Output Axis Diameter: Coupled            Output Axis Length: Coupled         </p>

### 3.5. Conclusions

We consider on the base of experiments for the measurement of torsion as evidence of the gravitational waves existence, a complementary electronic experiment, as part of another electronic experiment, to measure and detect al field torsion inside the framework of the gravitation theories [10], but with a modern focus using the curvature energy obtained in previous researches and consigned, and published in different physics and electronics journals [11].

However, the limitations of the our purely electronic devices only let see and interpret using the arguments of geometry, certain traces of electronic signals of the torsion evidence considering a magnetic field determined to certain voltage range and a movement of conic trajectory, searching generalize the case studied when is considered constant torsion to the variable torsion case (see the electronic components in the Table 1). However, this verifies the Conjecture 2.1, of this chapter and some theorems established in other studies in theoretical physics and mathematical physics. Likewise, using a Hall type sensor is detected the variation of the magnetic field produced from a magnetic dilation [9], which is moved along a conic spiral where the matter agitates the space and from which emanates gravity (from the kinematic tensor theory in duality with spinor waves), also interacts with the magnetic field of the proper dilation. The methods and results of the research are on themes parallel and related to the gravity (no gravity precisely) considering this method as analogous to detect gravity waves, but in this case detect waves of torsion in an indirect way.

### Acknowledgements

The electronics work has been very fine and we are very grateful with the help given for Rene Rivera-Roldán, Elec. Eng., Chief of Electronics Engineering Division of TESCHA, who supported with measurement equipment in circumstances of Covid-19, for the research.

### Technical Notation

$\kappa(\omega_1, \omega_2)$  – Spectral curvature or curvature energy. This is given in  $V / m^3$  [2].

$\mathbb{R}^3$  – Real ordinary space or 3-dimensional real space.

$\mathcal{L}_{\text{Hall}}^H$  – Topological vector space that represents the Hall type sensor space (or sensing space).

$\tau(\omega_1, \omega_2)$  – Torsion energy or spectra of torsion. Also given in  $V / m^3$  [1].

$\omega$  – Angular frequency given for  $\omega = \frac{2\pi}{T}$ ,  $T > 0$ .

$s(\omega_1, \omega_2)$  – Spectral spinor surface. This in the quantum case models the gravitational waves related with the field torsion.

$C(p)$  – Conic space of all conic spiral trajectories. This is defined as tangent subspace.

$\Omega_H$  – Vorticity field of all spiral trajectories of the inverted cone (vortex). We use an element of this field to analyse the torsion as observable. This vorticity field is conformed by the magnetic field.

$\gamma$  – Experimental constant. Taking as unit the magnetic momentum of the proton, we have that:  $\gamma = 5.64 \pm 2 \times 10^{31} \frac{\text{proton momentum}}{\text{Volt} \times \text{meter} \times \text{sec}}$ . In our research we are considering 1.

$H$  – Intensity of magnetic field.

$\Pi\left(\frac{t}{T}\right)$  – Rectangular pulse of period  $T > 0$ .

## References

- [1]. Dr. F. Bulnes, J. C. García-Limón, L. A. Ortiz-Dumas, V. A. Sánchez-Suarez, Detector of torsion as field observable and applications, *American Journal of Electrical and Electronic Engineering*, Vol. 8, Issue 4, 2020, pp. 108-115.
- [2]. F. Bulnes, I. Martínez, O. Zamudio, Fine curvature measurements through curvature energy and their gauging and sensing in the space, in *Advances in Sensors Reviews* (S. Y. Yurish, Ed.), Vol. 4, *IFSA Publishing*, Spain, 2016.
- [3]. F. Bulnes, I. Martínez, R. Cayetano, I. Valencia, B. A. Martínez, J. M. Beltrán, Curvature as obstacle to a photo-resistor sensor of illuminating and their minimal sensing region, *Sensors and Transducers*, Vol. 224, Issue 8, August 2018, pp. 51-62.
- [4]. F. Bulnes, I. Martínez, R. Cayetano, A. Rodríguez, I. M. Martínez, Curvature as obstacle to a photo-resistor sensor of illuminating and their minimal sensing region part II: Their transitory analysis and the non-dimensional  $\lambda$ , *International Journal of Sensors and Sensor Networks*, Vol. 6, Issue 4, 2018, pp. 43-52.
- [5]. F. Bulnes, J. C. García-Limón, Magnetic-dynamic system for torsion detection using perturbation of energy-space, <https://youtu.be/0s5H8w6JfyA>
- [6]. T. M. Boone, L. Folks, J. A. Katine, E. E. Marinero, N. Smith, B. A. Gurney, Magnetic sensitivity in mesoscopic EMR devices in I-V-I-V configuration, in *Proceedings of the Physics 65<sup>th</sup> Annual Device Research Conference*, 2007, pp. 247-248.
- [7]. M.-A. Paun, J.-M. Sallese, Hall effect sensors design, integration and behavior analysis, *Journal of Sensor and Actuator Networks*, Vol. 2, Issue 1, 2013, pp. 85-97.
- [8]. B. A. Gurney, J. R. Childress, M. S. Carey, R. Fontana, The evolution of read back sensors for magnetic recording, in *Proceedings of the 61<sup>st</sup> IEEE Device Research Conference*, 2003, pp. 107-110.
- [9]. F. Bulnes, J. C. García-Limón, V. Sánchez-Suárez, L. A. Ortiz-Dumas, M. Y. Flores-García, A. U. García, Galicia: International Research Group in Mathematics and Electronics, [https://youtu.be/mzcLnIO\\_hTE](https://youtu.be/mzcLnIO_hTE)

- [10]. F. Bulnes, Detection and Measurement of Quantum Gravity by a Curvature Energy Sensor: H-States of Curvature Energy, <https://www.intechopen.com/chapters/54948>
- [11]. F. Bulnes, Y. Stropovscky, I. Rabinovich, Curvature energy and their spectrum in the spinor-twistor framework: Torsion as indicium of gravitational waves, *Journal of Modern Physics*, Vol. 8, 2017, pp. 1723-1736.



## Chapter 4

# Optical Waveguides for Integrated Photonics: Mainstream Technologies Versus Low-cost Silica-titania Sol-gel Approach

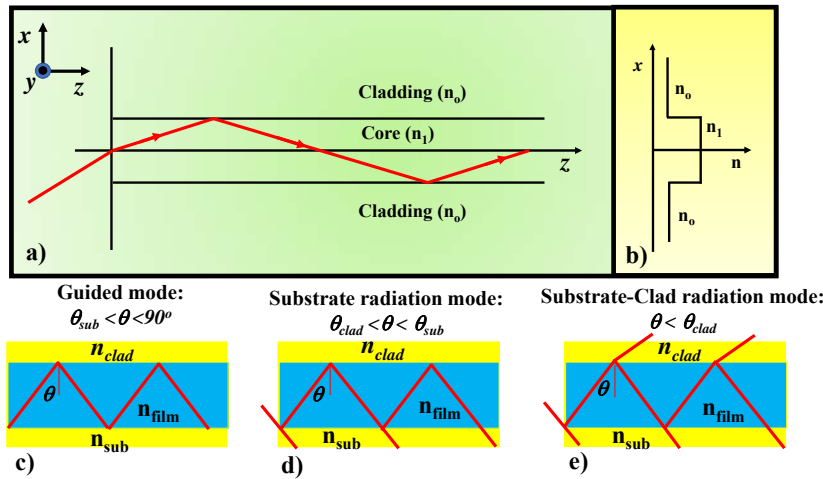
Muhammad Ali Butt, Andrzej Kaźmierczak, Stanisław Stopiński, Anna Jusza, Paweł Karasiński and Ryszard Pyramidowicz

### 4.1. Introduction

In general terms, a waveguide (hereafter represented as WG) can be described as any structure that guides the electromagnetic (EM) wave in a defined direction and within a defined wavelength range. Optical WGs may be classified with respect to their geometry, mode profile, refractive index (hereafter represented as RI) distribution, and material. The optical WG core is a longitudinally stretched high-index medium that is transversely enclosed in a low-index medium referred to as the cladding as shown in Fig. 4.1a. In the WG, an EM-wave travels in the longitudinal direction. The attributes of a WG are determined by the transverse profile of its dielectric constant ( $x, y$ ), that is independent of the  $z$ -coordinate. For a WG composed of optically isotropic material, the WG can be described simply with a single spatially reliant transverse refractive index (RI) profile,  $n(x, y)$ , as shown in Fig. 4.1b. In general, WG structures can be categorized into two types based on their RI profile i.e., step-index and graded-index. A WG structure in which the index profile changes rapidly between the core and the cladding is referred to as step-index WG [1], while in the graded-index WG [2], the index profile varies gradually. Most optical WGs support modes of two separate polarizations, with either the main electric (quasi-TE) or magnetic (quasi-TM) field component along the transverse (horizontal) direction. Typically, the optical WGs have to provide a single-mode operation for every polarization. This can be achieved by decreasing the size of the WG core until the point where the fundamental modes start to radiate [3].

Guided modes are referred to as the optical modes with an effective refractive index ( $n_{eff}$ ) greater than the largest cladding index, i.e.,  $\theta_{sub} < \theta < 90^\circ$  (see Fig. 4.1). The wave interferes with itself when it is reflected back and forth between the two interfaces. A

guided mode can only occur if a transverse resonance condition is met, resulting in constructive interference between the repeatedly reflected waves as shown in Fig. 4.1c. The modes with lower index radiate resulting in the leakage of the optical power to the cladding region. Substrate radiation modes can be classified as  $\theta_{clad} < \theta < \theta_{sub}$ . The total reflection arises just at the upper interface, which resulted in the refraction of the incident wave at the lower boundary from either the core or the substrate as shown in Fig. 4.1d. In the radiation modes, there is no total reflection at either interface, i.e.,  $\theta < \theta_{clad}$ . In both interfaces of the WG, the incident wave is refracted, and it can transversely reach infinity as shown in Fig. 4.1e. In evanescent modes, the field decays exponentially along the z-direction. The energy of an evanescent mode is emitted out from the WG transversely in a lossless WG structure.



**Fig. 4.1.** a) Fundamental structure of the optical WG, b) RI profile of the symmetric step-index slab WG. Light guiding conditions, c) guided mode, d) substrate radiation mode, e) substrate-clad radiation mode.

Optical communication and integrated photonic circuits strongly depend on the efficient transmission of EM waves. The use of a variety of dielectric, semiconductor, and polymer materials with excellent optical transparency is important for confining and guiding EM-wave with the minimal loss [4-6]. Especially polymer materials are desirable for manufacturing planar photonic integrated circuits (PICs) as they are less expensive than inorganic or semiconductor-based systems while allowing large-scale integration. Optical elements based on polymer platforms can be patterned and immediately combined with electrical controls using traditional microfabrication techniques. In recent times, several novel WG structures have been developed for different applications, among which biochemical sensing is one of the most intensively studied [7-9]. Glass WGs have also extremely attractive properties due to their simple technological processing, low optical losses, and compatibility with the glass fibers. Therefore, low-loss glasses, as well as reliable and low-cost WG fabrication processes, are highly desired. Manufacturing systems must be proficient in high yield, guaranteed duplicability within quantifiable

tolerances, and inherently low operating costs. It is possible to categorize the most prevalent WG manufacturing methods in terms of glass integrated optics. The guiding layer must have a greater RI than the substrate and cladding to provide effective light confinement. As a result, two methods of guiding layer fabrication can be used: thin film deposition and regional alteration of the RI of the bulk material. The thin film layer deposition procedures include processes like RF-sputtering and magnetron sputtering, chemical vapor deposition (CVD), plasma-enhanced CVD, flame hydrolysis deposition (FHD), spray pyrolysis deposition, pulsed laser deposition, and sol-gel coating. Ion exchange, ion implantation, and UV-irradiation may be used to execute local bulk material alteration, femtosecond (fs) laser writing is also appropriate for the direct inscription of a channel WG in the substrate.

## 4.2. Waveguide Structures

In the past few decades, several WG structures have been proposed and developed using different technological platforms for many eye-catching applications. Some of the WG constructions are quite typical, such as ridge, rib, and diffused WGs, which can carry standard dielectric WG modes. These WG structures can strongly confine the light, which is interesting for guiding purposes. However, for sensing applications, highly sensitive WG constructions are desired, in which light-matter interaction is strong. This led to the discovery of slot, hybrid, subwavelength grating (SWG), and plasmonic WGs. The geometries of the WGs discussed in this chapter are shown in Fig. 4.2. The chapter contains the fundamental description and discussion of these WG's structures only. The detailed analysis can be found in [10, 11].

### 4.2.1. Conventional Waveguide Structures

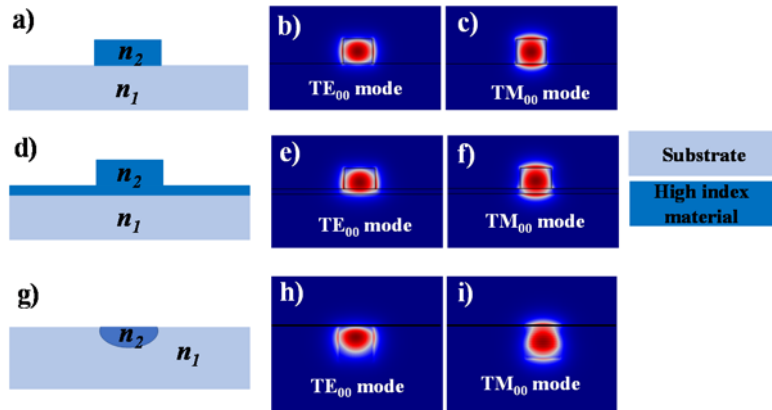
As it has been explained the conventional waveguides are typically deployed to the optical signal low-loss transmission without exposing the guiding light to the external environment. Three basic WG structures have been widely studied and used for data transmission in optical interconnects for on-chip communication systems. These WG structures include ridge, rib and diffused WG, which are highly attractive due to their strong light confinement.

#### 4.2.1.1. Ridge Waveguide

The ridge WG comprises a high index material core placed on a low index material substrate. The single-mode operation is typically required; hence the core should be properly designed to support only the fundamental mode. The principal requirement of the ridge WG is the guiding a single  $TE_{00}$  and  $TM_{00}$  mode. The mode is guided in the ridge WG structure when the  $n_{eff}$  is higher than the cladding and smaller than the core, otherwise, it will be radiated into the substrate. Furthermore, the higher the effective index, the stronger the mode confinement is. The rib WG geometry is shown in Fig. 4.2a. The E-field distribution of the fundamental TE and TM mode at  $\lambda = 1000$  nm guided by the rib WG (~core index 3.57) is shown in Fig. 4.2b and 4.2c, respectively. The numerical



simulations were conducted with the use of the 2D-finite element method (2D-FEM) implemented in the COMSOL environment.



**Fig. 4.2.** Conventional WG structures, a-c) Ridge WG geometry and fundamental TE and TM mode field distribution, d-f) Rib WG geometry and fundamental TE and TM mode field distribution, g-i) Diffused WG geometry and fundamental TE and TM mode field distribution.

#### 4.2.1.2. Rib Waveguide

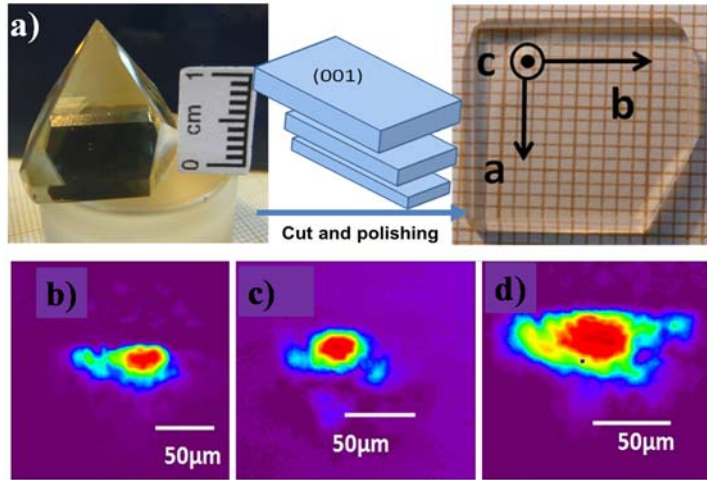
The rib WG is a special case of ridge WG where the high index material layer is not entirely etched. The guiding layer consists primarily of the slab with a ridge placed onto it [12]. Silicon-on-insulator (SOI) ridge WGs with sizes bigger than a few nanometers provide support for multiple modes that makes them highly unappealing to be integrated into the photonic integrated circuits (PIC). Though, there are some efforts to further develop monomodal large rib WGs [12]. One advantage of using large rib WG is that they can fairly enable the low-loss coupling to the optical fibers. Fig. 4.2d presents the schematic representation of the rib WG. While the E-field distribution of the fundamental TE and TM mode is shown in Figs. 4.2e and 4.2f, respectively. As previously, the numerical simulations were conducted via the 2D-FEM in the COMSOL software.

#### 4.2.1.3. Diffused Waveguide

Diffusion of appropriate dopants into a substrate might create a higher-index region, thus defining a diffused WG. The diffused WG structure is shown in Fig. 4.2g. The E-field distribution of the fundamental TE and TM mode at 1000 nm is shown in Figs. 4.2h and 4.2i, respectively. The numerical simulations were conducted via the 2D-FEM of COMSOL.

In [2], Butt et al. demonstrated the straight WG, S-bend and Mach-Zehnder (MZ) structure on  $\text{RbTiOPO}_4$  (RTP) crystal fabricated by  $\text{Cs}^+$  ion-exchange method at Universitat Rovira i Virgili, Spain.  $\text{KTiOPO}_4$  (KTP) and its isostructural RTP are renowned crystalline dielectric materials with superb attributes for non-linear optical applications. The bulk

RTP crystal and a c-cut and polished RTP (001) oriented substrate are presented in Fig. 4.3a. The near field mode intensity profile obtained at the output of the straight, S-bend and MZ structure is presented in Figs. 4.3b, 4.3c, and 4.3d, respectively. For a detailed ion exchange process in the RTP crystalline WG, we refer the readers to this work [2].



**Fig. 4.3.** a) (Yb, Nb): RTP/RTP (001) epitaxial layer [2]. Near field mode intensity profile of, b) Straight WG [2], c) S-bend [2], d) MZ-structure [2].

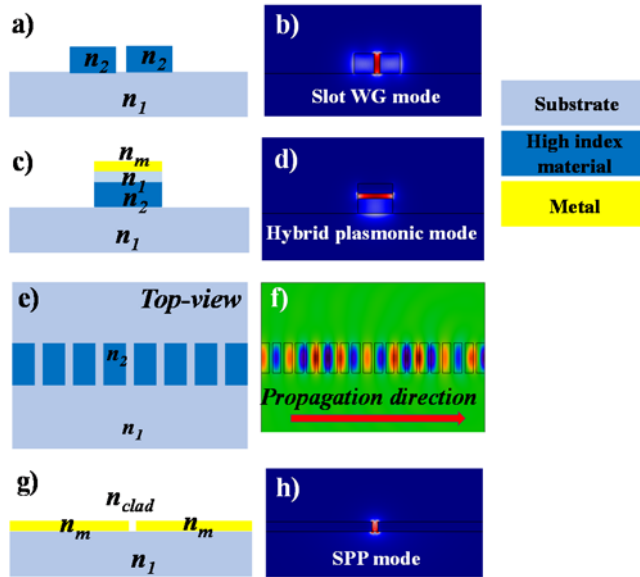
## 4.2.2. Unconventional Waveguide Structures

In this section, the four most widely studied unconventional WG structures such as slot, hybrid plasmonic, subwavelength grating and purely plasmonic WGs are presented. These structures are highly sensitive to the ambient medium. This distinct feature makes them an ideal candidate for sensing applications. Below, we provide a general introduction and fundamental knowledge about the unconventional WG structures and the E-field distribution of the propagating mode.

### 4.2.2.1. Slot Waveguide

In 2004, the first slot WGs were developed by Lipson's group. This particular WG arrangement is distinct from conventional WGs, in which light is guided in a high-index material by means of the total internal reflection (TIR) [13]. In a slot WG the light is guided by a low-index material slot that is just a few hundred nm wide, sandwiched between two high-index material strips. The interaction between eigenmodes of the two strips of high-index material results in the mode propagation in the slot WG, which is theoretically lossless and is a true eigenmode. The slot WG mode intensity is an order of magnitude higher than in a conventional WG, working on the TIR principle. In recent years, slot WG structures have been employed in several sensing applications [14, 15]. The slot WG geometry and the E-field distribution of the TE-polarized light calculated

for 1000 nm wavelength is shown in Figs. 4.4a and 4.4b, respectively. The numerical simulations were conducted using the 2D-FEM in the COMSOL software.



**Fig. 4.4.** Unconventional WG structures, a-b) Slot WG geometry and E-field distribution of the propagating mode, c-d) Hybrid plasmonic WG geometry and E-field distribution of the propagating mode, e-f) SWG WG structure and E-field distribution of the Bloch mode, g-h) Plasmonic WG geometry and SPP wave distribution.

#### 4.2.2.2. Hybrid Plasmonic Waveguide

The hybrid plasmonic WG (HPWG) structure has been developed to merge the characteristics of the dielectric and plasmonic WG to accomplish the high light confinement without undergoing a large loss (typical for plasmonic WGs) [16]. HPWG structures have lately drawn significant research interest, as they enable creating the next-generation and high-density integrated photonic circuits due to the long-distance plasmonic wave transmission with the sub-wavelength propagation mode size. These WG structures are made of a thin layer of low refractive index material sandwiched between the metal and a higher index dielectric layer. Unlike standard dielectric WGs where light is confined in the high index medium (core), in the case of HPWG, a substantial amount of light is confined in the low index spacer layer.

Owing to these desirable characteristics, these WGs have drawn a lot of attention, and several different applications of HPWG have been suggested, for instance, nanoscale lasers, Bragg gratings, ring resonators, and parametric amplifiers [17]. Moreover, they are compatible with Si photonics technology and can be integrated with Si WGs on the same chip. The schematic representation of the HPWG and the E-field distribution of the hybrid mode at 1000 nm is shown in Figs. 4.4c and 4.4d, respectively. The numerical simulations were conducted via the 2D-FEM in the COMSOL simulation tool. These WG structures

are highly sensitive to the ambient medium which make them most suitable for sensing applications as demonstrated in our previous works [17, 18].

#### 4.2.2.3. SWG Waveguide

Subwavelength gratings (SWGs) can be utilized to establish a microscopic-scale synthetic media with the desired macroscopic behavior [19]. For example, the periodic arrangement of high index material (e.g., silicon) and low index material (e.g., silica) may be used to define an SWG WG for the SOI platform, as shown in Fig. 4.4e. If the periodicity does not obey Bragg's coupling condition to other confined or radiative modes, such periodic structures conquer diffraction and act as a uniform medium. Antireflective coatings, planar mirrors, broadband mirrors, sensors, fiber-chip couplers, and modulators are only a few of the applications for SWG WGs.

In these WG structures, which has a core made of a periodic arrangement of Si and SiO<sub>2</sub> segments, light propagates in the form of Bloch modes, potentially travelling through the periodic WG segment with no diffraction losses. The  $n_{eff}$  of the medium can be locally devised by managing the pitch, width, and duty cycle of the SWG WG. Bloch-Floquet formalism depicts the propagation of EM waves in periodic media. For a given grating period ( $\Lambda$ ), propagation can be classified into three wavelength regions:

- (i) The sub-wavelength region in which the wavelength to period ratio is  $\lambda/\Lambda > 2 \cdot n_{eff}$ . This associate with a wavelength range longer than  $\lambda_{Bragg}$  for which the WG behaves as a conventional WG. In this situation, the periodic structure maintains a true lossless mode. The H<sub>z</sub>-field distribution of the Bloch mode is shown in Fig. 4.4f.
- (ii) Photonic bandgap (PBG) region, where Bragg reflection takes place, and no transmission occurs.
- (iii) The wavelength range smaller than the Bragg wavelength, where the Bloch wave is leaky, a portion of the energy is emitted out of the WG. The transmission loss is governed by reflection and diffraction at the segment borders produced by the high index change between air and Si.

There is a detailed study on the sensing application of ridge SWG WG, slot SWG WG and hybrid plasmonic SWG WG, presented in our previous works, which may be recommended as complementary reading [8, 14, 18].

#### 4.2.2.4. Plasmonic Waveguides

Surface plasmon polaritons (SPPs) are EM waves connected to the oscillations of conduction electrons propagating along with the metal-dielectric interface with an exponentially decaying field in the direction perpendicular to the interface. SPPs have been lauded as one of the most interesting energy carriers for resolving light's traditional diffraction limit and controlling light on a subwavelength scale. These characteristics pave the way for ultracompact light-guiding structures that permit subwavelength optical mode

confinement. As a result, several attractive optical components such as splitters, polarizers, sensors, (de)multiplexers might be developed.

Different types of plasmonic WGs have been suggested in recent years, comprising metal films, metal grooves, metal strips, metal wedge, insulator-metal-insulator (IMI) WG, and metal-insulator-metal (MIM) WG. These WGs have demonstrated a significant ability to direct subwavelength optical modes. The MIM structure is one of the most effective plasmonic WG structures that permits light to propagate on the nanoscale in the visible and near-IR ranges. A detailed study on MIM WG and its applications is presented in our review papers [20, 21]. These WGs are the plasmonic structures in which an insulator is encircled by two metal claddings as shown in Fig. 4.4g. The E-field distribution of the plasmonic mode at 1000 nm is shown in Fig. 4.4h. The numerical simulations were conducted via the 2D-FEM in the COMSOL simulation tool. In recent times, MIM WG structures have been widely studied for biochemical sensing applications [22-24].

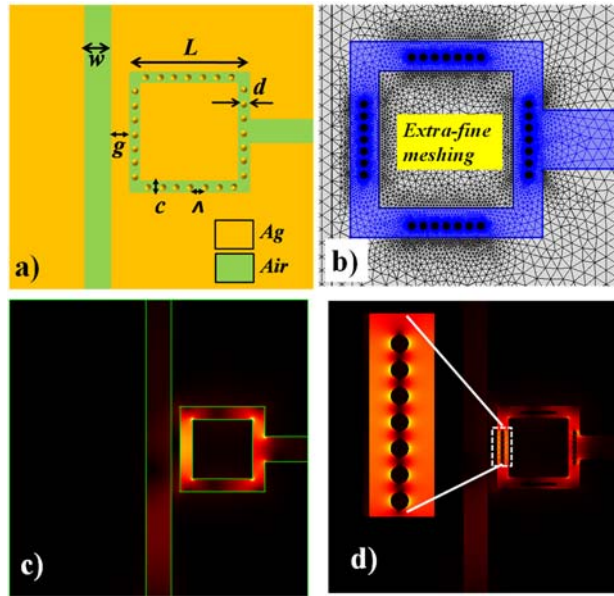
In our recent work [25], an attractive design of a MIM WG-based plasmonic sensor for RI sensing application was numerically investigated. To enhance the sensitivity of the plasmonic sensor, nanodots have been periodically placed in the square ring cavity as shown in Fig. 4.5a. This allows the intensification of the SPP wave at  $\lambda_{\text{res}}$  resulting in strong light-matter interaction. The COMSOL Multiphysics tool was used to simulate the sensor design. The extra-fine meshing built-in option has been applied to divide the device design into small finite elements as displayed in Fig. 4.5b. The E-field mapping in the standard and nanodots loaded cavity has been shown in Figs. 4.5c and 4.5d, respectively. A detailed description of the device performance can be found in [25]. This study was conducted at Samara National Research University, Russia.

### 4.3. Mainstream Photonic Integrated Circuits (PIC) Technologies

In the last three decades, rapid progress has been observed in the technology of photonic integrated circuits (PICs), benefiting directly from the potential of mature technologies used in the microelectronic industry. As a result, several market-ready PIC-based solutions have been developed addressing mainly telecom and sensing applications. Up to now, only two photonic technological platforms have reached maturity level allowing commercial use, namely silicon-on-insulator (SOI) and indium phosphide (InP), with increasing presence of silicon nitride (SiN) platform. These three technologies offer complementary features and might be considered as the mainstream of nowadays PIC technology. Below, the major features of each of these technologies are described.

#### 4.3.1. SOI WGs

The omnipresence of silicon (in the form of Silicon-on-Insulator: SOI) in integrated electronic CMOS circuits technology caused the use of standard SOI wafers and standard Si processing techniques for the fabrication of integrated WGs a very attractive option. This technology has been developed successfully by numerous research groups and fabrication foundries since the late 1990s.



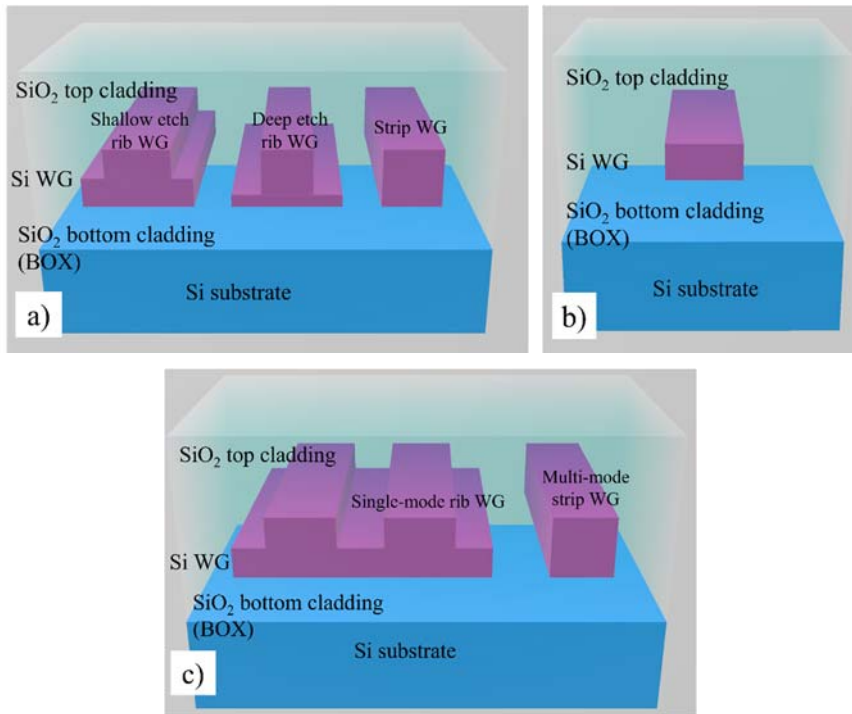
**Fig. 4.5.** a) Scheme of a MIM WG based square ring resonator (SRR) design with nanodots embedded in the cavity [25], b) Device model split into extra-fine triangular mesh elements [25], c) E-field mapping in the typical SRR at  $\lambda_{res}$  [25], d) E-field mapping in the nano-dots embedded SRR model at  $\lambda_{res}$  [25].

Taking into account a very high refractive index contrast between silicon ( $n_{Si} \approx 3.45$  for 1550 nm) and silicon dioxide ( $n_{SiO_2} \approx 1.45$  for 1550 nm) it is possible to fabricate densely integrated silicon WG structures having a width smaller than 500 nm and minimum bending radii smaller than  $2 \mu\text{m}$  [4]. This leads to possible dense integration of waveguide structures within SOI PIC providing numerous optical functionalities in a single chip.

A wide range of silicon WGs, of various geometries, can be fabricated using SOI PIC technologies. Three emblematic commercially available European SOI PIC technologies are compared hereafter to illustrate the possibilities of using SOI PICs. This includes close to square strip single-mode WGs in (i) 310 nm SOI technology offered by CEA Leti (France) [26] and single-mode rectangular strip WGs in (ii) 220 nm SOI technology offered by IMEC (Belgium) [4] as well as single-mode rib WGs in (iii)  $3 \mu\text{m}$  thick SOI technology offered by VTT (Finland) [27]. Schematics of the SOI WGs cross-sections offered by the above-mentioned foundries are shown in Fig. 4.6.

The SOI PIC platform offered by CEA Leti allows fabricating not only square single-mode strip WGs but also shallowly and deeply etched rib waveguides, while IMEC platform focuses on rectangular strip WGs. Both these platforms focus on dense integration of photonic components (single-mode WG width in the range of 300-400 nm and tight WG bends with a radius of the order of a few micrometers) and exploit the vertical optical signal coupling scheme between PIC and the optical fibers using vertical grating couplers (VGCs). The third platform (VTT) bases on a somehow different

approach, where tight integration is not so crucial, which in turn allows fabrication of WGs with a cross-section comparable to the diameter of the optical fiber. Consequently, the direct horizontal end-fire coupling to the optical fibers is possible (with the use of a vertical mirror when the vertical signal coupling is needed).



**Fig. 4.6.** Schematic representation of the SOI PIC WGs cross-section fabricated at, a) CEA Leti, b) Imec, and c) VTT.

Each platform has a wide range of developed integrated photonic devices such as straight and bent WGs, optical signal splitters (directional and MMI), ring resonators (both for sensing and modulation), arrayed waveguide gratings (AWG) and echelle grating multiplexers. Depending on the specific WG cross-section the propagation loss of SOI WGs ranges from 0.3 dB/cm (in the case of less dense integrated WGs) to 1.5 dB/cm (in the case of the most compact circuits).

The above-mentioned platforms offer three possible configurations of SOI PIC: (i) fully passive PIC including basic rib and strip silicon wave-guide structures, (ii) SOI PIC incorporating, apart from above mentioned passive components, also electrical connections and thermo-optic or PIN phase shifters (modulators) and finally (iii) the most advanced version incorporating all features of configuration (ii) with the integration of Ge photodiodes. It should be mentioned that the future development towards III-V active components hybrid integration is predicted by the manufacturers.

### 4.3.2. InP/InGaAsP WGs

The most significant advantage of the indium phosphide technology platform with respect to the development of photonic integrated circuits is its inherent capability of monolithic integration of active and passive components in a single chip [28]. This is a straightforward consequence of the direct bandgap of InP as well as ternary and quaternary alloys such as InGaAs and InGaAsP, which are used as active layers in detectors and amplifiers. However, the design and implementation of passive waveguides using this particular platform needs to take into account the geometry of consecutive epitaxial layers of InP/InGaAsP, which differ only slightly when the refractive index is of concern. Such a layer stack, which forms a diode from the electronic point of view, is necessary to achieve proper operation and functionality of amplifiers, electro-optic phase modulators and photodiodes. Therefore, achieving a high confinement factor in the vertical direction is not possible, which eventually results in a relatively large minimum WG dimensions and bending radius when compared to the silicon or silicon nitride platforms.

Simplified cross-sections of typical WG structures offered by two leading European foundries – SMART Photonics and Fraunhofer Heinrich Hertz Institute (HHI) in the framework of multi-project wafer runs are schematically depicted in Fig. 4.7 [29-31]. Generally, the indium phosphide substrate ( $n = 3.17$ ), which also forms the bottom cladding of the WG structures, is either n-doped or semi-insulating, depending on the specific technology. The guiding layer comprises a lattice-matched quaternary alloy – InGaAsP, with the energy band-gap equivalent to  $1.25 \mu\text{m}$  ( $n = 3.36$ ) or  $1.06 \mu\text{m}$  ( $n = 3.26$ ) for the SMART and HHI technology, respectively. Such a bandgap enables transmitting optical signals in the third telecom window, i.e., around  $1550 \text{ nm}$ . The top-cladding is formed by additional p-doped InP layers ( $n = 3.17$ ) in the SMART process, while by the air in the case of HHI.

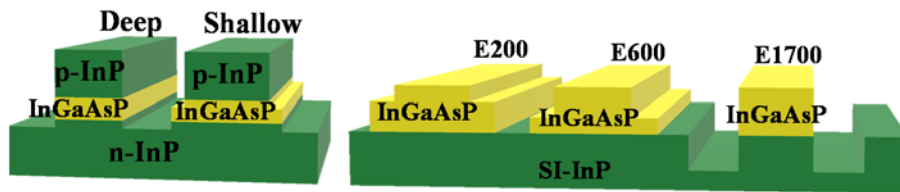


Fig. 4.7. Cross-section of typical InP WGs in the generic process of SMART Photonics (left) and Fraunhofer Heinrich Hertz Institute (right). The figure is based on [29-31].

The SMART Photonics platform provides two basic types of the WG cross-section. The first one is deeply etched, i.e., etching goes through the full guiding layer ( $500 \mu\text{m}$  thick) and beneath. In such a WG the light is strongly confined in the lateral direction, as the etched structure is surrounded by a polyimide with a low refractive index ( $n \approx 1.5$ ). This enables relatively tight bends – the minimum bending radius is of the order of  $100 \mu\text{m}$ . On the other hand, side-wall roughness causes significant scattering loss, which results in a total propagation loss coefficient of  $3 \text{ dB/cm}$ . The second cross-section type is shallowly etched. The etching depth of the InGaAsP layer is equal to merely  $100 \text{ nm}$ . Such a structure enables propagation with lower loss ( $L = 2 \text{ dB/cm}$ ), as there is no WG-polyimide



interface and thus scattering loss is strongly suppressed. However, the weak confinement of the optical field causes that the minimum bending radius is as high as 400  $\mu\text{m}$ .

Similar considerations apply to HHI WGs. This technology provides three different cross-sections denoted E200, E600 and E1700 (the number defines the etching depth in nanometers), which have weak, moderate, and strong field confinement, respectively. The attenuation coefficient is between 1.5 and 2.0 dB/cm, while minimum bending radii are 480  $\mu\text{m}$  (E600) and 250  $\mu\text{m}$  (E1700), for E200 the bends are not allowed in the generic process.

The basic function performed by InP/InGaAsP passive WGs in photonic integrated circuits is connecting with one another different active and passive elements of the device, e.g., amplifiers, phase modulators, photodiodes, multiplexers, or couplers. Secondly, the WG cross-section is used to realize more complex passive components, so-called composite building blocks, such as arrayed waveguide grating (AWG) demultiplexers, multi-mode interference (MMI) couplers, splitters and reflectors, ring resonator filters and others. Furthermore, slight modifications of the passive WG cross-section result in realizing new functional elements, such as distributed Bragg reflectors (by modifying the top-cladding layer), electrical isolation section (by removing some of the cladding layer), electro-optical phase modulator (e.g., by applying electrodes and InGaAs contact layers on top of the p-InP cladding), transition elements between two WGs with a different cross-section, WG crossings and others.

To summarize, the InP/InGaAsP passive WG technology platform offers a definite advantage of monolithic integration of all PIC components, however at the cost of higher loss and larger bending radii, resulting in slightly bigger footprints of the circuits, as well as an increased number of the technological processes required, when active circuits are of concern. The final performance parameters of the indium phosphide WGs are to some extent influenced by the monolithic technology approach, being a compromise between the necessity of providing simultaneously good waveguiding properties combined with the amplification, modulation, and detection functionalities as well.

### 4.3.3. SiN WGs

Another passive PIC technology platform bases on silicon nitride WGs. This technology can be seen as an auxiliary one to the two above mentioned. Considering a significantly lower refractive index of SiN ( $n_{\text{SiN}} \approx 2.1$ ) than in the case of silicon, this platform offers a substantially lower refractive index contrast between WG core and cladding. Consequently, the integration scale of SiN PICs is much smaller, which results in a larger WG cross-section and minimum bending radius. The biggest advantage of the SiN PIC platform lays in its operation wavelength spectrum, which covers the full visible (VIS) and NIR range. The biggest limitation of SiN PIC technology is its complete passiveness, without the possibility of monolithic integration of any active components (except thermo-optic phase modulators). Consequently, this PIC technology is often seen as a complementary solution for SOI and InP technologies. Indeed, SiN PICs are sometimes used as an interface layer between a SOI or InP chip and an array of optical fibers.

SiN PICs are available in two forms: (i) the conventional single-layer rectangular cross-section WGs, where most of the light energy is guided within the high-refractive-index region (similar to the ones fabricated in the SOI technology) and (ii) modified WG structure, where most of the energy is guided within the low-refractive-index region [32]. Configuration (ii) can have several different forms as shown in Fig. 4.8a including a rectangular shape of an outer section composed of SiN with the internal part of the WG composed of SiO<sub>2</sub>, a single thin SiN strip, two thin SiN strips separated by SiO<sub>2</sub> layer, or combination of thin and thick SiN strips separated by SiO<sub>2</sub> layer. The use of the configuration (ii) allows obtaining a lower WG effective refractive index and therefore wider single-mode WG cross-section, reducing both propagation loss and geometrical mismatch with optical fibers.

The SiN PICs basing on the (ii) approach has been commercialized by LioniX (The Netherlands) under the brand of Triplex. Exemplary schematics representation of SiN integrated WGs cross-sections are shown in Fig. 4.8b. This technology is also being developed at Warsaw University of Technology, by the teams of Center for Advanced Materials and Technologies CEZAMAT and Institute of Microelectronics and Optoelectronics [33].

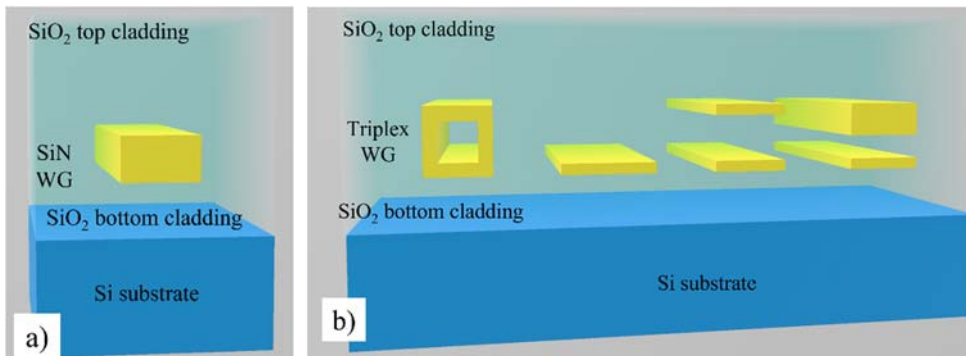


Fig. 4.8. Schematic representation of different SiN WGs cross-section (a) conventional cross section, (b) different WG cross-section available in Triplex platform.

It is worth noting, that this technology is also being developed at Warsaw University of Technology, by the teams of Center for Advanced Materials and Technologies CEZAMAT and Institute of Microelectronics and Optoelectronics, including co-authors of this work [33].

#### 4.4. Silica-titania Waveguides Developed Through the Sol-gel Dip-coating Approach

Ebelmen (a French chemist) is credited with developing a sol-gel synthesis of silicon tetra isoamyl oxide from silicon tetrachloride and isoamyl alcohol in the 1800s. The latter study involved the synthesis of boron amyloxyde, boron ethoxyde, and boron methoxyde with boron trichloride using isoamyl alcohol, ethanol, and methanol, respectively [34]. The

sol-gel procedure is a wet chemical method used to produce thin-film coatings. Its primary advantages are the procedure's low/lower/mitigated overall cost in comparison to more traditional procedures such as CVD and PVD, as well as the capability to tailor the composition and properties of thin films to meet the needs of the expected applications [35]. To be used effectively in integrated optics, thin-film coating techniques must meet the criteria of complete control over the thickness of the deposited film. As a result, thickness management is necessary for all thin-film development processes, including sol-gel. The ultimate layer thickness is largely determined by coating speed, angle of inclination, and sol concentration. Furthermore, the final heat-treated layer thickness can be influenced by sol viscosity, density, and liquid-vapour surface tension [36]. To obtain high optical quality sol-gel thin films, the coating process must be carried out in a cleanroom environment [37]. A three-step sol-gel process for producing organic dye-doped thin films with tailored porosity has been devised for applications in chemical sensing and optoelectronics [38]. Apart from that, sol-gel derived ceramic films are introduced in [39]. More substantive sol-gel method research can be found in these works [40-43].

The sol-gel method uses low-cost materials and equipment to fabricate simple integrated optical circuits, which makes it an appealing alternative to other technologies. Since channel WGs are a fundamental component of integrated optics, upgrading their fabrication technology is critical for manufacturing high-performance devices including splitters, couplers, deflectors, and amplifiers. For a long time, inorganic sol-gel derived materials have been studied extensively, with sol-gel materials like silica and titania being researched for optical applications. However, since the heat treatment temperature is usually very high (~1270 K), forming a thick layer of pure inorganic sol-gel film is challenging. The maximum thickness of a layer created by a single deposition is approximately 200 nm to 300 nm, and thicker layers tend to crack during the drying stage due to shrinkage. This restriction does not, however, prohibit the creation of multilayer films in which each layer is spun on and dried before the next is coated. However, inhomogeneities, defects and dust can be added between each layer, affecting the optical properties of the ultimate WG film. Merging the properties of different materials and producing molecular scale composite materials via sol-gel processing is an attractive way to resolve these issues. The use of organically modified silane precursors, which can create a thick single coating layer, is a great example. Another choice that has recently been introduced is to use the same sol-gel technique to integrate organic molecules into the inorganic matrix. Since the bulky organic components fill the pores between the inorganic oxide chains, there is absolutely no shrinkage when organic groups are integrated into the glass.

The sol-gel is a chemical method for creating optical materials from a liquid phase which offers complete control over the composition of the manufactured material [44]. Nevertheless, the evaporation of solvents during the heating process is causing the material to shrink during the layer fabrication process via sol-gel. This could result in the formation of inhomogeneity in the composition of the material used to create the optical layers and the induction of mechanical strain. Consequently, the RI of the entire layer may be non-uniform due to the differences in the reactivity of the used precursor, a separated

phase of one of the components (e.g.,  $\text{TiO}_2$  in this case of  $\text{SiO}_2\text{:TiO}_2$  material) can develop when layers are coated from multi-component materials. The resulting optical layers would be polycrystalline rather than amorphous, disqualifying it as a WG layer. We have been successful in resolving these critical technical issues. We are now able to deposit amorphous WG films with a high RI and low propagation losses [45]. By comparing the reflectance spectra of a substrate without a thin-film, optical homogeneity can be determined. Many scientists have used reflectance spectra to assess the homogeneity of the dielectric layer over the past [46, 47]. Each reflectance spectrum had to have at least three minima and three maxima, according to the methods referred to in the literature.

Because of the easy tunability of the RI and propagation losses of  $\sim 0.2$  dB/cm, silica-titania is an attractive glassy system for several applications, particularly planar WGs. That's why this technology is being developed and validated within the research project "Hybrid sensor platforms of integrated photonic systems based on ceramic and polymer materials" funded by the Foundation for Polish Science. The system bases on sol-gel derived  $\text{SiO}_2\text{:TiO}_2$  waveguide films technology, mastered at the Silesian University of Technology in the group of P. Karasiński. In this approach, the dip-coating method is used to form silica-titania optical WG layers with a  $\text{SiO}_2\text{:TiO}_2=1:1$  molar ratio on a BK7 glass substrate, which is then heated to  $500^\circ\text{C}$ . The main reagents providing as precursors of  $\text{SiO}_2$  and  $\text{TiO}_2$  are tetraethyl orthosilicate  $\text{Si}(\text{OC}_2\text{H}_5)_4$  (TEOS) and tetraethyl orthotitanate  $\text{Ti}(\text{OC}_2\text{H}_5)_4$  (TET), respectively. Water, ethanol, and hydrochloric acid are also used in the process to catalyze the hydrolysis and condensation reactions. Fig. 4.9 depicts a graphical image of the fabrication technique along with the images of substrates and reaction materials.

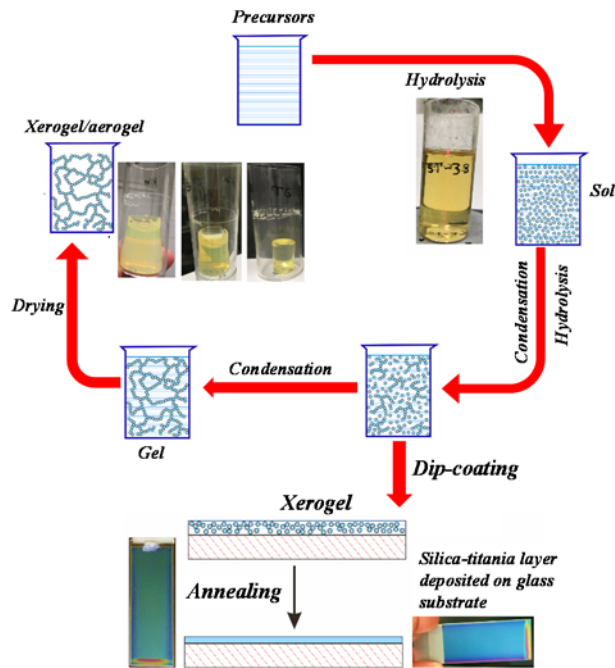


Fig. 4.9. Silica-titania WG development via the sol-gel dip-coating method.

## 4.5. Future of Silica-titania Waveguide Optical System

Optical WGs are integral part of PIC and the success of any optical WG system to be utilized in PIC can be assessed/evaluated using the following factors:

- Optical, physical, and chemical properties of the guiding layer;
- Cost of manufacturing;
- Durability;
- Ease of fabrication;
- Ease of integration (hybrid systems);
- Range of applications.

In this section, a few major physical, chemical, and optical properties of the silica-titania platform are compared with the existing WG technologies such as SOI, InP and SiN. These WG systems, considered as matured, have been commercially employed in several attractive applications such as telecommunication, sensors, and MEMS, among others. The silica-titania WG platform is relatively new, however, is displaying several attractive attributes such as low optical losses, low-cost, ease-of-fabrication and tunability of refractive index, which can solidify its way to compete with the existing expensive WG platforms in specific areas. The general comparison of SiO<sub>2</sub>:TiO<sub>2</sub> with SOI, InP and SiN platforms is shown in Fig. 4.10.

	SOI	InP	SiN	SiO <sub>2</sub> :TiO <sub>2</sub>
Refractive index	3.42	3.4	2	1.81-2.2
Spectral range (um)	1.1-6.5	NIR	VIS-NIR	VIS-NIR
Technological maturity	High	High	High	Under development
Fabrication method	Wafer bonding	LP MOCVD	LPCVD	Sol-gel
Implementation cost	High	High	High	Low
Propagation loss (dB/cm)	<0.1	>0.4	<0.1	~0.1
Cost efficiency	Very high	Moderate	Moderate	Very high
Available integration scale	Very high	Very high	Moderate	Moderate
Refractive index tailoring	No	No	Yes (only for SiO <sub>x</sub> N <sub>x</sub> )	Yes (1.2-2.2)
Chemical resistance	Low (tendency to oxidation)	Low (tendency to oxidation)	Moderate (tendency to oxidation)	Very high
Applications	Telecommunication, MEMS, Sensors	Telecommunication	Telecommunication, MEMS, Sensors	Sensors, special applications

**Fig. 4.10.** Comparison of physical, chemical, and optical properties of SiO<sub>2</sub>:TiO<sub>2</sub> platform with other optical technologies.

## 4.6. Examples

In this section, three basic and vital optical components based on the silica-titania platform are designed via finite element method (FEM). This will provide brief evidence to the readers about the capability of this platform to be utilized in photonic integrated circuits.

### 4.6.1. Numerical Modelling of Bandstop Filters Based on Silica-titania Platform

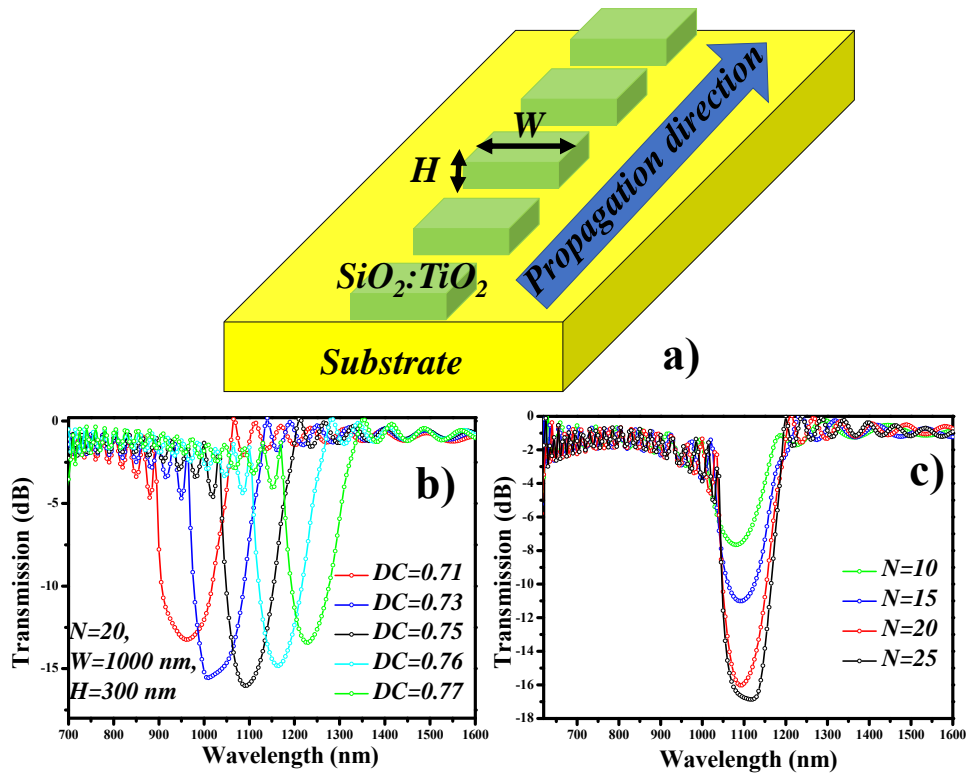
In this example, we presented a numerical analysis of an SWG WG based on a silica-titania platform conducted through 2D-FEM. The electromagnetic wave frequency domain (*emw*) has been chosen as a physics interface. The FEM splits oversized elements into smaller, simpler parts referred to as finite elements based on the meshing used. A precise meshing is necessary to obtain an accurate answer to the problem and the meshing is accomplished based upon the processing power of the computer. We used a built-in meshing function of “*very fine*” for the whole filter model that delivers an exact result. Scattering boundary conditions (SBCs) are applied at the outer edges of the simulation windows. Additionally, the “parametric sweep” function is also utilized to optimize the geometric parameters of the device design. The layer thickness and width of the silica-titania core layer is maintained at 300 nm and 1000 nm, respectively. By utilizing the effective index method, the  $n_{eff}$  of the core layer is determined which is around 1.6142. The segments of the high index material, i.e., silica-titania are periodically arranged on the substrate with the duty cycle (DC) of 0.71, 0.73, 0.75, 0.76, 0.77. The graphical illustration of the SWG WG is shown in Fig. 4.11a. The transmission spectrum of the proposed WG design is investigated for different DCs to locate the PBG as shown in Fig. 4.11b. The other geometric parameters such as  $W$ ,  $H$ ,  $N$  are maintained at 1000 nm, 300 nm and 20, respectively. Where  $N$  is the total number of the segment of the WG core. The unique feature of this WG structure is that it behaves as a standard WG when  $\frac{\lambda}{\Lambda} > 2n_{eff}$ , however, a photonic bandgap is observed when Bragg condition is satisfied. This attribute makes it an ideal candidate for optical bandstop filters and sensors. The  $\lambda_{Bragg}$  experiences a redshift when DC is increased from 0.71 to 0.77. This feature allows us to tune the stopband of the SWG WG filter to any desired range. The number of periodic segments ( $N$ ) plays an important role in defining the extinction ratio ( $ER$ ) of the filter. The  $ER$  boosts from ~8 dB to ~17 dB when  $N$  is increased from 10 to 25 as shown in Fig. 4.11c.

The  $H_z$ -field distribution at  $\lambda_{Bragg}$  and  $\lambda_{Bloch}$  at 1120 nm and 1500 nm is shown in Fig. 4.12a and Fig. 4.12b, respectively. The geometric parameters such as  $W$ ,  $H$  and  $N$  are fixed at 1000 nm, 300 nm and 25, respectively.

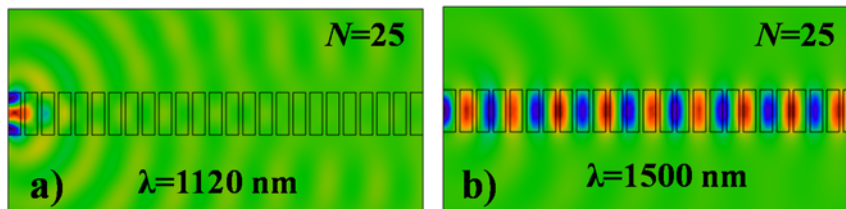
### 4.6.2. Numerical Modelling of 1×2 optical power splitter based on silica-titania platform

Y-splitters fit in highly valuable optical passive structures which are utilized for distributing signals from one port to two (or more) output ports. In this example, a 1×2 power splitter is designed on a silica-titania platform. The schematic representation

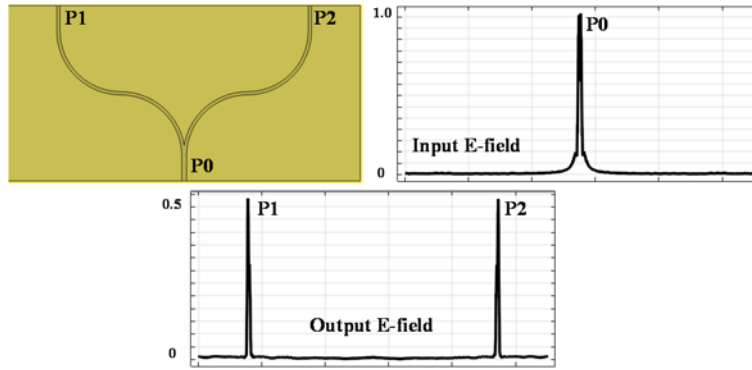
of the device is shown in Fig. 4.13a, where P0, P1 and P2 are the input port, output port 1 and output port 2, respectively. A  $1 \times 2$  Y-branch optical power splitter consists of a straight input WG, two S-bend sin arc WGs connected to the input WG by a linearly tapered WG at its leading edge, and two straight output WGs connected to the two S-bend WGs. The TE-polarized plane wave at 633 nm is launched at P0 which travels in a straight WG and splits in 50/50 power in both the output ports. The efficiency of the power splitter depends on the radius (R) of the bend WGs. The line integration of the E-field of the input and output ports is shown in Figs. 4.13b and 4.13c, respectively. For an ideal case, the input power should be equally divided between both output ports.



**Fig. 4.11.** a) Schematic illustration of a SWG WG based on  $\text{SiO}_2:\text{TiO}_2$  materials. Transmission spectrum of the SWG WG at different, b)  $DC$ , c)  $N$ .

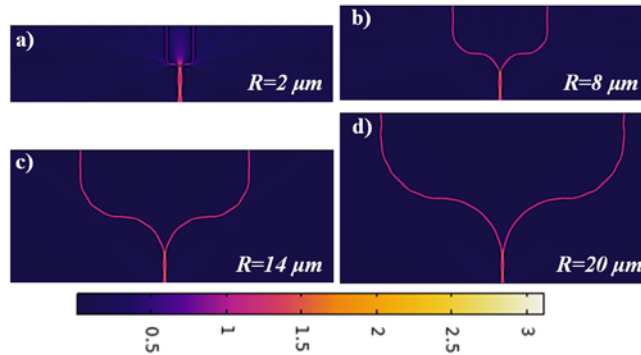


**Fig. 4.12.**  $H_z$ -field distribution in the SWG WG at, a)  $\lambda_{\text{Bragg}}=1120$  nm, b)  $\lambda_{\text{Bloch}}=1500$  nm.



**Fig. 4.13.** a)  $1 \times 2$  power splitter design, b) Line integration of the E-field of the input port, c) Line integration of the E-field of the output port.

Figs. 4.14 (a-d) shows the normalized E-field distribution in the  $1 \times 2$  power splitter structure at 633 nm. The efficient splitting depends on the appropriate radius ( $R$ ) of the bend WG. Two small  $R$  i.e.,  $2 \mu\text{m}$  can lead to high bending loss which results in the scattering of the light approaching the bends as shown in Fig. 4.14a. Figs. 4.14 (b, c and d) presents the balanced power splitting at  $R=8 \mu\text{m}$ ,  $14 \mu\text{m}$  and  $20 \mu\text{m}$ , respectively. This indicates that an efficient  $1 \times 2$  balanced optical power splitter can be designed on a silica-titania platform with a maximum chip size of  $\sim 60 \mu\text{m} \times 110 \mu\text{m}$  depending on the  $R$  of the S-bends.



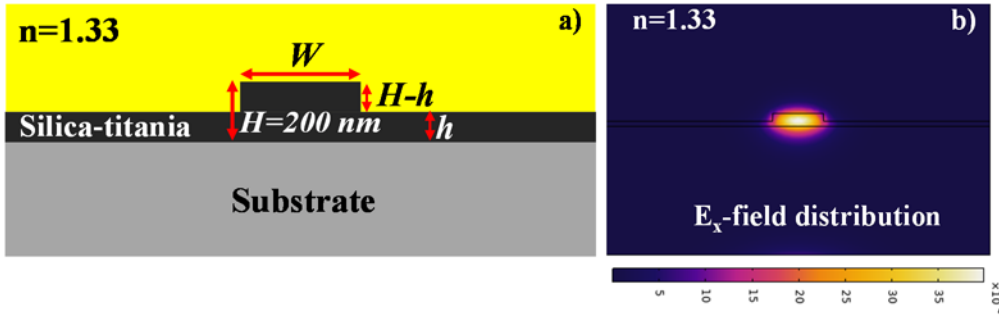
**Fig. 4.14.** Normalized E-field distribution in the  $1 \times 2$  power splitter of, a)  $R=2 \mu\text{m}$ , b)  $R=8 \mu\text{m}$ , c)  $R=14 \mu\text{m}$ , d)  $R=20 \mu\text{m}$ .

#### 4.6.3. Mode Sensitivity Analysis of a Rib Waveguide Based on Silica-titania Platform

In this example, the mode sensitivity analysis of a rib WG based on a silica-titania platform is performed for sensing applications. With the help of a single dip-coating process, a WG core layer ( $H$ ) of 200 nm can be obtained. That is why we have fixed  $H$  at 200 nm to simplify the analysis. The other geometric parameters such as the width of the



core, the height of the slab is denoted as  $W$  and  $h$ , respectively. The analysis is performed for the operational wavelength of 633 nm. The geometric representation of the rib WG is shown in Fig. 4.15a. To create the sensing environment, the WG is placed in a medium of refractive index ( $n$ ) = 1.33. The sensitivity of the WG is dependent on the geometric parameters of the WG which allows the propagation of evanescent field around the core of the WG structure. To develop a highly sensitive WG structure, the geometry of the WG is designed in such a way that the light-matter interaction is high. The  $E_x$ -field distribution of the propagating mode at 633 nm is shown in Fig. 4.15b.



**Fig. 4.15.** a) Graphical illustration of the rib WG based on silica-titania platform, b)  $E_x$ -field distribution of the propagating mode in the WG core at  $\lambda=633$  nm.

The effective refractive index ( $n_{\text{eff}}$ ) of the TE-mode at 633 nm is plotted for the ambient refractive index of 1.33 and 1.34 as shown in Fig. 4.16a. The shift in the  $n_{\text{eff}}$  of the mode depends on the degree of the confinement of the mode. If the mode is highly confined, then  $\Delta n_{\text{eff}}$  will be low. It is quite evident that the confinement of the propagating mode improves as  $h$  and  $W$  increase. The sensitivity of the WG which is calculated in terms of  $\Delta n_{\text{eff}}/\Delta n$  is plotted in Fig. 4.16b which demonstrates that the maximum sensitivity is obtained at  $W = 700$  nm. Therefore, one should choose the optimum dimensions of the WG which can provide the maximum sensitivity.

## 4.7. Final Remarks

In this chapter, several conventional and unconventional waveguide structures have been presented and discussed with respect of their properties and potential applications. Based on the geometric configuration of the WGs and electric field distribution of the propagating mode, these structures can be individually used for specific applications. Ridge, rib, and diffused waveguide structures have been classified as conventional waveguides due to their ability to confine a major part of the electric field in the high index part of the medium. Defined in such a way, these waveguides are the ideal candidates for developing the light-transmitting structures, providing e.g., the optical interconnects on a chip or between chips. Also, several other waveguide structures have been recently proposed due to their ability to confine the light in a low index part of the medium. These structures include slot, hybrid, SWG, and plasmonic waveguides. Due to their strong light-matter interaction, these waveguides seem to be ideal candidates for

sensing applications. In the end, we have proposed a cost-effective and easy-to-develop method for the realization of optical waveguides based on the silica-titania platform. The waveguides developed via the sol-gel dip-coating method have been analyzed numerically and verified experimentally. To solidify the idea, three examples were presented in which widely used optical components were designed based on this platform.

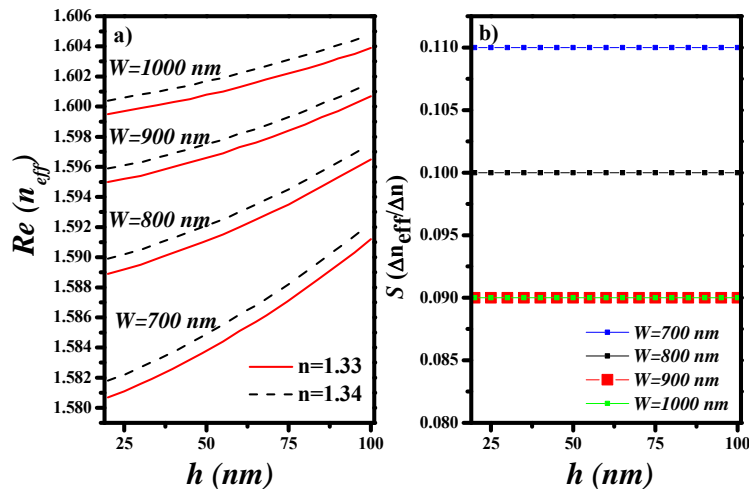


Fig. 4.16. a) Effective refractive index versus  $h$ , b) Mode sensitivity versus  $h$ .

## Acknowledgement

The work presented in this chapter was partially supported in the framework of “Hybrid sensor platforms of integrated photonic systems based on ceramic and polymer materials” project carried out within the TEAM-NET programme of the Foundation for Polish Science financed by the European Union under the European Regional Development Fund, POIR.04.04.00-00-14D6/18.

## References

- [1]. M. A. Butt, et al., Fabrication of Y-splitters, Mach-Zehnder structures on (Yb,Nb): RbTiOPO<sub>4</sub>/RbTiOPO<sub>4</sub> epitaxial layers by reactive ion etching, *Journal of Lightwave Technology*, Vol. 33, Issue 9, 2015, pp. 1863-1871.
- [2]. M. A. Butt, M. C. Pujol, R. Sole, A. Rodenas, G. Lifante, J. S. Wilkinson, M. Aguilo, F. Diaz, Channel waveguides and Mach-Zehnder structures on RbTiOPO<sub>4</sub> by Cs<sup>+</sup> ion exchange, *Optical Materials Express*, Vol. 5, Issue 5, 2015, pp. 1183-1194.
- [3]. M. A. Butt, E. S. Kozlova, S. N. Khonina, Conditions of a single-mode rib channel waveguide based on dielectric TiO<sub>2</sub>/SiO<sub>2</sub>, *Computer Optics*, Vol. 41, Issue 4, 2017, pp. 494-498.
- [4]. W. Bogaerts, et al., Nanophotonic waveguides in silicon-on-insulator fabricated with CMOS technology, *J. Lightwave Technol.*, Vol. 23, 2005, pp. 401-412.

- [5]. W. Gao, X. Yu, M. Fujita, T. Nagatsuma, C. Fumeaux, W. Withayachumnankul, Effective-medium-cladded dielectric waveguides for terahertz waves, *Optics Express*, Vol. 27, Issue 26, 2019, pp. 38721-38734.
- [6]. J. Halldorsson, N. B. Arnfinnsdottir, A. B. Jonsdottir, B. Agnarsson, K. Leosson, High index contrast polymer waveguide platform for integrated biophotonics, *Optics Express*, Vol. 18, Issue 15, 2010, pp. 16217-16226.
- [7]. S. N. Khonina, N. L. Kazanskiy, M. A. Butt, Evanescent field ratio enhancement of a modified ridge waveguide structure for methane gas sensing application, *IEEE Sensors Journal*, Vol. 20, Issue 15, 2020, pp. 8469-8476.
- [8]. N. L. Kazanskiy, S. N. Khonina, M. A. Butt, Subwavelength grating double slot waveguide racetrack ring resonator for refractive index sensing application, *Sensors*, Vol. 20, Issue 12, 2020, 3416.
- [9]. M. A. Butt, S. N. Khonina, N. L. Kazanskiy, Silicon on silicon dioxide slot waveguide evanescent field gas absorption sensor, *Journal of Modern Optics*, Vol. 65, Issue 2, 2018, pp. 174-178.
- [10]. K. Okamoto, Fundamentals of Optical Waveguides, *Academic Press*, Japan, 2006.
- [11]. M. A. Butt, S. N. Khonina, N. L. Kazanskiy, Optical elements based on silicon photonics, *Computer Optics*, Vol. 43, Issue 6, 2019, pp. 1079-1083.
- [12]. M. A. Butt, S. N. Khonina, N. L. Kazanskiy, Modelling of rib channel waveguides based on silicon-on-sapphire at 4.67  $\mu\text{m}$  wavelength for evanescent field gas absorption sensor, *Optik*, Vol. 168, 2018, pp. 692-697.
- [13]. V. R. Almeida, Q. Xu, C. A. Barrios, M. Lipson, Guiding and confining light in void nanostructure, *Optics Letters*, Vol. 29, Issue 11, 2004, pp. 1209-1211.
- [14]. M. A. Butt, S. N. Khonina, N. L. Kazanskiy, Ultrashort inverted tapered silicon ridge-to-slot waveguide coupler at 1.55  $\mu\text{m}$  and 3.392  $\mu\text{m}$  wavelength, *Applied Optics*, Vol. 59, Issue 26, 2020, pp. 7821-7828.
- [15]. M. A. Butt, S. N. Khonina, N. L. Kazanskiy, A highly sensitive design of subwavelength grating double-slot waveguide microring resonator, *Laser Physics Letters*, Vol. 17, Issue 7, 2020, 076201.
- [16]. M. A. Butt, S. N. Khonina, N. L. Kazanskiy, Device performance of standard strip, slot and hybrid plasmonic micro-ring resonator: A comparative study, *Waves in Random and Complex Media*, Vol. 31, 2021, pp. 2397-2406.
- [17]. M. A. Butt, S. N. Khonina, N. L. Kazanskiy, Highly sensitive refractive index sensor based on hybrid plasmonic waveguide microring resonator, *Waves in Random and Complex Media*, Vol. 30, Issue 2, 2020, pp. 292-299.
- [18]. M. A. Butt, N. L. Kazanskiy, S. N. Khonina, Modal characteristics of refractive index engineered hybrid plasmonic waveguide, *IEEE Sensors Journal*, Vol. 20, Issue 17, 2020, pp. 9779-9786.
- [19]. S. N. Khonina, N. L. Kazanskiy, M. A. Butt, Spectral characteristics of broad band-rejection filter based on Bragg grating, one-dimensional photonic crystal, and subwavelength grating waveguide, *Physica Scripta*, Vol. 96, 2021, 055505.
- [20]. N. L. Kazanskiy, M. A. Butt, S. A. Degtyarev, S. N. Khonina, Achievements in the development of plasmonic waveguide sensors for measuring the refractive index, *Computer Optics*, Vol. 44, Issue 3, 2020, pp. 295-318.
- [21]. N. L. Kazanskiy, S. N. Khonina, M. A. Butt, Plasmonic sensors based on Metal-insulator-metal waveguides for refractive index sensing applications: A brief review, *Physica E: Low-Dimensional Systems and Nanostructures*, Vol. 117, 2020, 113798.
- [22]. M. A. Butt, S. N. Khonina, N. L. Kazanskiy, An array of nano-dots loaded MIM square ring resonator with enhanced sensitivity at NIR wavelength range, *Optik*, Vol. 202, 2020, 163655.

- [23]. N. L. Kazanskiy, M. A. Butt, S. N. Khonina, Nanodots decorated MIM semi-ring resonator cavity for biochemical sensing, *Photonics and Nanostructures-Fundamentals and Applications*, Vol. 42, 2020, 100836.
- [24]. Y-X. Huang, et al., A plasmonic refractive index sensor based on a MIM waveguide with a side-coupled nanodisk resonator, in *Proceedings of the IEEE 20<sup>th</sup> International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA'14)*, Chongqing, 2014, pp. 1-5.
- [25]. N. L. Kazanskiy, M. A. Butt, Enhancing the sensitivity of a standard plasmonic MIM square ring resonator by incorporating nanodots in the cavity, *Photonics Letters of Poland*, Vol. 12, Issue 1, 2020, pp. 1-3.
- [26]. CEA-LETI, <https://www.leti-cea.com>
- [27]. T. Aalto, et al., Open-Access 3-um SOI waveguide platform for dense photonic integrated circuits, *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 25, Issue 5, 2019, 8201109.
- [28]. M. Smit, et al., An introduction to InP-based generic integration technology, *Journal of Semiconductor Science and Technology*, Vol. 29, 2014, 083001.
- [29]. L. Augustin, et al., InP-based Generic foundry platform for Photonic Integrated Circuits, *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 24, Issue 1, 2018, 6100210.
- [30]. F. M. Soares, M. Baier, T. Gaertner, N. Grote, M. Moehrle, T. Beckerwerth, P. Runge, M. Schell, InP-based foundry PICs for optical interconnects, *Applied Sciences*, Vol. 9, 2019, 1588.
- [31]. E. Kleijn, Passive components in indium phosphide generic integration technologies, PhD Thesis, *Eindhoven University of Technology*, 2014.
- [32]. C. G. H. Roeloffzen, et al., Low-loss Si<sub>3</sub>N<sub>4</sub> Triplex optical waveguides: Technology and Applications Overview, *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 24, Issue 4, 2018, pp. 1-21.
- [33]. M. Lelit, et al., Silicon nitride passive photonic platform for applications at visible wavelengths: design, fabrication and characterization, in *Proceedings of the European Conference on Integrated Optics (ECIO'20)*, 2020.
- [34]. M. Ebelmen, Sur les combinaisons des acides borique et silicique avec les ethers, *Annl. Chim. Phys.*, Vol. 16, 1846, pp. 129-166.
- [35]. T. Thor, J. Vaclavik, Sol-gel preparation of silica and titania thin films, in *Proceedings of the Optics and Measurement International Conference*, Liberec, Czech Republic, 2016, 101511A.
- [36]. H. Schroeder, Oxide layers deposited from organic solutions, in *Physics of Thin Films: Advances in Research and Developments*, *Academic Press*, New York, London, 1969, pp. 87-141.
- [37]. S. M. Attia, J. Wang, G. Wu, J. Shen, M. A. Jianhua, Review on sol-gel derived coatings: Process, techniques and optical applications, *J. Mater. Sci. Technol.*, Vol. 18, Issue 3, 2002, pp. 211-217.
- [38]. M. N. Logan, S. Prabakar, C. J. Brinker, Sol-gel-derived silica films with tailored microstructures for applications requiring organic dyes, *MRS Online Proceedings Library*, Vol. 346, 1994, pp. 115-120.
- [39]. C. J. Brinker, G. C. Frye, A. J. Hurd, C. S. Ashley, Fundamentals of sol-gel dip coating, *Thin Solid Films*, Vol. 201, Issue 1, 1991, pp. 97-108.
- [40]. A. Fidalgo, L. M. Ilharco, The defect structure of sol-gel-derived silica/polytetrahydrofuran hybrid films by FTIR, *Journal of Non-Crystalline Solids*, Vol. 283, Issues 1-3, 2001, pp. 144-154.
- [41]. R. Parin, M. Rigon, S. Bortolin, A. Martucci, D. D. Col, Optimization of hybrid sol-gel coating for dropwise condensation of pure steam, *Materials*, Vol. 13, 2020, 878.

- [42]. S. Acosta, A. Ayrál, C. Guizard, C. Lecornec, G. Passemard, M. Moussavi, Sol-gel derived silica layers for low-k dielectrics applications, *MRS Online Proceedings Library*, Vol. 612, 2000, 5261.
- [43]. M. Boudot, V. Gaud, M. Louarn, M. Selmane, D. Grosso, Sol-gel based hydrophobic antireflective coatings on organic substrates: A detailed investigation of Ammonia Vapor Treatment (AVT), *Chem. Mater.*, Vol. 26, Issue 5, 2014, pp. 1822-1833.
- [44]. C. J. Brinker, G. W. Scherer, Sol-gel Science: The Physics and Chemistry of Sol-gel processing, *Academic Press*, Boston, 1990.
- [45]. P. Karasinski, C. Tyszkiewicz, A. Domanowska, A. Michalewicz, J. Mazur, Low loss, long time stable sol-gel derived silica-titania waveguide films, *Materials Letters*, Vol. 143, 2015, pp. 5-7.
- [46]. J. C. Martinez-Anton, Determination of optical parameters in general film-substrate systems: a reformulation based on the concepts of envelop extremes and local magnitudes, *Appl. Opt.*, Vol. 39, Issue 25, 2000, pp. 4557-4568.
- [47]. J. Szczyrbowski, A. Czapla, On the determination of optical constants of films, *J. Phys. D: Appl. Phys.*, Vol. 12, 1979, pp. 1737-1751.

# **Chapter 5**

## **Detection, Classification and Re-identification of Road Users across Heterogeneous Video Surveillance Networks Using Artificial Intelligence and Calibration Techniques**

**Aleksej Makarov, Jakub Špaňhel, Miljan Vučetić, Jakub Sochor and Pavel Zemčik**

### **5.1. Introduction**

Detecting and tracking objects of interest is a classic problem in computer vision. Camera networks have been widely used in order to increase the robustness of tracking of a target over a large area. Fixed view video cameras have been used since decades for traffic and public safety applications. More recently, these cameras have been utilized in conjunction with pan-tilt-zoom (PTZ) mount plates for extended and adjustable Fields-of-View (FoV). Nowadays, drone-borne cameras are often exploited for the same purposes. In both traffic and public safety applications, it is often necessary to estimate metric (real-world) distances between and the dimensions of objects captured by separate cameras. Examples include speed monitoring and speed limit enforcement, detection of tailgating, compliance to stop lines, control of height and length limits for trucks etc. Moreover, estimates of distance, dimensions and velocity made by one smart sensor can be used to facilitate the target detection and re-identification within the field of view of another sensor.

There is a need encountered in practice to collect all visual scenes provided by different types of cameras within a network, in order to track objects of interest from one visual scene to another. Additionally, there is a need to extend the coverage when tracked objects leave the area covered by a fixed camera network. This limited observability can be overcome to a great extent by introducing moving sensors as part of a cooperative

heterogeneous network – modern drones, equipped with cameras and other sensors. The use of drones in such a framework is expected to boost coverage, reduce costs, provide flexibility and enable autonomous tracking [1]. However, it also encounters numerous challenges such as communication with multiple cameras, object feature mapping for seamless tracking, navigation, real-time processing, energy consumption, integration of embedded hardware, and top-down view angles.

In this work, on top of road user classification and metric calibration problems already addressed in literature [2-5], we explore the feasibility of a cooperative network consisting of multiple cameras and drones within an intelligent surveillance system. Motivated by recent advances in the Internet of Things (IoT), we propose a novel solution for dynamically extendable detection of objects in a field of view and mapping of object features from one sensor to another. We collect and share all the information about the position, size, and specific features of a detected and identified object among sensors. For instance, the features of an object captured by a camera can be forwarded to the closest inbound sensor [6], a drone or another camera, along with the predicted position of the object. To accomplish this, we use the state-of-the-art deep learning algorithms for calibration, object detection, classification and re-identification but also go beyond them, in order to design a solution applicable to the adopted context of heterogeneous vision nodes. Additional object features thus provided enable seamless integration between different network nodes. The object tracking capability can be augmented by means of the synergy of cameras and drones, enabling a wide range of services and perspectives in public security, traffic management, smart city applications, border control and law enforcement.

## 5.2. Object Detection

Most detection methods proposed for traditional fixed view surveillance networks relied on separation of the moving object from a fixed background [7-10]. With the advent of mobile smart cameras with increased computational capacity, these methods had to be replaced by novel feature-based techniques, which help identify not only the moving object, but also its nature. This work addresses the issues of detection, classification and re-identification of vehicles and pedestrians, which have been addressed in the past using artificial intelligence (AI) [11, 12]. Convolutional neural networks (CNNs) are nowadays used in many computer vision applications. They are a widely utilized tool in object detection, also known as *object recognition* [11], where objects are both localized and classified. CNN based detectors are generally composed of the following functional blocks [12]:

- Feature extractor (filters arranged in parallel or consecutive layers identify different features present in an image. Layers applied to input images provide simple features such as edges, lines, colors, while stacked layers are used to find higher level features, such as more shapes and objects. Each filter layer generates a new output image, called the feature map, indicating the locations and strength of a detected feature in the input image);

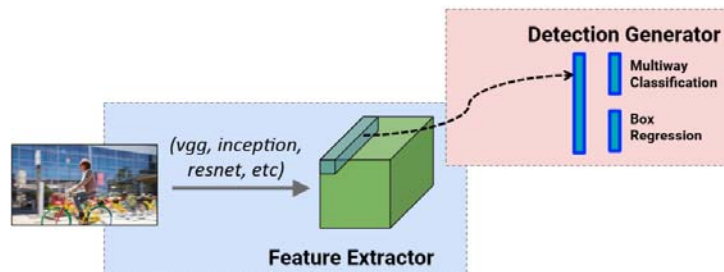
- Box proposer (or image region regressor, whose task is finding image region proposals from feature maps, typically top-left and bottom-right coordinates of bounding boxes enclosing and thus localizing the candidates for object detection);
- Proposal classifier (for each pre-determined category of objects of interest, assigning probability scores to proposals – bounding boxes; each proposal is subsequently classified to its most probable category).

The CNN architectures such as VGG-16 [13], Inception v2 [14], Inception v3 [15], ResNet-50, ResNet-101 [16] or MobileNet [18] can be used as feature extractors for various object detectors. These extractors can be used in conjunction with different proposal and classification blocks, depending on the choice of the overall detector architecture.

Broadly used are three different *meta-architectures* (i.e., families of detector architectures [18]) based on the following convolutional neural network structures:

- One-stage detectors (or single-shot detectors, SSD, YOLO) [19, 20];
- Faster R-CNN (Regions with CNN features) [21];
- Region-based Fully Convolutional Network (R-FCN) [22].

Single shot detectors are one-stage, feed-forward detectors based on a single CNN. Instead of performing separate steps for region proposal and proposal classification, they directly predict both the bounding box and class probabilities from input images [23]. They are consequently faster and implementable on real-time devices. Typical representatives are SSD [19], YOLO [20] or Multibox [24] or the Region Proposal Network (RPN) [21].



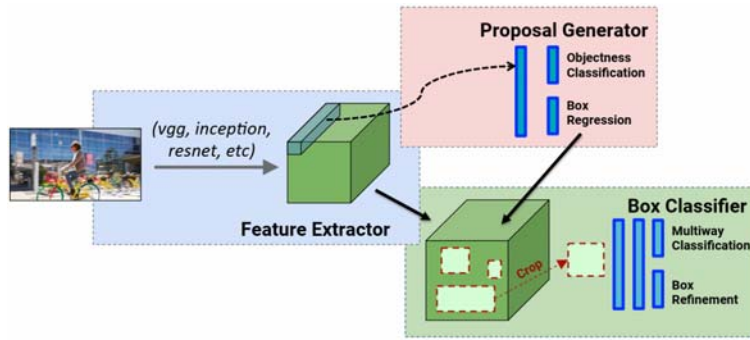
**Fig. 5.1.** A Feature extractor in conjunction with a one-stage detector [18].

The RPN are used as the first stage of Faster R-CNN, which are discussed below.

Detectors of the “Faster R-CNN” type perform the detection process in two stages. The Region Proposal Networks algorithm extracts intermediate-level features from the image in order to distinguish objects from the background. These features are associated with the corresponding pixels in the original image. After sliding a set of windows of different sizes and aspects across predetermined (so called “anchor”) image locations with the associated features, class-independent boxes are predicted. For instance, the originally proposed Faster RCNN used anchor boxes of 3 aspect ratios and 3 scales. These box



proposals are then used to crop the extracted features from the aforementioned feature map. The cropped features are passed to successive layers of the feature extractor for prediction of class-related boxes and their classes. Detectors based on this meta-architecture can be found in [25-29]. They are computationally more complex than the single stage detectors but provide higher accuracy.

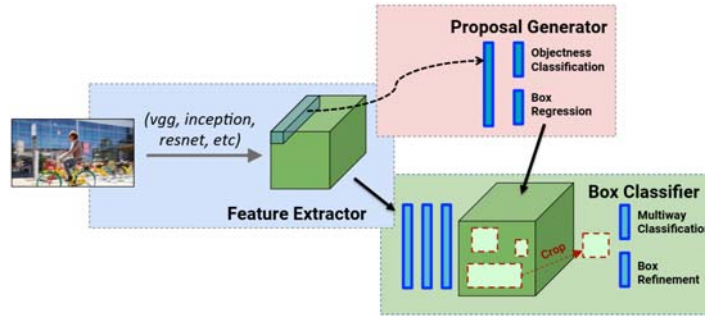


**Fig. 5.2.** The object detector meta-architecture consisting of a feature extractor and a Faster R-CNN [18].

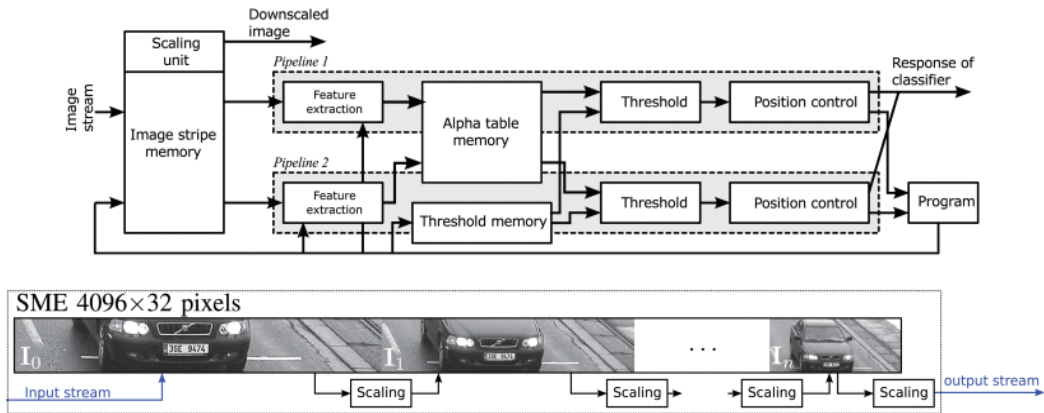
Finally, the region-based fully convolutional networks (R-FCN) [22] are similar to Faster R-CNN, except that the crops are passed to the penultimate convolutional layer of the network before prediction. The R-FCNs address the problem of improving the translational invariance in object detection in deep convolution networks. Such networks are good in recognizing objects regardless of their position and possible distortion, but the feature maps at the last convolution layer may become too small to detect translated positions of the object. In order to improve the object detection, the regions of interest are divided into a number of sub-regions arranged according to their positions (e.g., “upper left”). Each sub-region is checked against the corresponding sub-region learned for that position for all objects to be recognized (the so-called score bank), providing position-sensitive score maps. This change leads to less computation per region and faster prediction with a somewhat lower accuracy.

Despite the fact that the CNN based object detectors are nowadays the state of the art in quality of detection, more traditional methods based on machine learning achieve higher performance in image data throughput and can be well exploited for initial processing of the video data. These methods include WaldBoost [31] as well as Random Forest [31] based methods. Moreover, these methods can be accelerated not only in GPU [32] but also in embedded systems, or even in programmable hardware, such as FPGA (Field Programmable Gate Arrays) [31, 3]. For instance, due to their modularity and moderate resource consumption, they can be implemented in the hardware of smart sensors [34]. This approach allows for building heterogeneous systems, in which energy-efficient pre-processing of the bulk video data is performed using these traditional methods, while deeper analysis of the pre-selected objects is done by CNNs, reducing thus the computational demands posed on CPU or GPU, as well as the overall energy consumption.

Fig. 5.4 gives an example of multi-scale license plate detection in FPGA based on WaldBoost.



**Fig. 5.3.** The object detector meta-architecture consisting of a feature extractor and an R-FCN [18].



**Fig. 5.4.** Example of a WaldBoost multi-scale license plate detection circuit in FPGA [34].

Two implementation scenarios have been considered for our application. At a local level, a smart camera embedded implementation of the detectors was based on the Nvidia Jetson TX2 computing module. As a backup solution able to compensate for the limited computational capacity at the smart sensor level, the detectors were also implemented on the GeForce GTX 1080 Ti graphics processing unit (GPU) installed at the edge server level.

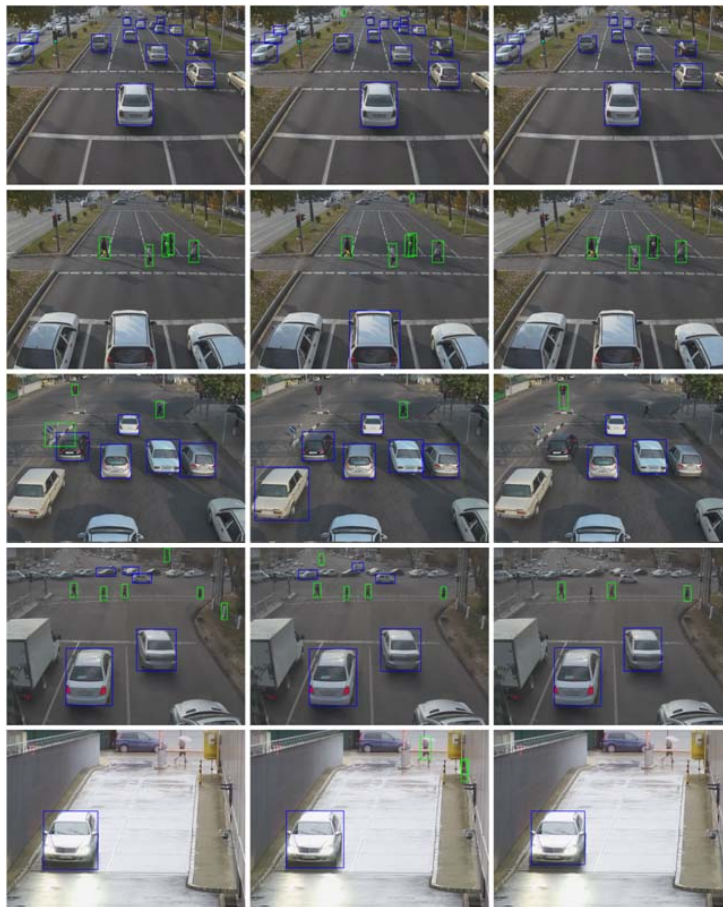
Three object detector architectures were tested. The feature extractors were implemented as MobileNet or Residual Networks with 50 (ResNet-50) and 101 layers (ResNet-101). The detector architectures were one-stage (SSD) and Faster R-CNN.

The training and testing datasets used for vehicle detection were COD20k [34, 31] and UA-DETRAC [35, 36], while the surveillance pedestrian image dataset SPID [37] was used for pedestrian detection.

The traffic surveillance COD20k dataset contains 19,126 car instances in 9,530 training images and other 4,457 cars in 1,127 testing images. 2D bounding box annotation and

viewpoint vector is provided for each vehicle instance in the dataset. The UA-DETRAC dataset contains annotations for 8,250 vehicle tracks of four different types (car, bus, van, other) on 140,000 frames in total, leading to a total number of 1.21 million of labelled bounding boxes of vehicles. The dataset is divided for train (84,000 frames, 5,900 vehicles, 578,000 boxes) and testing (56,000 frames, 2,300 cars, 63,200 boxes) purposes. The Surveillance Pedestrian Image Dataset (SPID) was built from a number of surveillances, covering a wide range of pedestrians at different scales and position from outdoor scenarios. SPID consists of 14,550 training and 15,439 testing images with 110,069 labelled pedestrians in total.

The detectors were experimentally evaluated for two different fixed camera settings. The first scenario uses a camera placed above the road facing the approaching vehicle (see the top three rows in Fig. 5.5). In this scenario, the camera sees a dashed crosswalk and the rear of moving vehicles. This is a typical setting for detecting traffic violations, such as running red light, speeding or failure to yield the right of way to pedestrians.



**Fig. 5.5.** Visual detection results of three different detectors. Left: Faster R-CNN based on ResNet-50. Center: Faster R-CNN based on ResNet-101. Right: SSD based on MobileNet. It can be seen that detection results differ from each other.

In the second setting, the camera is placed on building capturing arriving/departing vehicles into the underground garage. In this case, the camera sees front/rear part of vehicles. This scenario is focused on vehicle detection only and can be used for providing access to vehicles.

The object detection results of the Faster R-CNN detectors with the ResNet-50 and the ResNet-101 feature extractors, respectively, and those of the SSD detector used with the MobileNet feature extractor are shown in Fig. 5.5. The detection results differ due to the quality of the feature extractor used. Inference times for different detectors on GPU (GTX 1080 Ti) and Nvidia Jetson TX2 platforms were also evaluated. During testing, only the SSD detector based on MobileNet could run on both devices. The processing speed was 10.7 FPS on the Jetson TX2 platform and 278.9 FPS on the GTX 1080 Ti GPU.

### **5.3. Metric Calibration**

In surveillance applications, it is often of interest to measure or estimate the size of objects in the scene and their mutual distances. In our work, the size of detected objects is used as an additional classification feature, while measuring distances travelled over a time interval as well as those between contiguous road users are of specific importance to traffic law enforcement. Estimation of distances and dimensions of objects is contingent on a proper calibration of the scene. Various calibration methods have been reported in the literature, based on techniques such as stereoscopy [38], multiple scene views [39], target tracking [40, 41], chessboard pattern [42, 43] and 3D object-based camera calibration [44, 45].

Traffic cameras installed near traffic lights and school crossings are usually fixed view monocular cameras, hence not stereoscopic and unable to take different views of the same scene. Calibration against a chessboard pattern usually takes place in a laboratory. While being a tedious, time-consuming task by itself, this method often results in errors with respect to the laboratory measurements, due to misadjusted camera parameters upon installation. Tracking objects of known dimensions seems to be the most appropriate way to address the scene calibration problem and estimate the size of unknown objects in fixed scenes. This estimation can in return help to re-adjust the extrinsic parameters of the camera.

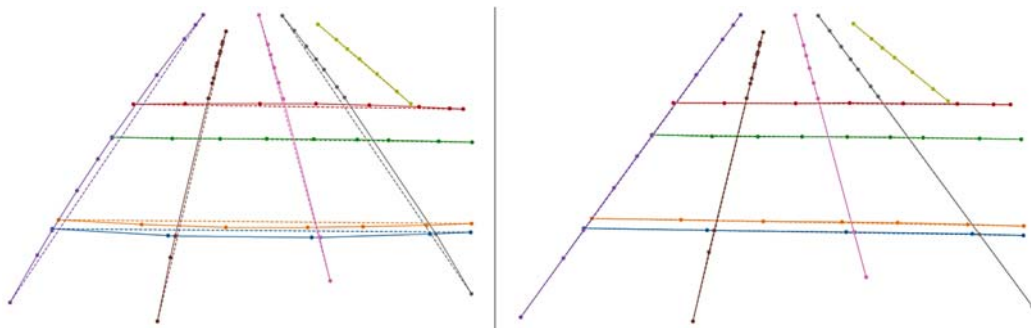
On the other hand, drone-borne cameras cannot easily benefit from objects of known dimensions. Namely, by the time an object of known dimensions has crossed the scene that should have been calibrated, the surveillance drone might have moved and another scene would be in its field of view. Fortunately, a drone can transmit its latitude, longitude, altitude, inertia measurement unit (IMU) and gimbal angle information, which provides a continuous knowledge of the extrinsic parameters of its camera. This knowledge helps in achieving fast calibration of ever-changing scenes viewed by drones.

Consequently, in a camera network containing sensors on heterogeneous supports, such as fixed mounts, pan-tilt-zoom (PTZ) platforms and drones, different methods should be used for calibration. In the case of fixed cameras used in this work, an approach based on

[46] with few modifications has been implemented. Our approach is based on vanishing points derived from consistent lines detected by the Hough transform. In the original approach, the dominant motion of the vehicles, i.e., the trajectories of several detected vehicles intersect at the first vanishing point. The lines containing the vehicle edges perpendicular to the vehicle movement but parallel to the ground plane intersect at the second vanishing point. The vanishing lines in the aforementioned directions define the ground plane, or more precisely, a plane parallel to the ground. The method deduces extrinsic camera parameters (i.e., the camera orientation and position up to scale). Under the reasonable assumption of a camera sensor with zero skew and unit aspect ratio, the focal length as the remaining intrinsic parameter can be computed as the geometrical mean of the real-world distances of the horizontal-plane vanishing points from the central point in the field of view (principal point), once metrically quantified by detecting objects of known dimensions, as shown later in the text.

In our modified approach, instead of the vehicle trajectories, the well-visible lane markings containing straight lines are detected by the Hough transform and used for computing the vanishing points. Although this method is more sensitive to the number and quality of lane markings, it can be performed without waiting for a sufficient number of vehicles to cross the field of view, which might be of use in PTZ and drone-borne cameras whose extrinsic parameters may quickly change.

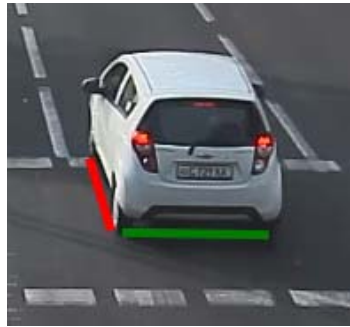
Assuming the vehicle trajectories follow a straight road, this calibration process can be used to calculate and compensate radial distortion, typical of wide-angle lenses used in surveillance cameras, as shown in Fig. 5.6.



**Fig. 5.6.** The left-hand side shows the lines containing lane separator markings (solid lines) and the lines connecting the end points of the detected collinear markings (dashed lines). The detected lines (solid) are obviously distorted with respect to the straight lines (dashed). The right-hand side shows the same annotated lines after correcting the radial distortion.

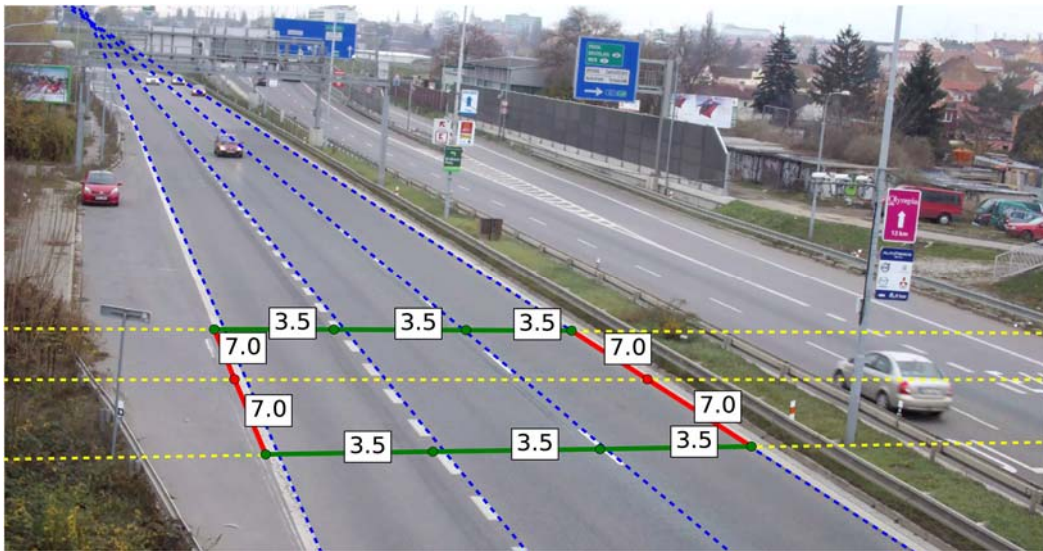
For metric quantification, the wheel base (the length between the front and the rear wheel centers) and the rear track (the distance between the rear wheel centers) of some known vehicles are used. These vehicles can be automatically detected, as shown later in this work. However, in our present implementation, manual annotation was used. Automatic segmentation of the wheel base and the rear track should be the subject of the future work.

Thus, in our approach, the vanishing points in the horizontal plane are computed as intersections of lines containing the lane separator markings (for the first vanishing point, VP1) and the stop line markings and vehicle rear tracks (the first vanishing point, VP2). The metric scale is then derived from the wheel base and rear track lengths of known vehicles. Knowing the intrinsic and extrinsic camera parameters, including the scale, allow us to compute all horizontal dimensions in the image in world units (meters).

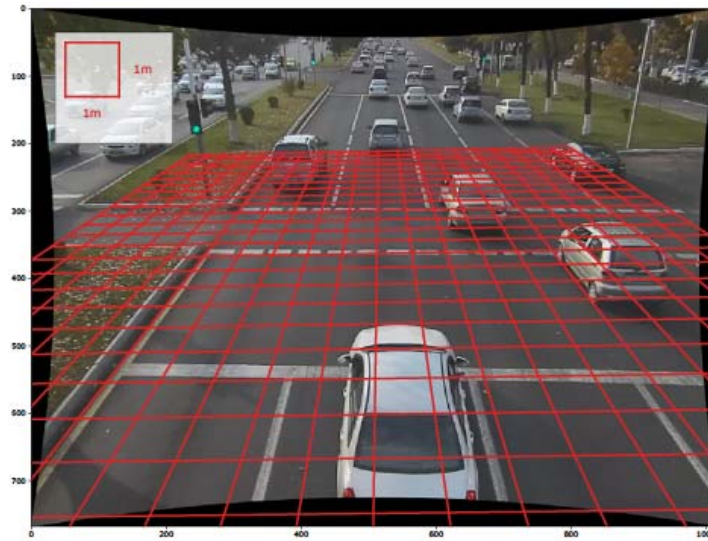


**Fig. 5.7.** Example of wheel base (red) and rear track (green) annotation on Chevrolet Spark vehicle.

An example of finding vanishing points and measuring distances in a scene is shown in Fig. 5.8. The calibrated ground plane as a result of the applied calibration method is shown in Fig. 5.9.



**Fig. 5.8.** The vanishing points in the horizontal plane found at intersections of vanishing lines collinear with lane separator markings (blue) and rear tracks (yellow). The metric calibration was performed from the known wheel base and rear track lengths and used for estimating distances on the road as illustrated (in meters).



**Fig. 5.9.** The final camera calibration obtained for a scene after distortion correction. The red grid represents the ground plane obtained from the calibration algorithm with dimensions (1×1 meter).

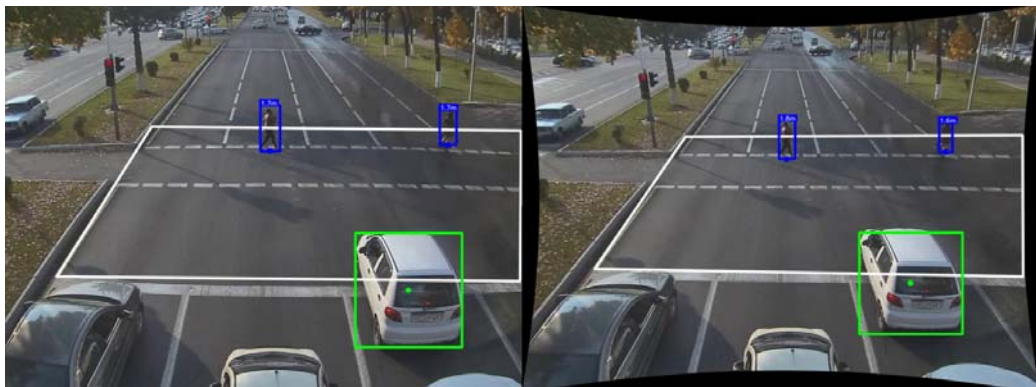
The above scene calibration technique can be used in numerous traffic monitoring and law enforcement applications, such as speed measurement, non-observance of prescribed distances (tailgating) or length- and width-based vehicle categorization.

#### 5.4. Height Estimation

The abovementioned method provides only the ground plane calibration and can measure only horizontal distances. This can be used in most traffic applications, such as speed limit enforcement, detection of tailgating or compliance to red light and stop lines. Extending the method to applications such as height clearance control for vehicles or detecting child pedestrians (useful for controlling traffic lights around schools) requires the detection of the third vanishing point at the intersection of vertical vanishing lines in the scene. However, vehicles usually do not provide clear vertical edges supporting the third vanishing point. In the case of sensor cells (pixels) with unitary aspect ratio, the third-dimension vectors can be approximated as oriented along the image columns, and scaled by rotating the orthogonal (row-wise) projection of the rear track length (annotated by green in Fig. 5.7. Example of wheel base (red) and rear track (green) annotation on Chevrolet Spark vehicle. In [46], the third vanishing point vector was derived mathematically as the cross product of the horizontal vanishing point vectors (originating at the principal point), scaled by a factor proportional to the focal length of the lens. Unfortunately, in some cases, the virtual points in the horizontal plane can be infinitely distant, so neither the focal length nor the vertical virtual point can have a finite value. Therefore, detecting vertical structures of known height with clear edges is necessary to achieve full three-dimensional calibration. Typical examples of such structures may be signposts, lampposts, traffic light or utility poles. However, the height of these posts can

vary a lot and needs to be known in advance. In [47] and [48], the vertical virtual point is derived from the pedestrians detected and tracked in the scene.

Instead of relying on the finiteness of the horizontal vanishing point vectors and focal length, our approach is also based on pedestrians detected by the previously described object detection methods (see Fig. 5.5). Human beings are not equally tall and no prior information can be collected about their height. Additionally, the average human height varies a lot from one country or ethnic group to another and depends on gender. Still, for applications where accuracy is not critical, such as classification to different height classes (for example, short-average-tall), one can assume the average height of pedestrians in the scene and classify them to a small number of classes. We tested this heuristic approach to detect unaccompanied child pedestrians. The expected average height by WHO standards is 176.5 cm for healthy men and 163 cm for women [49]. Since measurement errors of the order of 10 centimeters cannot be avoided anyway due to poor resolution per object in wide-angle scenes, object distances, body postures, footwear and carried loads, the gender estimation problem can be circumvented by assuming the average height of 1.7 meters for adult pedestrians of both sexes. The height of each pedestrian in a scene containing a group of people can be thus roughly estimated based on this average value. In order to avoid large variations in height estimates of the same person depending on their position in the scene, it is important to use undistorted images as previously discussed. An example of different height estimates for the same person in a radially distorted and corrected image is given in Fig. 5.10. For fixed-view cameras, the height estimates improve with time as the assumed average height approaches the statistical mean value for a larger population of pedestrians.

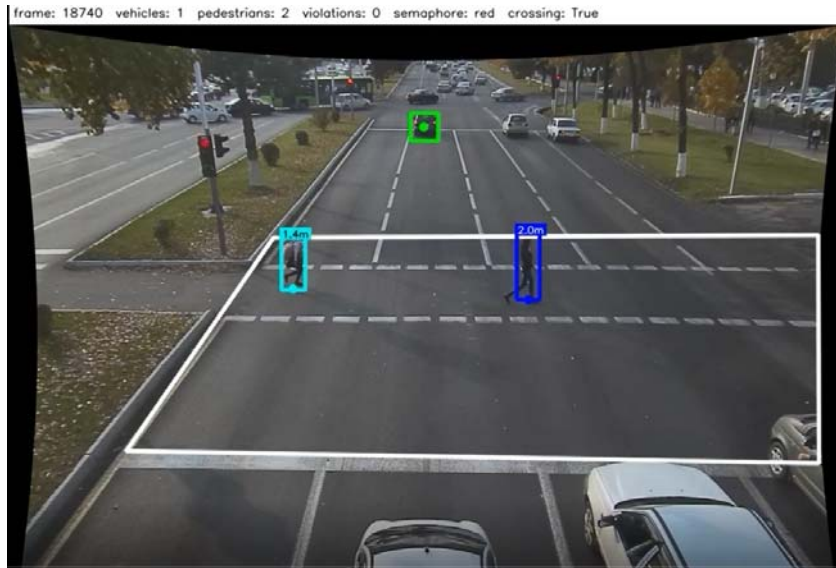


**Fig. 5.10.** Left: A distorted image providing height estimates for two persons in the field of view. Both are estimated to be equally tall (1.7 meters). Right: Undistorted version of the same image provides corrected height estimates for the same persons at the same positions, where the person on the left is estimated to be 1.8 meters tall, while the person on the right is 1.6 meters tall.

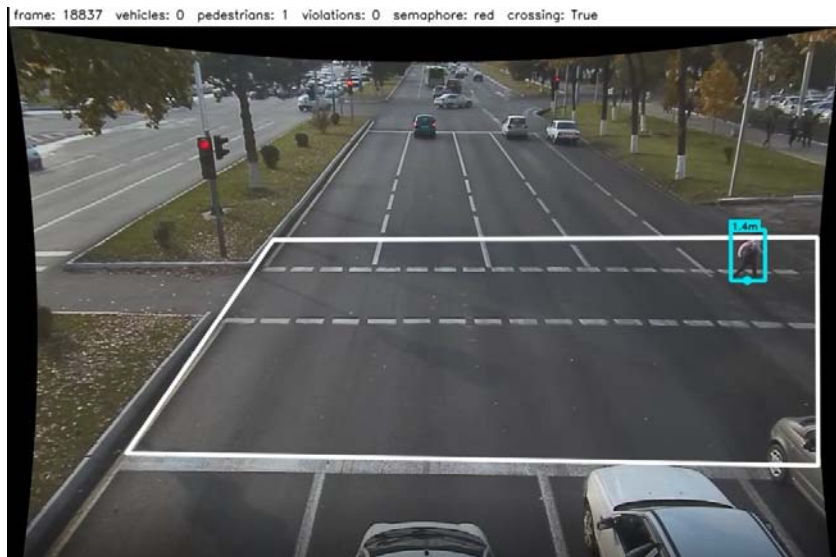
Based on this average height, pedestrians estimated to be shorter than 160 cm were automatically labeled as children. Namely, 95 % of 11-year-old boys and girls will be shorter than 155 cm, while 95 % of boys and almost 90 % of girls at the age of 15 and



over will be taller than 155 cm [50]. Detected unaccompanied child pedestrians are shown in Figs. 5.11-5.14. Figs. 5.11, 5.12 show the same child pedestrian at different positions of the undistorted scene. Overall the scene the child's height estimated to be equal. Fig. 5.11 demonstrates the capacity of the algorithm to detect and evaluate the height of two partially overlapping children.



**Fig. 5.11.** A child pedestrian detected (cyan) with the estimated height of 1.4 meters.



**Fig. 5.12.** The child pedestrian detected in Fig. 5.11 (cyan) at another position but with the estimated height as before, as expected in undistorted images.

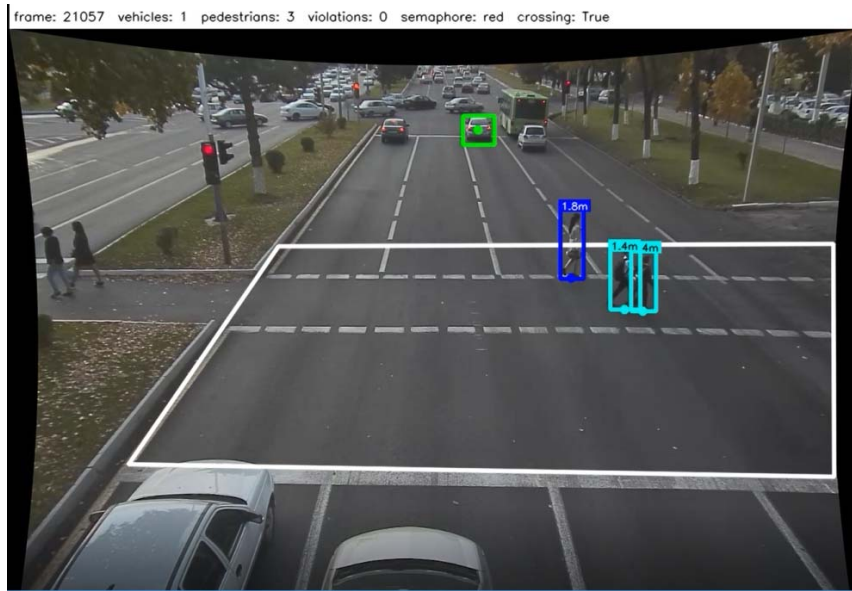


Fig. 5.13. Two overlapping child pedestrians (cyan) detected on a crosswalk.

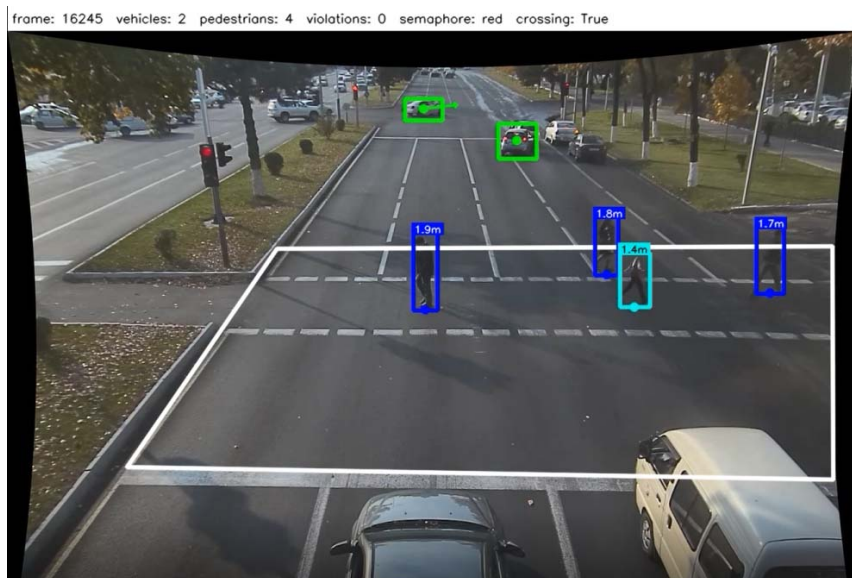


Fig. 5.14. A child at a crosswalk (cyan) with a group of adult pedestrians (blue).

## 5.5. Calibration with Known Extrinsic Parameters

The scene calibration and height estimation methods presented in the previous section were appropriate for cameras with unknown extrinsic parameters. On the other hand, drone-borne cameras can transmit their latitude, longitude, altitude, inertia measurement

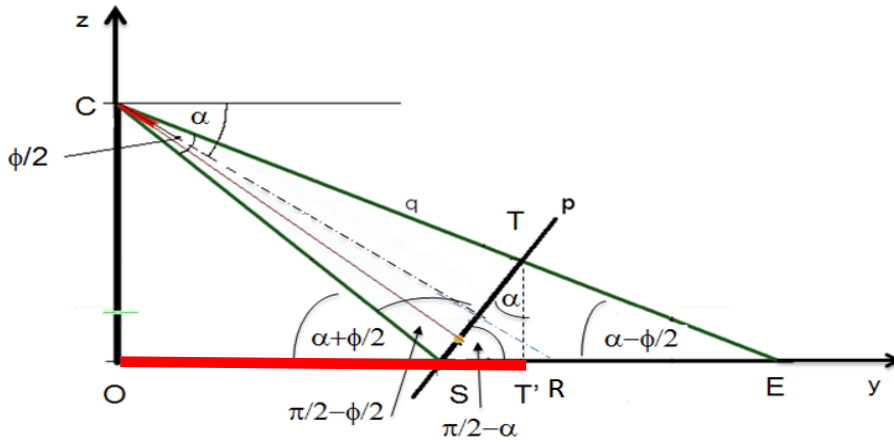
unit (IMU) and gimbal angle information, which provides a continuous knowledge of their extrinsic parameters. This knowledge helps in achieving fast calibration of ever-changing scenes in their fields of view.

The approach based on a priori knowledge of extrinsic camera parameters, as well as on recognizing objects of known dimensions will be presented hereafter. The same approach can be extended to other camera supports (e.g., PTZ), as long as the above parameters are available.

In this case, the extrinsic parameters known or derived in advance are:

- The angle (tilt) of the drone camera,  $\alpha$ ;
- The vertical angle of view  $\phi$ ;
- The camera altitude  $H$  with respect to the ground, which corresponds to the altitude of the drone.

As discussed in [40], the real-world position  $R$  (i.e., the horizontal distance from the orthogonal projection of the drone on the ground) and the height of an object can be computed from the above parameters and trigonometric relations in Fig. 5.15.



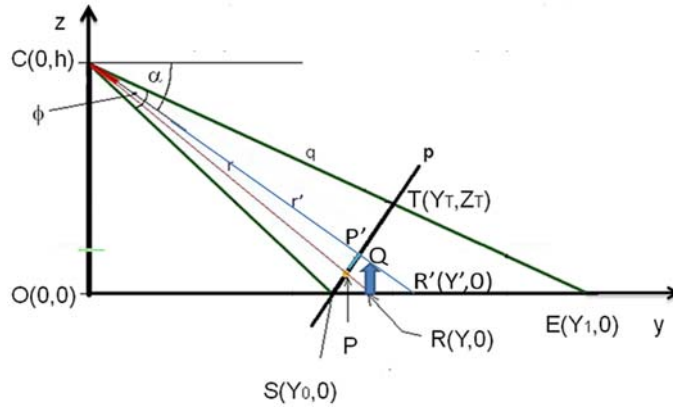
**Fig. 5.15.** Computing the real-world position  $R$  of a target on  $y$  axis from the camera tilt, angle of view and the assumed relative height of the camera with respect to  $R$ . Drone in  $C$ , camera angle  $\alpha$ , object at  $T$  (ground), the object horizontal distance  $OT'$ .

Assuming a camera sensor with  $N_r$  rows and  $N_c$  columns, tilted by angle  $\alpha$  towards the ground at altitude  $h$ , and the camera angle of view  $\phi$ , an object lying on the ground at pixel coordinates  $(n_p, n_c)$  is at distance  $Y_R$  in the direction of the horizontal projection of the camera axis  $y$ , and at orthogonal distance  $x$  away from line  $y$  [40]:

$$Y_R(n_p) = h \cdot \frac{\left(2 - \frac{n_p}{N_r}\right) \cos\left(\alpha + \frac{\phi}{2}\right) - \left(1 - \frac{n_p}{N_r}\right) \cos\left(\alpha - \frac{\phi}{2}\right)}{\frac{n_p}{N_r} \cdot \sin\left(\alpha + \frac{\phi}{2}\right) + \left(1 - \frac{n_p}{N_r}\right) \cdot \sin\left(\alpha - \frac{\phi}{2}\right)}, \quad (5.1)$$

$$x(n) = \left(2 \frac{n_c}{N_C} - 1\right) \tan\left(\frac{\omega}{2}\right) \cdot \sqrt{h^2 + Y_R(n_r)^2} \quad (5.2)$$

A height of an object depicted between points R (foot) and Q (top) can be derived from Fig. 5.16.



**Fig. 5.16.** The height  $RQ'$  of a linear vertical object posed on the ground can be computed from its projection  $PP'$  in the image plane.

From Fig. 5.16, the height  $H_o$  of the object can be calculated as [51]:

$$H_o = \frac{Z_{P'} - H}{Y_{P'}} \cdot (Y' - Y), \quad (5.3)$$

where  $Y$  is given by (5.1), while  $Z_{P'}$ ,  $Y_{P'}$  and  $Y'$  are given as

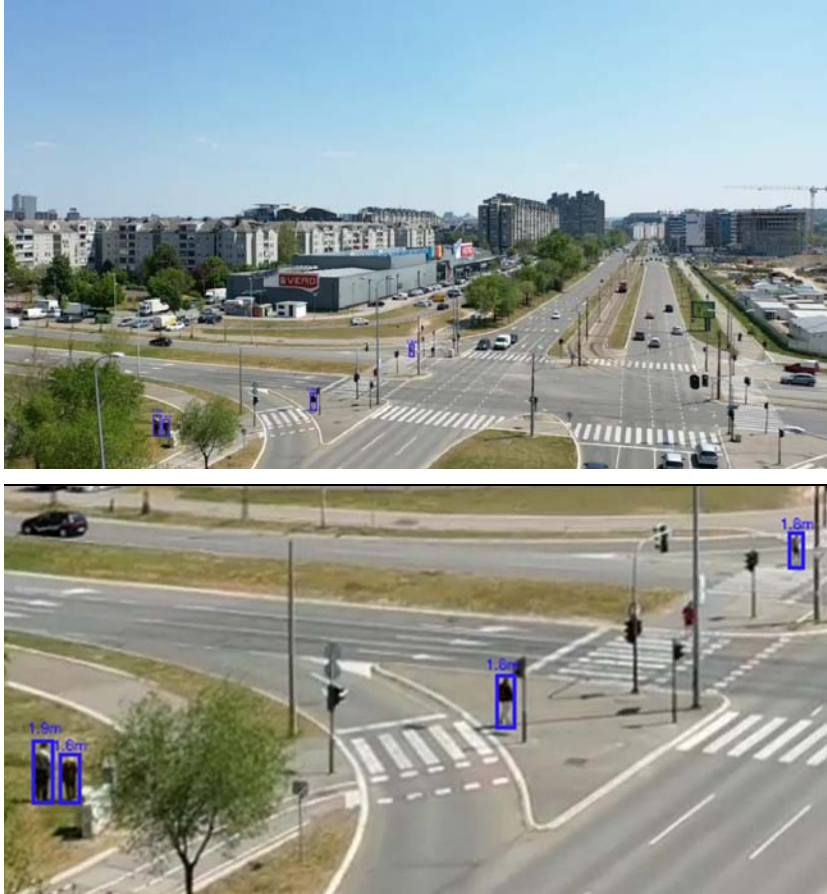
$$Z_{P'} = \frac{N_r - n_{p'}}{N_r} \cdot H \cdot \frac{1 - \cot(\alpha + \phi/2) \tan(\alpha - \phi/2)}{\cot \alpha + \tan(\alpha - \phi/2)}, \quad (5.4)$$

$$Y_{P'} = H \cdot \left[ \frac{n_{p'}}{N_r} \cot(\alpha + \phi/2) + \frac{N_r - n_{p'}}{N_r} \cdot \frac{1 + \cot(\alpha + \phi/2)}{\cot \alpha + \tan(\alpha - \phi/2)} \right], \quad (5.5)$$

and

$$Y'(n_{p'}) = H \frac{\left(2 - \frac{n_{p'}}{N_r}\right) \cos\left(\alpha + \frac{\varphi}{2}\right) - \left(1 - \frac{n_{p'}}{N_r}\right) \cos\left(\alpha - \frac{\varphi}{2}\right)}{\frac{n_{p'}}{N_r} \cdot \sin\left(\alpha + \frac{\varphi}{2}\right) + \left(1 - \frac{n_{p'}}{N_r}\right) \cdot \sin\left(\alpha - \frac{\varphi}{2}\right)} \quad (5.6)$$

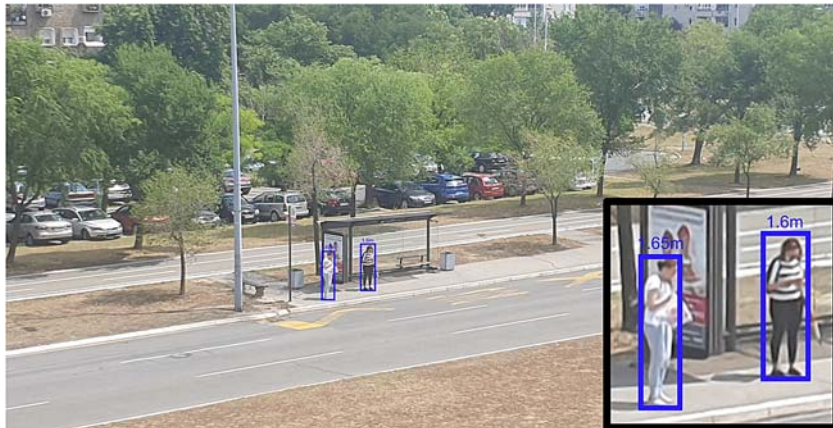
In our experiments, the values of altitude  $H$  with respect to the ground, tilt angle  $\alpha$  and angle of view  $\varphi$  were measured, processed and imported from the camera drone used (DJI Mavic 2 Pro). Some results showing the detected pedestrians and their height estimates in wide angle aerial images (up to 75 degrees) are given in Fig. 5.17 - 5.19.



**Fig. 5.17.** People detected at a crossing using a drone camera. Enlarged view of pedestrians with estimated heights in the bottom image.

Tracking humans within a camera network, including non-overlapping cameras distributed over a wide network has received recent interest in the literature [52-54]. However, the results achieved so far are limited to similar views, i.e., to similar extrinsic parameters of the deployed cameras. Tracking a person between cameras at significantly

different heights and tilts (e.g., from a fixed camera to a drone camera or vice versa) currently remains an unexplored task. The main challenges appear to be linked with the different angles of view, but also with a possible loss of resolution in images captured from distant drone cameras. The person height measurement explained above can be used as an additional feature to address the person re-identification problem in future heterogeneous networks.



**Fig. 5.18.** People detected at a bus stop using a drone camera. Enlarged view of pedestrians with estimated heights enlarged in the framed image in the bottom left corner.



**Fig. 5.19.** People detected on a sidewalk using a drone camera. Enlarged view of pedestrians with estimated heights enlarged in the framed circle in the left.

## 5.6. Vehicle Re-identification across a Heterogeneous Camera Network

In practice, there is often a need to generate a set of distinctive features in order to re-identify and monitor a vehicle of interest across a network of fixed or mobile

(e.g., drone-borne) non-overlapping cameras. A traditional way to do this is recognizing the vehicle license plate, but in the case of aerial images this method fails due to low resolution and difficult angles for capturing the plate [55]. Instead, vehicles can be recognized based on a chosen set of features and mapped from one scene to another. Vehicle classification can be coarse-grained (i.e., classified with respect to vehicle motion [56], category [57], color [58], size [59] etc.) or fine-grained (make, model, sub model, year [60, 61]). When possible, these features can be combined with the license plate recognition. The features set characterizing vehicles can be extracted using SIFT [62], SURF [63], HOG [64], CNN [65] and other techniques. The object detection and calibration based on previously extracted bounding boxes and estimated sizes, discussed in the previous sections, can be used to make the fine-grained classification faster and more robust.

In this work, we applied the Faster R-CNN detector fed by feature maps provided by the Inception v2 convolutional neural network. This choice has been made due to an appropriate balance between the computational efficiency and accuracy achieved among the detectors discussed in Section 5.2. This detector was trained using the Tensorflow Object Detection API [66, 67] on COD20k [34, 31] and UA-DETRAC [35, 36] datasets.

The fine-grained vehicle classifier has been implemented with the ResNet-50 CNN model. Again, the choice has been made based on good inference speed and classification accuracy. The classification task consisted of determining the vehicle make, model, sub model and year of production. The dataset used to train the ResNet-50 classifier was BoxCars116k [68, 69].

The BoxCars116k dataset was published by Sochor et al., for fine-grained vehicle recognition. The dataset was captured in traffic surveillance setting from 137 different cameras under various viewpoints. It contains 27,496 vehicles on 116,286 images of 45 different makes with 693 fine-grained classes (make, model, sub model, model year). The training subset contains 107 fine-grained classes with 11 653 video sequences containing a tracked vehicle, the so called “tracks”, which amounts to 51 691 images. The testing was performed on 11 125 tracks (39,149 images).

The training was conducted in 20 epochs with the learning rate of 0.0025 and 8 annotated images per batch. The random shift of bounding boxes and horizontal flip techniques [13] were used during the training.

An additional parameter which is important in re-identification of vehicles is color. Our classifier of vehicle colors is based on the MobileNet CNN model [18]. Eight color classes were annotated and recognized: black, blue, brown, green, grey, red, silver and white. The training set consisted of a small data set consisting of 19,751 images annotated with color. This training set, encompassing 44 fine-grained vehicle classes, was created due to the lack of color annotation in the original BoxCars116k dataset. Some annotated examples from this dataset are shown in Fig. 5.20, during the training, the model coefficients were updated using the categorical cross-entropy loss function [70] and the Adam optimizer, an adaptive learning rate optimization algorithm [71], with batches of 32 images. The overall training was conducted in 10 epochs. The color classifier was implemented in

parallel with the fine-grained classifier to provide the vehicle identification based on color, make, model, sub model and year.

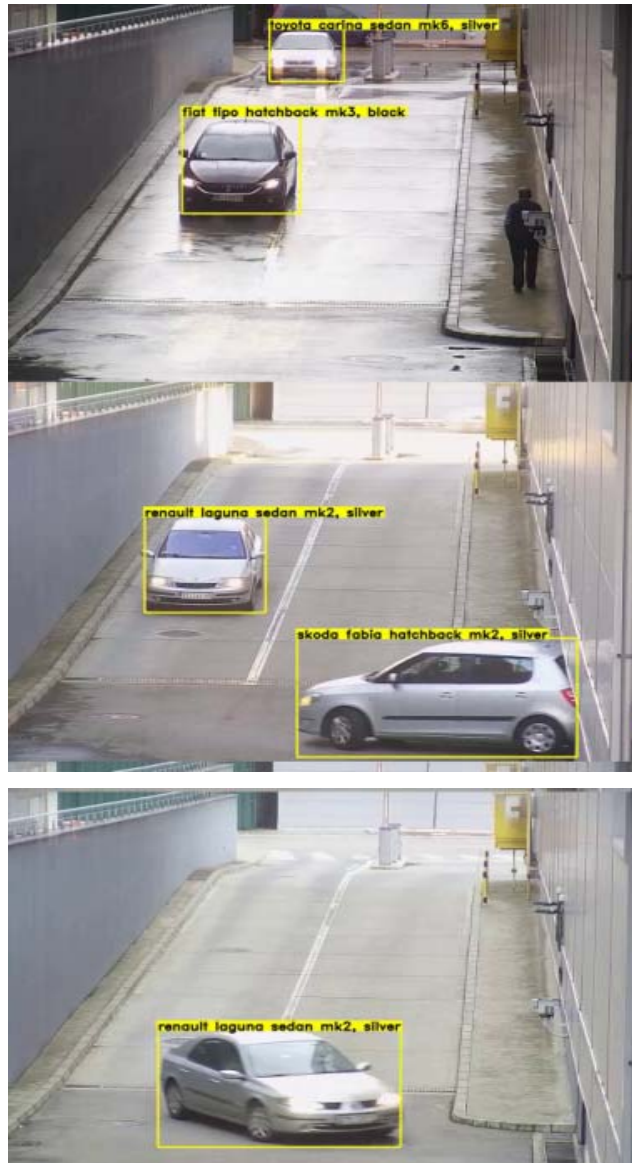


**Fig. 5.20.** Examples of vehicles from the color training dataset grouped by color annotation. The rows from top to bottom contain black, red, gray and blue annotated cars.

The ResNet-50 based classifier is computationally efficient for implementation in smart cameras. On the popular Nvidia Jetson TX2 GPU it runs in 83.3 milliseconds, allowing for real-time classification with a decent frame rate of up to 12 fps. The accuracy of 82.34 % on vehicles detected in a single image and 89.55 % on tracked vehicles within a video sequence. In case an external server is used on transmitted images, as for example with small drone cameras, higher frame rates can be achieved. For instance, if implemented on the Nvidia GeForce GTX 1080 Ti GPU, as much as 238 frames per second can be processed since the processing time per frame is only 4.1 milliseconds.

Typical results of the classifier consisting of parallel fine-grain and color recognition branches are illustrated in Fig. 5.21. These images were obtained using a fixed camera capturing vehicles on a ramp and turning to enter or exit a garage, which allowed for testing the classifier performance on different vehicle poses. Recognition failed only when a vehicle could not be captured in whole.





**Fig. 5.21.** Examples of the make-model-sub model-year and color recognition on video sequences taken by a fixed camera installed in front of a ramp, which allowed for different vehicle poses to be tested. All examples in the above images were correctly recognized.

In a multicamera scenario, a vehicle detected and recognized in the field view of say a fixed camera, is searched in the fields of view of other fixed and mobile sensors. For re-identification of the vehicle, beside the previously discussed and often poorly visible license plates, it is possible to use either the classification result obtained with the initial sensor or the feature map of the vehicle. The additional information that can be used for predicting the location of a vehicle of interest in the coverage area of a heterogeneous (fixed and mobile camera) network can be the coordinates of the last detection and the

velocity estimate [51], which can be broadcast to all sensors in the network. These data can be obtained from the known location of the camera having first detected the vehicle of interest and its position according to one of the two calibration methods discussed in Sections 5.3 and 5.5. Moreover, these calibration techniques can be used to as predict the expected bounding box size and reduce the total amount of search at the region proposal stage of the detector. Additionally, if the extrinsic camera parameters are known as is the case for drone cameras (e.g., DJI Mavic Pro 2), the field of view can be limited to the range of poses the classifier is able to correctly recognize. An illustrative example of the classification-based re-identification applied to high oblique aerial images [72] is shown in Fig. 5.22.



**Fig. 5.22.** Scenario with detecting re-identifying a vehicle in aerial image. The vehicle has been first classified for make-model-sub model-year and color labels in the scene presented in the top part of the figure, captured by a fixed camera. The same vehicle re-appeared in the aerial image in the bottom part of the figure and was positively re-identified as silver-colored Mercedes E Combi MK3 sub model.

## 5.7. Conclusions

In this work we applied detection, calibration and classification methods to localization, recognition and re-identification of road users in heterogeneous non-overlapping

multicamera networks. The heterogeneous camera networks were supposed to consist of both fixed and mobile sensors (such as PTZ or drone cameras). The effective and computationally efficient calibration, detection and classification techniques used in this work can be applied to recognize traffic and public order violations, including complex interactions between people or between vehicles and people, such as failure to yield the right of way to pedestrians. The estimated location and velocity of a road user within the field of view of one camera can be used to predict the closest inbound sensor covering its path and enable, upon re-identification, global tracking of a road user of interest across the coverage area of the camera network. The possibility to use drone cameras and exchange positional, kinematic and descriptive information on the object of interest has a potential to further improve video surveillance systems. This possibility has not yet been fully automated and implemented in this work, but will surely be in focus of our future research.

## Acknowledgements

The authors are grateful to Professor Adam Herout of the Brno University of Technology for his support, mentorship and contributions to the development of techniques and datasets used in this work.

## References

- [1]. S. Kumar, E. Yaghoubi, A. Das, B. Harish, H. Proença, The P-DESTRE: A fully annotated dataset for pedestrian detection, tracking, and short/long-term re-identification from aerial devices, *IEEE Transactions on Information Forensics and Security*, Vol. 16, 2020, pp. 1696-1708.
- [2]. V. Lukić, A. Makarov, G. Tao, B. Choubey, Vehicle speed estimation from tracking license plates, in *Proceedings of the 23<sup>rd</sup> Telecommunications Forum (TELFOR'15)*, Belgrade, 2015, pp. 429-432.
- [3]. O. Rahnama, A. Makarov, P. Torr, Real-time depth processing for embedded platforms, *Proceedings of SPIE*, Vol. 10223, 2017, 102230N.
- [4]. A. Makarov, V. Lukić, S. Vujić, B. Choubey, A stereoscopic rig for speed enforcement, in *Proceedings of the IEEE International Instrumentation & Measurement Technology (I2MTC'16)*, Taipei, Taiwan, 2016.
- [5]. A. Makarov, J.-M. Vesin, Zero-crossing based classification of intrusions, in *Proceedings of the 1<sup>st</sup> International Workshop on Face and Gesture Recognition*, Zurich, 1995.
- [6]. H. Aghajan, A. Cavallaro, Multi-Camera Networks: Principles and Applications, *Elsevier*, 2009.
- [7]. A. Makarov, J.-M. Vesin, F. Raymond, Intrusion detection algorithm robust to slow and abrupt lighting changes, *Proceedings of SPIE*, Vol. 2661, 1996.
- [8]. A. Makarov, Comparison of background extraction based intrusion detection algorithms robust to slow and abrupt lighting changes, in *Proceedings of the International Conference on Image Processing (ICIP'96)*, Lausanne, 1996.
- [9]. A. Makarov, Joint NLOS object detection using edge extraction and nonlinear order statistics, in *Proceedings of the Le Quinzième Colloque sur le Traitement du Signal et des Images (RETSI'95)*, Juan-les-Pins, 1995.

- [10]. A. Makarov, J.-M. Vesin, M. Kunt, Intrusion detection using extraction of moving edges, in *Proceedings of the International Conference on Pattern Recognition (ICPR'94)*, Jerusalem, 1994.
- [11]. J. Špaňhel, J. Sochor, A. Makarov, Detection of traffic violations of road users based on convolutional neural networks, in *Proceedings of the 14<sup>th</sup> Symposium on Neural Networks and Applications (NEUREL'18)*, Belgrade, Serbia, 2018, pp. 1-6.
- [12]. J. Špaňhel, J. Sochor, A. Makarov, Vehicle fine-grained recognition based on convolutional neural networks for real world applications, in *Proceedings of the 14<sup>th</sup> Symposium on Neural Networks and Applications (NEUREL'18)*, Belgrade, 2018, pp. 1-5.
- [13]. O. Russakovsky, et al., ImageNet large scale visual recognition challenge, *International Journal of Computer Vision*, Vol. 115, Issue 3, December 2015, pp. 251-512.
- [14]. A. Zisserman, K. Simonyan, Very deep convolutional networks for large-scale image recognition, in *Proceedings of the 3<sup>rd</sup> International Conference on Learning Representations (ICLR'15)*, San Diego, California, 2015.
- [15]. C. Szegedy, W. Liu, et al., Going deeper with convolutions, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'15)*, Boston, Massachusetts, 2015, pp. 1-9.
- [16]. C. Szegedy, V. Vanhoucke, et al., Rethinking the inception architecture for computer vision, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'16)*, Las Vegas, Nevada, 2016, pp. 2818-2826.
- [17]. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'16)*, Las Vegas, Nevada, 2016, pp. 770-778.
- [18]. A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, Mobilenets: Efficient convolutional neural networks for mobile vision applications, *arXiv Preprint*, arXiv:1704.04861, 2017.
- [19]. J. Huang, et al., Speed/Accuracy trade-offs for modern convolutional object detectors, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'17)*, Honolulu, Hawaii, 2017, pp. 7310-7319.
- [20]. W. Liu, et al., SSD: Single shot multibox detector, in *Proceedings of the European Conference on Computer Vision (ECCV'16)*, 2016, pp. 21-37.
- [21]. J. Redmon, S. K. Divvala, R. B. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'16)*, 2016, pp. 779-788.
- [22]. S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: Towards real-time object detection with region proposal networks, in *Proceedings of the Advances in Neural Information Processing Systems Conference (NIPS'15)*, Montréal, Canada, 2015, pp. 1-9.
- [23]. J. Dai, Y. Li, K. He, J. Sun, R-FCN: Object detection via region-based fully convolutional networks, in *Proceedings of the Advances in Neural Information Processing Systems Conference (NIPS'16)*, Barcelona, Spain, 2016.
- [24]. L. Jiao, et al., A survey of deep learning-based object detection, *IEEE Access*, Vol. 7, 2019, pp. 128837-128868.
- [25]. C. Szegedy, S. Reed, D. Erhan, D. Anguelov, S. Ioffe, Scalable, high-quality object detection, *arXiv Preprint*, arXiv:1412.1441, 2014.
- [26]. S. Bell, C. Lawrence Zitnick, K. Bala, R. Girshick, Inside-outside net: Detecting objects in context with skip pooling and recurrent neural networks, *arXiv Preprint*, arXiv:1512.04143, 2016.
- [27]. J. Dai, K. He, J. Sun, Instance-aware semantic segmentation via multi-task network cascades, in *Proceedings of the IEEE Conference on Computer Vision (CVPR'16)*, Las Vegas, Nevada, 2016.

- [28]. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *Proceedings of the IEEE Conference on Computer Vision (CVPR'16)*, Las Vegas, Nevada, 2016, pp. 770-778.
- [29]. A. Shrivastava, A. Gupta, R. Girshick, Training region-based object detectors with online hard example mining, in *Proceedings of the IEEE Conference on Computer Vision (CVPR'16)*, Las Vegas, Nevada, 2016, USA, 2016.
- [30]. S. Zagoruyko, A. Lerer, T.-Y. Lin, P. O. Pinheiro, S. Gross, S. Chintala, P. Dollar, A multipath network for object detection, *arXiv Preprint*, 2016.
- [31]. P. Zemčik, R. Juránek, M. Musil, P. Musil, M. Hradiš, High performance architecture for object detection in streamed videos, in *Proceedings of the 23<sup>rd</sup> International Conference on Field programmable Logic and Applications (FPL'13)*, 2013, pp. 1-4.
- [32]. R. Juránek, A. Herout, M. Dubska, P. Zemčik, Real-time pose estimation piggybacked on object detection, in *Proceedings of the International Conference on Computer Vision (ICCV'15)*, Santiago de Chile, 2015, pp. 2381-2389.
- [33]. A. Herout, R. Jošth, R. Juránek, J. Havel, M. Hradiš, P. Zemčik, Real-time object detection on CUDA, *Journal of Real-Time Image Processing*, Vol. 6, 2011, pp. 159-170.
- [34]. P. Musil, R. Juránek, M. Musil, P. Zemčik, Cascaded stripe memory engines for multi-scale object detection in FPGA, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 30, Issue 1, 2019, pp. 267-280.
- [35]. R. Juránek, A. Herout., M. Dubska, P. Zemčik, COD20K Dataset, Graph@FIT, Brno University of Technology, <http://www.fit.vutbr.cz/research/groups/graph/PoseEstimation/cod20k.html>
- [36]. L. Wen et al., UA-DETRAC: A new benchmark and protocol for multi-object detection and tracking, *Computer Vision and Image Processing*, Vol. 193, April 2020, 102907.
- [37]. S. Lyu, UA-DETRAC Benchmark Suite, QQ Group, University at Albany, State University of New York, <https://detrac-db.rit.albany.edu/>
- [38]. D. Wang, C. Zhang, H. Cheng, Y. Shang, L. Mei, SPID: Surveillance pedestrian image dataset and performance evaluation for pedestrian detection, in *Proceedings of the Computer Vision Workshops (ACCV'16)*, 2016.
- [39]. V. Lukić, S. Vujić, A. Makarov, Z. Popović, Stereoscopic vehicle speed measurement – System calibration and synchronization errors analysis, in *Proceedings of the International Conference on 3D Imaging (IC3D'11)*, 2011, pp. 1-6.
- [40]. R. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, *Cambridge University Press*, 2003.
- [41]. A. Makarov, V. Lukić, B. Choubey, Real-time vehicle speed estimation based on license plate tracking in monocular video sequences, *Sensors and Transducers*, Vol. 197, Issue 2, 2 February 2016, pp. 78-86.
- [42]. D. Perdomo, J. Alonso, C. Travieso, M. Ferrer, Automatic scene calibration for detecting and tracking people using a single camera, *Engineering Applications of Artificial Intelligence*, Vol. 26, Issue 2, 2013, pp. 924-935.
- [43]. G. T. Laureano, M. S. Veludo de Paiva, A. da Silva Soares, C. J. Coelho, A topological approach for detection of chessboard patterns for camera calibration, in *Emerging Trends in Image Processing, Computer Vision and Pattern Recognition, Elsevier*, 2015, pp. 517-531.
- [44]. L. Krueger, C. Woehler, Accurate chequerboard corner localisation for camera calibration, *Pattern Recognition Letters*, Vol. 32, Issue 10, 2011, pp. 1428-1435.
- [45]. Z. Zhang, Geometric calibration, in *Computer Vision, Springer*, 2014, pp. 333-338.
- [46]. M. Dubska, A. Herout, J. Sochor, Automatic camera calibration for traffic understanding, in *Proceedings of the British Machine Vision Conference*, 2014.
- [47]. M. Dubska, A. Herout, R. Juránek, J. Sochor, Fully automatic roadside camera calibration for traffic surveillance, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 16, Issue 3, June 2015, pp. 1162-1171.

- [48]. Y. Zheng, S. Peng, A practical roadside camera calibration method based on least squares optimization, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 15, Issue 2, 2014, pp. 831-843.
- [49]. Z. Zhang, T. Tan, K. Huang, Y. Wang, Practical camera calibration from moving objects for traffic scene surveillance, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 23, Issue 3, 2013, pp. 518-533.
- [50]. Average-height-around-the-world, BoneScience®, <https://thebonescience.com/blogs/journal/average-height-around-the-world>
- [51]. R. Kuczmarski, C. Ogden, S. Guo, et al., 2000 CDC growth charts for the United States: Methods and development, *Vital and Health Statistics*, Vol. 11, 2002, 246.
- [52]. A. Makarov, V. Lukić, O. Rahnema, Distance and speed measurements from monocular images, *Proceedings of SPIE*, Vol. 9897, 2016, 98970I.
- [53]. N. Martinel, C. Micheloni, Re-identify people in wide area camera network, in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'12)*, 2012, pp. 31-36.
- [54]. L. Liu, Y. Zhang, J. Chen, C. Gao, FUsing global and semantic-part features with multiple granularities for person re-identification, in *Proceedings of the IEEE International Conference on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Fusing Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom'19)*, 2019, pp. 1436-1440.
- [55]. K. Islam, Person search: New paradigm of person re-identification: A survey and outlook of recent works, *Image and Vision Computing*, Vol. 101, 2020, 103970.
- [56]. A. Makarov, V. Lukić, M. Španović, Deriving camera and lens settings for fixed traffic enforcement and ALPR cameras, in *Proceedings of the 20<sup>th</sup> Telecommunications Forum (TELFOR'12)*, Belgrade, 2012, pp. 670-676.
- [57]. M. Pavlović, N. Stojiljković, I. Gluvačević, M. Vučetić, M. Perić, Real-time moving object of interest detection in multi-sensor imaging system, in *Proceedings of the 7<sup>th</sup> International Conference on Electrical, Electronic and Computing Engineering (IcETRAN'20)*, Belgrade, 2020.
- [58]. P. Burnos, Alternative automatic vehicle classification method, *Metrology and Measurement Systems*, Vol. 17, Issue 3, 2010, pp. 323-334.
- [59]. J. Hsieh, L. Chen, S. Chen, D. Chen, S. Alghyaline, H. Chiang, Vehicle color classification under different lighting conditions through color correction, *IEEE Sensors Journal*, Vol. 15, Issue 2, 2015, pp. 971-983.
- [60]. Y. Liu, M. Reynolds, D. Huynh, G. Hassan, Study of accurate and fast estimation method of vehicle length based on YOLOs, in *Proceedings of the IEEE International Conference on Artificial Intelligence and Information Systems (ICAIS'20)*, 2020, pp. 118-121.
- [61]. M. Biglari, A. Soleimani, H. Hassanpour, A cascaded part-based system for fine-grained vehicle classification, *IEEE Transactions on Intelligent Transportation System*, Vol. 19, Issue 1, 2018, pp. 273-283.
- [62]. M. A. Manzoor, Y. Morgan, Vehicle make and model classification system using bag of SIFT features, in *Proceedings of the IEEE 7<sup>th</sup> Annual Computing and Communication Workshop and Conference (CCWC'17)*, 2017, pp. 1-5.
- [63]. J.W. Hsie, L.C. Chen, D.Y. Chen, Symmetrical SURF and its applications to vehicle detection and vehicle make and model recognition, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 15, Issue 1, 2014, pp. 6-20.
- [64]. Y. Wei, Q. Tian, J. Guo, W. Huang, J. Cao, Multi-vehicle detection algorithm through combining Harr and HOG features, *Mathematics and Computers in Simulation*, Vol. 155, 2019, pp. 130-145.

- 
- [65]. Y. Tan, Y. Xu, S. Das, A. Chaudhry, Vehicle detection and classification in aerial imagery, in *Proceedings of the 25<sup>th</sup> IEEE International Conference on Image Processing (ICIP'18)*, 2018, pp. 86-90.
  - [66]. Tensorflow, Object Detection, [https://www.tensorflow.org/lite/examples/object\\_detection/overview](https://www.tensorflow.org/lite/examples/object_detection/overview)
  - [67]. A. Margunov, How to Train Your Own Object Detector Using TensorFlow Object Detection API, Neptun, <https://neptune.ai/blog/how-to-train-your-own-object-detector-using-tensorflow-object-detection-api>
  - [68]. J. Sochor, BoxCars116k, Brno University of Technology <https://medusa.fit.vutbr.cz/traffic/data/BoxCars116k.zip>
  - [69]. J. Sochor, J. Špaňhel, A. Herout, BoxCars: Improving fine-grained recognition of vehicles using 3-D bounding boxes in traffic surveillance, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 20, Issue 1, 2019, pp. 97-108.
  - [70]. Z. Zhang, M. Sabuncu, Generalized cross entropy loss for training deep neural networks with noisy labels, in *Proceedings of the 32<sup>nd</sup> International Conference on Neural Information Processing Systems (NIPS'18)*, 2018.
  - [71]. D. Kingma, J. Ba, Adam: A method for stochastic optimization, in *Proceedings of the 3<sup>rd</sup> International Conference for Learning Representations*, San Diego, 2015.
  - [72]. S. Verykokou, C. Ioannidis, Oblique aerial images: A review focusing on georeferencing procedures, *International Journal of Remote Sensing*, Vol. 39, Issue 11, 2018, pp. 3452-3496.

# **Chapter 6**

## **Measurement of Microclimate as a Comfort Indicator**

**Bernhard Kurz and Christoph Russ**

### **6.1. Introduction**

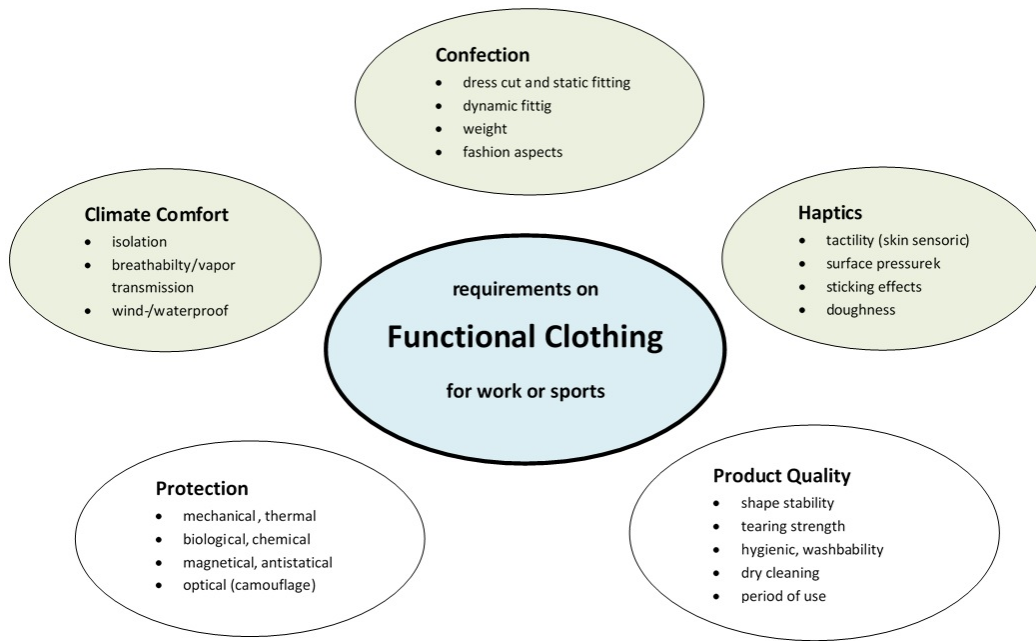
The requirements which are directed to modern apparel like work wear, protective wear, business and fashion wear, as well as to high quality seating or sleep systems, have changed fundamentally in the last decades. Today, apparel systems are expected to be multifunctional, e.g. protection against chemical, mechanical or thermal effects, but also skin-sensory or haptic aspects as well as odor, hygiene and cleaning requirements or biomechanical specifications about fit (see Fig. 6.1).

In addition to the above functions comfort aspects can be roughly divided into the areas of biomechanical and climatic comfort and amend the portfolio of requirements. It must be considered that both these comfort aspects interact and influence each other's physiological consequences. Poor shoe fit, for example, may increase the biomechanical risk for skin irritations. But in combination with a warm/humid shoe microclimate the risk of comfort loss may even accelerate the process because the mechanical properties of the skin get adversely affected (maceration). It is also known that in the initial phase of using apparel components or seating and sleep systems, at first the biomechanical conditions determine the wearing comfort, better their discomfort, and only after some time the climatic factors gain awareness and sometimes even take over the dominant role. Since comfort may ultimately be determined by no unpleasant sensation, any individual source for discomfort must have a significant influence on the overall perception (see Fig. 6.2). Thus, comfort has a direct impact on physical and mental performance, relaxation, and wellbeing.

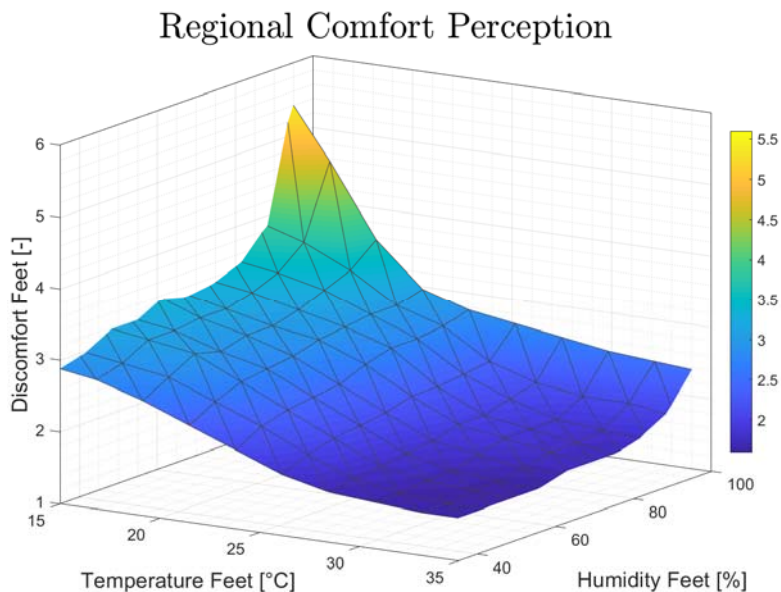
In addition to ergonomic and skin-sensory aspects, such as sweat-induced adhesion or moisture stains, climatic comfort plays a central role, since the global human



thermophysiological requirement is only met by an even heat balance. At the same time, the heat transport, breathability and moisture management of the garments ensure a pleasant local microclimate at the human/textile interface.



**Fig. 6.1.** Ergonomical requirements on clothing, green areas with comfort relation.



**Fig. 6.2.** Local discomfort perception in the feet region at various microclimates in the shoe.

Existing testing standards dealing with climatic functions were developed to focus on extreme situations, i.e., protection against rain from the outside but maximum vapor flow rate from the inside. With the successful rise of the outdoor or sportswear industries in the last decades the aim became to determine minimum and maximum heat transfer and vapor transfer rates.

From a climatic comfort point of view, however, extremes such as humidity saturation at the skin/textile interface may happen less than 10% of the using time in everyday life and else must be viewed as unrealistic and certainly unpleasant. This means 90 % of the wearing time are not directly considered and still open for investigation and proof of concept. This requires not only realistic simulation testing to obtain key figures, but also indicators that are comprehensible and meaningful to consumers, for example in the form of combined climate measures such as the Heat Index [6].

In the construction, design, and assessment of climaphysiological functions of apparel or body support systems, and incidentally also for actuator and operating handles, a multi-stage analysis process is consequently indicated by:

- Determination of (bio-)physical parameters for the evaluation of thermoregulatory properties (even heat balance);
- Quantification of the comfort-determining microclimate as a (human-) physiological criterion by standardized testing methodology;
- Use of validated predictive models for comfort prognosis;
- Selected, supplementary validation through tests with probands.

According to numerous studies from a wide range of private and public research institutions [1, 3, 5, 12, 15, 19, 24], the microclimate next to the skin represents the decisive, multidimensional indicator to assess the thermoregulatory function and the climatic and ergonomic wearing comfort. Thus, it must be determined as realistic as possible and requires suitable sensor technology as well as a reproducible and standardized heat and moisture source, without subjective influences from a human subject-based test procedure. The following explanations intend to provide scientific, economical, and practicable approaches.

## **6.2. Microclimate**

To define climate comfort in particular the physiological requirements of the thermoregulation as well as the individual sensations must be analyzed. When humans are asked whether they prefer warm or cool respectively dry or humid climate, the answers indicate a clear preference for dry. Warm or cool remains to be evenly distributed. This is although humans do not have any moisture receptors. Hence, the concept of warm or cold sensation is not only a matter of temperature but must be a combination of heat and humidity developments. This concludes that climate comfort is a matter of “dry” in the first place, warm or cold only ranks second.

To understand the central role of microclimate on the comfort perception, a look at the human thermophysiology can help. Its goal is the production of heat to stabilize the core temperature at about 37 °C. According to any thermal or physical load the heat production will vary. In case of too much heat the necessity of discharge will increase (e.g., sweating), otherwise the heat loss is decreased and additional heat production (e.g., trembling) will start. Heat loss is physiologically controlled by a change of blood circulation within the skin tissue and, in particular the heat increase, also through the evaporation of sweat. In both cases the apparel plays the central role because the heat discharge through the skin reaches about 90 % and just 10 % by breathing [9, 16, 21]. In cold environs apparel must support isolation to avoid too much heat loss. In warm environs the heat and mainly the humidity transmission must be enabled, regarding an evaporation fraction of about 70 % when temperature rises over 30 °C. Both heat transmission mechanisms will determine the temperature and humidity in the microclimate in the layer between skin and textile.

An applied rating method for global thermal comfort which can be found in various scientific literature [5, 8, 11] are the widely accepted ranges of skin temperature for non-extreme ambient climate conditions, limited wind speed and radiation, e.g., specified by the effective temperature [7] as well as energy consumptions below the continuous work capacity (approx. 250 W) which are defined as:

- Optimal range: 31,5 °C – 35,5 °C;
- Transition towards warm: 35,5 °C – 37 °C;
- Transition towards cold: 29 °C – 31,5 °C;
- Comfort threshold warm: above 37 °C (37.5 °C);
- Comfort threshold cold: below 29 °C.

Confirmed by many trials with human occupants on seats, in beds, and with diverse apparel products the human thermoregulation behavior measured as the microclimate close to the skin can be separated into three phases (see Fig. 6.3):

**Phase 1:** Thermal neutrality, characterized by:

- Low temperature changes with up to +/- 1 °C within the comfort range 31,5 °C – 35,5 °C;
- The heat balance is even, sweating as active body own method to influence the heat only happens in little doses (fine tuning);
- Dynamic regulating activity of low sensitive perspiration with slightly upward and downward developing absolute humidity levels (up to 25 g/kg) occur, which over time do not accumulate in the human/textile interface.

**Phase 2:** Tendency to feel cold, characterized by:

- Decrease of the skin temperature below the comfort zone; risk of a reduction of the body core temperature;

- Demand for heat supply increases; sensitive (active) sweating shuts down; humidity gradients decrease or remain stable;
- Lowest sweat emission, only insensitive sweating with very little absolute humidity and low to zero dynamic.

**Phase 3:** Tendency to overheat and sweat, characterized by:

- Increasing skin temperature over the comfort level; increasing tendency of the body core temperature;
- The demand for heat release increases; growing sweat activity with higher upward trends of humidity levels (over 25 g/kg);
- In the measurable view a medium to high accumulating absolute humidity in textile layers can be seen.

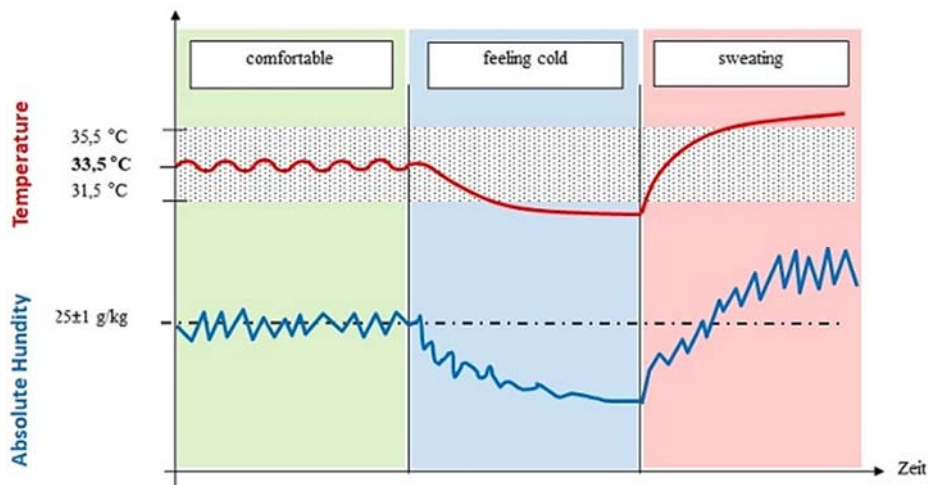


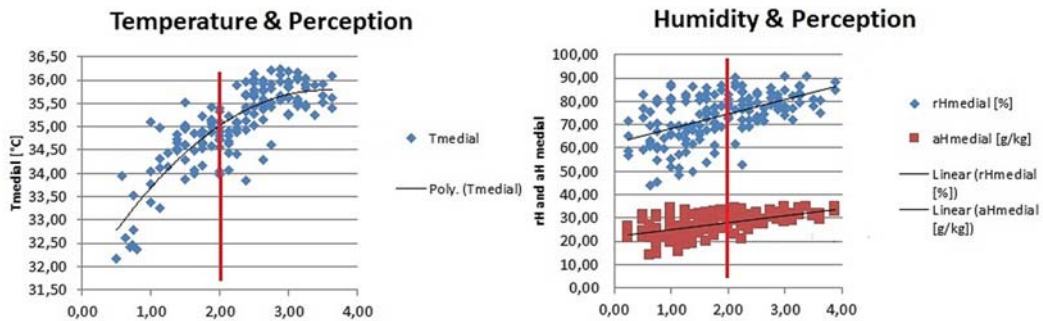
Fig. 6.3. Characteristics of thermoregulative phases.

Thus, the decisive conclusion for a measured assessment of climate comfort is to focus on the directly developing microclimate at the human/textile interface providing kind of a blueprint of the thermoregulation processes [12]. Combined with individual climate sensations the perceived comfort can be derived with high validity (see Fig. 6.4).

The investigation on regional comfort temperatures at diverse body parts can be found with Arens and Zang [2]. The common way to conduct such climate data in most cases is that they are recorded close to the skin or with one apparel interlayer and at different ambient temperatures as well as microclimate sensors.

Under these circumstances it must be realized that wearing different apparel systems under use of different measuring devices complicates the choice of suitable and comparable measurement positions. Moreover, the thermoregulating homiothermia

makes for the skin temperature to scarcely vary, which requires a precise and reproducible temperature resp. temperature distribution measurement, e.g., as mean skin temperature acc. to Ramanathan. Practical approaches to solve these challenges are an area related detection of the temperature and humidity allocation with cluster formation, and correlation analysis between perception and selected measurement sectors. Given a predominantly direct contact of a human and a test object, e.g., seat, laying or shoe systems, and only one defined interlayer of apparel the evaluation of thermal comfort can be obtained from the temperature values in the interface apparel to test object, e.g., directly on the seat or cushion surface [14, 17].

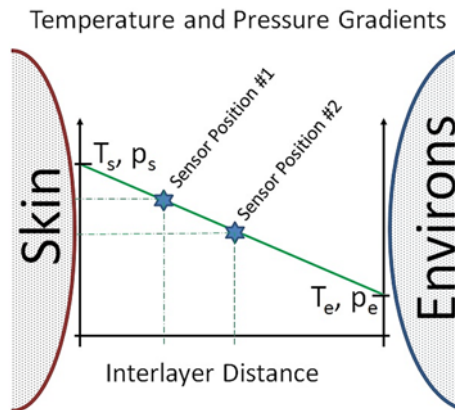


**Fig. 6.4.** Combined temperature and humidity perceptions of proband test in shoes at various climate and physical stress scenarios.

The so-called "skin-near" microclimate is ultimately determined by the heat and moisture released from the skin, the transfer properties of the air and textile layers next to the skin, and the ambient/room climate. Depending on the spatial extent of this layer system, different climate combinations occur with increasing distance from the skin surface. These are also determined by the heat and moisture transfers generated by the skin due to thermoregulatory circulation and sweating rates, which can be highly variable. Buffer and storage effects due to heat capacities, vaporous moisture accumulations (evaporative capacity) or condensation effects have an additional influence and act time-dependent, both in the formation of a stationary microclimate state (steady state) and during redrying.

In the case of stationary transport processes through a textile layer with known heat and moisture transfer resistance and without or after saturation of the storage capacities, heat and moisture transport are determined only by the temperature and moisture gradient. Assuming a temperature of the sweating skin of 35 °C and 100 % relative humidity, corresponding to a vapor pressure of 56 hPa, and an ambient climate of 20 °C and 50 % relative humidity, corresponding to 11 hPa, the microclimate conditions at each position can be determined or measured with suitable miniaturized sensor technology if the layer geometry (path length for transverse transport) is known (see Fig. 6.5). It must be taken into account that, above a certain sensor size, an additional cavity is created which also influences the thermal insulation, and a possible heat dissipation via the sensor cables may occur. Thus, not only the size and contacting of the climate sensors but also their orientation, i.e., active measurement window faced to the source or to the environment, as

well as the location position on the contact surface with the skin are in focus. The latter is of central importance, especially for measurements with single sensors on test persons, since small deviations from the desired sweat-producing area have an effect on the measured humidity values. In addition to reproducible and reliable measured values, there is also a demand for two-dimensional recording of temperature and humidity with miniaturized, dew-resistant climate sensors, which are preferably arranged in a fixed grid and with fixed orientation of the measuring window (no twisting).



**Fig. 6.5.** Microclimate at various positions along the gradients of temperature (T) and water vapour pressure (p).

Since thermoregulatory processes in humans are always subject to control dynamics (see Fig. 6.3), which also show inter- and intraindividual scatter, a standardized heat and humidity source should also be used instead of test subjects for reproducible analysis of microclimate states. Only then it will be possible to generate comfort-determining microclimates under steady-state conditions and with high physiological relevance, which are fundamental for reliable comfort predictions.

### 6.2.1. Sensor Technology

Usually, the measurement of the microclimate is realized with discrete combi-sensors for temperature and humidity [22, 23]. The inaccuracies that occur in sensor location and occupancy or coverage led to a scattering of the measured values and limit the quality of the information. A solution to avoid the difficulties with single sensors is a defined array of up to 60 sensors combined with an interpolation algorithm calculating the temperature and humidity distribution with all discrete sensor readings. To reduce wiring and technical complexity bus-compatible combi-sensors with I2C interface are built into a flexible, thin and compressible spacer mat, which enables two- and even three-dimensional recordings of the microclimates [13].

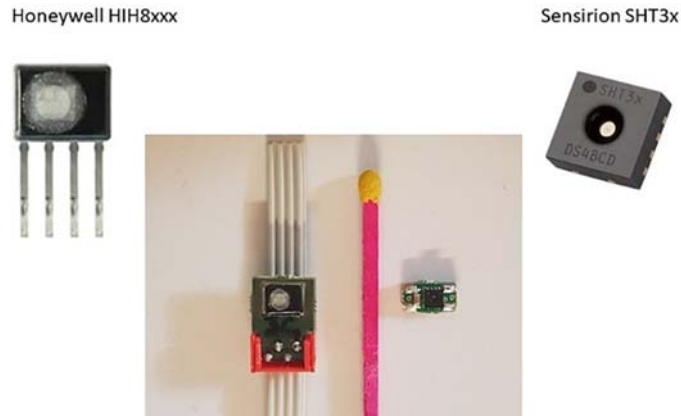
By laborious market analysis two suitable sensors with different specifications have been identified.

- Honeywell HIH8xxx:
  - I2C with up to 128 programmable addresses;
  - $\pm 2$  % humidity resolution, 0 ... 100 % rel. humidity;
  - $\pm 0,5$  °C temperature resolution, -40 ... 125 °C;
  - Measures:  $4 \times 5 \times 2$  mm<sup>3</sup>;
  - Read response time ca. 33 ms;
  - Condensation resistant, with or without filter membrane.
- Sensiron SHT3x:
  - I2C with 2 selectable addresses;
  - $\pm 1,5$  % humidity resolution, 0 ... 100 % rel. humidity;
  - $\pm 0,1$  °C temperature resolution, -40 ... 125 °C;
  - Measures:  $2,5 \times 2,5 \times 1$  mm<sup>3</sup>;
  - Read response time 4–15 ms;
  - Condensation resistant, with or without filter membrane.

The Honeywell HIH8xxx can be directly implemented in an array on one bus cable, has a sufficient but low resolution in temperature and is bigger than the Sensiron chip. The main disadvantage of the SHT3x is the restricted address range (just 2) for I2C operation with up to 60 sensors on one bus cable. To get it also operated an additional PIC (peripheral interface controller, Microchip corp.) is implemented for each SHT3x which serves as a bus respectively address manager. This solution offers a so-called flash-mode, where converting of all sensors is simultaneously initialized. The collection of the converted temperature and humidity values is then sequentially. This mode makes sure, that the complete sensor array is read within a few milliseconds, while the reading of 30 Honeywell sensors takes about 1 second.

To connect each sensor unit the SMD sensor chips are mounted on a tiny printed circuit board (PCB) which allows also the integration of a special crimp connector for the I2C wires and, in case of SHT3x, of the PIC (see Fig. 6.6). Before implementation the programming of different addresses to HIH8xxx as well as address managing routine to PIC of SHT3x is necessary.

By use of a 4-wire ribbon cable the sensor-PCBs can be arranged in any distance or matrix dimensions very quick. To finish the sensor array it will be prefixed on a spacer material and is sewed and covered by a highly breathable mesh material (see Fig. 6.7). The resulting sensor mat is about 4-5 mm thick and has for seat or backrest application a typical dimension of  $500 \times 500$  mm<sup>2</sup>. Compared with single sensor measurement the SWEATLOG mat takes neglectable influence on the microclimate, is highly flexible and confirms mechanical robustness (connector, cable, sensor-PCB) within many tests.



**Fig. 6.6.** I2C-sensors processed for measuring the microclimate distribution.



**Fig. 6.7.** SWEATLOG logger and sensor array (mat) with up to 60 HIH8xxx sensors at one I2C ribbon (4 wires).

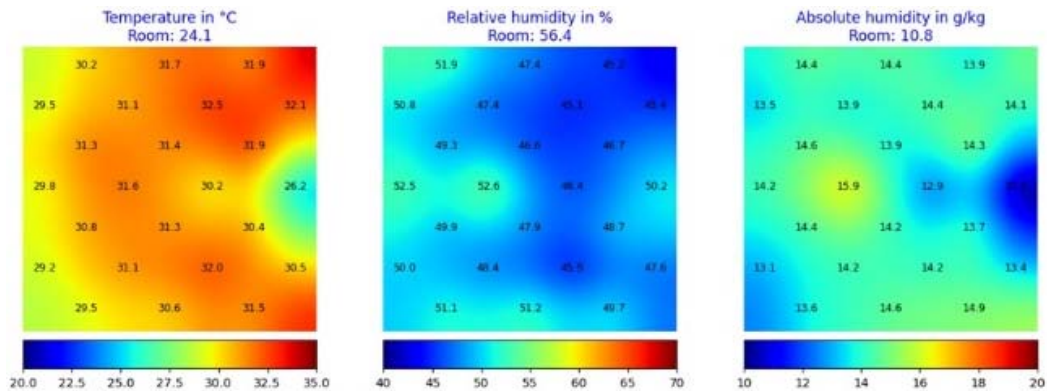
For application in shoes or other narrow spaces a sensor set with highly flexible wiring is available (up to 16 sensors). To achieve the demands on a valuable and reproduceable measurement the single sensors have to be fixed in a defined position, even for visualization, and orientation, e.g., directly on then test target or by use of a suitable grid.

The SWEATLOG sensor mats as well as wired sensor set are operated by a logger which works either as stand-alone or USB-controlled. In the USB mode the SWEATLOG-app allows an online monitoring and visualisation of the registered values as a heat and humidity map (see Fig. 6.8), otherwise these are processed from the data stored on the internal SD-card.

The visualizations are calculated by use of each sensor value with a special interpolation algorithm to get the continuous heat and humidity distributions. Additional clustering methods supports the processing of the characteristic values of thermodynamical or



physiological relevance [13, 17]. Fig. 6.8 shows the typical distribution of a proband sitting on a chair showing the contact zones of the buttock and the upper thighs impressively in the heat map, whereas the map of the relative humidity is nearly homogeneous and in the range of the ambient humidity, seeming that the occupant is not sweating. But a look at the absolute humidity map confirms an increase of the humidity in the microclimate as an indicator for additional slight sweating.



**Fig. 6.8.** Heat and humidity map of a proband on the seat of an office chair.

### 6.2.2. Sweating Source

As a consequence of the standardisation demands for climate tests, and of course due to cost extensive proband tests, artificial systems are applied. Those emit heat and humidity aiming to measure and calculate physical material parameters such as heat conduction and vapor transmission respectively storage. Others aim to simulate the human heat and humidity emission (thermoregulation) creating a certain microclimate with the test object, which allows a reliable prediction of the climate comfort [4, 10, 18, 20]. Such climate dummies can be found either as whole-body manikins or formed as single body parts. They are mostly of highly complex technology and sometimes come with instable processes. Standards have been available for decades like Sweating Guarded Hotplate (SGHP) respectively Hohenstein's Skin Model [26] but without a focus on microclimate.

An additional challenge, but not only related to the use of climate manikins, is the standardization of test conditions with view on valid comfort assessments. Unfortunately, no consistent ambient climate conditions, no consistent measurement spots and sensor qualities, no fixation of heat and particularly humidity emissions and last not least no standardized valuation variables can be found in the existing test procedures. So, the interpretation of the corresponding results only remains to be a relative statement based on the subjective experience of the examiner.

Apart from standardized test conditions the particular challenge for the setup and operation of such climate simulation systems is the best possible replication of the physiological human evaporation and the combination of heat and humidity emissions.

Only this way defined climatic conditions and energy rates will ensure realistic heat and moisture flows and stationary microclimate situations.

According to the necessary features of a standardized heat and humidity source the SWEATOR system has been developed [12]. This technology is based on a water-filled, heat controlled hollow body with a specific perforated surface, coated by a special water vapor permeable membrane. This sweating devices can be manufactured in different shapes (see Fig. 6.9), with different permeable membranes and without or with different surface perforations. The heating and water circulation are controlled by a touch screen control unit. Energy inputs from a few Watts to 400 W per square meter and sweating rates between 50 and 600 ml/m<sup>2</sup>\*h are possible.



**Fig. 6.9.** SWEATOR torso with test jacket, SWEATOR foot and SWEATOR head with safety helmet.

For human physiological simulations under normal conditions, settings of 30-50 W/m<sup>2</sup> or 90-120 ml/m<sup>2</sup>\*h are used. After pre-conditioning the device to the desired values, the SWEATOR manikin is applied to the confectioned test objects. In the boundary layer between SWEATOR surface and material probe the microclimate distribution is measured with single sensors resp. a sensor array at defined locations (see Section 6.2.1). The test trials take place under climatically defined room conditions including additional convections, if necessary. The SWEATOR technology thus enables realistic test conditions and non-destructive testing options with ready-made products.

The SWEATOR control unit records and stores the supplied energies, weight changes and microclimate conditions even within various textile layers or outlets of ventilation channels.

Regarding the thermal system self-loss, the moisture quantity, stored in the sample during the experiment, and the mass flow, converted by use of approximately constant

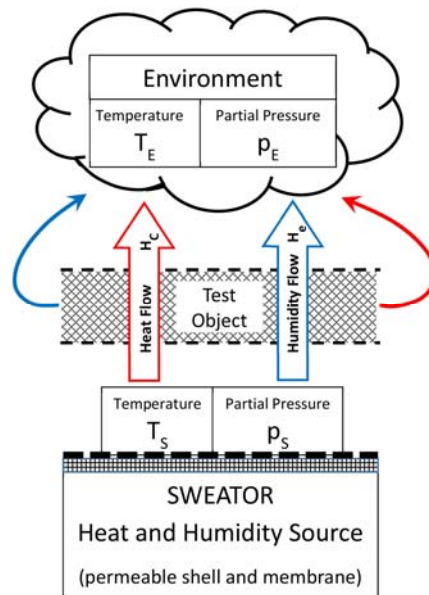
evaporation energy of approx. 2400 J/g, the most important values of heat flow  $H_c$  and moisture flow  $H_e$ , each with the unit Watt, are calculated. This means that the result variables:

- Moisture Vapor Transmission Rate MVTR [g/h];
- Evaporative Capacity EC [W];
- Storage from the vapor phase  $\Delta W_p$  [g];

can be determined. Including the measured microclimate parameters temperature, relative humidity or water vapor partial pressure between two measuring spots even the thermodynamic material parameters:

- Thermal resistance  $R_{ct}$  [ $K \cdot m^2/W$ ];
- Water vapor transmission resistance  $R_{et}$  [ $Pa \cdot m^2/W$ ];

are calculated (see Fig. 6.10).



**Fig. 6.10.** Heat and moisture transport in SWEATOR simulation.

A special feature of SWEATOR technology is the moisture release via a water vapor-permeable membrane, additionally influenced by targeted perforation of the release surface. In contrast to the sweat glands of the human skin, the moisture is released exclusively in the form of vapor and avoids the formation of "wet spots" that can occur with liquid humidity release. These are always an indication of overloading of the moisture removal system, do not describe a desired operating condition and massively change the thermal conductivity value. The maximum vaporous sweating rate is limited by the above-mentioned design features of the SWEATOR dummies, the occurring

sweating rate is determined by the partial pressure gradient and moisture transmission resistance. The simultaneous application of heat and moisture and the almost arbitrary shaping of the test specimen guarantee non-destructive product testing (clothing, seat, bed) with high thermophysiological relevance.

### 6.3. Applications

The following application examples with the SWEATOR heat and humidity source and the SWEATLOG system for measuring temperature and humidity distribution demonstrate application practicability as well as options for evaluating the climatic comfort of clothing components and seating and bedding systems.

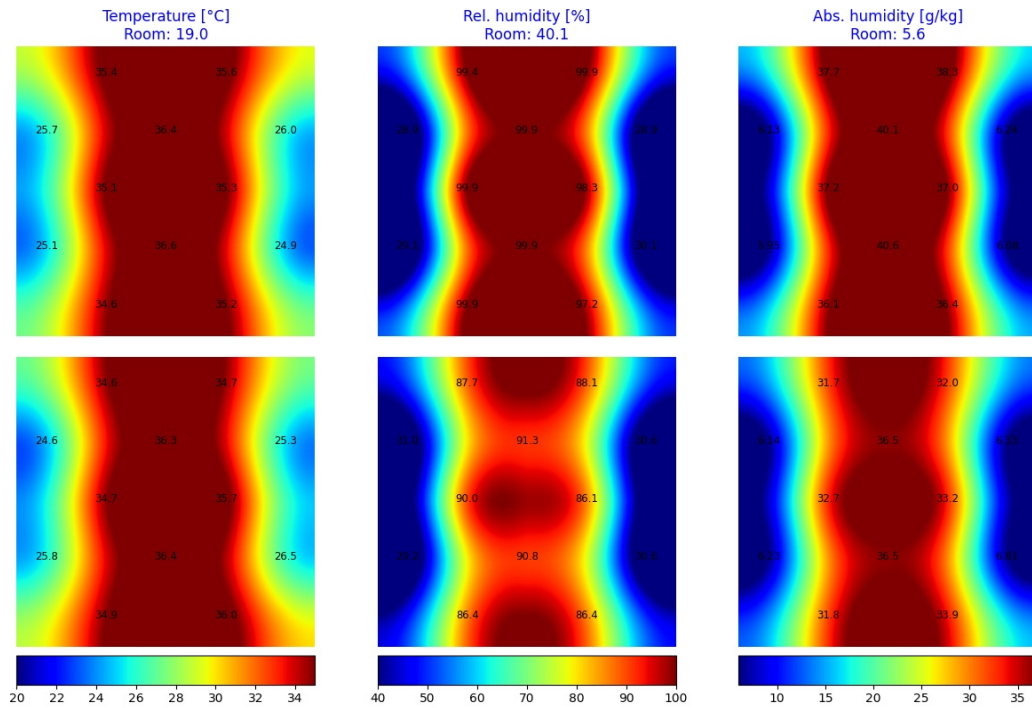
#### 6.3.1. SWEATOR and SWEATLOG

Fig. 6.11 shows the results of validation tests with SWEATOR torso and SWEATLOG sensor array. The sensor mat, constructed with a 4 mm spacer (white) and green cover fabric (see Fig 6.7), is placed on a reference mattress (100 mm PU cold foam) and repositioned for each test run. The torso is separately preclimatized to 37 °C core temperature and then placed on the test pad with measuring mat for 90 minutes. The measurements are repeated several times, in each case with the sensor openings oriented towards the torso or toward the pad. The final steady-state values of the evaluated cluster sizes in the central contact area of the torso sweating body, covered by 4 sensors, are summarized in Table 6.1.

The reproducibility of similar test setups is in the range  $\pm 0.12$  °C and  $\pm 0.4$  % relative humidity or  $\pm 0.25$  g/kg absolute humidity. Sensor orientation has virtually no effect on temperature but shows a reduced relative humidity of 87.6 % and 34.25 g/kg absolute humidity, respectively.

**Table 6.1.** Steady state values with mean value (MV) and standard deviation (SD) after 90 minutes in the central cluster (each test 4 times).

Test run	Temperature [°C]	Rel. humidity [%]	Abs. humidity [g/kg]
“to the source” #01	35.85	99.1	38.6
“to the source” #02	35.85	99.4	38.7
“to the source” #03	35.65	99.9	38.5
“to the source” #04	35.90	99.9	39.0
<i>MV / SD</i>	<i>36.81 / 0.11</i>	<i>99.57 / 0.39</i>	<i>38.70 / 0.21</i>
“to the pad” #11	35.99	87.8	34.5
“to the pad” #12	35.97	87.5	34.3
“to the pad” #13	35.75	87.6	33.9
“to the pad” #14	35.97	87.5	34.3
<i>MV / SD</i>	<i>35.92 / 0.11</i>	<i>87.6 / 0.14</i>	<i>34.25 / 0.25</i>



**Fig. 6.11.** Heat and humidity map by SWEATLOG mat, sensor window facing to source (above) resp. turned away from source, i.e., facing the pad (below).

The combination of standardized heat and humidity source (SWEATOR torso) with two-dimensional recording of the temperature and humidity distribution guarantees high reproducibility of the occurring microclimates. The influence of the sensor orientation between "to the source" and "to the pad" is accompanied by a difference of about 12 % relative humidity and is also highly reproducible. An influence of the green cover fabric on the climate measurands could not be determined.

### 6.3.2. Validation of Proband and SWEATOR Test

The focus of the SWEATOR simulation combined with SWEATLOG sensor mats is primarily set on the analysis of heat and moisture transport routes and, in particular, on the generation of realistic microclimates for direct comfort assessments. This requires, on the one hand, a secure mapping between regional microclimate conditions (seat, bed, jacket, etc.) and the corresponding comfort sensations of the users, and on the other hand, the appropriate choice of parameters for the simulation system.

For this purpose, a variety of studies on comfort correlations in seating systems, footwear and apparel components could be linked to the comfort areas. After evaluation and harmonization of the rating scales (e.g. [25]), the transition temperature and humidity were determined to reach discomfort in warm temperatures at about 35.5 °C respectively 25 g/kg [12].

The results in Fig. 6.12 emphasize the correlation between human subject and simulation test as well as to the perceptions of the various climate situations (test phase) appearing on car seats with active cooling systems during driving. As expected, the humidity perception is weak, but the temperature and especially the comfort perception, combining heat and humidity, presents a very good correlation.

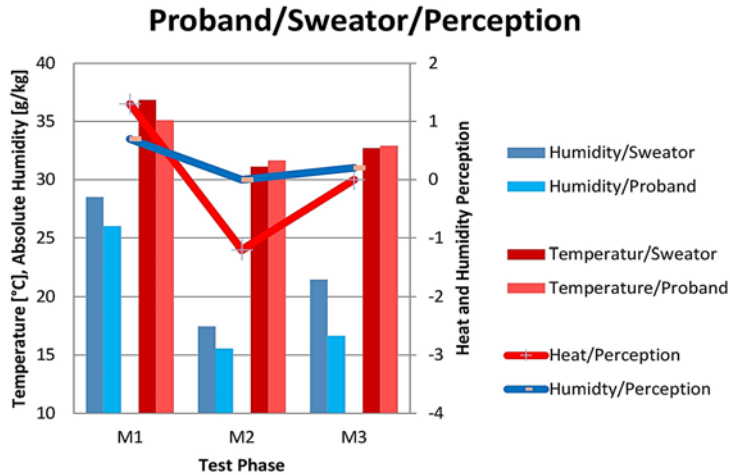
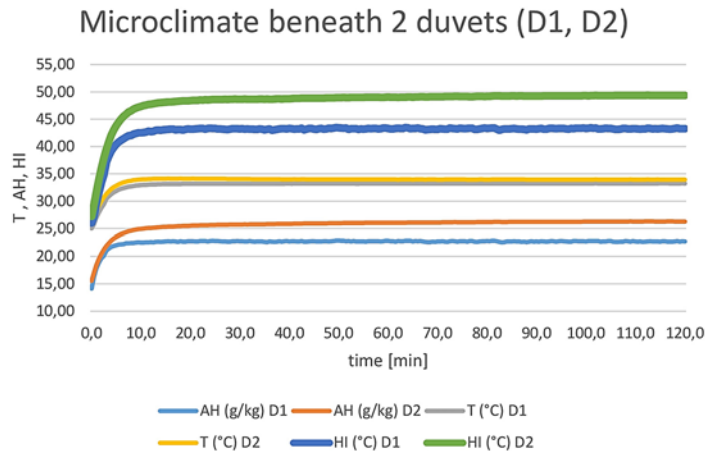


Fig. 6.12. Comfort correlation of proband and simulation results from seating tests.

### 6.3.3. Skin Integrity on Mattresses

Optimized lying systems, i.e., bedsteads with mattresses and covers, are a basic prerequisite for physical rest. In addition to special hygienic requirements, this system is of decisive importance in the care sector due to lengthy periods of lying and use. The focus is on skin integrity with the prevention of decubitus diseases, which can be a cause of continuous exposure to pressure in combination with a warm and humid climate. The weakened skin areas show degenerative changes with subsequent pressure ulcer formation even at low pressure levels. The key factors in preventing such diseases are therefore pressure reduction and pressure shifting (repositioning of patients) as well as optimized moisture management, possibly supported by technical means such as ventilation channels in the mattress or the use of special bed covers.

Fig. 6.13 shows the clustered microclimate variables temperature and absolute humidity under two different duvets on identical bedding systems. The Heat Index HI, determined by temperature and humidity as an indicator for climate perception and comfort perception, is shown as an additional climate qualifier. While the final values of the temperatures of both duvets hardly differ, already for the absolute humidity's a clear distinction is to be determined on the basis of the comfort limit of 25 g/kg. This is even more evident in the HI, which shows a significantly higher heat perception for duvet no. D2.



**Fig. 6.13.** Microclimate management (T: temperature, AH: absolute humidity) and Heat Index (HI) at 2 bedding systems.

## 6.4. Conclusion

Although microclimate is only one aspect when it comes to look at comfort determining factors, it decisively gains importance with increasing wearing or using cycles of apparel, seating and laying systems. Therefore, microclimate requirements already need to be considered in the planning and construction phase of apparel and body support systems. One of the challenges arising is the complex validation of climate comfort with test and evaluation standards to guarantee reproducible and comparable results as well as reliable forecasts. Extensive research work and data analysis at Munich University of Applied Sciences (MUAS) provide valid correlations between simulation results, tests with human occupants and their perception of comfort.

To gain a better understanding about climate comfort it seems relevant how intensely heat and vapor accumulate within the skin near microclimate because the microclimate directly affects the human heat sensation and under standard conditions allows to quantify it. Thermodynamic indicators should not only be conducted to reflect extreme conditions (e.g., maximum possible flow rate) but conditions under realistic wearing or using trials to reflect the 90 % of time in which the products are used.

The requirements on climate product testing can only be met economically with human-physiologically adapted simulation systems. Based on physiologically required heat and moisture release properties, simulations must also offer simple technology, practicability and result validity. The SWEATOR technology developed at MUAS in cooperation with Inside Climate GmbH covers a large part of the requirements mentioned and provides sweating manikins with different shapes adapted to human physiology for use on material samples and on ready-made textile products. Numerous measurement results confirm simple use, the valid determination of basic quality parameters such as Ret and Rct as well as microclimate simulation highly correlated to real using scenarios.

Apart from the simulation technique the decisive point, however, is a reproducible and valuable measurement of the appearing skin near microclimate, recorded in the boundary layer close to the skin and correlated to probands perception,

Regarding the demands on the microclimate sensor, i.e., precision, placement, orientation and at least reproducibility, special measuring methodologies have to be applied. The SWEATLOG system uses combined temperature and humidity sensors with I2C interface placed in a defined array and fixed in a flexible mat. The visualization as a heat and humidity map is realized by use of selected algorithms. Interesting thermodynamical as well as comfort related microclimate values are calculated within defined cluster areas.

The unique combination of SWEATLOG sensor mat and the SWEATOR system with anatomically formed and sweating manikin creates the basis for thermophysiological product analyses. Combined with differentiated evaluation methods and with correlations to human comfort perceptions, a reliable comfort prognosis for assessment as well as improvement of clothing or body support products like seats or beds can be achieved.

## Acknowledgements

Supporting institutions: *in clarification with collaborating companies.*

## References

- [1]. Apurba-Das, R. Alagirusamy, Science in Clothing Comfort, *Woodhead Publishing, Ltd.*, Cambridge, 2010.
- [2]. E. A. Arens, H. Zang, Thermal and Moisture Transport in Fibrous Materials (N. Pan, P. Gibson, Eds.), *Woodhead Publishing Ltd.*, Cambridge, 2006.
- [3]. D. Bektic, Tragekomfortprognose – Korrelationsanalyse ausgewählter Klimasimulationsverfahren, *Masterarbeit Munich University of Applied Sciences FK09*, 2012.
- [4]. E. McCullough, J. Huang, C. S. Kim, An explanation and comparison of sweating hot plate standards, *J. ASTM Int.*, Vol. 1, Issue 7, 2004, pp. 121-133.
- [5]. N. Djongyang, R. Tchinda, D. Njomo, Thermal comfort – A review paper, *Renewable and Sustainable Energy Reviews*, Vol. 14, 2010, pp. 2626-2640.
- [6]. Y. Epstein, D. S. Moran, Thermal comfort and heat stress indices, *Industrial Health*, Vol. 44, 2006, pp. 388-398.
- [7]. P. O. Fanger, Thermal Comfort, *McGraw-Hill Book Company*, New York, 1973.
- [8]. R. F. Goldman, B. Kampmann, Handbook on Clothing, 2<sup>nd</sup> Ed., *Biomedical Research*, 2007, pp. 2/1-2/19.
- [9]. G. Havenith, Clothing and thermoregulation, in Textiles and the Skin (P. Elser, K. Hatch, W. Alberti-Wigger, Eds.), *Current Problems in Dermatology*, 2003, pp. 35-49.
- [10]. L. Hes, J. Williams, Laboratory measurement of thermo-physiological comfort, in Improving Comfort in Clothing (G. Song, Ed.), *Woodhead Publishing Ltd.*, Cambridge, 2011, pp. 114-137.
- [11]. J. Huang, Thermal parameters for assessing thermal properties of clothing, *Journal of Thermal Biology*, Vol. 31, 2006, pp. 461-466.
- [12]. B. Kurz, Ch. Russ, Climate comfort and product testing, *Technical Textiles*, Vol. 4/5, 2020, pp. 172-174.



- 
- [13]. B. Kurz, Ch. Russ, M. Kurz, Comfort prognosis by microclimate simulation and measurement, in *Proceedings of the 7<sup>th</sup> International Conference on Sensors and Electronic Instrumentation Advances (SEIA'21)*, Palma de Mallorca, 2021, pp. 88-89.
- [14]. B. Kurz, S. Langenmeir, C. Zimmermann, W. Uedelhoven, M. Rottenfusser, Klimamanagement im Schuh, *Orthopädienschuhtechnik*, Vol. 11, 2012, pp. 42-51.
- [15]. M. Kurz, Ch. Russ, B. Kurz, Simulation and measurement of comfort related microclimate, *Sensors & Transducers*, Vol. 252, 2021, pp. 35-41.
- [16]. J. Mecheels, Körper-Klima-Kleidung: Grundzüge der Bekleidungsphysiologie, *Verlag Schiele und Schön*, Berlin, 1991.
- [17]. M. Morena, B. Kurz, R. Krah, How users perceive the climate comfort of vehicle seats, *ATZworldwide*, Vol. 114, 2012, pp. 16-21.
- [18]. A. Psikuta, L. C. Wang, R. M. J. Rossi, Prediction of the physiological response of humans wearing protective clothing using a thermophysiological human simulator, *J. Occup. Environ. Hyg.*, Vol. 10, Issue 4, 2013, pp. 222-232.
- [19]. G. Song (Ed.), Improving Comfort in Clothing, *Woodhead Publishing Ltd.*, Cambridge, 2011.
- [20]. G. Song, S. Mandal, Testing and evaluating the thermal comfort of clothing ensembles, in *Performance Testing of Textiles* (L. Wang, Ed.), *Woodhead Publishing of Elsevier*, Cambridge, 2016, pp. 29-64.
- [21]. K. H. Umbach, Die physiologische Funktion der Bekleidung, in *Funktionstextilien* (P. Knecht, Ed.), *High-Tech-Produkte bei Bekleidung und Heimtextilien*, 2003, pp. 43-56.
- [22]. L. Wang (Ed.), *Performance Testing of Textiles*, *Woodhead Publishing of Elsevier*, Cambridge, 2016.
- [23]. R. Wernecke, J. Wernecke, Humidity measurement in gases, in *Advances in Sensors: Reviews* (S. Y. Yurish, Ed.), Vol. 7, *IFSA Publishing*, 2019, pp. 139-163.
- [24]. C. Zimmermann, W. Uedelhoven, B. Kurz, K.-J. Glitz, Thermal comfort range of a military cold protection glove: database by thermophysiological simulation, *Eur. J. Appl. Physiol.*, Vol. 104, 2008, pp. 229-236.
- [25]. ASHRAE, ANSI/ASHRAE 55 – Thermal Environmental Condition for Human Occupancy, *American Society of Heating, Refrigerating and Air Conditioning Engineers*, 2010.
- [26]. DIN EN ISO 11092:2014-12, Textilien – Physiologische Wirkungen – Messung des Wärme- und Wasserdampfdurchgangswiderstands unter stationären Bedingungen (Sweating Guarded Hotplate Test), 2014.

# Chapter 7

## From Food to Mining: Dual Energy XRT Applications

**Rebecca Wagner, Christine Bauer, Markus Firsching  
and Johannes Leisner**

### 7.1. Introduction

With the discovery of X-rays by Wilhelm Conrad Röntgen in 1895, a multitude of applications in the medical and industrial sectors became possible [1-5]. Driven by the use of medical diagnostics, the devices for generating X-rays and taking images became more powerful and compact. As early as 1931, X-ray technology was established and used for the inspection of weld seams [6].

Until the mid-1980's, image acquisition was done mostly with X-ray film that had to be developed with the help of chemical processes. However, digital images were already possible with gaseous ionization detectors, but these were not suited for real time applications [7]. Starting from 1985, digital X-ray sensors based on scintillators and photodiodes were developed. This technology made computer-supported real-time acquisition and evaluation of X-ray image data possible [8].

The fact that chemical elements have a specific attenuation of X-rays depending on energy and on their atomic number was also first used for medical applications by acquiring two X-ray images with different X-ray spectra. By comparing these two X-ray images, information on the material and areal density can be obtained. This procedure is called dual energy X-ray transmission (DE-XRT). For medical applications, the contrast of bone to tissue could be significantly increased [9].

In the industrial sector, a common task is to separate desired from undesired objects, such as food in good condition from contaminated food or valuables from trash. X-ray transmission (XRT) methods are useful for these applications, especially when the sorting criterion is hidden inside the object or not visible on the surface due to a dusty

environment. Consequently, the dual energy X-ray technology was established for the industrial sectors of security control of freight and luggage [10], quality assurance of food [11] and sorting of primary and secondary raw materials [12, 13]. The following sections give an introduction to the DE-XRT technology, an overview of existing applications and the limitations of the method.

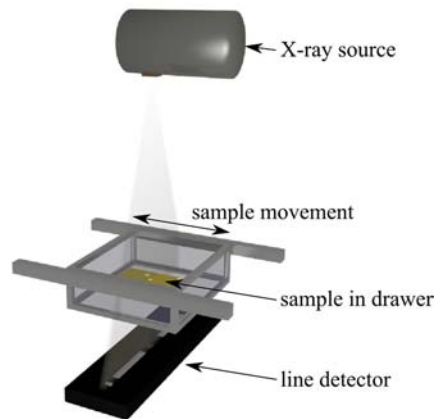
## 7.2. Dual Energy X-ray Transmission: The Method

### 7.2.1. Dual Energy-X-ray Transmission Measurements

In a sorting application, XRT measurements are typically performed on samples on a chute or a conveyor belt. The objects slide down the chute or are moved by the belt between X-ray source and detector. Under these circumstances, the use of a line detector is convenient. It records an image line by line. The resulting X-ray radiograph shows a projection of the object's X-ray attenuation integrated along the beam path.

For DE-XRT, projections at two different X-ray spectral energies are required. In a sorting application, a dual energy detector is the most appropriate option to obtain them. It consists of two detection layers. The first layer records the low-energy (LE) X-ray image. Simultaneously, it acts as a pre-filter for the second layer as it attenuates primarily the low-energy X-rays. The second layer then records the high-energy (HE) image. Usually, an additional filter is placed between both layers to enhance the deviation between LE and HE image.

Most of the examples shown in this chapter were measured in a setup consisting of a drawer system (an open box, looking like a drawer, mounted on a linear stage for translational movement, Fig. 7.1) equipped with a Comet MXR-225HP/11 X-ray source, capable of up to 225 kV acceleration voltage and a tube power of up to 1800 W. The detector and the measurement parameters were adapted to the various samples. Details can be found in Table 7.1.



**Fig. 7.1.** Schematic of the drawer system used for most of the presented DE-XRT measurements.

**Table 7.1.** Measurement parameters for the examples presented in this chapter.

Sample	Detector	Tube voltage [kV]	Prefilter	Current [mA]	Velocity [mm/s]	Exposure [ms]
Metal slices (Fig. 7.2, Fig. 7.3)	DT X-Scan 0.8iL-21 410 DE-USB-C2 (HE-filter: 0.6 mm Cu) Pixel pitch 0.8 mm	220	–	0.9	340	2.67
Coal (Fig. 7.4)	Perkin Elmer XRD 0820 CN14 (flat panel) Pixel pitch 0.2 mm	60 (LE)/ 100 (HE)	1 mm Al (LE)/ 0.9 mm Ti (HE)	2.8 (LE)/ 1.5 (HE)	–	200
Berry Mix (Fig. 7.5)	Hamamatsu C10800-09FCM-C Pixel pitch 0.4 mm	100	–	5.0	170	2.67
Paella (Fig. 7.7)						
Vegetables (Fig. 7.8)						
Salad/ Sushi (Fig. 7.9)						
Pizza (Fig. 7.10)	Hamamatsu C11800-08U Pixel pitch 0.4 mm	100	–	3.6	170	2.67
Clothes (Fig. 7.11)	Hamamatsu C10800-09FCM-C Pixel pitch 0.4 mm	100	–	5.0	170	2.67

### 7.2.2. Basis Material Decomposition

The differentiation of materials using DE-XRT relies on the dependence of X-ray attenuation on the materials' effective atomic number  $Z$  and on the X-ray energy  $E$ . This dependence is described by the Lambert-Beer law, that gives the intensity  $I$  of the X-rays after transmission through an object

$$I = I_0 \exp(-\mu'(E)p), \quad (7.1)$$

where  $I_0$  is the unattenuated intensity and  $\mu'$  the mass attenuation coefficient. The areal density  $p$  along the X-ray path is defined as the mass per area or as density times material thickness for materials with homogenous density.

In case of a superposition of materials, their attenuations  $\mu'_i p_i$  sum up, resulting in

$$I = I_0 \exp(-\sum_i \mu'_i(E)p_i) \quad (7.2)$$

When two monoenergetic measurements are performed at a lower energy  $E_{LE}$  and a higher energy  $E_{HE}$ , this results in a linear equation system

$$I_{LE} = I_{0,LE} \exp(-\sum_i \mu'_i(E_{LE}) p_i), I_{HE} = I_{0,HE} \exp(-\sum_i \mu'_i(E_{HE}) p_i) \quad (7.3)$$

The intensities are measured and the mass attenuation coefficients of chemical elements can be found in tables. For other materials that are not pure elements, but of known composition, the mass attenuations can be calculated from the weight fractions  $w_i$  of the constituting chemical elements using

$$\mu' = \sum_j w_j \mu'_j \quad (7.4)$$

These materials can further be described by an effective atomic number calculated from their constituents' partial chemical densities  $\rho_j$  [14]:

$$Z_{\text{eff}}^k = \frac{\sum_j Z_j^k \rho_j}{\sum_j \rho_j}, k \approx 3 \quad (7.5)$$

Thus, the equation system (7.3) can be solved for the areal densities  $p_i$  of two known materials.

For non-monochromatic X-ray sources, the emitted source spectrum  $S(E)$  and the detector efficiency  $D(E)$  have to be taken into account. For further calculations, only the product of these two functions is relevant, which can be obtained from simulations or measurements [15]. It determines the detected unattenuated intensity (open beam) via

$$I_0 = \int S(E) D(E) dE \quad (7.6)$$

The combination of equations (7.2) and (7.6) yields

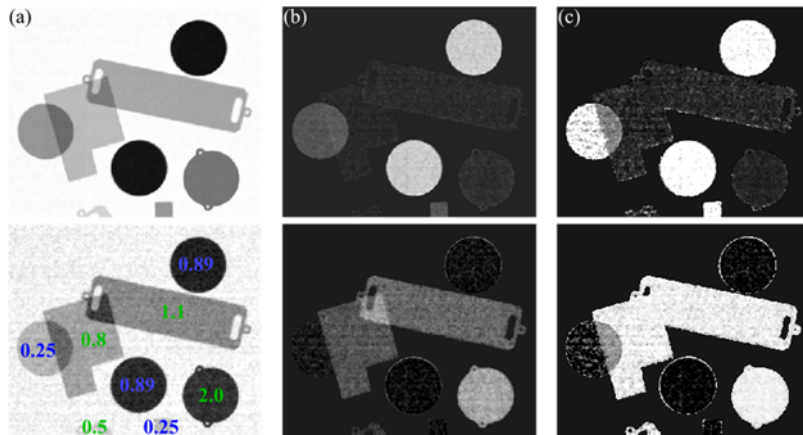
$$I = \int S(E) D(E) \exp(-\sum_i \mu'_i(E) p_i) dE \quad (7.7)$$

Although complicated by polychromatic spectra, the equation system resulting from two measurements can still be solved. The method to do so is called basis material decomposition (BMD) [16]. It results in the areal densities  $p_i$  of two basis materials that have to be chosen by the user. From them, the virtual concentration  $C_i$  can be calculated using

$$C_i = \frac{p_i}{p_1 + p_2}, \quad (7.8)$$

with  $i = 1$  or  $2$ .

This procedure is performed on every pixel of the pair of LE and HE projections, resulting in a new set of images: the so-called basis material images showing the areal densities  $p_1$  and  $p_2$  and the virtual concentration images depicting  $C_1$  and  $C_2$ , where  $C_2 = 1 - C_1$ .



**Fig. 7.2.** (a) LE (top) and HE (bottom) X-ray projections of Al- and Ti-slabs labelled with their respective thicknesses in mm (green Al, blue Ti). Darker color means higher attenuation. A differentiation of the two materials is not possible; (b) Areal density and (c) concentration images showing Ti (top) and Al (bottom) allow a discrimination of the two elements. Here, brighter color means higher areal density and concentration of the respective material.

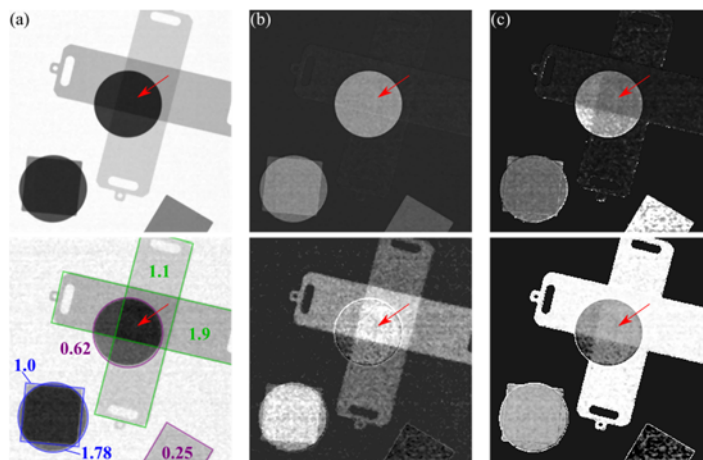
An example of a BMD analysis is shown in Fig. 7.2. It depicts a measurement of Al ( $Z = 13$ ) and Ti ( $Z = 22$ ) metal sheets of different thicknesses. The measurement parameters can be found in Table 7.1. To improve the signal to noise ratio, averaging over three projections was done and a  $3 \times 3$  median filter was applied before BMD. The X-ray projections (Fig. 7.2a) give an impression of the attenuation of the metal sheets. Darker color means higher attenuation  $\mu'p$ . However, an object with high mass attenuation coefficient  $\mu'$  and low areal density  $p$  or thickness can cause the same gray value as an object with low mass attenuation coefficient  $\mu'$  and high areal density  $p$  or thickness. A differentiation between the two kinds of metal is thus not possible based on the projection data. However, in the areal density images (Fig. 7.2b) resulting from BMD, the materials are separated. Here, brighter gray values correspond to higher areal density. Essentially, each material is visible in only one of the basis material images. Only traces of it remain in the other image. While the areal density images of the two basis materials show that each slab has a different thickness, the concentration images (Fig. 7.2c) calculated from them are independent of the thickness, except where both materials overlap. Consequently, pure materials appear white in the respective concentration image, while overlapping materials appear gray.

### 7.2.3. Limits of BMD

So far, it has been assumed that the investigated samples consist of two known materials. An identification of unknown materials by BMD is not possible. Only in the special case, where objects of up to two different materials appear without overlap in the X-ray projections, an estimate of the atomic number  $Z$  may be found. This can be done by testing mass attenuation coefficients  $\mu'$  of different basis materials in a BMD, in an attempt to achieve a separation of the two materials into two different areal density images similar to Fig. 7.2.

A further restriction is the limitation to two materials. When a third material is present, its attenuation coefficient can be expressed as a linear combination of the other two materials (in absence of absorption edges). This is due to the fact that, in the relevant energy range, the attenuation is determined by two effects only: the Compton effect and the photo effect. Therefore, even a measurement with a third energy will not necessarily enable the differentiation of more than two materials.

In a BMD analysis, a third material will consequently appear as a superposition of the two chosen basis materials. Fig. 7.3 shows a measurement of different metal sheets to illustrate that a Ti ( $Z = 22$ ) slab cannot be distinguished from overlapping Al ( $Z = 13$ ) and Zn ( $Z = 30$ ) samples. The measurement parameters can be found in Table 7.1. A  $3 \times 3$  median filter was applied to the X-ray projections before BMD.



**Fig. 7.3.** (a) LE (top) and HE (bottom) X-ray projections of the Al-, Ti- and Zn-slabs labelled with their respective thicknesses in mm (green Al, blue Ti, purple Zn); (b) Areal density images showing Zn (top) and Al (bottom). (c) Concentration images showing Zn (top) and Al (bottom). In none of the images it is possible to distinguish Ti from overlapping slices of Al and Zn (red arrow).

This fact has to be taken into account in different applications. In many cases the objects of interest are of complex composition. For instance, while the task may be to determine the iron content of silicate type rocks, these rocks may also contain copper and other minerals. Consequently, the results of a BMD with the basis materials glass and iron would not give the true concentrations of these elements, but rather the fraction of heavy ( $Z$  around 26) and light ( $Z$  around 11.6) materials.

A further issue is the increase of noise when the difference in the atomic number  $\Delta Z$  decreases. As a consequence, it is a challenge for BMD to differentiate between materials with small difference of effective atomic numbers. The difference in atomic numbers  $\Delta Z$  that can still be discriminated depends on the noise level of the input images and the spectral properties of the X-ray system. In particular, these are the kV settings and the pre-filtering of the X-ray beam, or, in the case of a dual energy detector, the spectral difference

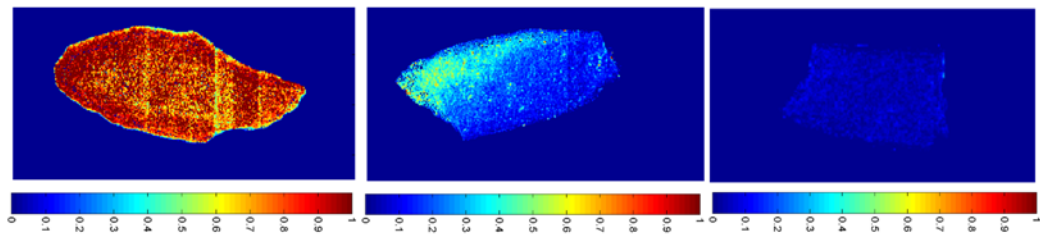
between its energy channels. As a consequence, dual energy detectors are typically optimized for a specific energy range (kVs) in which both, the overall image quality and the material discrimination, is best. Systems and setups that allow individual kV settings for low and high-energy allow more flexibility in this regard, however they are more complex to adjust and to operate at the optimal working point.

### 7.3. Mining

Today the ore grade decreases in most existing mines as well as in new deposits. As a consequence, the mass of material to be processed increases in order to obtain the same amount of the end product. DE-XRT allows to calculate a concentration image for a given target material class against a waste material class. Therefore, it is a powerful method in many applications in the mining business, as it can be used as the sensor technology for the sorting of pieces in a material stream according to their effective atomic number. This kind of pre-concentration is especially effective when the difference in atomic number between the valuable material and the waste material is large. The separation of waste material reduces the amount of material that needs to be processed and thus decreases the energy consumption of the subsequent crushing and milling stages [17].

The DE-XRT technology allows the sorting of pieces of material either by a threshold (no piece below or above a given threshold are kept) or to reach a target concentration in the resulting fraction. Examples of beneficiation include, but are not limited to coal [12], copper and iron ore [18, 19] or the detection of enclosed diamonds in kimberlite [20].

In the case of coal beneficiation, DE-XRT can separate coal (carbon) from mineral material, which usually correlates to the ash content left after burning. Therefore, a high ash content is an indicator for low coal quality. Fig. 7.4 shows an example of the concentration images of coal and country rock using the basis material carbon ( $Z = 6$ ) and aluminum ( $Z = 13$ ). The measurement parameters are given in Table 7.1. As expected, the mineral content of the country rock is close to 100 % (left image). The center image shows the spatial distribution of a piece of coal of low quality with mineral content up to 50 % in the top left part. A piece of high-quality coal is shown in the right image, the mineral content is very low in the entire piece.



**Fig. 7.4.** Shown is the ash-content in each image. The higher the value is, the higher is the ash content. On the left, a rock is shown that is corresponding to almost 100 % ash content, in the middle, the spatial distribution of ash content of a medium quality piece of coal is shown and on the right a piece of high-quality coal can be seen.



## 7.4. Food Safety

On the German market, the number of product recalls is steadily increasing every year (except in 2017) [21]. Around one third of the total amount of recalls takes place in the food sector and is caused by faulty packaging or foreign bodies, which entered the food during the processing chain before packaging. Sources for foreign bodies can be either intrinsic or extrinsic [22]. Intrinsic sources attribute to the food itself and result from imperfect processing like bones in chicken, fish bones in fish or pits and stalks in fruits. Extrinsic sources are foreign bodies that are caused by unsanitary conditions, defect tools etc. They consist of a wide range of materials like stones, glass, metal, steel or plastics. As product recalls degrade the credibility and reputation of the producer and harm the trust of the consumer in the respective product, there is a need for a reliable and stable noninvasive technique to detect foreign bodies. As large as the variety of possible foreign body materials is the range of techniques that are used to detect those objects. Besides X-ray food inspection, the application of other noninvasive techniques like magnetic field systems, metal detection, nuclear magnetic resonance, microwave- or near-infrared techniques are common [23-26]. However, X-ray technology is the preferred option for a wide range of application in the food sector like the inspection of apples, grain, packaged foods, poultry and nuts [11, 27, 28].

Food has a complex chemical composition. It contains mostly carbon, hydrogen, oxygen and nitrogen. The effective atomic number  $Z_{\text{eff}}$  is usually around 7 or 8 but varies slightly depending on the amount of water, carbohydrates, proteins, fats and other ingredients. Therefore,  $Z_{\text{eff}}$  was optimized for each of the presented examples.

### 7.4.1. Calibrated Test Bodies

Calibrated test objects of different materials and sizes can be bought commercially to obtain a first insight about type and size of foreign bodies that may be detectable in food. For this purpose, twelve different test cards by RONDOTEST® [29] were acquired. Each test card includes five test objects either spherical or cubic with different diameters  $d$  or edge length  $d$ , respectively. The investigated materials were aluminum, synthetic bone, rubber (EPDM), glass, steel (AISI304) and Teflon (PTFE) (Table 7.2).

In order to include a variety of food, both frozen and fresh convenience food was investigated. The studied samples were:

- A mix of frozen berries containing blueberries, strawberries, blackberries, redcurrants, sour cherries and raspberries, (contaminant: Al);
- A frozen paella with rice, vegetables, seafood, chicken and salmon (contaminants: glass/steel);
- Frozen fried vegetables, namely savoy, carrots, pepper, onion, ginger, cauliflower with soy sauce (contaminant: PTFE);
- Frozen vegetables, namely corn, peas and cauliflower with butter (contaminant: EPDM);

- A salad with olives and tuna with individually packed dressing (contaminant: synthetic bone);
- Different sushi with individually packed soy sauce, ginger and wasabi (contaminant: synthetic bone).

**Table 7.2.** Overview of the used test cards with their respective materials, shapes and thicknesses.  $Z$  and  $Z_{\text{eff}}$  indicate the (effective) atomic numbers that were used for the BMD algorithm.

Material	$Z / Z_{\text{eff}}$	Shape	Thickness [mm]				
			$d_1$	$d_2$	$d_3$	$d_4$	$d_5$
Aluminum	13	Spherical	0.5	0.8	1.0	1.2	1.5
Aluminum		Spherical	2.0	2.5	3.0	3.5	4.0
Synth. Bone	11.3	Cubic	2.0	3.0	4.0	5.0	6.0
Synth. Bone		Cubic	6.0	7.0	8.0	9.0	10.0
EPDM Rubber	6	Spherical	2.36	3.0	3.5	4.0	4.5
EPDM Rubber		Spherical	4.0	5.0	6.0	7.0	8.0
Glass	11.6	Spherical	0.5	0.8	1.0	1.2	1.5
Glass		Spherical	1.5	2.0	2.5	3.0	3.5
Steel AISI304	26	Spherical	0.3	0.4	0.5	0.6	0.8
Steel AISI304		Spherical	0.6	0.8	1.0	1.2	1.5
Teflon PTFE	10	Spherical	2.0	2.5	3.0	3.5	4.0
Teflon PTFE		Spherical	4.0	5.0	6.0	7.0	8.0

Some of the above-mentioned foods were already investigated in previous measurements but combined with different test objects [30]. Also, tofu balls, nuts and wafer treats have been examined successfully [31].

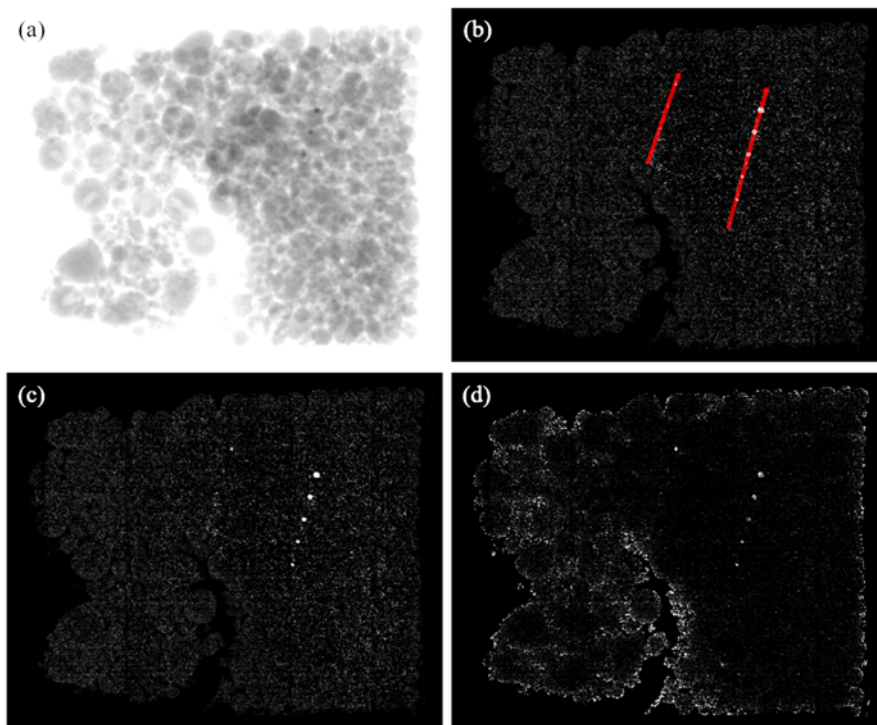
The first examples show food contaminated with a variety of extrinsic foreign test bodies, namely aluminum, glass, PTFE, steel and EPDM (Table 7.2). In Fig. 7.5a, the high-energy X-ray image of a mix of frozen berries contaminated with ten aluminum test bodies can be seen.

The aluminum beads are almost unrecognizable by the bare eye without knowing their exact position due to the highly heterogeneous background induced by the berries. Accordingly, a detection of the test bodies by automated image processing is very challenging. After applying the BMD, the resulting images of areal density and the concentration of aluminum can be seen in Fig. 7.5c and Fig. 7.5d, respectively. In those images the berries ( $Z_{\text{eff}} = 7.4$ ) vanish almost completely and the tests beads can clearly be seen. The segmentation of the five bigger spheres (2.0 mm – 4.0 mm) is feasible. The detection of the smaller beads is limited by the spatial resolution of the imaging system [30].

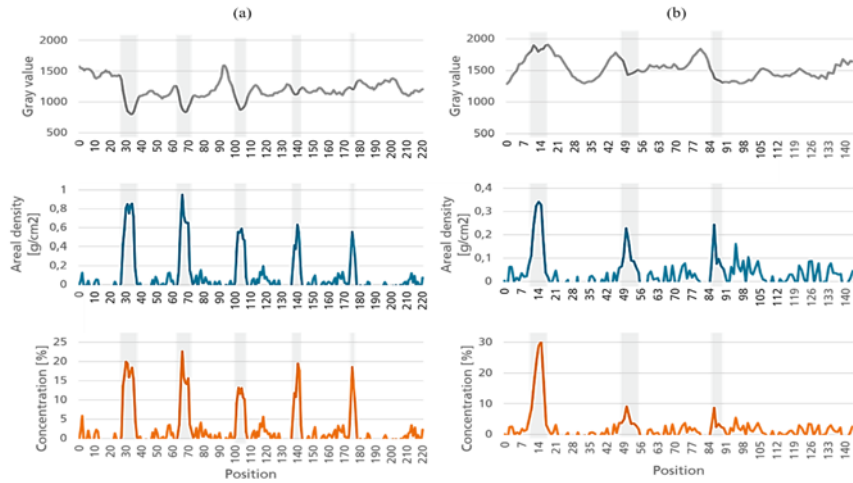
To display the advantages of DE-XRT for detecting foreign bodies, line trackers of five pixels in width (Fig. 7.5b) were used to obtain gray value, areal density and concentration of the two test cards. The resulting graphs can be seen in Fig. 7.6. The gray value shows

great fluctuations in dependence of its position. While some are caused by the test bodies, most of them are due to overlapping of berries causing a highly heterogeneous background. By applying a threshold only, the three biggest beads may be identified based on their gray value. However, the areal density and concentration line profiles reveal very clearly the position of the five bigger beads as well as three of the smaller ones (1.0 mm – 1.5 mm).

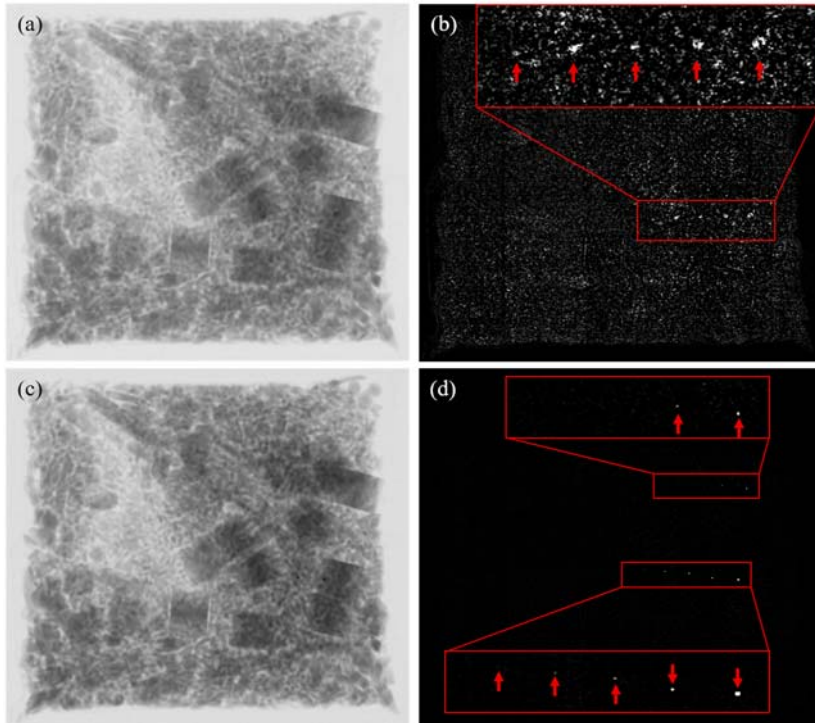
A frozen paella sample was successively contaminated with glass (Figs. 7.7a and b) and steel test bodies (Figs. 7.7c and d). The food itself is very heterogeneous due to the variety of ingredients like small rice corns, big squared pieces of fish and bulky pieces of chicken. Due to that, the test bodies of glass and steel are very difficult to see in the high-energy XRT images. In contrast, the four bigger glass bodies (2.0 mm – 3.5 mm) can be seen in the BMD image, while the bead of 1.5 mm can only be seen with some effort. The smaller beads (< 1.5 mm) disappear in the noisy background that is caused by the close atomic numbers of the food ( $Z_{\text{eff}} = 7.4$ ) and the glass contaminants ( $Z_{\text{eff}} = 11.6$ ). In the case of a bigger difference  $\Delta Z$  in the atomic number of two materials, smaller contaminants can be detected. This can be seen in Fig. 7.7d, where the test bodies of steel ( $Z_{\text{eff}} = 26$ ) in the size of 0.6 mm to 1.5 mm are clearly visible against the background ( $Z_{\text{eff}} = 7.4$ ) despite their small diameter.



**Fig. 7.5.** (a) High-energy XRT image of a frozen berry mix contaminated with aluminum; (b) Areal density image of aluminum with two trackers of five pixels in width. The corresponding line profiles are shown in Fig. 7.6; (c) Areal density image of aluminum; (d) Concentration image of aluminum.

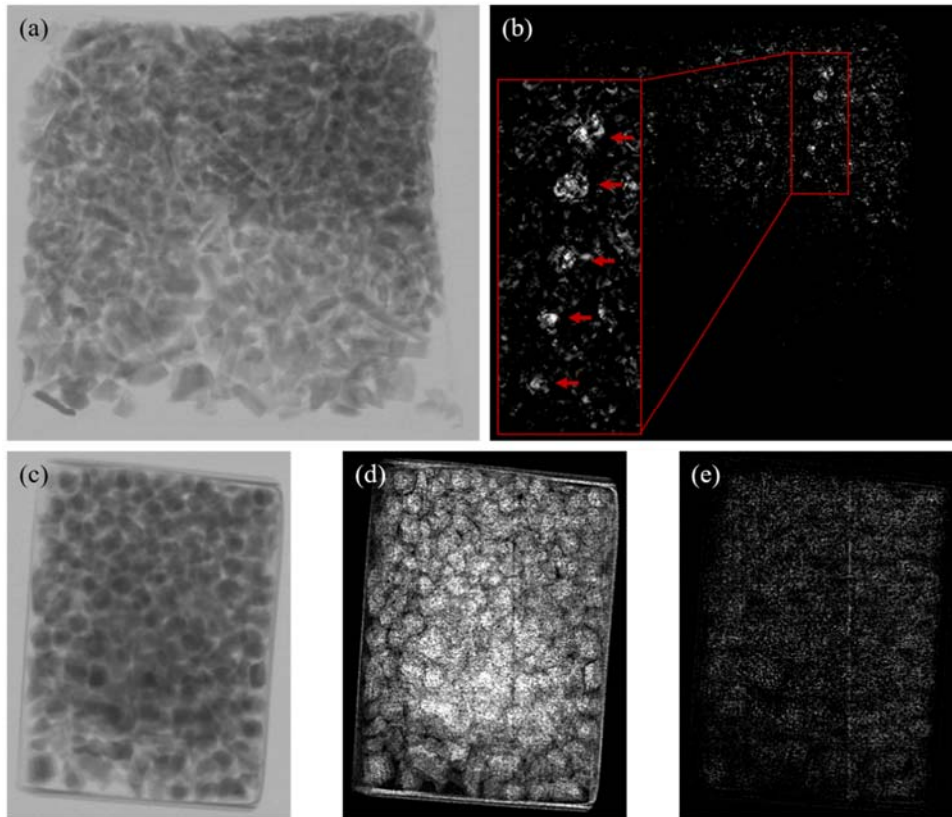


**Fig. 7.6.** (a) Gray value in the high-energy X-ray projection (gray), areal density (blue) and concentration (orange) of the five bigger aluminum test bodies (2.0 mm – 4.0 mm); (b) Gray value, areal density and concentration for the smaller aluminum test bodies (0.5 mm – 1.5 mm). The approximate position of the test bodies is indicated by light gray pillars.



**Fig. 7.7.** (a) High-energy XRT image of a frozen paella contaminated with glass; (b) Areal density image of glass. The inset has a magnification of 2.5. The red arrows display the rough position of the glass beads; (c) High-energy XRT image of a frozen paella contaminated with steel; (d) Areal density image of steel. The insets have a magnification of 2.5.

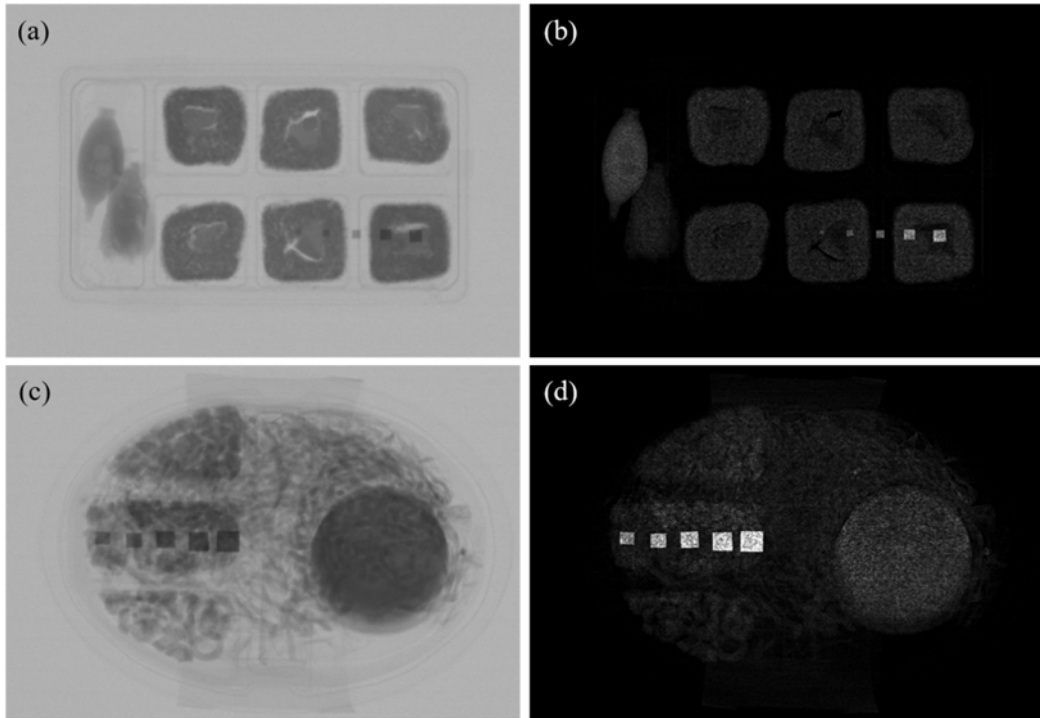
When the difference  $\Delta Z$  in the atomic number of both materials is very low, the benefit of BMD decreases. This can be shown with the extrinsic contaminants PTFE ( $Z_{\text{eff}} = 10$ ) and EPDM ( $Z_{\text{eff}} = 6$ , Fig. 7.8). In the BMD image of PTFE and vegetables ( $Z_{\text{eff}} = 7.5$ ) the test beads are very difficult to detect because their chemical composition is very close: PTFE mainly contains carbon and fluor. In the EPDM image, the beads cannot be located at all, because its effective atomic number is even closer to the food ( $Z_{\text{eff}} = 7.5$ ) than in the case of PTFE. The small difference in attenuation caused by EPDM is lost in noise.



**Fig. 7.8.** (a) High-energy XRT image of frozen fried vegetables contaminated with PTFE; (b) Areal density image of PTFE. The inset has a magnification of 2.5. The red arrows display the rough position of the PTFE beads; (c) High-energy XRT image of frozen buttered vegetables contaminated with EPDM; (d) Areal density image of vegetables; (e) Areal density image of EPDM.

The last examined samples were contaminated with intrinsic foreign bodies which may be found naturally in fish, poultry and meat products: (fish-) bones. Cubic test bodies made of synthetic bone were hidden in convenience food, namely sushi and a tuna salad. The high-energy XRT and areal density images of the synthetic bone ( $Z_{\text{eff}} = 11.3$ ) can be seen in Fig. 7.9. Four of the smaller bone cubes are overlapping with the sushi while one is placed close to them. Here, it is not possible to distinguish this test cube from the food

by simple thresholding. Additionally, the two smallest cubes are also difficult to determine. In the BMD image it is much easier to discriminate the synthetic bone cubes from the background material. A similar effect can be observed for the salad.



**Fig. 7.9.** (a) High-energy XRT image of sushi contaminated with synthetic bone; (b) Areal density image of synthetic bone; (c) High-energy XRT image of a salad contaminated with synthetic bone; (d) Areal density image of synthetic bone.

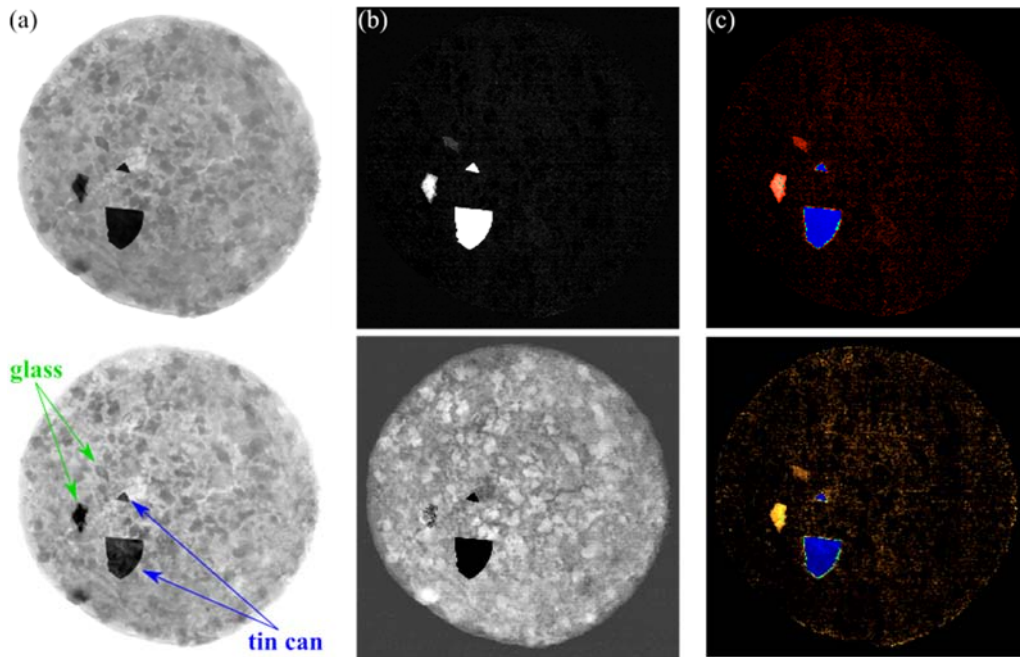
#### 7.4.2. Realistic Foreign Bodies

To illustrate a case with more realistic foreign bodies, frozen pizza was studied. It was contaminated on purpose with pieces of various materials that would be harmful when eaten by a consumer. These objects included plastic foil, aluminum foil, different kinds of broken glass and pieces cut from the lid of a tin can. Only thicker pieces of glass and metal were visible in X-ray projections, while thin glass and the foils were invisible (Fig. 7.10a). The measurement parameters are given in Table 7.1.

As expected, the plastic foil could not be detected by BMD analysis due to the low attenuation in comparison to the pizza and the chemical similarity. The aluminum foil, while visible in the Al material image when lying next to the pizza, could not be detected when placed on the product. Although the contrast in effective atomic numbers between pizza (~7.8) and Al (13) is sufficient, the foil is too thin. All other materials could be detected by BMD analysis.

A suitable  $Z_{\text{eff}}$  was determined for each contaminant by minimizing its visibility in the areal density image showing the pizza. As explained in Section 7.2.3, BMD is limited to two basis materials. A third material will consequently appear as a superposition of these materials. Thus, when using  $Z = 7.8$  to represent the food and  $Z = 10$  for glass, metal parts with their much higher  $Z$  will be visible in the glass image, but lead to a negative areal density in the pizza image (Fig. 7.10b). In contrast, in an analysis using pizza and metal as basis materials, the glass will be visible in both material images, similar to the example of Ti in Fig. 7.3.

The different effects of the choice of basis materials on the material images obtained by BMD analysis can be used to combine these results to RGB images of contaminants. These show materials of low, medium and high atomic number in red, green and blue, respectively (Fig. 7.10c).



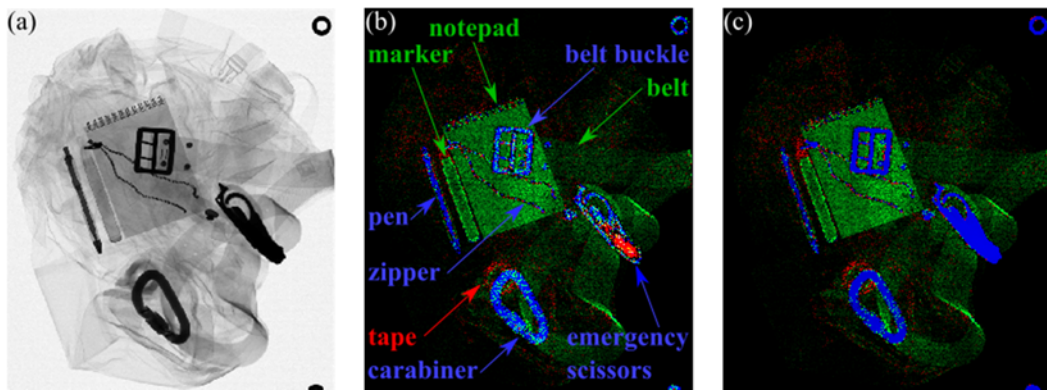
**Fig. 7.10.** (a) LE (top) and HE (bottom) X-ray projections of pizza contaminated with pieces of tin can and glass; (b) Areal density images showing glass (top) and pizza (bottom); (c) Results of BMD analysis with different basis materials combined to RGB images showing areal density (top) and concentration (bottom) of the contaminants.  $Z_{\text{eff}}$  increases from red to blue.

## 7.5. Laundry

Many companies and organizations give their work wear to laundries for professional cleaning. Pens, papers or sweets forgotten in pockets can lead to stains on the clothes and have to be removed before washing. Searching pockets manually, however, carries the risk of injuries due to sharp objects like forgotten scissors, syringes or scalpels.

Finding objects left in clothes by DE-XRT is challenging because of the variety of the contaminants, which cover a wide range of effective atomic numbers (from food and plastics with  $Z_{\text{eff}} \approx 7$  to gold jewelry with  $Z_{\text{eff}} \approx 79$ ), but also due to the clothes themselves. Work wear can be more inhomogeneous than food. One piece of cloth may be made of different kinds of fabric, contain reflective stripes, zippers and press studs made of metal. In a sorting application, these components of clothes must be distinguished from contaminants of a similar material. Consequently, BMD analysis alone is not suitable for an automatic sorting decision. However, it may help an operator to recognize objects and decide whether they have to be removed. For an automatic decision BMD can be combined with image processing and/or artificial intelligence.

As a test case, clothing worn by emergency medical services was equipped with typical accessories. The parameters used for DE-XRT measurements can be found in Table 7.1.



**Fig. 7.11.** (a) LE X-ray projection of emergency medical service trousers; (b) RGB image obtained from areal density images showing different equipment and cloth components (red: low  $Z$ , green: medium  $Z$ , blue: high  $Z$ ); (c) RGB image with objects of too low transmission colored in blue.

Fig. 7.11 shows the X-ray projection of trousers. The gray value varies due to the folding so that some of the equipment cannot be distinguished from clothes based on thresholding. However, an RGB image combined from areal density images reveals paper and different pens that could cause paper fuzz with the corresponding loss of information and ink stains after washing. Also, a roll of adhesive tape is visible which could cause new stains and thus the need for additional cleaning of the clothing after washing. While the discrimination of tape and cloth is challenging due to the similar effective atomic number ( $Z_{\text{cloth}} = 7.1$ ,  $Z_{\text{tape}} = 6$ ), a different challenge for BMD analysis is posed by metal parts. For example, there is hardly any intensity transmitted through the blades of the emergency scissors. As a consequence, BMD yields an incorrect result and they appear to be of a low  $Z$  material in Fig. 7.11b. However, the low transmission in the X-ray projection can be used as additional information to correct the RGB image by labeling the corresponding areas as high  $Z$  material (Fig. 7.11c). An alternative would be to measure with higher tube voltage. However, that would further reduce contrast between cloths and other objects of low and medium  $Z$  materials.

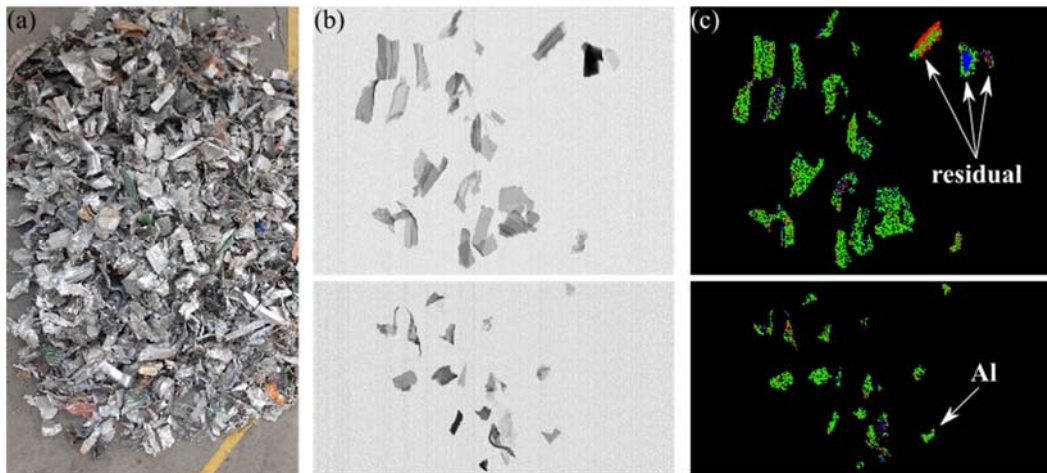


## 7.6. Metal Discrimination

The availability of raw materials is an important factor for economy. However, while the demand increases, raw materials become scarcer. As a consequence, recycling and the concept of a circular economy become crucial to supply the industry with secondary raw materials.

One important resource is aluminum. As a secondary raw material it is found, for instance, in major household appliances, which also contain various other materials like plastics, or copper. In recycling, these items are shredded as a preprocessing step before the resulting flakes have to be sorted. The aim of the sorting process is to obtain a high-grade aluminum class and a low-grade residuals class.

Fig. 7.12a shows a photograph of shredded end-of-life aluminum trash. In contrast to the previous examples, DE-XRT measurements of the flakes (Fig. 7.12b) have been performed at Binder+Co AG. The X-ray tube voltage was set to 160 kV and 2 mA tube current was used. The DE line detector had 896 pixels and a pixel pitch of 1.6 mm. Fig. 7.12c shows RGB images composed of the concentration images obtained by BMD analysis. The red, green and blue color show the concentration of C, Al and Cu, respectively. Most flakes appear mainly green. However, many also contain red and blue pixels. While some of it may be due to noise, the flakes can also be made of a material with an effective  $Z$  different from Al or they may consist of Al with some kind of coating. Trained personnel can use the BMD results to create ground truth data for a machine learning approach that attempts an automated sorting of trash [32].



**Fig. 7.12.** (a) Photograph of end-of life aluminum trash flakes; (b) LE X-ray projections of flakes from different trash mixtures; (c) Composition of concentration images obtained by BMD (red:  $Z_{\text{eff}} = 6$ , green:  $Z_{\text{eff}} = 13$ , blue:  $Z_{\text{eff}} = 29$ ). The data are used with kind permission from Binder+Co AG.

## 7.7. Summary and Outlook

Sorting of objects with respect to a specific feature is required in many processes and a variety of established techniques is available. Especially when the feature is inside the object (packed food) or the sorting takes place in a dusty environment (mining, recycling), X-ray technology can be an advantage.

While common X-ray projections do not differentiate between thin objects of highly attenuating material and thick objects of lowly attenuating material, DE-XRT allows to distinguish objects based on a material property, i.e., the mass attenuation coefficient. The preceding sections show some examples that illustrate the advantages of the method. The discrimination of different materials works the better, the more different the X-ray attenuation is. The smaller the difference in this parameter, the more important is the image quality of the X-ray projections. This is determined mainly by noise. While noise may be reduced by longer exposure time, this option is limited for inline applications. In these edge cases the optimization of the setup for the application (tube settings, detector scintillators, filters, etc.) becomes crucial.

Further improvements of sorting processes are possible by combining DE-XRT with other technologies like RGB imaging [33], machine learning [32] or X-ray fluorescence. An extension to differentiation of three materials is possible by multi energy XRT if one of the materials has an absorption edge in the detected spectral range [34].

## Acknowledgment

We thank W. Holub for providing medical service laundry equipped with accessories. We further acknowledge permission to use XRT data provided by Binder+Co AG.

## References

- [1]. T. Götttsche, Einführung in die Radiologie: Diagnostik und Interventionen; 116 Tabellen, *Georg Thieme Verlag*, 2006.
- [2]. T. M. Borman, B. C. Stoel, Review of the uses of computed tomography for analyzing instruments of the violin family with a focus on the future, *Journal of the Violin Society of America: VSA Papers*, Vol. 22, Issue 1, 2009, pp. 1-12.
- [3]. S. Carmignato, W. Dewulf, R. Leach, Industrial X-ray Computed Tomography, *Springer*, 2018.
- [4]. W. Sun, S. B. Bown, R. K. Leach, An Overview of Industrial X-ray Computed Tomography, *NPL Report*, 2012.
- [5]. N. Kotwaliwale, K. Singh, A. Kalne, S. N. Jha, N. Seth, A. Kar, X-ray imaging methods for internal quality evaluation of agricultural produce, *Journal of Food Science and Technology*, Vol. 51, Issue 1, 2014, pp. 1-15.
- [6]. H. E. Martz, C. M. Logan, D. J. Schneberk, P. J. Shull, X-ray Imaging: Fundamentals, Industrial Techniques and Applications, *CRC Press*, 2016.
- [7]. K. Watanabe, Photoionization and total absorption cross section of gases. I. Ionization potentials of several molecules. Cross sections of NH<sub>3</sub> and NO, *The Journal of Chemical Physics*, Vol. 22, Issue 9, 1954, pp. 1564-1570.

- [8]. V. Perez-Mendez, Digital Radiography: Present Detectors and Future Developments, *Lawrence Berkeley Lab.*, CA, USA, 1990.
- [9]. W. R. Brody, D. M. Cassel, F. G. Sommer, L. A. Lehmann, A. Macovski, R. E. Alvarez, N. J. Pelc, S. J. Riederer, A. L. Hall, Dual-energy projection radiography: initial clinical experience, *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology*, Vol. 137, Issue 2, 1981, pp. 201-205.
- [10]. R. F. Eilbert, K. D. Krug, Aspects of image recognition in Vivid Technologies' dual-energy x-ray system for explosives detection, *Proceedings of SPIE*, Vol. 1824, 1993, pp. 127-143.
- [11]. R. P. Haff, N. Toyofuku, X-ray detection of defects and contaminants in the food industry, *Sensing and Instrumentation for Food Quality and Safety*, Vol. 2, Issue 4, 2008, pp. 262-273.
- [12]. L. von Ketelhodt, C. Bergmann, Dual energy X-ray transmission sorting of coal. *Journal of the Southern African Institute of Mining and Metallurgy*, Vol. 110, Issue 7, 2010, pp. 371-378.
- [13]. M. Mesina, T. De Jong, W. Dalmijn, Automatic sorting of scrap metals with a combined electromagnetic and dual energy X-ray transmission sensor, *International Journal of Mineral Processing*, Vol. 82, Issue 4, 2007, pp. 222-232.
- [14]. B. J. Heismann, J. Leppert, K. Stierstorfer, Density and atomic number measurements with spectral X-ray attenuation method, *Journal of Applied Physics*, Vol. 94, Issue 3, 2003, pp. 2073-2079.
- [15]. E. Y. Sidky, L. Yu, X. Pan, Y. Zou, M. Vannier, A robust method of x-ray source spectrum estimation from transmission measurements: Demonstrated on computer simulated, scatter-free transmission data. *Journal of Applied Physics*, Vol. 97, Issue 12, 2005, 124701.
- [16]. M. Firsching, F. Nachtrab, N. Uhlmann, R. Hanke, Multi-energy X-ray imaging as a quantitative method for materials characterization, *Advanced Materials*, Vol. 23, Issues 22-23, 2011, pp. 2655-2656.
- [17]. J. Lessard, J. de Bakker, L. McHugh, Development of ore sorting and its impact on mineral processing economics, *Minerals Engineering*, Vol. 65, 2014, pp. 88-97.
- [18]. C. Bauer, R. Wagner, B. Orberger, M. Firsching, A. Ennen, C. Garcia Pina, C. Wagner, M. Honarmand, G. Nabatian, I. Monsef, Potential of dual and multi energy XRT and CT analyses on iron formations, *Sensors*, Vol. 21, Issue 7, 2021, 2455.
- [19]. C. Bauer, R. Wagner, B. Orberger, M. Firsching, C. Wagner, O. Boudouma, K. Siahcheshm, Dual and multi energy XRT and CT analyses applied to copper-molybdenum mineralizations in porphyry deposits, *Materials Proceedings*, Vol. 5, Issue 1, 2021, 27.
- [20]. M. Firsching, F. Nachtrab, J. Mühlbauer, N. Uhlmann, Detection of enclosed diamonds using dual energy X-ray imaging, in *Proceedings of the 18<sup>th</sup> World Conference on Nondestructive Testing*, Durban, South Africa, 16-20 April 2012, pp. 16-20.
- [21]. Bundesanstalt für Arbeitsschutz und Arbeitsmedizin website, <https://www.baua.de/DE/Angebote/Publikationen/Berichte/ProdSG-2020.html>
- [22]. M. C. Edwards, M. F. Stringer, Observations on patterns in foreign material investigations, *Food Control*, Vol. 18, Issue 7, 2007, pp. 773-782.
- [23]. M. Graves, A. Smith, B. Batchelor, Approaches to foreign body detection in foods, *Trends in Food Science & Technology*, Vol. 9, Issue 1, 1998, pp. 21-27.
- [24]. M. T. Mohd Khairi, S. Ibrahim, M. A. Md Yunus, M. Faramarzi, Noninvasive techniques for detection of foreign bodies in food: A review, *Journal of Food Process Engineering*, Vol. 41, Issue 6, 2018, e12808.
- [25]. J. A. Abbott, Quality measurement of fruits and vegetables, *Postharvest Biology and Technology*, Vol. 15, Issue 3, 1999, pp. 207-225.
- [26]. S. N. Jha, T. Matsuoka, Non-destructive techniques for quality evaluation of intact fruits and vegetables, *Food Science and Technology Research*, Vol. 6, Issue 4, 2000, pp. 248-251.
- [27]. R. Haff, D. C. Slaughter, Y. Sarig, A. Kader, X-ray assessment of translucency in pineapple, *Journal of Food Processing and Preservation*, Vol. 30, Issue 5, 2006, pp. 527-533.

- [28]. T. Schatzki, R. P. Haff, R. Young, I. Can, L. C. Le, N. Toyofuku, Defect detection in apples by means of X-ray imaging, *Transactions of the ASAE*, Vol. 40, Issue 5, 1997, pp. 1407-1415.
- [29]. Rondotest GmbH & Co. KG Website, <https://rondotest.de/>
- [30]. C. Bauer, R. Wagner, J. Leisner, Foreign body detection in frozen food by dual energy X-ray transmission, *Sensors & Transducers*, Vol. 253, Issue 6, 2021, pp. 23-30.
- [31]. C. Bauer, R. Wagner, J. Leisner, Potential of dual energy X-ray transmission on food safety, in *Proceedings of the Conference on Sensors Engineering and Electronics Instrumentation Advances (SEIA'21)*, Palma de Mallorca, Balearic Islands, Spain, 14-16 September 2021, pp. 90-93.
- [32]. S. Rüger, J. Goschenhofer, A. Nath, M. Firsching, A. Ennen, B. Bischl, Deep-learning-based aluminum sorting on dual energy X-ray transmission data, in *Proceedings of the Sensor-Based Sorting & Control (SBSC'22)*, Aachen, Germany, 13-14 April 2022, to be published.
- [33]. M. Firsching, S. Rüger, W. Benner, M. Vogelgesang, A. Ennen, Material value estimation for recycling of Waste Printed Circuit Boards (WPCBs) by a deep-learning-assisted approach on X-Ray images, in *Proceedings of the Sensor-Based Sorting & Control (SBSC'22)*, Aachen, Germany, 13-14 April 2022, to be published.
- [34]. C. Bauer, R. Wagner, M. Firsching, B. Orberger, J. Lucic, A. Ennen, N. Wörlein, J. L. Dubos, J. Banchet, J.-M. Milazzo, N. Uhlmann, Recycling of Mn-processing dusts: Quality control through 3D computed tomography, in *Proceedings of the XXX International Mineral Processing Congress (IMPC'20)*, Cape Town, South Africa, 18-22 October 2020, pp. 3551-3562.



# **Chapter 8**

# **Industrial Measurements, Communications and Protocols**

**José Miguel Costa Dias Pereira**

## **8.1. Introduction**

Industrial measurements and protocols have specific demands that have been evolving over the time. There is a large number of specialized documentations that describes in detail the evolution registered in these areas. However, a global and integrated vision of the main steps that were registered in the evolution of industrial measurement and communication protocols is oftentimes missing.

In the first part of the chapter, several concerns related with industrial measurements are underlined. It can be referred that industrial measurement problems start, themselves, in the specificities of the quantities that are under measurement. Almost all industrial quantities are affected by multiple influence variables whose variations, over time, must be accounted in order to obtain metrological comparability and traceability of measurement results. Suitable data processing techniques must be used to identify the main influence variables, that affect the measurement result of a given quantity, and to evaluate the associated compensation coefficients in order to convert the measurements results to a predefined influence variables value, usually called as reference values.

Sensor selectivity is also a problem that must be properly addressed because the output signal from sensors is always affected by other variables that are not under measurement. This problem has a particular relevance in chemical and biological sensors and sensor selectivity can compromise significantly measurement accuracy. This drawback can be minimized by using multiple sensing devices, with different measurement principles, that

---

José Miguel Costa Dias Pereira  
ESTSetúbal/IPS, Instituto Politécnico de Setúbal, Setúbal, Portugal  
Instituto de Telecomunicações, Lisboa, Portugal

assure different cross-sensitivity coefficients, or by using intelligent data processing techniques, like the ones based on artificial neural networks.

It is also important to refer that smart sensing capabilities, including adaptability, self-testing, self-calibration, auto-ranging and errors correction capabilities, among others, contribute to improve the accuracy and the performance of industrial measurement systems and to promote interoperability of sensing units and industrial devices. Implementation of these capabilities in industrial instruments, from different manufacturers, is very important, regarding instrumentation and measurement solutions for industrial applications, since it promotes devices interoperability and consequently cost reduction in maintenance and operation of I&C systems.

It is important to refer that security and safety issues are also very important topics, especially in industrial hazards areas that exhibit fire, explosion, toxic release and environmental damage risks. However, a detailed approach of these important topics is also out of the scope of this chapter.

In the second part of the chapter, a survey of the main common industrial communication protocols, from current loops to fieldbuses, is presented and their main advantages and limitations are underlined. The survey includes only the main protocol solutions for continuous industries processes, starts from the need to define standardized levels in pneumatic and electrical signals, and continues towards more advanced protocols, such as, highway addressable remote transducers (HART) and fieldbuses.

The chapter finishes with a summary of the main conclusions.

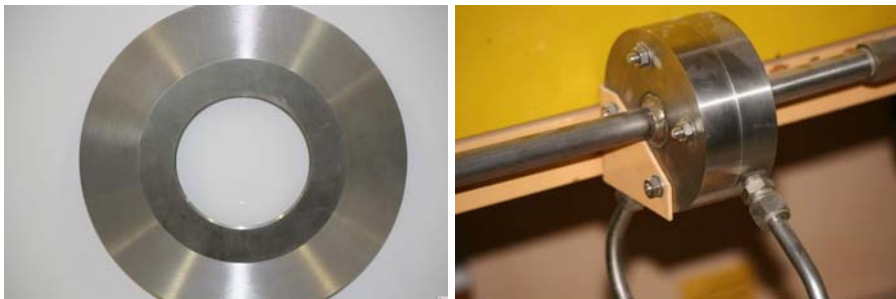
## **8.2. Industrial Measurements**

This section highlights some pitfalls of industrial measurements that can have impact on the transmission network. As a matter of fact, there are specific issues that can affect the way transmission data is performed, particularly, in which concerns signal conditioning and local data processing tasks. Static and dynamic behavior of sensing units and measuring signal characteristics, including signal to noise ratio, are important topics since the transmission bandwidth usage must be judiciously performed. Moreover, the choice of the optimum transmission rate for each signal must be set wisely in order to minimize the negative impact of protocol overheads. Thus, it is important to refer that the multiplicity of measuring variables, the new capabilities of smart sensors, the geographical dispersion of the measurement points and the real time (RT) requirements of the industrial processes, among others, demand a suitable choice of the industrial communication network and protocol.

The working conditions of industrial instruments are not the same of laboratory instruments. Those instruments are typically affected by the harsh measuring conditions characterized by large temperature and humidity variations, water, dust, vibrations and chemical attacks, such as corrosion. These conditions imply a proper choice of the protection index of the equipment and network devices.

By other way, the large number of measurement variables that are interdependent and exhibit, sometimes, large variation ranges, implies the usage of a large number of sensing units for temperature, pressure, level, flow, force, pH, conductivity, ORP and turbidity, sensing units which corresponds to a large number of network measuring nodes. These large number of measurement variables that are affected by multiple influence variables, implies also a large number of sensing units in order to provide metrological comparability of measurement results. Suitable advanced data processing techniques that include the usage of artificial neural networks (ANN) [1-3] must be used to identify the main influence variables and to evaluate the associated influence coefficients. Otherwise, it is not possible to perform a correct metrological comparability of measurement results.

Other advanced data processing techniques, such as fuzzy, fuzzy neural and PCA [4] techniques must also be used to minimize sensor cross-sensitivities effects on measurement accuracy. These effects are particularly important for chemical sensors that present negative values of cross sensitivity coefficients because they can hide real values of variables that affect process safety and security. Another characteristic of industrial measurements is related with the usage of industrial measurement methods. An indirect measurement method involves a combination of different direct measurements, each of which may be affected by different influence variables. As an example, of an indirect measurement method it can be referred the flow measurement, in a pipe, using the orifice method. The pressure drop in the orifice is used to access flow intensity. As a matter of fact, there exist a non-linear relationship between flow intensity and the square root of the differential pressure but this value is affected by a large number of influence variables that include fluid density, pipe diameter, fluid velocity, viscosity, orifice geometric coefficients, and temperature, because temperature variations affect fluid dynamic viscosity. Moreover, the mathematical relationships between the variable under measurement and the indirect variables are not exact and there are also problems related with measurement errors' propagation. Other important issue related with industrial measurements are neutrality errors that means the values of variables under measurement are modified, in a higher or lower degree, by the measurement instrument. Continuing with the same example of flow measurement, an orifice flowmeter, like the one represented in Fig. 8.1, creates a pressure drop that affects, more or less, the flow intensity under measurement.



**Fig. 8.1.** Orifice plate and assembling in a flow pipe.



Another hard limitation in industrial measurements is signal to noise ratio. Signal amplitude, like the voltage delivered by a thermocouple is in the order of a few millivolts and the noise sources in industrial environments are multiple. Typical noises sources are associated with power lines, motors, switching power supplies, voltage regulators, electronic speed variators and switchers, among others. To minimize conductive, inductive, capacitive and electromagnetic noise coupling mechanisms installation rules and circuits' design must consider the following rules: avoid galvanic connections and ground loops, whenever possible; reduce loops areas; avoid proximity and use appropriate electrical and magnetic shields and avoid conductor length higher than  $\lambda/20$ , where  $\lambda$  represent the electromagnetic signal wavelength ( $\lambda = c/f$ ).

Another issue that can dictate the protocol type is related with timing and synchronization of measurements and control tasks. Sometimes, the control applications require time synchronization between acquisition and control tasks and if this is the case the control loops that can actuate in several network nodes must be synchronized and the equipment must share a common network clock. Regarding critical issues in industrial measurements a topic that deserves a particular attention is related with grounding and shielding. Ground loops must be avoided because they can cause inaccurate measurements, intermittent measurement results and can even damage signal conditioning circuits. Differential transmission, galvanic isolation, low resistance and a single ground connection must be used whenever possible.

### **8.3. Industrial Communication Protocols**

This section describes three main continuous process protocols that are still used nowadays in many manufacturing units. From the past to the future, the protocols that are highlighted includes current loop, HART and field buses, taken Foundation Fieldbus (FF) as an example of the last family of protocols.

#### **8.3.1. Current Loops**

The 4-20 mA analogue signaling [5] has been the most common way to transfer an instrumentation signal from or to a field device. Typically, 4 mA is associated with the minimum value of the measured physical variable (X) and 20 mA is associated with its maximum value.

The minimum value of the current (live zero signaling) is used to sense potential transmission problems like broken wires. Usually, the transmitter receives its operating voltage from the loop and does not require any additional electrical connection. Often the receiving devices that can be connected to a single transmitter in a given loop, has an input resistance of 250  $\Omega$  that translates the 4-20 mA variation in 1-5 V variation. The reason to avoid voltage signaling, is that unless additional sense wires are used, the signal would be lost over distance due to line resistance and this will cause inaccurate measurements.

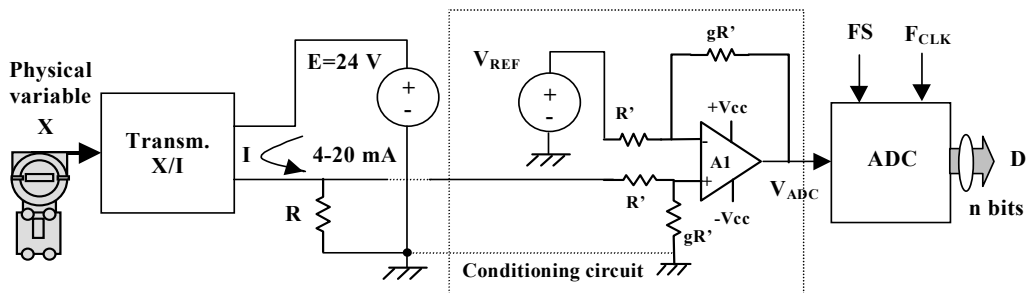
Fig. 8.2 represents a current-loop transmitter associated with its power, a typical conditioning circuit (difference amplifier) and an analogue-to-digital converter (ADC). Assuming a linear relationship between the physical variable ( $X$ ), under measurement, and the transmitter current ( $I$ ), the output code ( $D$ ) of the analogue-to-digital (ADC) is given by:

$$D = \text{round}\left(\frac{V_{ADC}}{Q}\right), \quad (8.1)$$

where  $\text{round}()$  represents the rounding to an integer function,  $V_{ADC}$  and  $Q$  represent the ADC input voltage and the ADC quantization step, respectively, being  $Q$  defined by:

$$Q = \frac{FS}{2^n}, \quad (8.2)$$

where  $FS$  represent the ADC full-scale amplitude and  $n$  its number of bits.



**Fig. 8.2.** Current-loop transmitter, conditioning and ADC circuits ( $R' \gg R$ ).

Considering an ADC with minimum and maximum input voltages equal to  $V_{min}$  and  $V_{max}$ , respectively, the reference voltage ( $V_{REF}$ ) and the gain ( $g$ ) of the conditioning circuit to obtain a maximum measurement resolution is given by:

$$V_{REF} = 4 \cdot R(k\Omega) - \frac{V_{min}}{g}, \quad g = \frac{V_{max} - V_{min}}{16 \cdot R(k\Omega)}, \quad (8.3)$$

where  $R(k\Omega)$  is the receiver resistor value in  $k\Omega$  units, typically equal to  $0.25 k\Omega$ .

The measurement resolution ( $\Delta X$ ) is then given by:

$$\Delta X = \frac{X_{max} - X_{min}}{2^n}, \quad (8.4)$$

where  $X_{max} - X_{min}$  represents the transmitter span.

Expressions (8.1) to (8.4) are still valid if actuators are considered instead of transmitters but in that case digital-to-analogue converters (DAC) with an output current variable between 4 and 20 mA are considered, instead of ADC's.

The most common solutions used nowadays include 8- or 12-bits A/D or D/A converters which means that the measurement resolution is less than 0.5 % and 0.03 % of transmitter or receiver span, respectively. Obviously higher resolution values can be obtained if converters with a higher number of bits are used. However, incrementing the number of bits is limited by the signal-to-noise relation (SNR) of the transmitted signal that have a typical value of 50 dB. Expression (8.5) gives the relationship between the minimum numbers of bits of the converters ( $n_{min}$ ) and the SNR (Signal to Noise Ratio):

$$n_{min} = \frac{SNR(dB) - 1.76}{6.02} \quad (8.5)$$

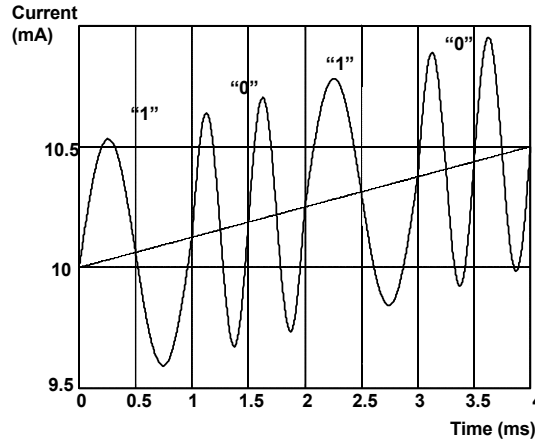
As disadvantages related with this signaling mode it can be referred: the wiring costs are too high since it is required, at least, a pair of wire for each process value; every current loop transmitter has a maximum loop resistance, which means a minimum wire gauge for a maximum loop length; it is basically a simplex transmission mode and it suffers from the low noise immunity and accuracy limitations that are associated with the usage of analog signals. However, it is important to refer that current loop values out of the transmitter current range (4-20 mA) can be used to signalize transmitter alarms and faults and out of range measurements. This means measurement values that are lower than the lower range value (LRV) or higher than the upper range value (URV).

### 8.3.2. HART

Highway addressable remote transducer (HART) systems [6-8] can be considered the selected industrial solution for the transition between the old 4-20 mA current loop systems and the future, fully digital, fieldbus solutions [9-10]. Compatibility with 4-20 mA systems, use of existing analogue wiring and multidrop capability are some of the more important benefits of HART systems. The digital capability of HART systems enables the access to all instrument parameters and diagnostics, supports multivariable instruments and provides on-line device status. Standardization of HART protocol enhanced by the use of a device description language [11] is another key factor that provides interoperability of HART devices and allows communication between host and field devices without any custom programming. It is also important to note that although new fieldbus solutions are available today, the existing industrial networks need a smooth transition to the new fully digital technologies. Simultaneously the power and adaptability of HART communication technology have still new challenges today to create extended meter reading networks via spread spectrum radio and satellite communications and can serve a large segment of the process automation industry market while different standards in the fieldbus area will stabilize [12].

The HART protocol uses the Bell 202 FSK standard [13] to superimpose digital signals on the 4-20 mA analogue signal. Signal frequencies of 1200 Hz and 2200 Hz are used to represent the binary values “1” and “0”, respectively. Communication with smart field instruments [14-15] is assured by superimposing previous low-level sinusoidal signals with the analogue signal, being the average value of the digital signal null. This means that the digital data doesn't interfere with the analogue signal and both share simultaneously, and in a bi-directional way, the same physical support. The main physical

characteristics of the HART signal are exemplified in Fig. 8.3 where a binary sequence is superimposed to a linear analogue signal variation between 10 mA and 10.5 mA. Minimum and maximum peak-to-peak amplitudes of the digital signal are equal to 0.8 mA and 1.2 mA, respectively.



**Fig. 8.3.** HART signal time variation associated with the binary sequence “1010”.

HART messages are coded as a set of 8-bit (byte) characters that are transmitted serially. Each byte contains a start bit, an odd parity bit and a stop bit. The generic message format, represented in Fig. 8.4, contains the following main fields: preamble, start character, source and destination addresses, command, byte count, status, data and checksum.

P	STR	ADDR	COM	BCNT	ST	D	CHK
---	-----	------	-----	------	----	---	-----

**Fig. 8.4.** Hart frame fields: P – preamble, STRT – start character, ADDR – addresses, COM – command, BCNT – byte count, ST – status, D – data, CHK – checksum.

The HART signal derivative is used to trigger the capture of the digital message that starts with the preamble field containing a series, between five and twenty, hexadecimal characters.

Error information of the HART communication is contained in the status field of the reply messages returned by the HART devices. This field, consisting of two bytes, reports any outgoing communication error, the status of the received command and the operational state of the HART device.

The checksum byte that is equal to the exclusive-or of all the bytes that precede it in the message, starting and including the start character, is used to test transmission integrity and to evaluate the bit error ratio (BER) of the HART transmission line.

It is important to refer that the HART protocol has the capability to connect multiple field devices to the same pair of wires in a multidrop network configuration, as depicted in Fig. 8.5.

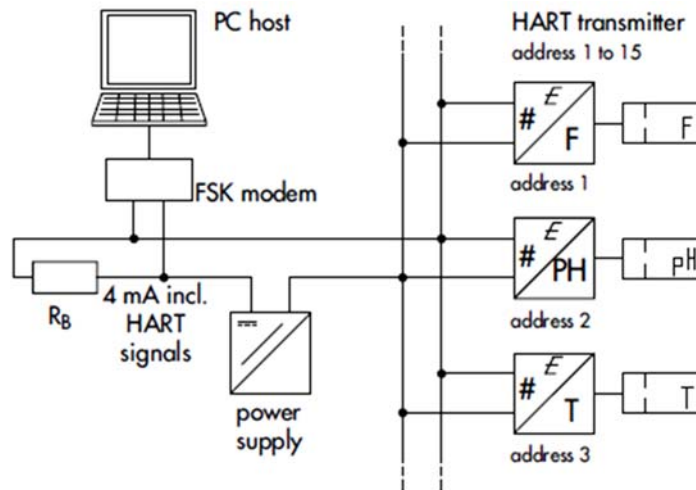


Fig. 8.5. HART multidrop network configuration.

However, if a multidrop association of HART devices is used the loop current of all HART devices is fixed to 4 mA and only digital HART communication is used. In the previous figure, resistance  $R_B$  represents the loop resistance value for successful HART communication, and the value of resistance must be in range between  $230 \Omega$  and  $1.1 \text{ k}\Omega$ .

Regarding the main advantages of the HART protocol, it can be referred: HART uses a mixed analog and digital signal communication mode being possible to use current loop devices in the same loop; maintenance costs are reduced because diagnostic tests can be performed remotely and the same is true for operation costs since instrument's configuration and tests can be performed remotely; cabling cost are substantially reduced because up to 15 instruments can share the same pair of wires that is used for current loop signals and there are substantial interoperability gains since a set of universal and command practice commands are recognized by different HART devices from different manufacturers. Regarding limitations, it can be referred: HART uses a low baud rate of 1200 b.p.s. (bits per second); the process value update speed is only equal to 2 and 4 messages/s in master-slave and burst modes, respectively; there is no capability to recover from network failures and mesh network topologies can't be implemented.

Regarding HART, it is important to refer that there are significant improvements that can be achieved with wireless HART protocol. Wireless HART, particularly version 7 of the protocol, surpasses a large number, not all, the limitations of wired HART protocol. Wireless HART protocol uses a standard similar to IEEE 802.14 (ZigBee) and the usage of frequency hopping techniques, on a per packet basis, enables a significant improvement in transmission reliability. Different mesh, star, combined mesh and star topologies can

be implemented in the same network and alternative paths can be automatically created if a network communication failure occurs since wireless HART devices can act as a router. Regarding main advantages, it can be referred: no wires are needed; the protocol can be used in moving measurement platforms; the protocol exhibits self-configuring and self-healing capabilities and a channel blacklisting capability can be implemented to remove bad channels usage in the network. As limitation, it can be referred: low message speed, similar to the traditional HART protocol; battery power limitation that can be surpassed avoiding transmission of redundant data and usage of adjustable transmission power levels and limited coverage range between two devices that is in the order of 250 m LOS (line of sight). Obviously, this last limitation is minimized by using HART routers and directional antennas.

### **8.3.3. Fieldbus (FF)**

The usual solution, still considered nowadays, for industrial control and instrumentation systems is based on distributed control systems (DCS). However, after the development of smart sensing and actuating (SSA) devices [16-17], new solutions based on instrumentation and control systems interconnected by industrial area networks (IAN) [18-19] became a promising alternative to implement new and to update present instrumentation and DCS. In which concerns systems' operation and maintenance, these networks, and associated devices, have a large number of advantages that are well-known and described in the literature [18]. One of the main advantages that is always referred is device interoperability of field bus devices. However, this is a very ambitious objective and there are multiple protocols to interconnect fieldbus devices from different manufacturers. In this context, two of the most promising fieldbus solutions, for continuous processing applications, are Foundation Fieldbus (FF) and Profibus Process Automation (P-PA).

Foundation Fieldbus [20-25] is a fully digital solution designed to interconnect industrial control and instrumentation devices, such as proportional, integral and derivative (PID) controllers, programmable logic controllers (PLC), and transducers and actuators. Basically, FF defines a standardized hardware and software for IAN implementation of instrumentation and control systems (ICS). The protocol proposes two hierarchy levels. The lower hierarchy level, known as H1, is used to connect field devices and runs at 31.25 kbit/s over a single wire pair that can be used simultaneously to power instrumentation without requirements of external power supplies. The second hierarchy level, known as H2 or high-speed Ethernet (HSE), is used to interconnect different fieldbus networks (H1) and mainly supports supervision, management and other system's global tasks. In this paper, only the first hierarchy level (H1) is considered. Regarding architecture, the FF H1 architecture is based on a simplified version of the open system interconnection reference model (OSI Reference Model) [26-27]. The protocol architecture includes the following layers: physical, communication stack and user application.

The physical layer is equivalent to layer 1 of the OSI model. It performs bidirectional tasks that convert digital bit streams into electrical signals that are transmitted over the physical communication support.

The communication stack, immediately above the physical layer, concatenates layers 2 to 7 of the OSI model. It includes the data link layer (DLL) and the application layer that is divided into two sub-layers: the fieldbus access sub-layer (FAS) and the fieldbus message specification (FMS) that corresponds to the OSI application layer, just underneath the user application layer.

The user application layer, not included in the OSI model, defines a set of standardized software blocks that perform modular functions according to users' requirements.

Regarding the physical layer and medium, FF signals are encoded using the Manchester biphasic-L coding [28]. In this coding technique the clock information is sent together with the data signal. The transmitting devices deliver  $\pm 10$  mA current pulses at 31.25 kbit/s rate. These current pulses produce on the FF line terminator ( $50 \Omega$  resistor) a voltage of  $\pm 0.5$  V that is applied to all receptor devices. The maximum number of network addresses for each FF (H1) link is 240 but the practical limit depends on instruments power consumption and cable characteristics. Using normalized AWG 20 wires inserted in a shielded twisted pair cable it is possible to run a network with a maximum length, including spurs, of 1900 m. However, typically, the maximum number of network addresses for each FF (H1) link is much lower than 240, namely, 16 and 6 devices, for non-intrinsically safe or intrinsically safe networks [29-32], respectively. By its turn, FF communication stack contains the data link layer (DLL), the fieldbus access sub-layer (FAS), and the fieldbus message specification (FMS).

The DLL controls the transmission of the messages onto the FF network. It uses a deterministic and centralized bus scheduler called the link active scheduler (LAS). The LAS manages a list of transmitting time slots for all devices that need to access the bus cyclically in a scheduled mode. When a scheduled communication takes place, the LAS sends a control data (CD) token to that device, which broadcasts or publishes data for all other devices that are configured to act as subscribers or clients. On the other way, FF devices transmit unscheduled messages between the transmissions of scheduled messages that have the highest priority.

The FAS sub-layer uses the scheduled and unscheduled features of the DLL to provide a service for the FMS. There are several types of FAS services, described by virtual communication relationships (VCR), in particular: client/server, report distribution and publisher/subscriber VCRs.

The FMS sub-layer describes the communication services, message formats and protocol behavior needed to build the messages for the user application layer.

Concerning the user' layer that is not normalized as an OSI layer, the FF applications are implemented through modular software blocks. These blocks are standard software routines that implement FF functions. The main block types include: resource blocks, used

to describe the characteristics of the fieldbus devices; transducer blocks, used to read data from sensors and output data to actuators; and function blocks, used to define the tasks performed by each block.

As an example, Fig. 8.6 represents a set of blocks that are associated with two devices. Device 1 contains a set of blocks that are associated with a thermocouple transducer and device 2 contains a set of blocks that are associated with a control pneumatic valve actuator.

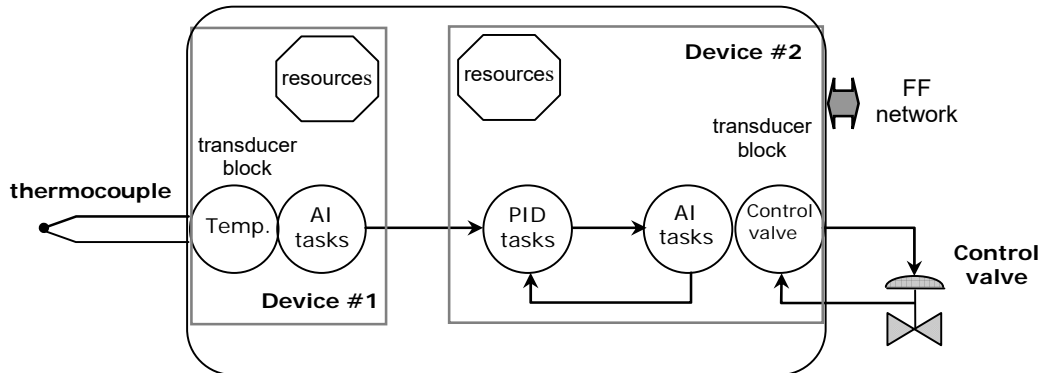


Fig. 8.6. Temperature and control valve block diagrams.

Each block has its own diagram and can be linked, in a convenient way, to perform the user control strategy. The software links can be established between blocks of different devices. Fig. 8.7 exemplifies the block diagram of an analog input block (AIB).

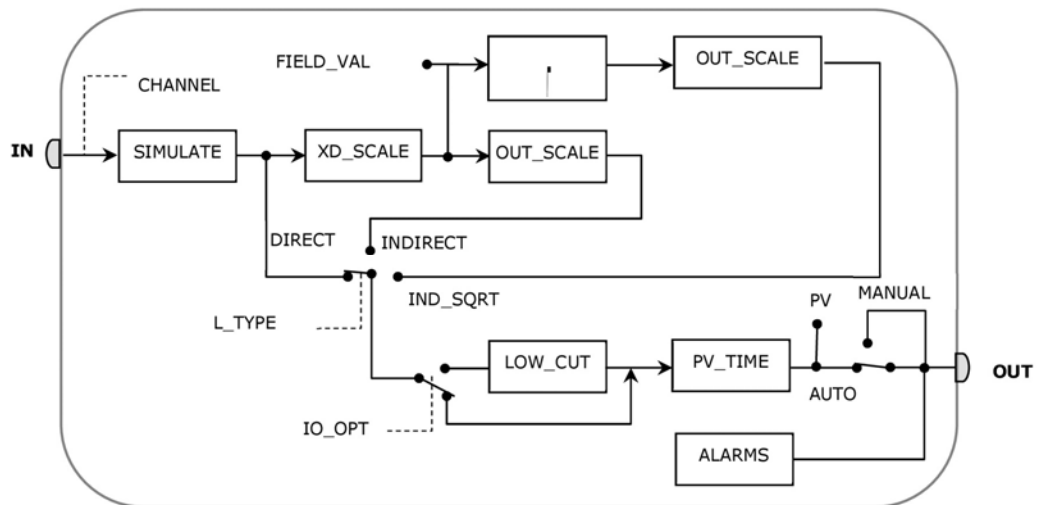


Fig. 8.7. Analog input block diagram.



The software configuration of this block is performed through the following parameters: channel type and number, which identify the transducer's type and its sequence number in the block, respectively; *l\_type*, which defines the input data processing method (direct, indirect and indirect square root, usually used for linearization of flow meters' measurements); *xd\_scale* and *out\_scale*, used for calibration purposes and physical to engineering units conversion; *blk\_mode*, that defines block's operation mode (auto or manual); *low\_cut*, which defines the cut-off frequency of the digital filter that attenuates low frequency components of the measuring signal; and *pv\_time*, which defines the time constant of the exponential filter that is used to attenuate high frequency components of the input signal.

Regarding main advantages of FF, it can be referred: it assures an open system architecture being devices' interoperability extend to the higher levels' OSI layers and to user's layer; wiring cost reduction since multidrop networks' configuration can be implemented in the same pair of wires; increased maintenance, diagnostic, calibration and plug&play capabilities. Moreover, FF, not P-PA, can implement true field control system (FCS) solutions, being FF H1 networks deterministic, isochronous, synchronized and with real-time capabilities, as long as the process time constants be much (10 times) than FF macrocycle period. As disadvantages, or limitations, it can be referred: FF cycles times are typically several hundreds of ms; gateways between fieldbus and corporate and office networks are still required since protocols are not Ethernet based; FF is a half-duplex token-based protocol and expensive and specific FF communicators are required for calibration purposes.

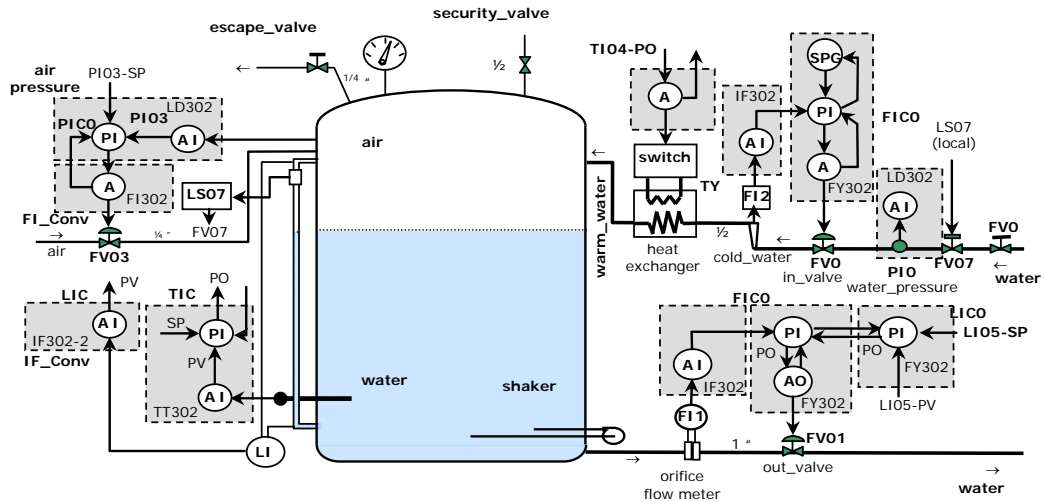
Another fieldbus protocol that is largely applied in industrial application is P-PA. This protocol has a physical layer based on Manchester bus powered (MBP) that is intrinsically safe, however, P-PA is a master-slave protocol and cannot implement FCS solutions. Control tasks in P-PA must reside in the host and most Profibus control hosts don't accept P-PA inputs directly. Instead, a linking device or coupler must be used to convert P-PA into P-DP that uses different physical layer protocol.

Furthermore, P-PA does not provide a real-time distributed clock being all communications and tasks' synchronization centralized in the host. Instead, FF is appropriate for real-time control on the bus with or without host.

#### **8.4. Case Study: FF Prototype**

To test some FF devices available in our laboratory, and to evaluate their performance, a small process was built and several instruments were installed, configured and tested. Fig. 8.8 depicts the process and instrumentation (P&I) diagram of the implemented prototype. It includes a closed 240 liters tank, analog and digital instrumentation, and the interfaces between both instrumentation types. The analog instrumentation includes an orifice flow meter with the associated transmitter (FI1) that is connected to the FF bus through channel 1 of the current to FF converter (IF302) [33]; a differential pressure transmitter (LI04) that measures the water level inside the tank and is connected to channel

2 of the same current to FF converter; and an input water flow meter based on a magnetic resistive rotameter (FI2), which is connected to channel 3 of the IF302 converter. The power switch (TY04) used to control water temperature through the heat exchanger, is controlled through channel 1 of a FF to current converter (FI202) [34] and the channel 2 of the same converter is connected to valve (FV03) that controls the air pressure inside the tank.



**Fig. 8.8.** Prototype P&I (AI – analog input; AO – analog output; LD302 – FF level device; TT302 – FF temperature device; FY302 – FF control valve; FI – FF to current converter; IF – current to FF converter; TY – water heater; SP – set point; PV – process variable; FV – analog control valve).

The digital instrumentation includes a FF temperature device (TT302) [35], 2 FF pressure devices (LD302) [36] and a FF valve (FY302) [37]. Configuration of all devices was performed according to the process characteristics and to the control strategy to be implemented. This strategy is defined by the parameters of each system's block diagrams and by the software links established between them.

According to system's design, it possible to control several variables namely, pressure, temperature, level and flow, and to study the mutual interdependence among those variables.

#### 8.4.1. Experimental Results

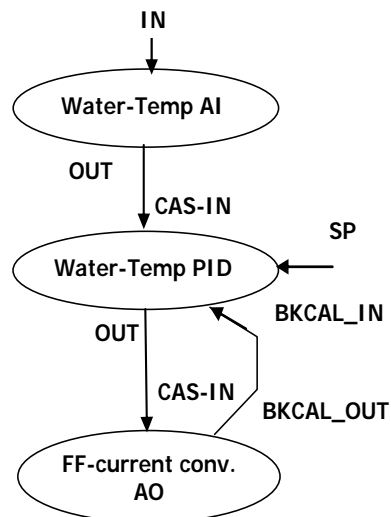
There are a lot of experimental tests that can be realized with the proposed FF prototype. With the results from those tests, it is possible to evaluate system's static and dynamic characterizations under different working conditions. Mainly, there are four system's variables, pressure, temperature, water flow and water level, whose dynamic behavior depends on system's design, control parameters and control strategy. It is important to

underline that there are interdependences between system's variables whose mutual correlations coefficients can be experimentally quantified.

Taking as example the aim to control water temperature inside the tank, this section will present and discuss some experimental results.

Water temperature control is performed by a water exchanger available in our laboratory. This water heater is powered by a 24 V supply and its electrical power is equal to 19 kW. However, in regular working conditions, it delivers a constant and continuous power being inadequate to perform the desired water temperature control. To surpass this limitation, a pulse width modulated (PWM) signal, controlled by a 4-20 mA current signal generated by the FF system through a FF to current converter, was used to command a power switch that switches-on and off the water heater. In this way, the output power obtained from the water heater can be continuously adjusted by the duty cycle value of the PWM signal. If the PWM signal duty cycle is zero the power switch is always switched off, and, conversely, if the duty cycle is one the power switch is always switched on, and the power delivered by the water heater is equal to 0 and 19 kW, respectively. The power switch driver circuit is controlled by a comparator circuit with hysteresis (Schmitt trigger); otherwise, the power switch could be continuously switching on and off, which would damage it.

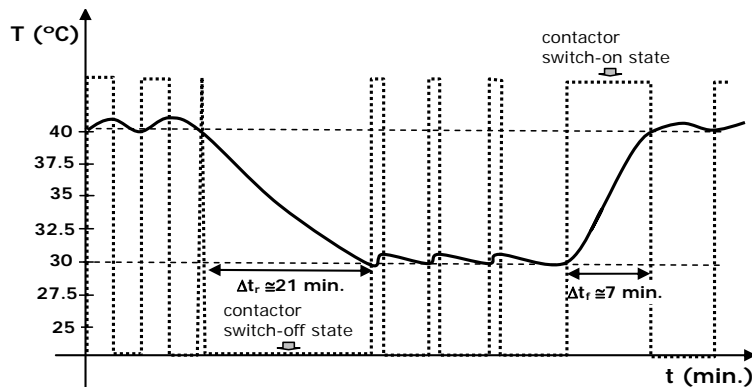
The strategy used to control water temperature is represented in Fig. 8.9. The control tasks were implemented in the FF temperature device (TT302) and the main software blocks that were involved include: an analog input block that captures the water temperature from a Pt100 sensor, a PID software block whose set point is the desired water temperature, and a FF to current converter to control the duty cycle of the PWM signal that controls the power delivered by the water heater.



**Fig. 8.9.** Graphical representation of the water temperature control strategy.

It is also important to refer that the PID block runs in cascade mode. With this set-up, the operator cannot access the PID set point and only a supervisory block can do that. For testing purposes, the PID block was set to run in proportional and integral mode and its proportional gain and reset time were tuned to the middle point of the temperature range (25 °C). It is also important to underline that the temperature control test was performed with a constant water level in the tank. Otherwise, the results will be strongly affected by the water volume contained in the tank. If the water volume is higher or lower the temperature time constants will be higher or lower, respectively.

Fig. 8.10 represents the water temperature control results for a temperature set point change between 30 °C and 40 °C. From the figure it is clearly perceptible that the system exhibits a non-linear behavior in which concerns water temperature control. This behavior was expected since the water temperature depends strongly on input water and ambient temperatures and the temperature control is performed exclusively by a high temperature heater source. Consequently, the temperature rising time constant is lower than the temperature falling time constant. Really, in this case, if the temperature set point is lower than ambient or water temperature, the water temperature settling time will be infinite since the system does not contain any cooling source, for example Peltier cells, to improve water temperature reduction rate. Consequently, as represented in Fig. 8.10, the falling time constant is almost equal to three times the rising time constant for a 10 °C step variation of the temperature set point.



**Fig. 8.10.** Water temperature control results.

The proposed prototype is a very suitable solution for teaching purposes for STEM students that require laboratory expertise in recent industrial instrumentation digital fieldbus technologies.

## 8.5. Conclusions

This chapter highlighted three specific protocols for continuous process industrial applications. Obviously, it does not make sense to present all solution that have been used,

over time, to promote industrial communication. Some protocol solutions for manufacturing processes, such as CAN and DeviceNet, among others, were not highlighted because they are not common in industrial continuous processes, as well IEEE 1451 standard because it has, not yet, a clear acceptance in industrial communication applications.

By its turn, the three protocols that were presented still have a large implementation in several industrial applications and this fact justifies their choice. A case study related with FF digital protocol that includes the development of a didactic prototype that uses a hybrid digital and analog system was presented in the end of the chapter.

## References

- [1]. M. Pereira, O. Postolache, P. Girão, Using a dual-channel FDC device and ANN techniques to improve measurements accuracy, *Sensors & Transducers Journal*, Vol. 62, Issue 12, 2005, pp. 462-472.
- [2]. P. Girão, M. Dias Pereira, O. Postolache, Multisensor data fusion and its application to decision making, in *Advanced Mathematical & Computational Tools in Metrology VII*, Vol. 72, *World Scientific Publishing Co Pte Ltd*, 2006, pp. 47-59.
- [3]. S. Haykin, *Neural Networks – A Comprehensive Foundation*, *Prentice Hall*, 1999.
- [4]. L. C. Paul, A. Al Suman, N. Sultan, Methodological analysis of Principal Component Analysis (PCA) method, *IJCEM International Journal of Computational Engineering & Management*, Vol. 16 Issue 2, March 2013, pp. 32-38.
- [5]. ANSI/ISA-S50.1, *Compatibility of Analog Signals for Electronic Industrial Process Instruments*, ISA, USA, 1982.
- [6]. R. Bowder, HART – A Technical Overview, 2<sup>nd</sup> Ed., *Emerson Process Management*, February 2002.
- [7]. R. Bowden, HART, A Technical Description, *Fisher-Rosemount*, 1996.
- [8]. HART Smart Communication Protocol Specification, Rev. 5.8, *HART Communication Foundation*, April 1998.
- [9]. Technical Overview, FD-043, *Fieldbus Foundation*, 1998.
- [10]. Field Devices Installation and Maintenance Manual, *Smar*, *Fieldbus Foundation*, 1998.
- [11]. Romilly's HART and Fieldbus Web Site, <http://www.romilly.co.uk>
- [12]. HART communication foundation, *HART Line Journal*, Issue 1, 2000.
- [13]. Bell System Technical Reference: PUB 41212, Data Sets 202S and 202T Interface Specification, *Bell Telephone Company*, July 1976.
- [14]. P. W. Chapman, *Smart Sensors*, ISA, 1995.
- [15]. R. Frank, *Understanding Smart Sensors*, *Artech House*, 1996.
- [16]. G. van der Horn, J. L. Huijsing, *Integrated Smart Sensors, Design and Calibration*, *Kluwer Academic Publishers*, 1998.
- [17]. R. van de Plassche, *Integrated Analog-to-Digital and Digital-to-Analog Converters*, *Kluwer Academic Publishers*, 1994.
- [18]. D. Caro, *Wireless Networks for Industrial Instrumentation*, ISA, 2004.
- [19]. G. Held, *Understanding Data Communications*, 5<sup>th</sup> Ed., *SAMS Publishing*, 1996.
- [20]. Technical Overview, FD-043, *Fieldbus Foundation*, 1998.
- [21]. *Fieldbus Installation and Planning Guide*, AG-165, *Fieldbus Foundation*, 1998.
- [22]. J. Berge, *Introduction to Fieldbuses for Process Control*, *The Instrumentation, Systems, and Automation Society*, 2002.
- [23]. Network Management Specification FF-801, *Fieldbus Foundation*, 1998.

- [24]. Specification – System Architecture – FF-800-1.3, *Fieldbus Foundation*, 1998.
- [25]. Specification – Fieldbus Message Specification – FF-870-1.3, *Fieldbus Foundation*, 1998.
- [26]. ISO/IEC Standard 7498-1:1994, *USA*, 1994.
- [27]. H. Zimmermann, OSI reference – The ISO model of architecture for open systems interconnection, *IEEE Transactions on Communications*, Vol. 28, Issue 4, April 1980, pp. 425-432.
- [28]. A. Tanenbaum, *Computer Networks*, 4<sup>th</sup> Ed., *Prentice Hall*, 2002, pp. 274-275.
- [29]. *Guide to Intrinsic Safety*, *Elcon Instruments*, 1996.
- [30]. *A User’s Guide to Intrinsic Safety – Application Note AN9003*, *Measurement Technology Limited*, 1999.
- [31]. B. Rankin, Electrical equipment in hazardous areas: Field inspections, in *Proceedings of the Hazardous Areas Conference*, Vol. 2, 2007, pp. 2-16.
- [32]. *Guide to Intrinsic Safety*, *Elcon Instruments*, 1996.
- [33]. *The Fieldbus Book – IF302 Current to Fieldbus Converter*, *Smar*, 2002, pp. F132-F137.
- [34]. *The Fieldbus Book – FI302 Fieldbus to Current Converter*, *Smar*, 2002, pp. F138-F143.
- [35]. *The Fieldbus Book – TT302 Fieldbus Temperature Transmitter*, *Smar*, 2002, pp. F92-F99.
- [36]. *The Fieldbus Book - LD302 Fieldbus Pressure Transmitter*, *Smar*, 2002, pp. F56-F65.
- [37]. *The Fieldbus Book - FY302 Fieldbus Valve Positioner*, *Smar*, 2002, pp. F118-F125.



# **Chapter 9**

# **Industrial Ethernet and Time Sensitive Networking for Industry 4.0**

**José Miguel Costa Dias Pereira**

## **9.1. Introduction**

One of the main foundations for Industry 4.0 (I4.0) is the availability of all relevant information in real-time (RT) by connecting all instances involved in the product value chain. The ability to extract information from a large quantity of data (Big Data) is also crucial for the success of I4.0. Thus, the framework of I4.0 depends on digitalization, networking and advanced data processing (Data Analytics). Usage of industrial IoT (IIoT), together with advanced processing tools, will give access to a huge quantity of knowledge, enabling the find out of hidden information that can be gathered from data analysis, leading to an optimization in systems' engineering, maintenance and operation costs.

The mains supporting vectors associated with I4.0 includes: interoperability; virtualization; virtualization; decentralization; RT capability; service orientation and modularity. Interoperability refers to the capability of all components to connect, communicate and operate together via the IIoT. Virtual models based on sensor data are used to create models of smart factories (SF) that can be used for process simulation proposes and performance optimization. Decentralization of different manufacturing units (MU), eventually located in different locations, is performed in order to maximize the organizational goals and yields. Regarding networking and RT requirements, I4.0 efforts are directed to fulfill RT requirements, particularly in the floor level that, by its turn, must be integrated to the different working levels of an organization. According to the RT demands of the different tasks, a communication network and protocols must be developed to accommodate the services requirement associated with the different tasks that can range from hard RT demands to surveillance monitoring videos. Service oriented capabilities are also provided by the IIoT being internet of services crucial in I4.0.

---

José Miguel Costa Dias Pereira  
ESTSetúbal/IPS, Instituto Politécnico de Setúbal, Setúbal, Portugal  
Instituto de Telecomunicações, Lisboa, Portugal



Flexibility and modularity are also required for the successful implementation of SF and for I4.0. The SF configuration can be changed dynamically with a minimum disruption of the production processes or any shutdown periods.

Regarding RT, it is important to refer that full-duplex industrial Ethernet solutions, at the floor level, removes traffic collisions and enables communication at the Medium Access Control (MAC) address layer with RT and deterministic data time delivery capabilities. At the same time, higher OSI protocol layers compatibility enable the integration with Internet Protocol (IP) networks that include business LAN and WAN without need of gateways.

Regarding chapter organization, in Section 9.2 includes some definition about the main concepts related with IIoT and I4.0, Section 9.3 is dedicated to industrial Ethernet protocols and Section 9.4 includes the main principles associated with time sensitive networking (TSN). The last section, Section 9.5, includes a summary of the main conclusions of this chapter.

## **9.2. IIoT and Industry 4.0**

### **Main Characteristics and Differences between M2M, IIoT and Industry 4.0.**

The main techniques and technologies that are nowadays applied in machine-to-machine (M2M) network applications are basically the same that will be applied in IIoT. In both cases, devices are used to capture data, such as physical measurements, that is transmitted over a wired, wireless or hybrid network, and processed by an application that runs remotely. However, even if it seems similar, the main and substantial difference between M2M technologies and IIoT, is related with the scale of integration. In IIoT the network integration enables that a huge quantity of data can be analyzed online, in distributed cloud storage systems, and that a large number of services are integrated to optimize process efficiency. Large amounts of data, from different equipment, running in the same or different manufacturing units (MU), can be processed together improving the overall performance and efficiency of a global production system. This capability can justify the implementation of IIoT due to the cost reductions and products' quality improvements that can be achieved. It is important to refer that advanced algorithm that are used to process the Big Data (Big Data Analytics) [1-2] can detect hidden correlations between data that can't be found by processing, locally, captured data for a single device, or from a group of devices, eventually working in different MU. In this scenario, diagnostic data from sensors, actuators and PLC (Programmable Logic Controller), as well as historical data trends of the different industrial variables, can provide the required data to establish a successful predictive maintenance, reducing maintenance costs, avoiding superfluous preventive maintenance tasks and costly unscheduled production shutdowns. Predictive analytics tools can be used to identify the precursors of unplanned adverse events, determine the most profitable way to handle aging and degrading of assets and to optimize the usage of maintenance resources [3-5]. In summary, it can be referred that IIoT promotes the integration of devices, systems and resources using, as much as possible, standard communications and protocols, to integrate companies' distributed systems and

to increase the entire value chain of their business. Regarding IIoT protocols, even if the purpose is always standardization, it is obvious that the different real-time and security requirements in the MU floor and in the cloud are completely different. Thus, different protocols, more or less compatible with the OSI model, are implemented to provide the desired global interconnectivity between MU and services. Concerning transmission modes, an increased usage of Wireless and 5G technologies, to provide remote and centralized data visualization, and to support monitoring and maintenance activities, is obviously expected. IPv6 and network security are also two important topics that must be considered for the success of IIoT.

Regarding the Industry 4.0, it can be referred that IIoT is an enabling force for its implementation. Industry 4.0 is a blend of digitalization and new technologies focused to increase industrial process flexibility, reliability and efficiency, based on unprecedented visibility of every production equipment including the supply chain and, eventually consumers. The selection of IIoT technologies is a required condition, or step, to achieve the Industry 4.0 desiderate. This mean basically that IIoT is the enabling technology that connects devices, machines and people and Industry 4.0 implements the corporate philosophy based on digitalization and new technologies that includes lean initiatives and automation solutions in every MU.

The synergies that result from IIoT and Industry 4.0 gives a set of globally competitive results from which stands out a performance increase of visibility, flexibility, efficiency, zero downtime and maximum overall effectiveness of industrial equipment.

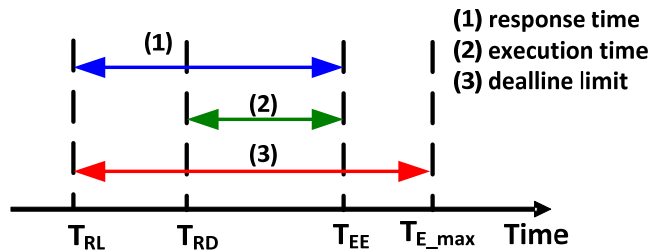
### **9.3. Industrial Ethernet**

Ethernet, as defined by IEEE standard 802.3, is a non-deterministic protocol and, thus, not suitable for applications with RT demands. The medium access control (MAC) of Ethernet is based on the well-known Carrier Sense Multiple Access with Collision Detection (CSMA/CD). The random occurrence of collisions is not compatible with the achievement of deterministic latency times to exchange data in the network. However, Industrial Ethernet should ensure RT demands, in terms of latency and jitter, and must enable data transmission of non-RT data that share the same transmission support [6-8]. The first approaches to reach these objectives were materialized by using proprietary solutions, such as, Profinet IRT, Ethercat, and Ethernet Powerlink [8-11], EtherNet/IP and SERCOS III. However, these solutions are not open and require special interfaces and gateways to connect different network types from different manufacturers. This section starts with a brief introduction of real time definitions.

#### **9.3.1. Real Time**

Definitions about RT are sometimes not consistent in the literature. As a matter of fact, RT is not necessarily associated with fast execution times but is more related with determinism and synchronization of tasks' execution in a given system. Considering a job that must be processed in a Computer Processing Unit (CPU), Fig. 9.1 represents the

timings that are associated with its execution. The release time ( $T_{RL}$ ) represents the time when the job is available to the system, the ready time ( $T_{RD}$ ) represents the time when the job can start execution, the end of execution time ( $T_{EE}$ ) represents the time when the job finish execution and the deadline ( $T_{E\_max}$ ) represents the time when the job execution must be finished. According to these time marks, the execution time is the time interval between  $T_{EE}$  and  $T_{RD}$  and the response time is the time interval between  $T_{EE}$  and  $T_{RL}$ .



**Fig. 9.1.** Time marks of a job execution ( $T_{RL}$  – release time;  $T_{RD}$  – ready time;  $T_{EE}$  – end of execution time;  $T_{E\_max}$  – deadline time).

The deadline limit that is associated with a given job is directly related with its RT requirements. According to the RT requirement, it is common to consider four RT categories, namely, Hard Real Time (HRT), Soft Real Time (SRT), Firm Real Time (FRT) and Weakly Hard Real Time (WHRT). HRT requirements are associated with mission and safety critical systems where failures can result in catastrophic events, such as, loss of life or property. Systems that have HRT requirements use typically hardware redundancy techniques and, if considered for industrial applications, must be redundant in terms of hardware, software and networking, using, if necessary, RT operating systems (RTOS) to assure the right outputs within stringent deadlines. Examples of HRT requirements in the industrial field are automatic control and robotics applications that must preserve the security of equipment and safety of the people that work nearby those system. Another example, in process industries, can be found in the ethylene production by naphtha cracking technology where pressure values of the process can achieve values over 1000 bar and where timing requirements are very critical. By its turn, in SRT systems deadline overruns, within certain limits, are not catastrophic and are not associated with critical failures. Real time performance of SRT systems is usually measured by a Quality of Service (QoS) index that can be used to evaluate system performance. Concerning industrial applications, SRT requirements could, for example, be associated with the temperature control of a fluid that is contained in a tank, since the temperature adjustment is usually a very slow process and deadline overruns, within certain limits, are not critical and don't cause any reduction of system performance. In FRS systems or applications, with high tolerance for deadlines overruns, there are no critical problems that arise when deadlines are not fulfilled but the usefulness of tasks' completion decreases, and could be of no value, if the task is not completed within a maximum period of time. In the industrial field, considering an example about the measurement of the water temperature contained in a boiler, and supposing that a set of four temperature sensors are used to take an average value of the water temperature, if the temperature acquisition from one of the sensors is

not obtained before the pre-defined deadline, the deadline overrun has no impact in the system performance, since the average value of the temperature acquired from the others sensors is still meaningful to evaluate the water temperature. Regarding WHRT systems or applications, they can tolerate a maximum number of deadlines overruns within a certain time period. For example, in the industrial field, a WHRT task could be associated with a hypothetical wireless measurement system that tolerates a maximum number of transmission errors. If model predictive control is used, the real measurement data that is not received can be replaced by simulated data during a predefined maximum time period but, after that period, system critical faults can appear. Obviously, in this case the choice of a wireless transmission system could not be the best choice or, at least, other redundant mechanisms should be implemented.

Concerning RT requirements in industrial systems, it is also important to underline that as the hardware price of the instruments decrease, and their computing power increase, industrial systems are enlarging their distributed computing characteristics, increasing the implementation of control functions tasks in the field devices, which implies an increase of the performance of communication networks and protocols, namely in which concerns their RT performance. Thus, RT demands of the industrial applications must fulfil latency and jitter requirements not only in local computing tasks but, above all, those requirements and tasks' scheduling (synchronization) must be fulfilled in the industrial data networks that supports the communication between network nodes.

Nowadays, industrial Ethernet solutions can assure cycle times of some tens of microseconds [9-11], much lower than the cycle times of fieldbus solutions whose typical values are above several tens of ms. Thus, fieldbus solutions, such as FF, can only be used for low-speed sensing and drive control systems, and are not suitable for applications that require cycles times lower than ms, namely, applications related with motion control, high-speed devices interfacing and transient fault detection, among others. For this type of applications, industrial Ethernet, and other specific automation protocols are commonly used.

### **9.3.2. Collision Avoidance and Synchronization**

Collision avoidance in Ethernet networks can be achieved by using fully Ethernet switches instead of Ethernet hubs or bridges. The usage of fully Ethernet switched networks enables full-duplex communication, between all network nodes, and as long as the switch provides the required number of ports according to the maximum network traffic, there are no collisions, nor data loss and, at the same time, the communication bandwidth is doubled, since for a given link, between two nodes, the communications in both senses can occur simultaneously. Moreover, besides the segmentation of the collision domains provided by full-duplex switches, these network devices can handle prioritization according to the different QoS levels for the different types of frames that are transmitted in the network. Ethernet frames also include a set of bytes to support virtual local area networks (VLAN) that minimize broadcast traffic in the LAN. In this way, as long as the switching time over the network is predictable, it is possible to avoid missing of hard deadlines improving RT performance of the data network.

Besides collision avoidance and prioritization, industrial Ethernet solutions must synchronize network nodes using clock synchronization protocols. One of these protocols is the Precision Time Protocol (PTP) [12, 13] that can work on the top of different protocols, not only Ethernet, and enables that different clock in a heterogeneous network, including, obviously, network nodes, with different accuracies, resolutions and stability, to be synchronized to a grandmaster clock. With a network that supports PTP it is possible to achieve a sub  $\mu\text{s}$  timing precision that fulfils the requirements of most industrial measurement, control and automation applications.

The proprietary Ethernet solutions that are used nowadays can support the PTP but, as already referred, the different solutions don't assure interoperability of equipment and the integration of the different network levels, from manufacturing unit to office networks, can only be achieved with the usage of network interfaces and gateways which introduce additional costs, limits network flexibility and, above all, compromises network performance in terms of latency and jitter.

### **9.3.3. Industrial Internet**

In the following paragraphs a brief review of three RT Ethernet protocols is performed. RT networks must provide a guarantee of service in order to assure a deterministic and consistent operation. So, it is important not only validity of data but also timeliness. In RT systems jitter should be measurable so system performance can be guaranteed. This implies, typically, network clock synchronization [12] in order to achieve distributed RT processing that is much more than RT processing. Traditional Ethernet and fieldbuses, such as Foundation Fieldbus (FF), cannot meet cycles times below some tens of ms that are not enough to control high-speed devices and electronic ranging applications, such as fault detection. To implement deterministic Ethernet, all collisions must be avoided and the easiest way to achieve that is through the usage of hardware, namely, fast and intelligent switches to implement Switched Ethernet.

#### **Profinet**

Profinet IRT (Isochronous RT) was developed for system requiring sub-microseconds synchronization, typically, high performance motion control systems. The benchmark for such systems is 1 ms cycle time and 1  $\mu\text{s}$  jitter accuracy that are verified by Profinet-IRT. Profinet achieves its deterministic performance because, like all Ethernet switches, it uses its MAC address to communicate with other Ethernet devices over the LAN. By not using IP address, Profinet can reduce latency and jitter substantially. Determinism is also achieved by splitting data transmissions in three types: cyclic I/O data transmitted in RT; acyclic data transmissions used for parameter data and diagnostic information, which is not in real time; and alarms, which is another RT transmission channel for urgent maintenance required status alarms. Each Profinet-IRT device has a special ASIC (Application-Specific Integrated Circuit) for handling node synchronization and incorporate an intelligent 2 or 4 port switches. The Profinet switch in every node contains

a schedule for RT and non-RT traffic and provides full-duplex links for all ports. It is still important to refer that Profinet enables wireless communication with WLAN and Bluetooth.

### **EtherCAT**

EtherCAT is the motion -control RT solution from Beckhoff. It processes 1000 I/O in 30  $\mu$ s but requires full-duplex. It is based on the master-slave principle and can interoperate with normal TCP/IP and other Ethernet-based solutions such as Profinet. The EtherCAT master processes RT data via dedicated hardware and software. The EtherCAT master prioritizes EtherCAT frames over normal Ethernet traffic, which is transmitted in gaps. The two EtherCAT communication methods used “ether Type” or UDP/IP encapsulation. EtherCAT is internally synchronized by a distribution clock algorithm although external synchronization is achievable with IEEE 1558. Although, it is important to refer that EtherCAT is a fast RT deterministic solution in not used with UDT/IP or intermediate switches or routers between master and slave.

### **Ethernet Powerlink (EPL)**

EPL is a hard-RT protocol based on fast Ethernet that does not require special ASIC. EPL can deliver a cycle time of 200  $\mu$ s with jitter under 1  $\mu$ s. EPL uses cyclic communication with time-slot division and mater-slave principle. Powerlink cycle includes a start period, a cyclic period, an asynchronous period and an idle period. The start period the EPL broadcasts the start of cycle (SoC) frame that is used to synchronize the slaves. The slave responds to the master but any slave can access data, thus, interslave communication can occur. After successful polling the master broadcast the end-of-cycle (EoC) frame, also for synchronization and cycle validation purposes. Unlike Profinet, EPL does not ProfNet employ switch to avoid collisions or to provide network synchronization. Unlike Profinet where normal Ethernet and RT devices can co-exist, in EPL RT traffic must be protected from non-RT communication through bridges or routers. However, EPL does not requires special ASIC and employs only standard Ethernet hardware

### **EtherNet/IP**

EtherNet/IP is an application layer protocol that run on the top of TCP/IP. It was developed by Allen-Bradley and ODVA (open DeviceNet Vendors association). EtherNet/IP uses standard physical, data link, network and transport layers, using Common Industrial Protocol (CIP) over TCP/IP. This protocol establishes communications from one application node to another through CIP connection over TCP/IP, and multiple CIP connections can be established over a single TCP connection. The standard Ethernet and switches are used which means that an unlimited number of node can exist over EtherNet/IP. However, EtherNet/IP is not cycle-based but time-based meaning that it only requires that control command is received by field stations in time. RT delivery is safeguarded by three mechanisms, namely: UDP, quality of service (prioritization) and PTP IEEE 1588. To achieve RT capability, accessible bandwidth is limited to avoid contention and latency which mean that there is a limited performance range of RT services.

## SERCOS III

Serial RT communication system (SERCOS) is predominantly used in motion control-based automation systems. SERCOS III uses Ethernet physics (100 Mbps) and Ethernet telegram but is based on a time slot mechanism in which bandwidth is reserved for synchronous RT services and for asynchronous data services. SERCOS III works without need of hubs and switches. Each station has a special integrated ASIC or FPGA with two communication ports being possible to use line or ring topologies at the expenses of network topology flexibility. SERCOS III network nodes must be deployed in a daisy chain or a closed ring. Data is processed while passing through a device, using different types of telegrams for different communication types. At the same time, due to the full-duplex capacity of the Ethernet connection, a daisy chain acts as a single ring, being possible to implement easily redundant data transfer based on a ring topology. Based on the usage of double communication ports, on every node, devices are capable of communicating with each other within one communication cycle, with no need to route their data through the master.

### 9.4. Time Sensitive Networking

To promote network standardization, real time capabilities and interoperability, in an integrated Ethernet network, a novel version of the IEEE 802 standards is being developed [14-20]. To fulfil real time demands in factory networks, several mechanisms are being considering for TSN [21]. One of these mechanisms, very similar to the one that is used with FF, but running at a much higher bit rate, is the usage of TDMA within each data stream macro cycle (MC). The Time-Aware Scheduler (TAS), together with network timing synchronization mechanisms, assures a deterministic end-to-end latency and a maximum jitter uncertainty for time-sensitive data. Without the implementation of the TAS, low-priority traffic, without any real-time requirements, could prevent the on-time delivery of time-sensitive data. The TAS assigns specific time-slots, in each MC, to carry time-sensitive data throughout the data network without need of any token-based mechanism. As represented in Fig. 9.2, TDMA enables data transmission of scheduled data in a deterministic and periodic way. Each MC contains several time-slots, some of them assigned to time-sensitive data and others assigned to best-effort data that has not end-to-end hard deadline requirements in terms on latency and jitter. In this way, as long as there is no timing overlap between time slots, it is possible to assure a timely transmission of RT data without collisions.

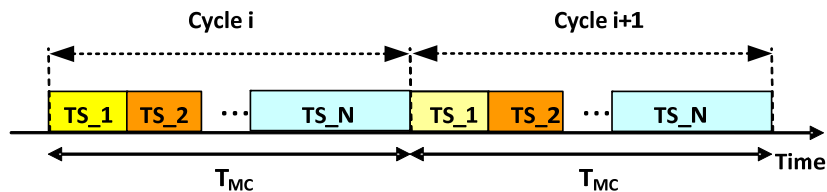
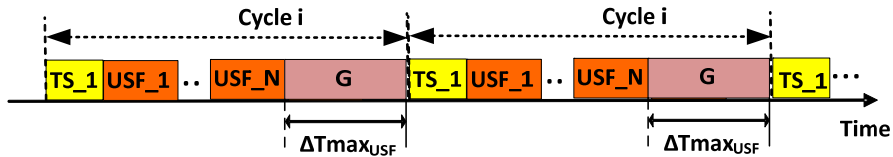


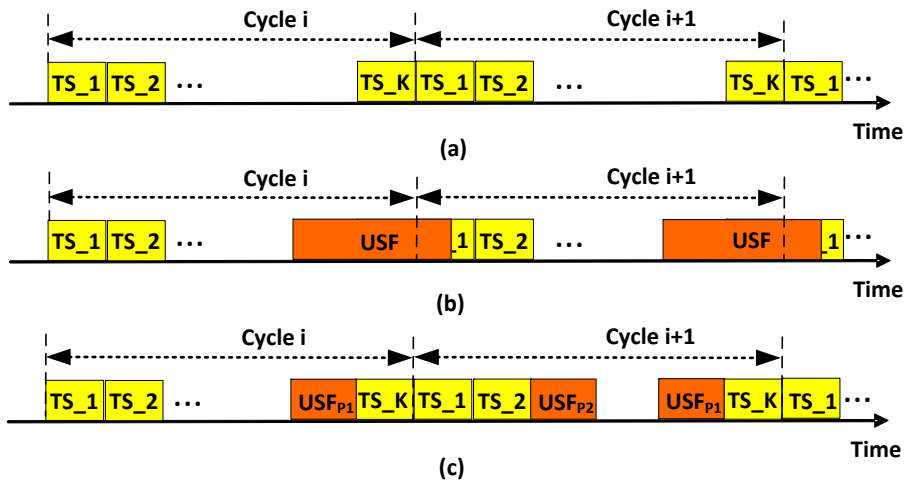
Fig. 9.2. Assignment of time slots using a periodic TDMA mechanism within the TSN macrocycle.

Other mechanism that must be considered to avoid that the best-effort data overlaps the scheduled time-critical data is based on the usage of guard-bands. As represented in Fig. 9.3, these guard bands, whose time duration must be higher than the maximum unscheduled packet duration ( $\Delta T_{max_{USF}}$ ), avoids the occurrence of timing overlaps associated with non-critical data with TS reserved for critical data. However, used in this way, this mechanism has an inherent waste of transmission bandwidth because there is always a non-efficient usage of the channel data transmission capability whenever the unscheduled data packet has a time duration lower than  $\Delta T_{max_{USF}}$ .



**Fig. 9.3.** Usage of guard bands to prevent unscheduled data packets to extend into a TS used for critical data.

In order to avoid the loss of useful bandwidth caused by the usage of guard bands, IEEE proposes another mechanism that is based on the usage of framelets. This mechanism enables that best-effort data frames are segmented into framelets that can have a minimum size of 64 bytes, being possible to interrupt, without data loss, best-effort frames to transmit scheduled data with real time requirements. This mechanism is denominated frame pre-emption. Fig. 9.4 depicts the segmentation of an unscheduled frame (USF) in two framelets ( $USF_{P1}$  and  $USF_{P2}$ ) in order to assure the correct timing requirements associated with the scheduled data (TS1, TS2 and TSK), being possible to reduce the guard band bandwidth waste to a maximum size of 64 bytes.



**Fig. 9.4.** Frame pre-emption mechanism: (a) time slots associated with time sensitive data; (b) overlap caused by unscheduled frame; (c) usage of the framelet segmentation mechanism to prevent timing overlaps and to optimize network performance.



In the network nodes TSN also implements Time-Aware Gate (TAG) mechanisms that uses the priority code point (PCP) of the IEEE 802.1Q frame header to prioritize the traffic queue that is transmitted in a specific time slot within the MC. Each output port of any switch (Out\_P) has an associated gate schedule control that manages the traffic queues according to their priorities and assigned TS [22-24]. In the example represented in Fig. 9.5, all input gates of the transmission multiplexer are closed, except gate 1 that is open to transmit time-sensitive data buffered in traffic queue 1. Regarding data transmission priorities, different centralized or hybrid approaches [21] can be used to assign the correct data Class of Service (CoS) priority in order to ensure that data frames are transmitted according to their latency and jitter requirements.

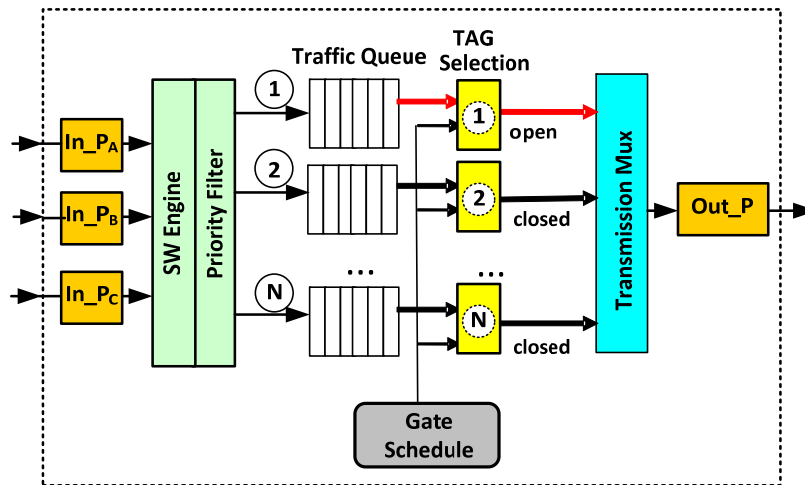


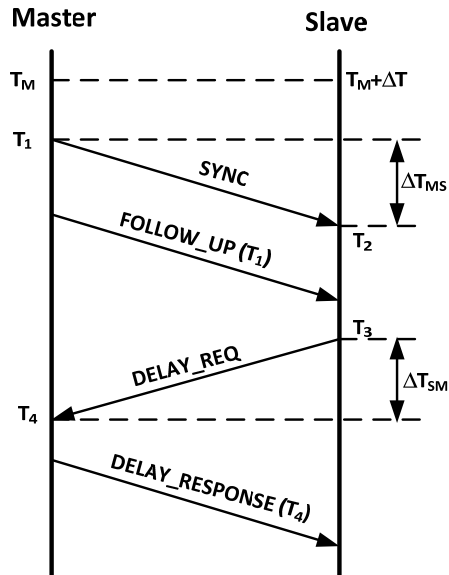
Fig. 9.5. Block diagram of a TSN switch with gate schedule capabilities.

Another issue that is of major importance in TSN is synchronization. Clock distributions mechanisms must be implemented in order to achieve a common time base for all network devices, end devices and network switches that transmit time sensitive data. In every device connected to the network the time marks that are used to define the start and the end of each time slot must be synchronized to assure that the time-sensitive data is transmitted according its schedule TS. Different types of time synchronization protocols, such as the PTP [12, 13], previously referred, can be implemented in order to assure clock accuracy values, between all the network devices, better that  $1 \mu s$ . The White Rabbit (WR) project proposes enhancements of the IEEE 1588 PTP that supports clocks' synchronization in the sub-nanosecond level [25] that is more than enough for TSN requirements. It is important to refer that the time accuracy values are measured against the network reference clock, usually referred as Grand Master Clock (GMC) that synchronizes all the network devices clocks after measuring transmission message delays between networks nodes. As represented in Fig. 9.6, there are four different messages that are used for network synchronization purposes, namely: synchronization, follow-up, delay request and delay response messages. With this set of messages, it is possible to compensate clock offsets caused by propagation delays, random encapsulation times and

internal queuing delays of the synchronization messages. Considering the offset timing values that are associated with each pair of synchronization messages, the synchronization of the slave clock is based on the average time offset of the timing stamps of the messages that are exchanged between the master and slave clocks:

$$\Delta T = \frac{\Delta T_{MS} + \Delta T_{SM}}{2} \quad (9.1)$$

being  $\Delta T_{MS}$  the time offset that is associated with the master to slave synchronization messages and  $\Delta T_{SM}$  the time offset that is associated with the slave to master synchronization messages.

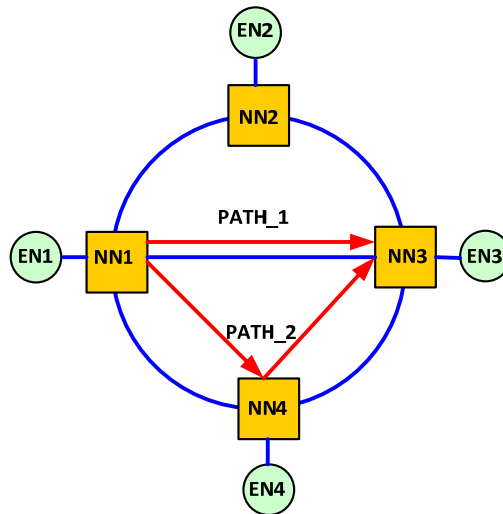


**Fig. 9.6.** Messages used to perform clock synchronization with PTP (SYNC – synchronization message; FOLLOW\_UP – follow-up message; DELAY\_REQ – delay request message and DELAY\_RESPONSE – delay response message).

Standard IEEE 1588-2008 [12] enables that the GMC performs up to ten synchronization cycles per second in order to assure a tight synchronization of the network clocks.

Concerning network reliability, that is another issue of major importance for almost all TSN applications, IEEE is also developing standards to improve redundant routing mechanisms [26]. Taking in mind that replication only makes sense for critical data and not for best effort data, data replication mechanisms, through redundant path, can be implemented in the end devices, or if they don't support redundant network connections, they can be implemented in the network switches. As an example, Fig. 9.7 represents a message replication that is implemented in the first network node (NN1) of the TSN. The time critical data frame is replicated and transmitted over two redundant paths that are connected to NN1. The first frame received in the destination switch (NN3) is forward to

the destination end node (EN3) and the second one is discarded, being this process transparent to the higher OSI layers. Even if it possible to improve reliability by expanding this mechanism to multiple redundant paths, it is important that each redundant path fulfil latency and jitter requirements of the running applications and an obvious compromise must always be achieved since network performance is reduced if redundancy is not used judiciously. To minimize network congestion IEEE P802.1CB [17] enables a conditional replication of packets according to destination address, traffic class and data routing information.



**Fig. 9.7.** Message replication of critical data in TSN (EN – end node; NN – network node; PATH\_1 – direct path; PATH\_2 – redundant path).

**Table 9.1.** IEEE TSN Standards.

IEEE Standard/Draft	Title	Main Purpose
<b>802.1AS</b>	Network Time Synchronization	Common clock for all network nodes
<b>802.1Qbv</b>	Scheduled Traffic	No collision of scheduled Ethernet frames
<b>802.1Qci</b>	Filtering & Policing	Network security
<b>P802.1CB</b>	Seamless Redundancy	Active switch-over with no data loss
<b>P802.1Qcc</b>	Stream Reservation	Centralized network configuration and management
<b>802.1Qbu / 802.3br</b>	Frame Pre-emption	Bandwidth optimization using framelets

## 9.5. Conclusions

IIoT and Industry are not a panacea for all the problems related with industrial networks. As a matter of fact, at the plant floor level there are challenges that don't have an easy solution, particularly when remote access of industrial data through the Internet can be risky in terms of confidentiality, security and safety. If by one hand, the spread of the data through the Internet promotes the access of information between different entities that are involved in a given manufacturing process, by the other hand, dangers associated with data security increase. It is also important to underline that there are costs associated with the introduction of IIoT and that the returns of investment are not achieved in the short term. In this context RT requirements of industrial data must be fulfilled not only in device level but, above all, in network level. The trends that are verified nowadays, and particularly will increase in a near future, are towards decentralization and spread of internet connected devices, even at the sensor level. The scenario of a global world in the field is also expected. By other hand, if nowadays, centralization of industrial production is verified in localized manufacturing units is a new different MU can operate collaboratively in different geographical locations. Regarding, unified RT networking goal the first implementation solution started a few years ago with Profinet, EtherCAT and Ethernet Powerlink. It started to be possible to include in the same network link deterministic and non-deterministic data. However, those solution are not interoperable and each one uses different definition in the higher layers of OSI model. Gateways and network separation, in terms of RT requirements, are still required. The solution to RT networking from the floor plant to the cloud is being normalized by IEEE, namely with the definition of standards to TSN that intends to achieve interoperability at all OSI layers with IT and OT devices from different manufacturers. Let wait for the next years to analyze the evolution in TSN and Internet (IPV6).

## References

- [1]. P. Simon, Too Big to Ignore: The Business Case for Big Data, *Wiley and SAS Business*, 2015.
- [2]. V. Granville, Developing Analytic Talent: Becoming a Data Scientist, 1<sup>st</sup> Ed., *Wiley and SAS Business*, 2016.
- [3]. Using Advanced Analytics to Predict Equipment Failure, <https://www.gartner.com/doc/3065317/using-advanced-analytics-predict-equipment>
- [4]. Channeling Streaming Data for Competitive Advantage, [https://www.sas.com/en\\_us/whitepapers/channeling-data-streams-107736.html](https://www.sas.com/en_us/whitepapers/channeling-data-streams-107736.html)
- [5]. The Forrester Wave: Big Data Streaming Analytics, <https://www.sap.com/documents/2016/06/92d1323c-797c-0010-82c7-eda71af511fa.html>
- [6]. P. Doyle, Introduction to real-time Ethernet I, *The Extension: A Technical Supplement to Control Network*, Vol. 5, Issue 3, 2004, pp. 1-4.
- [7]. P. Doyle, Introduction to real-time Ethernet I, *The Extension: A Technical Supplement to Control Network*, Vol. 5, Issue 4, 2004, pp. 1-4.
- [8]. N. Crockett, Introduction industrial Ethernet is ready to revolutionize the factory – Connecting the factory floor, *Manufacturing Engineer*, Vol. 82, Issue 3, 2003, pp. 41-44.
- [9]. PROFIBUS and PROFINET International, Specifications and Standards, <https://www.profibus.com/download/specifications-standards/>

- 
- [10]. Beckhoff EtherCAT, Application Notes, <http://www.beckhoff.com/english/default.htm?ethercat/default.htm>
  - [11]. ETHERNET Powerlink, Powerlink Technology, <https://www.ethernet-powerlink.org/powerlink/technology>
  - [12]. Clock Synchronization in Telecommunications via PTP, <http://ieeexplore.ieee.org/document/4319093/>
  - [13]. Std 802.1Q-2014 – IEEE Standard for Local and Metropolitan Area Networks – Bridges and Bridged Networks, <https://standards.ieee.org/findstds/standard/802.1Q-2014.html>
  - [14]. Std 802.1AS-2011 – IEEE Standard for Local and Metropolitan Area Networks – Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area, <https://standards.ieee.org/findstds/standard/802.1AS-2011.html>
  - [15]. Std 802.1Qbv-2015 – IEEE Standard for Local and metropolitan area networks – Bridges and Bridged Networks – Amendment 25: Enhancements for Scheduled Traffic, <https://standards.ieee.org/findstds/standard/802.1Qbv-2015.html>
  - [16]. Std. for Local and metropolitan area networks – Bridges and Bridged Networks – Amendment 28: Per-Stream Filtering and Policing, <http://ieeexplore.ieee.org/document/8064221/>
  - [17]. Std 802.1CB-2017 – IEEE Standard for Local and metropolitan area networks – Frame Replication and Elimination for Reliability, <https://standards.ieee.org/findstds/standard/802.1CB-2017.html>
  - [18]. Std 802.1Qcc – IEEE Draft Standard for Local and metropolitan area networks – Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks Amendment: Stream Reservation Protocol (SRP) Enhancements and Performance Improvements, <https://standards.ieee.org/develop/project/802.1Qcc.html>
  - [19]. Std 802.1Qbu-2016 – IEEE Standard for Local and metropolitan area networks – Bridges and Bridged Networks – Amendment 26: Frame Preemption, <https://standards.ieee.org/findstds/standard/802.1Qbu-2016.html>
  - [20]. Std 802.3br-2016 – IEEE Standard for Ethernet Amendment 5: Specification and Management Parameters for Interspersing Express Traffic, <https://standards.ieee.org/findstds/standard/802.3br-2016.html>
  - [21]. TSN – Time Sensitive Networking, *Belden Inc.*, 2016.
  - [22]. S. Craciunas, R. Olivier, Combined task- and network-level scheduling for distributed time-triggered systems, *Real-Time Systems*, Vol. 52, Issue 2, 2016, pp. 161-200.
  - [23]. T. Abdelzaher, K. Shin, Combined task and message scheduling in distributed real-time systems, *IEEE Transactions on Parallel and Distributed System*, Vol. 10, Issue 11, 1999, pp. 1179-1191.
  - [24]. S. Baruah, L. Rosier, R. Howell, Algorithms and complexity concerning the preemptive scheduling of periodic, real-time tasks on one processor, *Real-Time Systems*, Vol. 2, Issue 4, 1990, pp. 301-324.
  - [25]. O. Ronen, M. Lipinski, Enhanced synchronization accuracy in IEEE1588, in *Proceeding of the IEEE International Symposium on Precision Clock Synchronization for Measurement, Control, and Communication (ISPCS'15)*, Beijing, China, 2015, pp. 76-81.
  - [26]. S. Tongji, F. Luo, J. Xu, An analysis of frame replication and elimination for time-sensitive networking, in *Proceeding of the VI International Conference on Network, Communication and Computing (ICNCC'17)*, Kunming, China, 2017, pp. 166-170.

# **Chapter 10**

## **Towards a Digital Electronic Conditioning Approach Dedicated to High Sensitivity Giant Magnetoimpedance Sensor**

**Papa Silly Traore**

### **10.1. Introduction**

Magnetic sensors are widely used for measuring the magnetic field in a multitude of fields such as biomedical applications, non-destructive testing, space applications, military, etc. They are also widely used in many industrial fields for indirect measurement (for non-contact field measurement) of non-magnetic quantities such as electric current or linear and angular position. There is a wide variety of sensors with different physical principles. A non-exhaustive list may include, for example, Hall Effect sensors, Fluxgates, inductive sensors, Anisotropic MagnetoResistance (AMR), Giant MagnetoResistance (GMR) sensors, etc. [1].

Since more than two decades, the investigation of Giant Magneto-Impedance (GMI) effect has mobilized the efforts of many research groups around the world. This continuously growing interest in the phenomenon is obviously justified by the will to develop a new category of magnetic sensors with outstanding potential. Actually, these GMI sensors combine great benefits such as the high detection sensitivity and high resolution, the potential of miniaturization as well as the expected large bandwidth (from dc to at least several kilohertz).

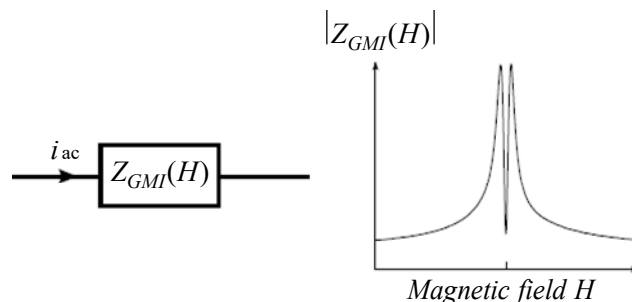
An important progress has so far been made in the theoretical understanding and formulation of the GMI effect and in the fabrication of materials. However, this new technology is still actually not very mature. Despite the great potential, several issues have to be overmastered so as to allow for a widespread use of these sensors in a large palette of industrial applications. In addition to the high sensitivity and high-resolution magnetic field measurements (for nondestructive testing for example), these industrial applications

may also include the use of the magnetic sensor for the indirect and contactless measurement of nonmagnetic physical quantities like electrical current, linear or angular positions, etc. The general objectives of the work in this field, both at the level of the materials and at the level of the physical parameters of implementation of the transducer, generally aim at a better understanding of the physical phenomena involved in order to understand their impact on performance and to decline avenues for optimization, particularly in terms of sensitivity and noise. From this point of view, it has been shown that the ultimate performance in noise have not yet reached and that those in the white noise region are mainly limited by the noise of the analog conditioning electronics of this sensor.

The present chapter mostly concerned the exploitation of GMI sensor for the realization of high sensitivity digital magnetic sensors. Section 10.2 describe the physical basis for modelling the GMI effect. For simplicity, we focus on single domain wires with uniform circumferential anisotropy, thus avoiding any difficulties associated with the details of the domain structure and domain-wall dynamics and of non-uniform anisotropy distribution. Section 10.3 is deal with the use of the GMI effect to realize a sensor. Section 10.4 is devoted to the high sensitivity digital GMI sensor. A detailed description of the different elements of the design is given so as to establish comprehensive and theoretical rules for this implementation. The resulting performances of the developed sensor are given together with discussions. The results mainly illustrate the potential of the design, in terms of sensor response, dynamic range and linearity, sensitivity and equivalent magnetic noise level.

## 10.2. Origin and Context of the Giant Magneto Impedance Effect

The GMI characterizes the impedance variation  $Z_{GMI}$  of a magnetic conductor subjected to a magnetic field  $H$  and traversed by an alternating excitation current  $i_{ac}$  of a few hundreds of kilohertz. By way of illustration, Fig. 10.1 shows a typical variation of impedance as a function of the applied magnetic field.



**Fig. 10.1.** Illustration of the GMI effect.

This is a long-known phenomenon. It was established as early as 1935 by E.P. Harrison et. al., who observed variations of the imaginary part of the impedance for frequencies of

a few tens of kilohertz on Nickel Iron wires [2, 3]. Forgotten, K. Mohri and L. V. Panina who working on ultra-soft amorphous ferromagnetic materials renamed the phenomenon, Giant Magneto-Impedance in 1994, because of the wide impedance dependence measured on amorphous wires based on Cobalt and FeCoSiB [4, 5]. This phenomenon has therefore continued to attract the interest of the scientific community for the development of sensors. These can be designed in different shapes and according to different geometries. One distinguishes the homogeneous structures integrating the ribbons (or thin layers), the wires (or microwires) then the non-homogeneous structures where one meets the sandwiches, the electro-deposited wires (or composite wires) and the configurations where the material is surrounded by a coil. A detailed synthesis of these different structures is carried out in reference [6]. Thus, the most convincing results had been measured on microwires. For example, the work of K. R. Pirota et al. showed that a variation of 600 % could be obtained at 15 MHz with a CoFeSiB microwire [7]. J. Velluer et. al. also presented a 700 % variation with NiFeMo microwires [8]. Promising to produce magnetometers, it has aroused widespread enthusiasm, particularly for their marketing. A number of sensors using microwires have already been marketed [9]. The work presented in this chapter will be mainly focused on amorphous ferromagnetic microwires. Most of the principles listed here can nevertheless be transposed to other geometric structures.

Generally, the complex electrical impedance of any dipole is defined as the ratio of the complex magnitudes of the voltage across it to the current flowing through it. The impedance  $\underline{Z}$  of the GMI wire can therefore be expressed by:

$$\underline{Z} = R + jX = \frac{\underline{V}_{ac}}{\underline{I}_{ac}}, \quad (10.1)$$

where R and X designate respectively the real and imaginary parts of the impedance;  $\underline{V}_{ac}$  and  $\underline{I}_{ac}$  represent the complex magnitudes of the voltage,  $v_{ac}$ , across the GMI device and the current,  $i_{ac}$ , flowing through it, respectively. The quantities to be handled in this chapter will be of a complex nature. For the sake of readability, their complex notations are implicit (for example, the complex magnitude of the impedance ( $\underline{Z}$ ) is denoted Z). The complex magnitudes of current and voltage are such that:

$$\underline{i}_{ac} = \underline{I}_{ac} \times e^{-j(2\pi f_p t)} \quad \text{where} \quad \underline{I}_{ac} = I_{ac} e^{-j\phi_1}, \quad (10.2)$$

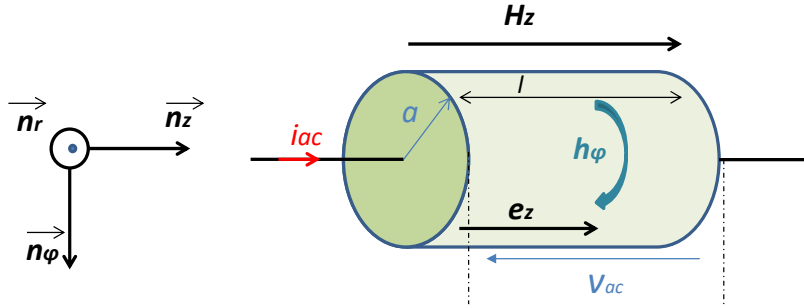
$$\underline{v}_{ac} = \underline{V}_{ac} \times e^{-j(2\pi f_p t)} \quad \text{where} \quad \underline{V}_{ac} = V_{ac} e^{-j\phi_2}, \quad (10.3)$$

where  $I_{ac}$  and  $V_{ac}$  are the current and voltage magnitudes of the GMI wire;  $\phi_1$  and  $\phi_2$  are the phases at the origin of the current and the voltage, respectively and  $f_p$  represents the frequency of the excitation current  $i_{ac}$ .

In GMI materials, the relation between  $v_{ac}$  and  $i_{ac}$  can be strongly nonlinear, it is nevertheless possible to approach them in the case of a linear approximation for low amplitudes of the  $I_{ac}$  current. Fig. 10.2 shows the geometry of a ferromagnetic wire (of radius  $a$  and of length  $l$  assumed to be very large compared to the radius) traversed by a



high frequency (HF) excitation current and subjected to an external axial magnetic field  $H_z$ .



**Fig. 10.2.** Representation of the GMI element geometry and the different quantities involved.

Thus, the excitation current  $i_{ac}$  of the material generates, according to the ampere theorem, a variable circumferential magnetic field  $h_\phi$ . The circulation of this magnetic field is at the origin of a variable flux creating an axial electric field  $e_z$  (Faraday's law) proportional to the voltage at the terminals of the GMI device according to the following relationships:

$$i_{ac} = 2\pi \cdot a \times h_\phi, \quad (10.4)$$

$$v_{ac} = l \times e_z \quad (10.5)$$

The study of the GMI effect is based on understanding the links between the electric and magnetic fields at the surface of the sample. These are given by the surface impedance tensor  $[\xi_s]$  [10] from a combination of Maxwell-Faraday, Maxwell-Ampère and Maxwell magnetic flux equations. The surface impedance tensor is given by:

$$\begin{pmatrix} e_\phi \\ e_z \end{pmatrix} = [\xi_s] \times \begin{pmatrix} h_z \\ h_\phi \end{pmatrix} \Rightarrow \begin{pmatrix} e_\phi \\ e_z \end{pmatrix} = \begin{pmatrix} \xi_{\phi\phi} & \xi_{\phi z} \\ \xi_{z\phi} & \xi_{zz} \end{pmatrix} \times \begin{pmatrix} h_z \\ h_\phi \end{pmatrix}, \quad (10.6)$$

where  $e_\phi$  and  $h_z$  denote respectively the circumferential electric field and the axial magnetic field at the surface of the GMI element.

Using equations (10.2), (10.3) and (10.4), the impedance of the material becomes:

$$Z = \frac{l}{2\pi a} \times \frac{e_z}{h_\phi} = \frac{l}{2\pi a} \left( \xi_{zz} + \xi_{z\phi} \frac{h_z}{h_\phi} \right) \quad (10.7)$$

This expression is reduced to the term  $\xi_{zz}$  if the component of the axial magnetic field  $h_z$  or the component  $\xi_{z\phi}$  is cancelled. It has also been shown that in the case of current

excitation and voltage detection at the terminals of the GMI device, the impedance is only a function of the term  $\xi_{zz}$  given by [10]:

$$\xi_{zz} = \Gamma \rho \frac{J_0(\Gamma a)}{J_1(\Gamma a)}, \quad (10.8)$$

where  $\Gamma = \frac{1-j}{\delta}$  and  $\delta$  represent respectively the wave vector linked to the radial propagation of the electromagnetic wave and the penetration depth of this wave in the material (cf. equation (10.10));  $\rho$  is the resistivity of the material;  $J_0$  and  $J_1$  denote the first-order Bessel functions. It is then possible to deduce the expression of the impedance  $Z$  of the material by:

$$Z = R_{dc} \times \frac{\Gamma a J_0(\Gamma a)}{2 J_1(\Gamma a)}, \quad (10.9)$$

where  $R_{dc} = \frac{\rho \times l}{\pi \times a^2}$  is the static resistance of the ferromagnetic conductor.

The GMI effect can then be likened to a classical phenomenon of electrodynamics called the skin effect provided that a relative effective permeability,  $\mu_{reff}$ , is introduced into the expression of the skin thickness of the current flowing through the conductor, such as:

$$\delta = \sqrt{\frac{\rho}{\pi \times f_p \times \mu_0 \times \mu_{reff}(H_z, i_{ac}, f_p)}} \quad (10.10)$$

The high relative permeability of soft ferromagnetic materials causes the appearance of the skin effect at lower frequencies than for a non-magnetic material of similar resistivity. Thus, the variation of impedance as a function of the field linked to the variation of the depth of penetration, depends on the frequency of excitation of the current [11, 12].

### 10.3. From Ferromagnetic Wire to Making a Sensor

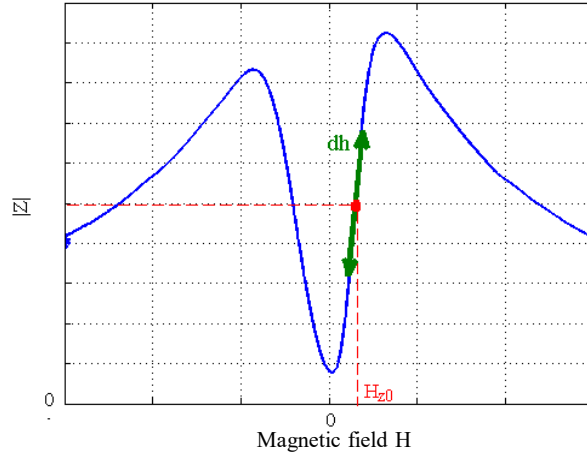
This section details the implementation of a GMI wire as a sensitive element of a sensor. Supplied by a HF excitation current  $I_{ac}$ , of frequency  $f_p$  then subjected to a magnetic field  $H_z$ , the impedance of the sensitive element can be expressed as:

$$Z(f_p, H_z) = |Z(f_p, H_z)| e^{j\phi(f_p, H_z)}, \quad (10.11)$$

where  $|Z|$  et  $\phi$  represent respectively the modulus and the phase of the impedance. They are a function of the frequency and the applied magnetic field.

### 10.3.1. GMI Sensor Conditioning Strategy

As shown in Fig. 10.3, the classical characteristic of the GMI effect is strongly nonlinear and does not allow the positive and negative fields to be distinguished. The linearity of the sensor is then obtained by fixing an operating point,  $H_{z0}$ , such that  $H_z = H_{z0} + dh$ . This operating point is generally chosen in the quasi-linear zone and presenting the maximum sensitivity.



**Fig. 10.3.** Illustration of the polarization of the GMI effect for sensor operation.

Subject to a variation  $dh$  of magnetic field around this operating point, the variation  $dZ$  of the impedance can be written as follows:

$$Z + dZ = \left| Z(f_p, H_{z0}) \right| e^{j\phi(f_p, H_{z0})} + d \left[ \left| Z(f_p, H_z) \right| e^{j\phi(f_p, H_z)} \right] \quad (10.12)$$

By developing the term related to the variation of impedance, then considering infinitesimal  $dh$ , the previous expression becomes [3]:

$$\begin{aligned} Z + dZ = & \left| Z(f_p, H_{z0}) \right| e^{j\phi(f_p, H_{z0})} + \frac{\partial \left| Z(f_p, H_z) \right|}{\partial H_z} \Bigg|_{H_z=H_{z0}} \times dh e^{j\phi(f_p, H_z)} + \\ & + j \frac{\partial \phi(f_p, H_z)}{\partial H_z} \Bigg|_{H_z=H_{z0}} \times dh \left[ \left| Z(f_p, H_z) \right| e^{j\phi(f_p, H_z)} \right] \end{aligned} \quad (10.13)$$

This last relation thus reveals the expressions of the variations of the modulus and the phase of the impedance. By expressing the impedance in the form of a module and a phase, we obtain:

$$Z + dZ = \sqrt{\left( \left| Z(f_p, H_{z0}) \right| + \frac{\partial |Z(f_p, H_z)|}{\partial H_z} \Big|_{H_z=H_{z0}} \times dh \right)^2 + \left( \left| Z(f_p, H_z) \right| \frac{\partial \phi(f_p, H_z)}{\partial H_z} \Big|_{H_z=H_{z0}} dh \right)^2} \times e^A, \quad (10.14)$$

where

$$A = j\phi(f_p, H_{z0}) + \arctan \left( \frac{\left| Z(f_p, H_z) \right| \frac{\partial \phi(f_p, H_z)}{\partial H_z} \Big|_{H_z=H_{z0}}}{\left| Z(f_p, H_z) \right| + \frac{\partial |Z(f_p, H_z)|}{\partial H_z} \Big|_{H_z=H_{z0}}} \right) \quad (10.15)$$

Most works on GMI sensors exploit amplitude modulation [13-16]. Phase modulation has so far been used very little [17-19]. In the work presented in this manuscript, only amplitude modulation is considered. Thus, by neglecting the phase modulation regarding the amplitude modulation, equation (10.14) becomes:

$$Z + dZ = \left( Z_0 + \frac{\partial |Z(f_p, H_z)|}{\partial H_z} \Big|_{H_z=H_{z0}} \times dh \right) \times e^{j\phi_0}, \quad (10.16)$$

where

$$Z_0 = |Z(f_p, H_{z0})|, \quad \phi_0 = \phi(f_p, H_{z0}) \quad (10.17)$$

The voltage appearing at the terminals of the GMI wire is thus given by:

$$v_{GMI}(t) = Z_0 \left( 1 + \frac{1}{Z_0} \times \frac{\partial |Z(f_p, H_z)|}{\partial H_z} \Big|_{H_z=H_{z0}} \times dh \right) \times I_{ac} \sin(2\pi f_p t + \phi_1) \quad (10.18)$$

This last relation shows the expression of a voltage modulated in amplitude by the magnetic field  $dh$  to be measured. The detection chain must therefore imperatively integrate a demodulator in order to obtain an image output voltage of the measured magnetic field such that:

$$v_s(t) = G_d \times I_{ac} \frac{\partial |Z(f_p, H_z)|}{\partial H_z} \Bigg|_{H_z=H_{z0}} \times dh, \quad (10.19)$$

where  $G_d$  is the gain of the demodulator.

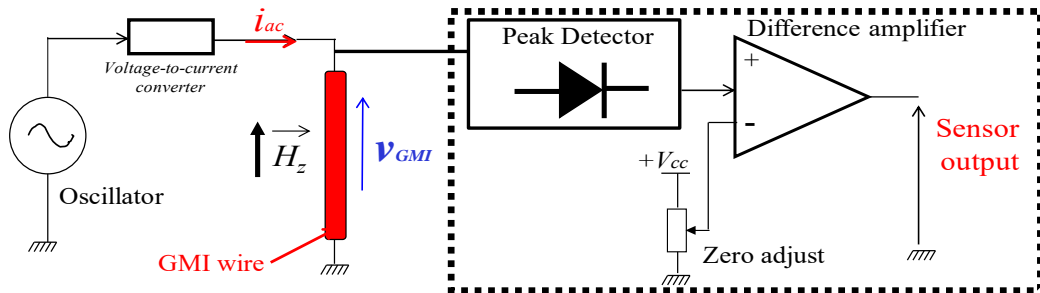
Thus, around the static component  $H_{z0}$  and considering a linear behavior, the impedance variation of the sensitive element is characterized by the voltage sensitivity in V/T. The latter can be evaluated to the nearest gain of the demodulator, by:

$$S_v = I_{ac} \frac{\partial |Z(f_p, H_z)|}{\mu_0 \partial H_z} \Bigg|_{H_z=H_{z0}} \times dh = I_{ac} \times S_\Omega, \quad (10.20)$$

where  $S_\Omega = \frac{\partial |Z(f_p, H_z)|}{\mu_0 \partial H_z}$  is the intrinsic sensitivity of the sensing element in  $\Omega/T$ .

As mentioned earlier, a DC magnetic bias must be applied to ensure linearity. This bias current, generally applied in the region exhibiting the maximum sensitivity, thus causes an intrinsic shift (offset) at zero magnetic field. To solve this problem, a differential amplifier to suppress the impedance value at null field is generally used [20, 21]. Moreover, this intrinsic offset could also be affected by a variety of influencing parameters such as temperature [22, 23] and stress [24, 25]. A regular adjustment of this offset is then often necessary.

In summary, the implementation of a GMI sensor requires a HF excitation system polarizing the sensitive element then a system of amplitude demodulation of the fundamental of the GMI signal in order to extract a voltage proportional to the magnetic field to be measured. Fig. 10.4 shows a typical configuration of a GMI sensor.



**Fig. 10.4.** Typical configuration of GMI sensor.

### 10.3.2. Noise Performance of a GMI Sensor

In applications requiring measurements of very low magnetic fields, an important feature, that characterizes the smallest magnetic field that a magnetic sensor can measure, is the equivalent magnetic noise level. Expressed in  $T/\sqrt{Hz}$ , it is defined by the ratio between the voltage noise spectral density, expressed in  $V/\sqrt{Hz}$ , and the sensor sensitivity, expressed in  $V/T$  according to (10.21)

$$b_{ec} = \sqrt{\frac{e_{n_{total}}^2}{S_v^2} + b_m^2}, \quad (10.21)$$

where

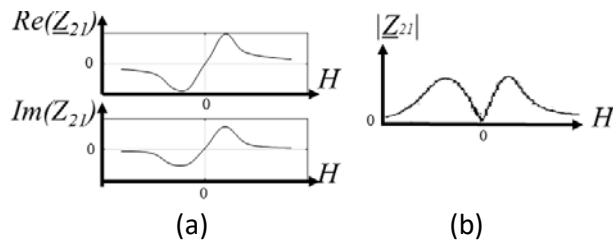
- $e_{n_{total}}$  designates the electronics conditioning noise spectral density expressed in  $V/\sqrt{Hz}$ ;
- $b_m$  represents the noise of the sensitive element in  $T/\sqrt{Hz}$ ;
- $S_v$  is the voltage sensitivity expressed in  $V/T$ .

In the context of high sensitivity applications, this equivalent magnetic noise level needs to be improved. This improvement could be theoretically done by decreasing to voltage noise spectral density of the sensor (including electronic conditioning noise and the intrinsic noise of the sensing element) while increasing the GMI sensor sensitivity.

To enhance the sensor sensitivity, a new magnetic configuration called off-diagonal configuration could be used. A promising implementation of the linear GMI sensors makes use of the off-diagonal component, noted  $Z_{21}$  of the impedance tensor. This component can be obtained using a pick-up coil wound around the GMI wire. A voltage,  $v_{coil}$ , is induced in this coil by the axial magnetization produced by a crossed-magnetization process [26, 27]. Both the real,  $Re\{Z_{21}\}$ , and imaginary,  $Im\{Z_{21}\}$ , parts of this impedance component are intrinsically almost linear and asymmetric (with odd symmetry) with respect to the zero-field point [28, 29] as seen in Fig. 10.5(a). The highest intrinsic sensitivity of this component is roughly obtained near the zero-field. On another hand, additional degree of flexibility is introduced for improving the voltage sensitivity (until some limits). This could actually be performed by increasing the number of turns of the pick-up coil [20]. Moreover, a few recent studies have shown that, in the case of amorphous wires with transverse anisotropy as the wire used in this study, the fluctuation of sensor response with the temperature change of the off-diagonal sensor were less important when compared to the classical diagonal implementation.

However, to take advantage of these important characteristics of the off-diagonal component, the sensitive element must be combined with adequate electronics for the amplitude demodulation. Most of the practical realizations of the off-diagonal sensor use a diode-based peak-detector since it is very easy to implement. However, such diode detectors suffer from the diode's nonlinearity and their temperature dependence. Most

importantly is the fact that a peak-detector does not allow to preserve the asymmetric characteristics of the off-diagonal component in the *final sensor output*. Actually, this *final sensor output* is still symmetric (even function) with respect to the zero-field since the peak-detector measures the absolute amplitude of the coil voltage which reflects only the modulus of the  $Z_{21}$  component (Fig. 10.5(b)). It does not allow distinguishing the positive and negative measured fields since the phase information of the coil voltage is lost. In other words, it is this phase information that allows obtaining the odd symmetry of the *final sensor response*. Indeed, when the measured field is reversed, the phase of the coil voltage is shifted by  $180^\circ$ . If the sensor is based on measuring the only amplitude of the voltage across the sensitive element, this phase reverse will not be preserved in the final sensor response. This is why, when using a peak-detector with the off-diagonal realization, a bias field and an offset canceling device, with additional noise contribution, are still required, at least in the implementation step, to obtain a linear sensor which has a *final output* of odd symmetry with respect the zero-field.



**Fig. 10.5.** Characteristic of the off-diagonal configuration.

On the other hand, advances in the area of noise analysis of GMI sensors have been made for different types of sensitive elements (wire, thin films, etc.). Several studies have shown that the voltage noise spectral density in the white noise region was mainly due to the analog conditioning electronics and not to the intrinsic noise of the sensing element [20]. Actually, the common point between all these GMI sensors was the use of an analog electronics to realize the functions involved in the conditioning of the sensor. In this case, the obtained results have shown that the equivalent magnetic noise level, in the white noise region, is limited by the noise from the analog electronics. In other words, the ultimate intrinsic noise performance of the GMI element has not been yet achieved. Consequently, it is worth exploring new conditioning techniques to reduce the noise contribution from the electronics and to advance towards reaching the ultimate noise performance.

#### 10.4. Hardware Architecture of a Digital off Diagonal GMI Sensor

In this section, a novel concept of digital electronic conditioning for the off-diagonal GMI sensors is presented. It is, for the first time in GMI sensors, implemented using Field-Programmable-Gate-Array (FPGA). The idea behind this concept is to perform the demodulation in digital domain. This is why the voltage across the pick-up coil is directly sampled by a high-speed Analog-to-Digital Converter (ADC). The digital processing data

is then achieved in real-time. The proposed design is based on the FPGA Zynq Z-7010 available with the commercial, low-cost and relatively new Red Pitaya board. Compared with traditional analog synchronous detection, the digital architecture has decisive advantages such as lower noise from the conditioning electronics, the ease-of implementation and the flexibility of reconfiguration and optimization with only software, as well as the potential for integration.

The principle of this new FPGA-based digital conditioning electronics is explained. A detailed description of the different elements of the design is given so as to establish comprehensive and theoretical rules for this implementation. The resulting performances of the developed sensor are given together with discussions. The results mainly illustrate the potential of the design, in terms of sensor response, dynamic range and linearity, sensitivity and equivalent magnetic noise level.

The hardware architecture of the off-diagonal GMI sensor using an FPGA implementation is shown in Fig. 10.6. The main component is the FPGA Zynq available on the Red Pitaya board. The Red Pitaya is a relatively newly available and innovative low-cost single-board with credit-card size. It also includes two Digital-to-Analog Converters (DAC) and two Analog-to-Digital Analog Converters (ADC). Only one ADC is shown in Fig. 10.6. The converters as well as the FPGA are operating in the proposed concept with a frequency reference clock  $f_s$  ( $f_s=125$  MHz).

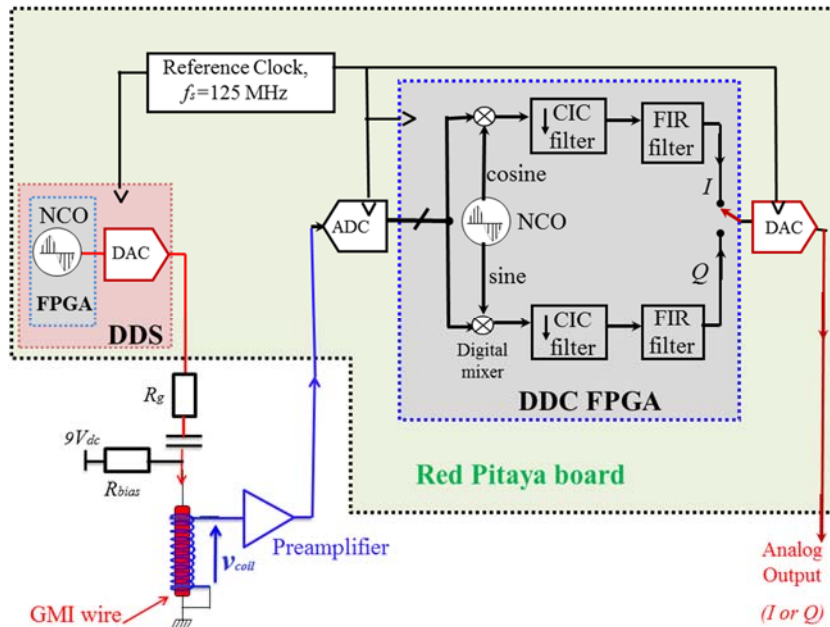


Fig. 10.6. Hardware architecture of the off-diagonal GMI sensors using an FPGA [32].

The FPGA is used to implement the digital functions like the Direct Digital synthesizer (DDS) and the Digital-Down-Converter (DDC) necessary for sensor implementation. The



implemented DDS, which is composed of a Numerically-Controlled Oscillator (NCO) associated to a DAC, supplies the high frequency voltage. This voltage is first converted into a current,  $i_{ac}$ , through the resistor  $R_g$ , before being supplied to the GMI sensitive element (cf Fig. 10.4). It is also to note that a  $dc$  current source,  $I_{dc}$ , formed by the  $dc$  voltage,  $V_{dc}$ , and the resistor  $R_{bias}$ , is supplied to the sensitive element. This  $dc$  current is necessary to enhance the off-diagonal response in amorphous wires having a circumferential anisotropy [15]. It also allows reducing the intrinsic magnetic noise of the wire in both diagonal and off-diagonal configurations [30].

When the sensing element is subjected to an external measured magnetic field, the voltage across the pick-up coil is amplitude-modulated by the measured magnetic field. This voltage is first amplified by a low-noise noninverting preamplifier based on the op-amp OP27. This op-amp has an input voltage noise spectral density of  $3\frac{nV}{\sqrt{Hz}}$  and an input current noise spectral density of  $0.4\frac{pA}{\sqrt{Hz}}$ . The preamplifier output is digitized by a 14-bit ADC at a sampling frequency  $f_s = 125$  MHz (corresponding to the frequency of reference clock). This ADC is the LTC2145 from Linear Technology [31]. It has signal-to-noise ratio (SNR) of about 73.1 dB. The digital data are then applied to the DDC implemented in the FPGA for digital quadrature demodulation. This demodulation includes a digital mixing of the data with the digital sine and cosine signals provided by a local NCO. After mixing; decimation and low-pass filtering are performed using a Cascaded-Integrator-Comb (CIC) followed by a Finite-Impulse Response (FIR) filter which outputs the ‘In-phase’ ( $I$ ) and ‘Quadrature’ ( $Q$ ) digital signals. If necessary, one or the other of these signals can be converted into the analog domain using the second DAC of the board.

#### 10.4.1. The FPGA Zynq SOC and Design Process

FPGA are configurable integrated circuits that can be used to design digital functions. One of the main features of FPGAs is their capacity for implementing configurable designs that combine the strengths of hardware and software. The configuration of a FPGA is generally specified using hardware description languages such as VHDL or Verilog. The FPGA structure is organized as an array of configurable logic blocks (CLB) and programmable routing resources used to provide the connectivity between the CLB, FPGA I/O pins and other resources such as on-chip memory and embedded hardware multipliers. Each CLB include look-up tables (LUTs), flip-flops, multiplexers and arithmetic carry logic. In the Red Pitaya board, the FPGA Zynq Z-7010 is used.

For GMI sensors application, the DDS and DDC are implemented in the FPGA using Vivado Design suite which ensures the enhancement of the overall productivity for designing, integrating, and implementing systems using Xilinx FPGA. This design suite also includes the Integrator Property (IP) catalog for creating complex system. The design presented here is created by the VHDL language and the Tool Command Language (TCL) to link the IP.

### 10.4.2. FPGA Implementation of the Direct Digital Synthesizer (DDS)

The implemented DDS is based on the Xilinx logiCORE IP DDS compiler [32]. As it can be seen in Fig. 10.7, the core (or the NCO) of the DDS is mainly composed of a phase accumulator, a phase-to-amplitude converter (generally a sine look-up table) as well as of a frequency register. The desired frequency,  $f_{DDS}$  (expressed in Hz), at the output of the NCO is obtained by [32]:

$$f_{DDS} = \frac{f_s * FTW_{DDS}}{2^N}, \quad (10.22)$$

where:

$f_s$  is the frequency of the reference clock (125 MHz in our case);

$FTW_{DDS}$  is the Frequency Tuning Word (a binary number which is written in the frequency register);

$N$  is the number of bits of the phase accumulator (12 bits in the proposed design).

To generate sine wave of a fixed frequency  $f_{DDS}$ , the binary number ( $FTW_{DDS}$  in the frequency register) is inputted to the phase accumulator. This binary number corresponds to a constant phase increment. At each clock cycle, this phase increment is added to the phase accumulator which computes a phase argument for the phase-to-amplitude converter. This converter provides the digital value of amplitude that corresponds to the sine of the current phase argument. The digital value of amplitude is then converted into the analog domain using the first DAC clocked at frequency  $f_s$ , and available on the Red Pitaya board. The analog output voltage is then used for the current excitation of the GMI wire.

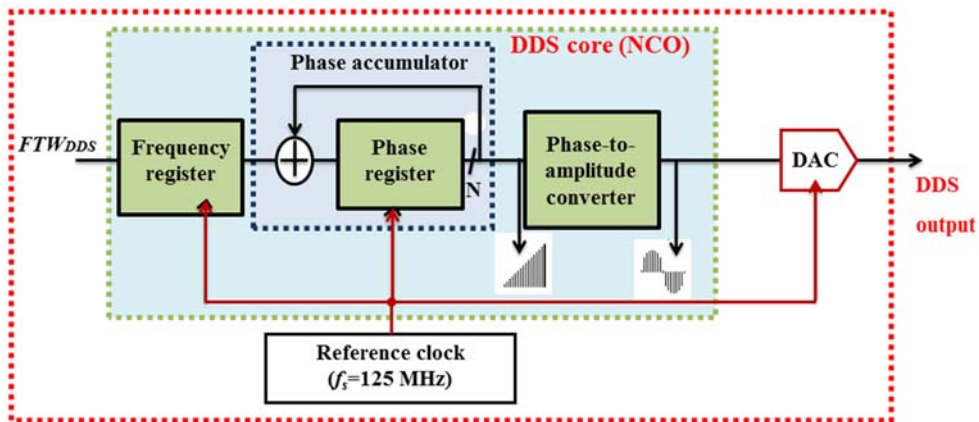


Fig. 10.7. Block diagram of the implementation of the DDS using the FPGA [32].

Thanks to the high degree of flexibility, the frequency, phase and amplitude of the DDS are fully and digitally programmable. This is a great advantage for the GMI sensors since

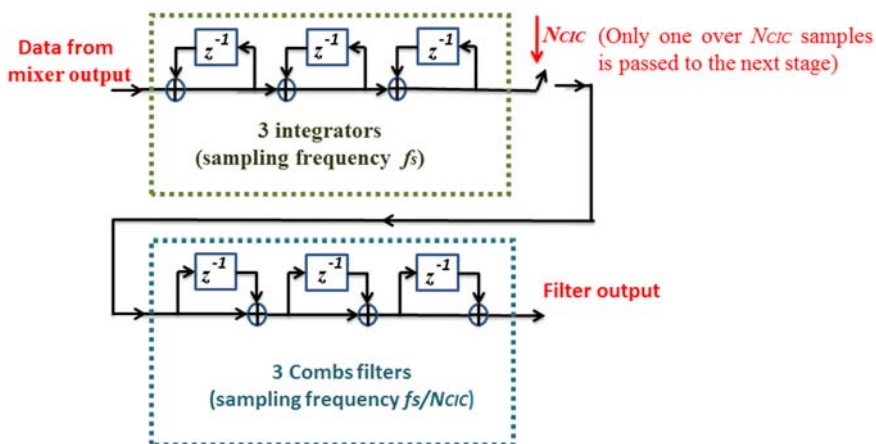
the GMI characteristics depends on a large number of parameters such as the frequency and the amplitude of the excitation current. Therefore, the programmability features of the DDS allow implementing real-time procedures for quick optimization of the GMI response for a given sensitive element.

### 10.4.3. Implementation of the Digital-Down-Converter (DDC)

This DDC is mainly composed of Numerically-Controlled Oscillator (NCO), a Cascaded-Integrator-Comb (CIC) decimation filter and a Finite Impulse Response (FIR) low-pass filter. The quadrature demodulation is performed by multiplying (using the digital mixers) the digital data from the ADC with the sine and cosine digital signals of the local NCO. In the same way as for the DDS, the implementation of this NCO was realized using the Xilinx logiCORE IP DDS compiler. The frequency, the amplitude and the phase are programmable with high flexibility and resolution.

The ease of programming the phase of the NCO with high resolution is an important feature when considering synchronous detection. Actually, accurate adjustment between the phase of the amplitude modulated (AM) input signal from the sensor and the phase of NCO is mandatory in order to increase the final sensor output. This high resolution (less than 0.001 degrees in the current design) of phase programming is straightforward to achieve and easy to implement in this digital design.

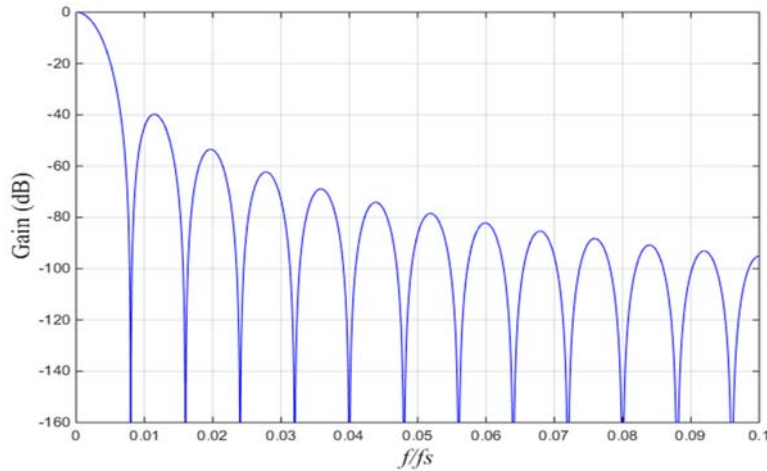
The outputs of the digital mixers are applied to the input of a CIC decimation filter. Its name is derived from its structure which consists of an equal number of integrator sections operating at a high sampling frequency,  $f_s$ , combined with a comb section operating at a low sampling rate  $f_s/N_{CIC}$ , where  $N_{CIC}$  is the decimation rate. In the current design, a third-order CIC decimation filter with a programmable decimation rate,  $N_{CIC}$ , is used. This CIC decimation filter is implemented using LogiCORE IP CIC compiler [32]. Fig. 10.8 shows the block diagram of its implementation.



**Fig. 10.8.** Block diagram of the 3<sup>rd</sup> order CIC decimation filter.  $z^{-1}$  is the usual delay element for a digital filter [32].

One of the advantages of CIC filters is that their implementation requires only integrators and Combs. Since no multipliers are needed, the real-time implementation is simplified and the calculation speed is enhanced. These important advantages make them very used in applications that require high decimation rates like telecommunications applications.

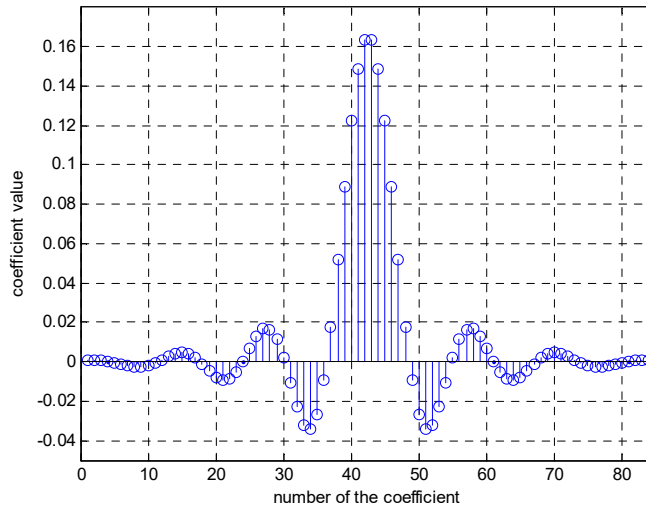
Fig. 10.9 shows a typical frequency response of a 3<sup>rd</sup> order CIC filter with a decimation rate of  $N_{CIC} = 125$ . This typical response clearly shows that this category of filters does not have a flat response in the passband. Moreover; they can introduce some aliasing in the passband because of their attenuation which is not high enough in the passband. These two main drawbacks are undesirable in many applications such as the GMI sensors. This problem can be addressed using a compensation FIR filter.



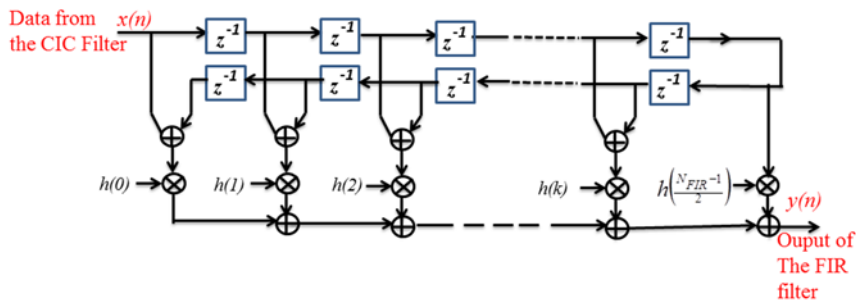
**Fig. 10.9.** Frequency response of a 3<sup>rd</sup> order CIC filter with a decimation rate of  $N_{CIC}=125$ . The x-axis is the normalized frequency (with respect to the sampling frequency  $f_s$ . Only a part of the Nyquist bandwidth  $[0, f/f_s = 0.5]$  is represented where  $f$  is the frequency expressed in Hz [32].

Thereby, the output data of the CIC filter are applied to the input of a FIR filter in order to compensate for this drop of the CIC filter and to improve the flatness of the final bandpass. This compensation filter, which is the last digital stage of the DDC, is based on LogiCORE IP FIR compiler [32]. It was implemented using  $N_{FIR}$  programmable coefficients,  $h(k)$ , where  $k$  is the number of the coefficient. Fig. 10.10 shows an example of the impulse response of the filter obtained with  $N_{FIR} = 83$ .

As it can be seen in this figure, the filter was synthesized to obtain a symmetric impulse response (with an odd number of coefficients). In this way, only  $\left(\frac{N_{FIR} + 1}{2}\right)$  coefficients are actually loaded in the DDC to optimize the computation time and the storage capacity of the hardware. Fig. 10.11 shows the real-time implementation of this symmetrical impulse response for  $N_{FIR} = 83$ .



**Fig. 10.10.** The Impulse Response of the FIR filter with a number of coefficients  $N_{FIR} = 83$ .



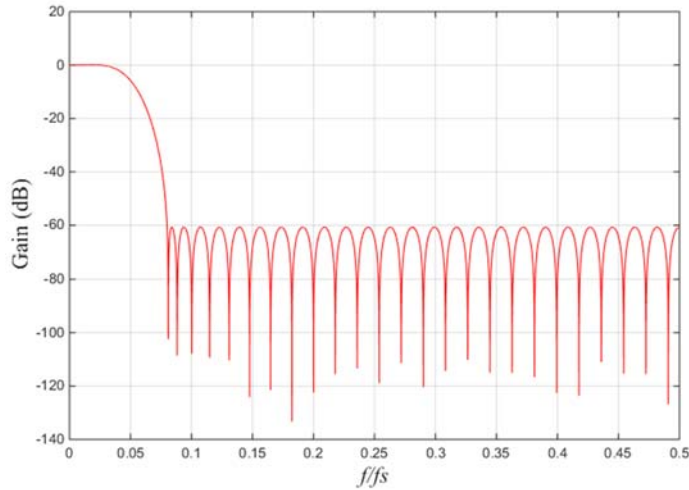
**Fig. 10.11.** FIR filter implementation with symmetrical coefficients for an odd number,  $N_{FIR}$ , of filter taps ( $N_{FIR} = 83$ ).  $z^{-1}$  is the usual delay element for a digital filter.

Compared to a classical structure of a FIR filter that needs  $N_{FIR}$  multiplications and  $(N_{FIR}-1)$  additions, the implementation of symmetrical impulse response requires  $\left(\frac{N_{FIR}+1}{2}\right)$  multiplications and approximately  $N_{FIR}$  additions. This reduction in the multiplications number allow to develop fast and efficient filtering methods. It also presents a good choice for the realization of high sensitivity GMI sensors especially for the reduction of noise due to the processing chain [33].

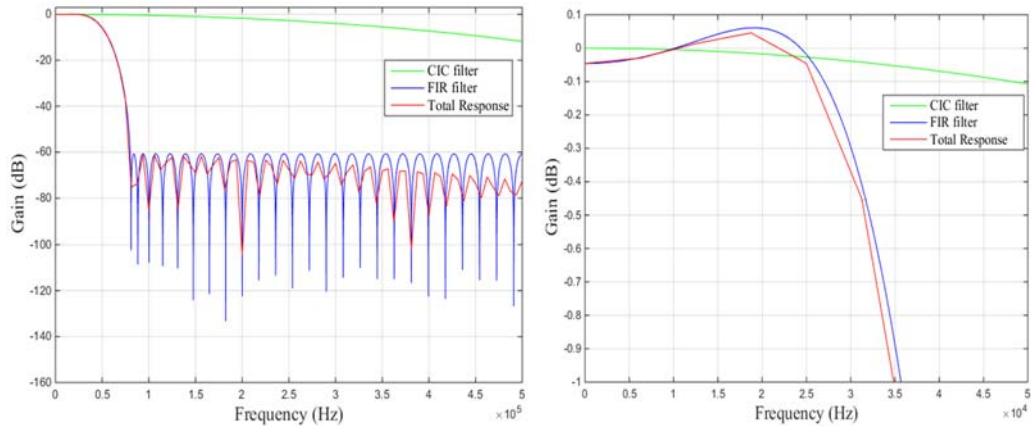
The frequency response of this filter is illustrated in Fig. 10.12. It can be seen that it is almost flat in the filter bandwidth. Furthermore, the out-of-band-pass rejection of the filter is more than 60 dB.

The final frequency response of the cascade of the CIC and the FIR filters is shown in Fig. 10.13. It was obtained with a decimation rate of  $N_{CIC}$  of 125 for the CIC filter. No decimation was performed by the compensation FIR filter. A zoom in the low frequency

band that illustrates the compensation process is also shown (right figure). It can be clearly seen that the final frequency response is mainly determined by the FIR filter.



**Fig. 10.12.** The frequency response of the FIR filter. The x-axis is the normalized frequency  $f/f_s$ .



**Fig. 10.13.** Frequency responses of the CIC (with  $N_{CIC} = 125$ ) and the FIR filters, together with the total response. The upper figure shows the total frequency response over the entire final Nyquist band of  $f_s/(2N_{CIC}) = 500$  kHz. The lower figure shows a zoom to illustrate the FIR filter compensation of the slight droop due the CIC filter [32].

Furthermore, it is worth mentioning that, in addition to the possibility to easily program the final bandwidth of the sensor by adjusting the decimation rate, the digital concept using decimation filter allows also enhancing the SNR of the ADC. The operation is known as Processing Gain ( $PG$ ). For example, for an initial sampling frequency,  $f_s = 125$  MHz, and a final sampling frequency,  $f_s/N_{CIC}$  after decimation, the SNR improvement, or  $PG$ , of the ADC is given by:

$$PG = 10 \times \log \left( \frac{f_s}{\left( \frac{f_s}{N_{CIC}} \right)} \right) = 10 \times \log(N_{CIC}) \quad (10.23)$$

For example, a decimation factor  $N_{CIC} = 125$  yields a processing gain  $PG \approx 21$  dB. In this case, the final SNR of the ADC in the band of  $f_s/N_{CIC}$  is expressed as:

$$SNR = 73.1 + 21 = 94.1 \text{ dB} \quad (10.24)$$

This theoretical value of SNR will however not be achieved in the practice since these equations take only into account the noise from the ADC (quantization noise, thermal noise, etc.). If others sources of noise, like the noise from the sensitive element and the preamplifier, are more important than the noises from the ADC, the processing gain will not be efficient. However, some improvement in the SNR could be realized thanks to the effective rejection of the out-of-band noise achieved by the proper design of the digital filters.

#### 10.4.4. Results and Discussions

A complete design of the digital sensor has been realized. This sensor includes a GMI sensing element based on a CoFeSiB wire (from Unitika, Ltd) of 25 mm long and 100  $\mu\text{m}$  diameter. It is a sample with high permeability, transverse anisotropy, nearly zero magnetostriction and very low coercive force,  $H_c = 0.6$  Oe.

The residual magnetic flux density is about  $B_r = 4$  kG, and saturation magnetic flux density is  $B_s = 8$  kG. A 400-turn pick-up coil was wound around the GMI wire. This sensitive element was associated with the implemented digital conditioning electronics described in Section 10.3 as shown in Fig. 10.14. In this illustration, only the Red Pitaya board (in which are implemented the digital conditioning electronics) and the sensing element are shown.

Fig. 10.15 shows an example of the “ $P$ ” and “ $Q$ ” outputs of the DDC as a function of the applied magnetic field. The frequency of the used ac current,  $i_{ac}$ , in the GMI wire was of 1 MHz and its RMS value about 5 mA<sub>rms</sub>. The dc current,  $I_{dc}$ , was about 5 mA. The decimation ratio of the CIC filter was  $N_{CIC} = 125$ .

Unlike the frequently used peak-detectors in GMI sensors, the use of the digital quadrature demodulation (DDC concept), which is easily implemented using the FPGA, clearly allows preserving the intrinsic interesting features of the off-diagonal GMI component in the final sensor response. As it was mentioned in the introduction of the paper (Section 10.2), these features are mainly the linearity and the asymmetric behavior (an odd function) with respect to the zero-field. In the example of the sensor response of Fig. 10.15, the error of linearity was about 0.055 % of a full scale (FS) of  $\pm 61$   $\mu\text{T}$  of the measured field. Furthermore, it is observed that the positive maximum of the response is obtained at a field of about 100 A/m while the negative maximum is at about 70 A/m for

the studied wire. Most probably, the difference could be attributed to the presence a finite remanence at zero magnetic fields in amorphous CoFeSiB wires produced by in-rotating-water spinning process as it is the case of the used wire.

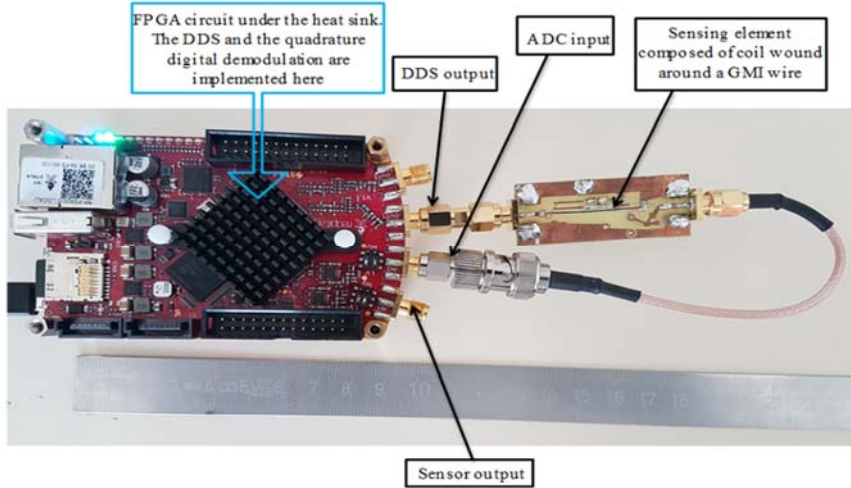


Fig. 10.14. Illustration of the measurement chain. The amplifier OP27 is not represented [32].

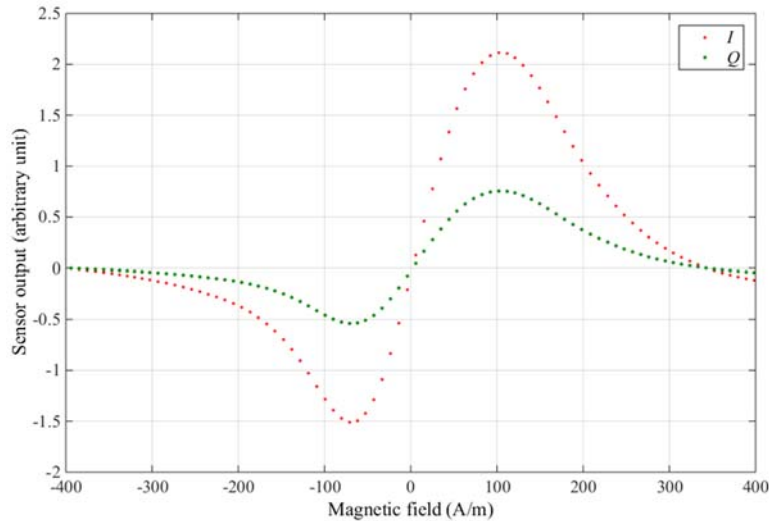


Fig. 10.15. Example of the  $I$  and  $Q$  outputs of the FPGA-based off-diagonal GMI sensor [32].

This combination of the off-diagonal configuration and the digital quadrature demodulation allows using the sensor without the need of a bias magnetic field, and by far, without using any more electronics for an offset canceling. Moreover, the quadrature demodulation guarantees obtaining the phase information easily from the combination of  $I$  and  $Q$  components, while this same information is actually lost when using a



peak-detector. Some investigations to explore the possibility to realize GMI sensors based on the phase information, rather than on the amplitude information, have already been conducted [17, 18]. The implementation of the analog electronics for phase measurement is by far not a trivial task. The reconfigurable FPGA design of conditioning electronics of the GMI sensors could therefore exhibit great advantages since it is straightforward to use it for having the two kinds of information without any hardware modification.

From another point of view, in general, when using a synchronous detector for amplitude demodulation (AM), the adjustment of the phase difference between the AM input and the local oscillator is necessary in order to maximize the output of the detector. This is very important in the case of the GMI sensors for optimizing and increasing the sensitivity. This is easily performed with high accuracy and resolution in the FPGA design through a digital adjustment of the phase of the NCO signals of the DDC (the phase can be adjusted with a resolution of less than 0.001 degree in the current design).

While this digital implementation presents this set of advantages, it is still necessary to evaluate its expected performance in terms of equivalent magnetic noise, noted here  $e_{mn}$ . This quantity, expressed in  $T/\sqrt{Hz}$ , determines the ultimate detection level of a given magnetic sensor. It is defined by:

$$e_{mn} = \frac{e_n}{T_r}, \quad (10.24)$$

where  $e_n$  is the voltage noise spectral density (VNSD) of the sensor expressed in  $V/\sqrt{Hz}$ , and, and  $T_r$  is the sensitivity (or the transfer function) of the sensor in open-loop. It is expressed in  $V/T$ .

In parallel to the design of the new digital electronics, a first analytical model has been developed in order to evaluate the expected noise performance of the whole sensor in the white noise regime only. This model takes into account all the white noise from the sensitive element and noise from the electronics. It has been demonstrated that the VNSD was given by [33-37]:

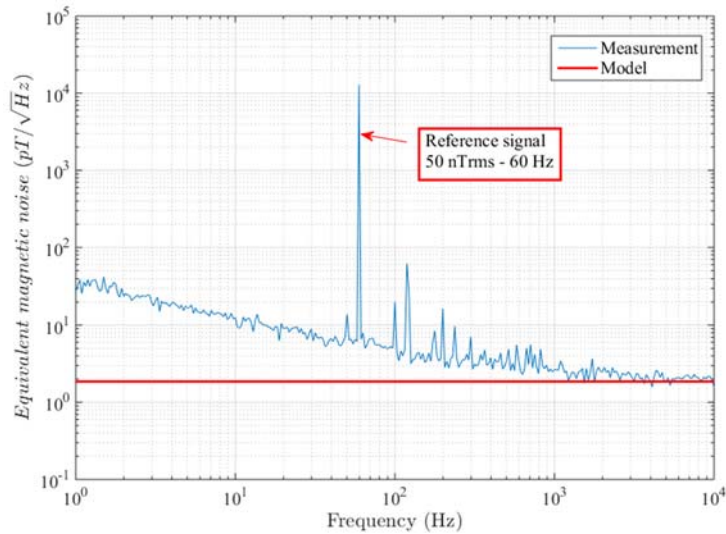
$$e_n = \sqrt{\left[ \left( e_{n_{exc}}^2 + e_{n_{off}}^2 + e_{n_{mag}}^2 + e_{n_{OP27}}^2 \right) G_{OP27}^2 + e_{n_{ADC}}^2 \right] G_{DDC}^2 + e_{n_{DDC}}^2} \quad (10.25)$$

where:

- $e_{n_{exc}}$  is the VNSD of the excitation part (DDS and the two voltage-to-current converters formed by  $R_g$  and  $R_{bias}$ );
- $e_{n_{off}}$  and  $e_{n_{mag}}$  are respectively the VNSD related to the off-diagonal configuration and the magnetic noise (white noise);
- $e_{n_{OP27}}$ ,  $e_{n_{ADC}}$ ,  $e_{n_{DDC}}$  are the VNSD of the preamplifier, the ADC and the DDC, respectively;

- $G_{OP27}$  and  $G_{DDC}$  are the gain of the preamplifier and digital gain of the DDC, respectively.

This model was validated by the measurement. The experimental conditions were: amplitude of  $i_{ac} = 5 \text{ mA}_{\text{rms}}$ , excitation frequency = 1 MHz,  $I_{dc} = 5 \text{ mA}$ ,  $G_{OP27} = 1$ ,  $G_{DDC} = 40$ ,  $N_{CIC} = 125$ . The final sensor bandwidth at -3 dB was adjusted to about 43 kHz. The voltage VNSD was measured. During this measurement, a reference *ac* magnetic field (50 nT<sub>rms</sub>, 60 Hz) was applied. Only the sensitive element and the source of the magnetic field were placed in a shielding enclosure. The measured VNSD,  $e_n$ , was then divided by the sensitivity (or the transfer function,  $T_r$ ) of the whole sensor. This sensitivity was experimentally evaluated to be about 660 kV/T. With this value, the obtained equivalent magnetic noise,  $e_{mn}$ , is shown in Fig. 10.16. In this same figure, the equivalent magnetic noise predicted by the model is also given.



**Fig. 10.16.** Measured and predicted equivalent magnetic noise of the FPGA-based off-diagonal GMI sensor in open-loop. The setup parameters were: excitation current: 1 MHz/5 mA<sub>rms</sub>,  $I_{dc} = 5 \text{ mA}$ ,  $G_{OP27} = 1$ ,  $G_{DDC} = 40$ ,  $N_{CIC} = 125$ . The final sensor bandwidth at -3 dB was adjusted to about 43 kHz [32].

Performance of 1.8 pT/√Hz is obtained in the white noise region (from 1 to 10 kHz) by the measurement and by the model. This noise level is promising and well-comparable with the levels achieved in other works where 1 pT/√Hz to several pT/√Hz are, by far, usually obtained depending on a variety of conditions like the excitation parameters, the type and length of the sensitive element as well as the kind of analog electronic conditioning.

These preliminary results, in terms of noise performance, show that at least the advantages of the digital design explained throughout the paper are well-guaranteed while preserving a promising noise level.

A main limitation of the current design is the value of the sensitivity,  $T_r$ , (in V/T) that could be achieved. This value is limited to a maximum allowed by the input dynamic range of ADC (only 2 V peak-to-peak). This limits, for a given VNSD,  $e_n$ , the equivalent magnetic noise level  $e_{mn}$ . In analog conditioning, based on a peak-detector for example, the achieved dynamic, and consequently the sensitivity, could be by far several times higher than 2 V. If with the digital concept, the obtained equivalent magnetic noise  $e_{mn}$  is still comparable with the one obtained with analog systems, one might argue that the total VNSD of the whole digital electronic is well lower than the one that could have obtained with analog electronic conditioning. This is interesting since it opens doors for further improvements of the digital systems by using for example ADC with higher dynamic and higher resolution in order to obtain better equivalent magnetic noise level than the levels presented in this paper.

Furthermore, a noise level of  $33 \text{ pT}/\sqrt{\text{Hz}}$  at 1 Hz was obtained by the measurement. It is to note here that, at this step, the theoretical model does not allow predicting the noise performance at low-frequency. This noise, known as  $1/f$  noise could be attributed to the intrinsic magnetic noise from the sensitive element [38, 39].

From another point of view of performance, it is to note that the length of the sensitive element will obviously determine the spatial resolution of the sensor. Typically, wires of one centimeter to a few centimeters are generally used in a number of works. The use of shorter sensitive element allows a better spatial resolution. However, in this case, the intrinsic sensitivity  $T_r$  is decreased. At fixed voltage noise spectral density (VNSD),  $e_n$ , the decrease of  $T_r$  results in an increase of the equivalent magnetic noise according to equation (10.24). However, for the digital design, similar performance in terms of this equivalent magnetic noise should be obtained with shorter wires. Actually, the loss of sensitivity  $T_r$  in this case could be recovered by increasing the gain of the preamplifier (which is currently used in unit-gain configuration for impedance matching purposes). The use of higher gain for the preamplifier will not degrade voltage noise spectral density (VNSD),  $e_n$ , since the dominant noise source is the ADC.

The main characteristics of the developed off-diagonal GMI sensors using FPGA are summarized in Table 10.1.

**Table 10.1.** Performance of the FPGA-based off-diagonal GMI sensor in open loop [32].

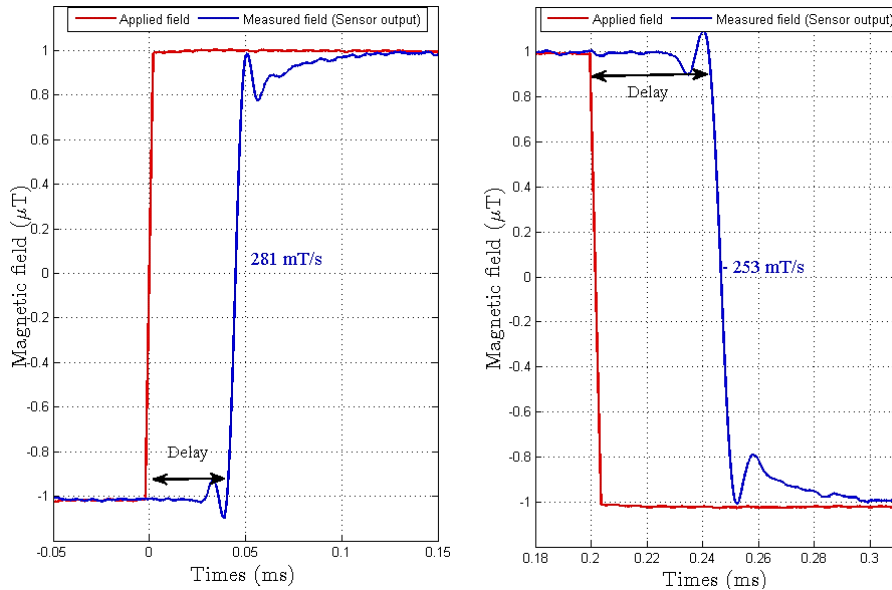
Sensitivity ( $T_r$ )	660 kV/T
Maximum Digital excursion	$\pm 61 \mu\text{T}$
Linearity	0.055 % FS (Full Scale)
Bandwidth	DC to 43 kHz
Slew rate	$\uparrow$ 281 mT/s   $\downarrow$ -253 mT/s
White noise level	$1.8 \text{ pT}/\sqrt{\text{Hz}}$
Noise level at 10 Hz	$11 \text{ pT}/\sqrt{\text{Hz}}$
Noise level at 1 Hz	$33 \text{ pT}/\sqrt{\text{Hz}}$

It is to note that the slew rate of the sensor was also measured. This slew rate is defined as the maximum change of an applied measured magnetic field (by second) that the sensor can reproduce at his output. It is determined by measuring the output in response to a fast modification of the magnetic field as shown in Fig. 10.17. In our case, this measured slew rate was about  $281 \text{ mT/s}$  and  $-253 \text{ mT/s}$  for a rising field and a falling field respectively.

The Fig. 10.17 also shows a time-delay (about  $45 \mu\text{s}$  in this example) between the applied magnetic field and the sensor output. This group delay is inherent to any digital processing system. If the GMI sensor is intended to operate in closed-loop, precautions should be taken to avoid the instability of the closed-loop. The delay is mainly introduced by the CIC filter and the FIR compensation filter. In our design, this delay is a function of the order and the decimation rate of the CIC filter, the number of filter coefficients of the FIR filter and the frequency reference clock  $f_s$  according to the equation:

$$\tau = \frac{N_{CIC}}{f_s} \left( 3 + \frac{N_{FIR} + 1}{2} \right) + \frac{3}{f_s} \quad (10.26)$$

For  $f_s = 125 \text{ MHz}$ ,  $N_{CIC} = 125$  and  $N_{FIR} = 83$ , the calculated value of the delay is  $\tau \approx 44 \mu\text{s}$ , which is in agreement with the measurement in Fig. 10.17.



**Fig. 10.17.** Illustration of the delay of the digital system and of the slew rate of the FPGA-Based off-diagonal GMI sensor. The figure shows the responses of the sensor due to rising (left figure) and falling (right figure) magnetic field variations [32].

Finally, the obtained, dynamic range is comparable (roughly in the same order on magnitude) with several reported realizations of off-diagonal GMI sensors [13, 20]. This dynamic range as well as linearity could also be improved using a closed-loop. In this

context, future works should be focalized on the study of the impact of the delay (Fig. 10.17) on the stability of the closed-loop of the sensor. The final underlying idea is to proceed towards a fully digital high sensitivity magnetometer (closed-loop sensors) [40, 41].

## 10.5. Conclusions

We present in this chapter a digital electronic conditioning approach in order to improve the equivalent magnetic noise of a GMI sensor. The proposed approach is based on an FPGA available with the low-cost and small size Red Pitaya board. The design is reconfigurable with high flexibility. The excitation system of the sensitive element was based on the use of a DDS. The amplitude demodulation of the GMI was performed in the digital domain. This digital demodulation was easily implemented based on a synchronous quadrature detection that allows to preserve the intrinsic characteristic of the off-diagonal configuration (odd characteristic and an almost linear response near the zero-field, without the use of bias magnetic field and without additional offset canceling devices). The performance of the sensor in terms of equivalent magnetic noise was presented and discussed. Obtained results is promising when compared to the best reported in the state of the art of GMI sensors. This leaves possibilities for further improvement of these performances since the results suggested that these digital electronics had by far lower noise than the frequently used analog electronics.

## References

- [1]. P. Ripka, *Magnetic Sensors and Magnetometers*, Artech House Publisher, 2001.
- [2]. E. P. Harrison, G. L. Turney, H. Rowe, Electrical properties of wires of high permeability, *Nature*, Vol. 135, 1935, 961.
- [3]. E. P. Harrison, G. L. Turney, H. Rowe, H. Gollop, The electrical properties of high permeability wires carrying alternating current, *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, Vol. 157, 1936, pp. 451-479.
- [4]. L. Panina, K. Mohri, Magneto-impedance effect in amorphous wires, *Applied Physics Letters*, Vol. 65, Issue 9, 1994, pp. 1189-1191.
- [5]. K. V. Rao, F. B. Humphrey, J. L. Costa-Kramer, Very large magnetoimpedance in amorphous soft ferromagnetic wires, *Journal of Applied Physics*, Vol. 76, Issue 10, 1994, pp. 6204-6208.
- [6]. B. Miloud, Contribution à l'étude théorique de l'effet de la magnétoimpédance (GMI), Mémoire de Magister en Electrotechnique (MD Thesis), *Université El-Hadj Lakhdarbatna*, Soutenue le 29/04/2010.
- [7]. K. R. Pirota, L. Kraus, H. Chiriac, M. Knobel, Magnetic properties and giant magnetoimpedance in a CoFeSiB glass-covered microwire, *Journal of Magnetism and Magnetic Materials*, Vol. 221, 2000, pp. L243-L247.
- [8]. J. Velleuer, A. G. Munoz, H. Yakabchuk, C. Schiefer, A. Hackl, E. Kisker, Giant magneto impedance in electroplated NiFeMo/Cu microwires, *Journal of Magnetism and Magnetic Materials*, Vol. 311, 2007, pp. 651-657.
- [9]. L. Landau, E. Lifshitz, *Electrodynamics of Continuous Media*, Pergamon Press, Oxford, 1960.
- [10]. M. Knobel, K. Pirota, Giant magnetoimpedance: concepts and recent progress, *Journal of Magnetism and Magnetic Materials*, Vol. 242, 2002, pp. 33-40.

- [11]. M. Knobel, M. Vázquez, L. Kraus, Giant magnetoimpedance, *Handbook of Magnetic Materials*, Vol. 15, 2003, pp. 497-563.
- [12]. L. Ding, Étude et amélioration des performances en bruit de capteurs à MagnétoImpédance Géante. Application à la magnétométrie à très haute sensibilité, Thèse Doctorat (PhD Thesis), *Université de Caen*, 2008.
- [13]. B. Dufay, Optimisation d'un magnétomètre haute sensibilité à base de MagnétoImpédance Géante (MIG), Thèse Doctorat (PhD Thesis), *Université de Caen*, 2012.
- [14]. Geliang Yu, Xiongzhu Bu, Bo Yang, Yunlong Li, and Chao Xiang, Differential-Type GMI Magnetic Sensor Based on Longitudinal Excitation, *IEEE Transactions on Magnetics*, Vol. 11, No. 10, Oct. 2011, pp. 2273-2278.
- [15]. T. Uchiyama, K. Mohri, Fellow, L. V. Panina, K. Furuno, Magneto Impedance in sputtered amorphous films for micro magnetic sensor, *IEEE Transactions on Magnetics*, Vol. 31, Issue 6, November 1995, pp. 3182-3184.
- [16]. S. Yabukami, H. Mawatari, Y. Murayama, T. Ozawa, K. Ishiyama, K. I. Arai, High frequency carrier type thin-film sensor using low noise crystal oscillator, *IEEE Transactions on Magnetics*, Vol. 40, Issue 4, July 2004, pp. 2670-2672.
- [17]. E. C. Silva, L. A. P. Gusmão, C. R. H. Barbosa, E. C. Monteiro, Electronic approach for enhancing impedance phase sensitivity of GMI magnetic sensors, *Electronics Letters*, Vol. 49, Issue 6, 14 March 2013, pp. 396-397.
- [18]. S. Yabukami, K. Kato, Y. Ohtomo, T. Ozawa, K. I. Arai, A thin film magnetic field sensor of sub-PT resolution and magneto cardiogram (mcg) measurement at room temperature, *Journal of Magnetism and Magnetic Materials*, Vol. 321, 2009, pp. 675-678.
- [19]. E. C. Silva, L. A. P. Gusmao, C. R. H. Barbosa, E. C. Monteiro, F. L. A. Machado, High sensitivity giant magnetoimpedance (GMI) magnetic transducer: Magnitude versus phase sensing, *Meas. Sci. Technol.*, Vol. 22, 2011, 035204.
- [20]. B. Dufay, S. Saez, C. P. Dolabdjian, A. Yelon, D. Ménard, Characterization of an optimized off-diagonal GMI based magnetometer, *IEEE Sens. J.*, Vol. 13, Issue 1, Jan. 2013, pp. 379-388.
- [21]. G. Yu, X. Bu, B. Yang, Y. Li, C. Xiang, Differential-type GMI magnetic sensor based on longitudinal excitation, *IEEE Transactions on Magnetics*, Vol. 11, Issue 10, Oct. 2011, pp. 2273-2278.
- [22]. M. Malatek, P. Ripka, L. Kraus, Temperature offset drift of GMI sensors, *Sens. Actuators A*, Vol. 147, May 2008, pp. 415-418.
- [23]. A. V. Semirov, D. A. Bukreev, A. A. Moiseev, V. A. Lukshina, E. G. Volkova, S. O. Volchkov, G. V. Kurlyandskaya, Temperature dependence of the magnetic properties and magnetoimpedance of nanocrystalline Fe<sub>73.5</sub>Si<sub>16.5</sub>B<sub>6</sub>Nb<sub>3</sub>Cu<sub>1</sub> ribbons, *Technical Physics*, Vol. 56, Issue 3, 2011, pp. 395-399.
- [24]. S. L. Zhang, J. F. Sun, D. W. Xing, F. X. Qin, H. X. Peng, Large GMI effect in Corich amorphous wire by tensile stress, *Journal of Magnetism and Magnetic Materials*, Vol. 323, June 2011, pp. 3018-3021.
- [25]. J. Nabias, A. Asfour, J. P. Yonnet, Investigation of bending stress effect on the diagonal and off-diagonal impedances for GMI sensor implementation, *IEEE Transactions on Magnetics*, Vol. 53, May 2017, 4002308.
- [26]. S. Sandacci, D. Makhnovskiy, L. Panina, K. Mohri, Y. Honkura, Off- diagonal impedance in amorphous wires and its application to linear magnetic sensors, *IEEE Transactions on Magnetics*, Vol. 40, Issue 6, 2004, pp. 3505-3511.
- [27]. A. S. Antonov, I. T. Iakubov, A. N. Lagarkov, Nondiagonal impedance of amorphous wires with circular magnetic anisotropy, *Journal of Magnetism and Magnetic Materials*, Vol. 187, Issue. 2, 1998, pp. 252-260.

- [28]. N. A. Usov, S. A. Gudoshnikov, Giant magneto-impedance effect in amorphous ferromagnetic wire with a weak helical anisotropy: theory and experiment, *J. Appl. Phys.*, Vol. 113, 2013, 243902.
- [29]. S. Gudoshnikov, N. Usov, A. Nozdrin, M. Ipatov, A. Zhukov, V. Zhukova, Highly sensitive magnetometer based on the off-diagonal GMI effect in Co-rich glass-coated microwire, *Physica Status Solidi A: Applications and Materials Science*, Vol. 211, 2014, pp. 980-985.
- [30]. C. Dolabdjian, B. Dufay, S. Saez, A. Yelon, D. Menard, Is low frequency excess noise of GMI induced by magnetization fluctuations, *Key Eng. Mater.*, Vol. 605, 2014, pp. 437-440.
- [31]. LTC2145, Low Power Dual ADCs Datasheet, Linear Technology, *McCarthy*, USA, 2011.
- [32]. P. S. Traoré, A. Asfour, J. P. Yonnet, C. Boudinet, Digital electronic conditioning approach for the high-sensitivity off-diagonal GMI sensor, *Sensors and Actuators A*, Vol. 271, 2018, pp. 290-302.
- [33]. P. S. Traoré, A. Asfour, J. P. Yonnet, Ch. Dolabdjian, Noise performance of SDR-based off-diagonal GMI sensors, *IEEE Sensors Journal*, Vol. 17, October 2017, pp. 6175-6184.
- [34]. P. S. Traoré, A. Asfour, J. P. Yonnet, C. Boudinet, Introduction of real-time digital signal processing techniques for the high sensitivity GMI sensors, *International Journal of Applied Electromagnetics and Mechanics*, Vol. 59, 2019, pp. 455-463.
- [35]. P. S. Traoré, A. Asfour, J. P. Yonnet, Off-diagonal GMI sensors with software defined radio detector: Implementation and performance, *IEEE Transactions on Magnetics*, Vol. 53, 2017, 4000907.
- [36]. P. S. Traoré, A. Asfour, J. P. Yonnet, Noise analysis of a high sensitivity GMI sensor based on a Field Programmable Gate Array, *Sensors and Actuators A*, Vol. 331, 2021, 112972.
- [37]. P. S. Traoré, A. Asfour, J. P. Yonnet, Design approach for an off-diagonal Giant Magneto Impedance Sensor based Field Programmable Gate Array, *Sensors & Transducers*, Vol. 245, Issue 6, October 2020, pp. 20-26.
- [38]. B. Dufay, E. Portalier, S. Saez, C. Dolabdjian, D. Seddaoui, A. Yelon, D. Menard, Low frequency excess noise source investigation of off-diagonal GMI-based magnetometers, *IEEE Transactions on Magnetics*, Vol. 53, Issue 1, 2017, 4000106.
- [39]. E. Portalier, B. Dufay, C. Dolabdjian, D. Seddaoui, A. Yelon, D. Menard, Evaluation of the imaginary part of the magnetic susceptibility,  $\chi''$ , and application to the estimation of the low frequency,  $1/f$ , excess noise in GMI sensors, *IEEE Transactions on Magnetics*, Vol. 35, Issue 8, 2017, 4001805.
- [40]. P. S. Traoré, M. I. Corrêa, Digital magnetometer based on giant magneto impedance effect: A feasibility study, *Sensors & Transducers*, Vol. 252, Issue 5, pp. 12-18, October 2021.
- [41]. P. S. Traoré, S. Konan, A. Asfour, J. P. Yonnet, Analysis of the impact of the processing time delay on the stability of a digital GMI magnetometer, in *Proceedings of the IEEE Sensors Conference (SENSORS'21)*, 2021, pp. 1-4.

## **Chapter 11**

# **R-testbench – How to Build a Reliable and Open-source Software Solution for the Remote Control of Measurements Instruments and Experimental Test Benches?**

**Alexandre QUENON, Evelyne DAUBIE,  
Véronique MOEYAERT and Fortunato DUALIBE**

### **11.1. Introduction**

Since the beginning of scientific research, instrumentation has played a critical role to prove or reject hypotheses, such as the famous Michelson and Morley experiment about the luminiferous ether [1], and to build theories describing a huge range of phenomena, e.g., the CMS experiment carried out at CERN to build and test the Standard Model of Physics [2]. Recently, with the rise of the Internet of Things (IoT) and the Industry 4.0, instrumentation has been pushed further, adapted to comply with sensors and their communication protocols.

Three aspects must generally be considered: the reliability and accuracy of the results, the physical and health safety of the operators, and the increasing demand for open-source solutions from both industrial and academic actors.

Regarding the reliability and the precision, it has become essential to develop an automated control of the instrumentation. This might be due to several factors, such as the size of the experimental facilities and measurement devices [2], the amount and the velocity of specific and accurate test sequences run sequentially [3], or the period of time of the events to record, which can be either very long or very short [4].

Considering the human and environmental safety, some experiments require to be carried out in extreme, possibly hazardous conditions. For instance, the characterization of radiation-hardened electronics, which necessitates the generation of ionizing radiation, dangerous or even lethal if not carefully handled, to determine the aging of the circuits.



This issue can be addressed by preventing the operators from the exposure to the hazardous conditions, hence the remote control of the instrumentation.

Finally, open-source software tools are becoming more and more popular in both industry and academia. This can be explained by two main facts: on the one hand, the price of the proprietary tools; on the other hand, the benefits of faster development, expansion, and adoption generally observed when open-source licensing is used [5]. Of course, open-source solutions are not defect-free: discontinuity, lack of documentation, and security issues are common problems, among others.

The present chapter is proposed as a guide to build a reliable and open-source software solution to control remotely measurements instruments and experimental test benches. The design process is illustrated with the development of R-testbench, a toolkit written in Python that implements such a remote control and is currently available on GitHub [6].

This chapter is divided in five main parts. Firstly, existing solutions of instruments control are reviewed. Secondly, the specifications of the system are presented and discussed according to the requirements. Then, the design of the software architecture, as well as the implementation of the code are reviewed. Afterwards, data management optimization is considered. Finally, the validation process is presented and discussed.

## **11.2. State of the Art: an Overview of the Existing Solutions**

Designing a new solution from scratch is rewarding but it also takes time and money. Hence the need to analyze the existing solutions to evaluate the need of creating a new tool.

The state of the art is classified into two categories: proprietary solutions and open-source alternatives. Even though the open-source access is a requirement, proprietary tools can offer interesting features for a fair comparison.

### **11.2.1. Proprietary Solutions**

LabVIEW from National Instruments [7] and MATLAB from MathWorks are the leaders in proprietary solutions for instrumentation. The former offers a Graphical User Interface (GUI) to control test benches from a computer, whereas the latter has become famous for its “programming language” dedicated to matrices manipulations as well as mathematical algorithms and problems solvers.

Both provide some graphical programming capabilities. However, this is rarely sufficient to control complex test benches. In this case, the proprietary scripting languages that are imposed by these proprietary solutions are G and MATLAB, respectively. These languages are generally known by specific communities, such as electronics engineers for G and MATLAB. But this is often not common for most software developers or experts from other technical areas.

Finally, neither LabVIEW nor MATLAB has released any open-source library at the time of writing this chapter. The main consequence is an increase in the cost of the solution.

However, both companies offer the compatibility with numerous communication protocols and standards, as well as a technical support for their customers.

### 11.2.2. Open-source Alternatives

Three serious open-source solutions can be found. Two of them are based on scripting languages for numerical computing: Octave [8, 9], and Scilab [10, 11]. The third one is a Python package, namely PyMeasure [12].

Both Octave and Scilab require (very) low-level programming skills and are built with their own scripting language. This is their main drawback, as not so many people develop a deep knowledge of both instrumentation and low-level programming.

The PyMeasure Python package is an interesting solution that offers high-level programming with an interpreted language, which facilitates the creation of scripts to control instruments in the lab. Nonetheless, it misses the support of common files formats to save the results.

### 11.2.3. Comparison and Discussion

The main advantage of the proprietary solutions is the graphical programming, even though, as discussed earlier, it is not adequate to all situations. Most of the tools require low-level programming to control the instruments. Finally, the number of supported files types is quite low in open-source solutions. The main features of the compared software tools are summarized in Table 11.1.

**Table 11.1.** Comparison of the main characteristics of the existing software solutions for instruments control.

Tool	License	Programming		Com. protocols	Supported files	Additional features
		Level	Language			
LabVIEW	Proprietary	High, low	MathScript	VISA, and more	Many	Graph. program.
MATLAB	Proprietary	Low	MATLAB	VISA, and more	Many	Graph. program.
Octave	GPL	Low	Octave	VISA, except PXI	txt, bin, CSV, HDF5, mat	
Scilab	GPL	Low	Scilab	VISA	SOD (HDF5)	
PyMeasure	MIT	High	Python	VISA	CSV	

### 11.3. System-level Specifications

The review of the state of the art tends to indicate that there is an interest to create a new open-source tool that fits the main requirements of instrumentation and offers additional features that are currently either only available in proprietary solutions or not available at all.

The new solution that we propose is “*R-testbench*”, a software library that allows the user to control instruments and test benches remotely. The development of the program has been driven by four main requirements:

1. The reliability of communication between the control unit and the instruments;
2. The adaptability and compatibility with several platforms that can run the tool, any brands and models of instruments, and communication busses;
3. The user interface;
4. The open source access.

This section describes the aforementioned specifications and the technical choices that precede the design of the software architecture and the implementation.

#### 11.3.1. Reliability

The reliability of the software is a necessary condition to guarantee the reliability and the quality of the experimental data acquired from a test bench. In this respect, two important features are required. On the one hand, a standard that rules the communications between the computer and the instruments should be used. On the other hand, an error management system that checks if the process works as expected should be deployed.

##### 11.3.1.1. Reliability of Control-instrument Communications

The best solution to ensure the reliability of the control-instrument communications is to use a standard that is largely adopted by the instruments’ manufacturers. Such a standard has been existing since 1995: this is the Virtual Instrument Software Architecture (VISA). VISA offers an interface-independent framework to control instruments by software [13], compatible with many data busses (e.g., GPIB and USB). The standard allows the user to open a communication channel to send commands to and get data from an instrument, as depicted in Fig. 11.1. The VISA standard is directly implemented by the main manufacturers of measurement instruments and released under the form of libraries [14-16].

In addition to the interface-independent system, a generalized “language” common to all instruments would be of high interest to design a software tool to control test benches. A first attempt of generalization was performed with the Standard Commands for Programmable Instruments (SCPI) [17], which offers a common syntax to program all

instruments in a unified style. However, it does not prevent the user from dealing with the commands specific to each instrument, such as the data format.



**Fig. 11.1.** Illustration of the VISA standard overseeing the communications between a control unit (e.g., a computer) and an instrument, connected with any type of interfaces supported by VISA.

### 11.3.1.2. Reliability of Source Code (Core Development)

The reliability of the source code, which is the fundament of the core development of the software tool, strongly depends on the programming language itself, as well as its standard and third-party libraries, and on the programming techniques used to write the code.

There are numerous programming languages available. Some of them are well known and commonly used by the IT community, such as C++ and Python, whereas some others are quite marginal or purpose-specific, e.g., Oz.

Programming languages can be standardized by several means. For example, C++ was standardized by the International Organization for Standardization as ISO/IEC 14882:2017 [18], for the 2017 release. Python, in contrast, is standardized de facto by the core developers through the Python Enhancement Proposals (PEPs) [19], which constitute a set of documents describing and tracking the design changes in the language. However, it does not offer the same confidence level and quality of maintenance as a certificate delivered by an international organism. Nevertheless, the current trends in software developments are an increase in the quality of maintenance and tracking of modifications in core libraries of programming languages, powered by versioning tools and platforms.

In addition to the quality of maintenance, a reliable programming language includes at least a robust standard library, which provides a set of basic data types and functions, data containers, streams managers (such as terminal or file streams), and specific functions (e.g., mathematics or multithreading). Some languages can also rely on powerful third-party libraries that are so famous that they are used in many projects. Examples are the Boost library in C++ [20] and the numpy package in Python [21].

Last but not least, programming techniques are part of the reliability of the source code. A good example is the use of the Resource Acquisition Is Initialization (RAII), created by Bjarne Stroustrup and Andrew Koenig for exception-safe resource management, i.e., to avoid the use of a resource that is not available and to avoid memory leak [22]. An important feature to design a reliable software is an error management mechanism, which raises an error that can be caught to fix the issue on the software level or to warn the user if something goes wrong at execution time.

### **11.3.2. Adaptability**

The adaptability must take three different aspects into account: (1) the platforms, i.e., both hardware and software, on which the tool is run, (2) the brand and model of the controlled instruments, and (3) the transmission busses, protocols or standards that can be used to communicate with the instruments.

#### **11.3.2.1. Portability**

The adaptability to several platforms is known as the portability. It is the ability of a software tool to be executed with different operating systems (OS) running on different hardware components.

The portability is strongly related to the nature of the language, i.e., compiled or interpreted. A compiled language requires to be “converted” to machine-language instructions. Consequently, the compilation process must consider the target deployment platform, which can request stronger developer’s efforts when a cross-platform solution is desired. Conversely, an interpreted language is run immediately, without compilation, by an interpreter, which makes deployment easier, but can, however, generate bugs that will have to be fixed late in the development process. In terms of performance, compiled solutions tend to be better than interpreted ones.

A common issue that infers portability restrictions is missing libraries for a specific OS or a small size hardware platform such as Raspberry Pi. Such issues can sometimes be solved by compiling the missing library to target the desired platform. However, this requires extra efforts and consumes times.

#### **11.3.2.2. No Brands or Models Restrictions**

As discussed earlier (Section 11.3.1.1), the VISA and SCPI protocols offer a first attempt of generalization of instruments communication by software. Using both of them should therefore remove any brand or model restriction.

#### **11.3.2.3. Communication Bus Multi-compatibility**

The VISA standard also offers a compatibility with several types of communication busses, such as GPIB and PXI. However, it does not include recent standards that could be interesting to control sensors remotely, like Bluetooth or ZigBee.

### **11.3.3. User Interface**

VISA libraries for software communication with instruments already exist, as mentioned earlier. However, using them require low-level programming skills and a good knowledge on the method to open a communication channel. Consequently, a strong work on the user interface is required to provide an intuitive user experience.

The proposed method consists in using the object-oriented paradigm, which allows the encapsulation of the operating details into objects that are easy to implement and to use. In addition, the main demand in laboratories is the possibility to create scripts dedicated to specific test benches, experiments, and characterizations. Moreover, the acquired raw data are rarely sufficient for features extraction. Therefore, adding post-processing capabilities, such as data manipulation and filtering, within the tool is a plus.

Finally, the user experience can be enhanced by offering automatic recognition of the instrument connected to the control unit. This would avoid an exhausting process of configuration to choose the brand, family, and model on instruments, as well as the data bus that is used for the communication.

#### **11.3.4. Open-source Access**

The source code of the R-testbench library is released on a public repository under an open-source license, the Open Software License 3.0 (OSL-3.0) [23], on the well-known and widely used GitHub platform: <https://github.com/Arkh42/rtestbench>

In addition, the source code is entirely Python code, which is itself open source [24]. The main dependencies are the pyvisa library, which is distributed under the MIT License [25], and the numpy and the pandas libraries, which are disclosed under the 3-Clause Berkeley Software Distribution License (BSD-3-Clause) [26].

#### **11.3.5. Summary of the Specifications**

In summary, the main system specifications are listed below:

- Standards compliance with VISA, SCPI;
- Standardized programming language;
- Portable on multiple OS and hardware;
- No instruments models or brands restrictions;
- Communication bus multi-compatibility;
- High-level user interface;
- Automatic instruments recognition.

### **11.4. Software Architecture and Implementation**

Based on the aforementioned specifications, the software architecture can be designed, and the necessary functions can be implemented.

### **11.4.1. Selection of the Programming Language**

The discussion in Section 11.3.2.1 regarding the portability tends toward the selection of an interpreted programming language. Especially, data acquisition and processing is often made through scripts that are written with more flexibility if no compilation is required.

In addition, the compilation process might require a time-consuming work to find compatible libraries between several operating systems, and therefore increase the pressure on maintenance.

As a conclusion, Python has been selected for this project. Firstly, it allows to use encapsulation that is useful to separate the usage from the internal process. Secondly, this programming language provides great scripting capabilities, making it very convenient to use in laboratories and test facilities. Finally, Python has been granted with powerful scientific computation libraries, such as the popular NumPy and pandas libraries. Consequently, using Python allows the user to unify the coding of the data acquisition and the data processing under the same programming language, a considerable asset to increase the production.

### **11.4.2. Design of the Software Architecture**

The objective of R-testbench is to allow the user to control and automate a test bench that involves several instruments used for data acquisition. This is done from a unique control interface, a laptop or a Raspberry Pi, for example. In addition, the nature of the connections between the control interface and the instruments can change, and that has to be taken into account.

The software design has a specific consideration for the user interface. To do so, two major features must be added. The first one consists in hiding the type of connection (e.g., GPIB or USB) used to control an instrument, from the software side. This is necessary to avoid script modification if, for any reason, the type of connection changes between two runs of experiments. The second feature is an automatic identification of the connected instrument, including the brand and the model. This is mandatory to offer simple and meaningful commands in Python, i.e., functions and methods that will be automatically “translated” by R-testbench into the right instrument-compatible commands.

This section presents the software architecture that has been designed to implement the aforementioned features and the specifications described in Section 11.3.

#### **11.4.2.1. A Front End to Call the VISA Libraries**

To ensure the reliability of the communication with the instruments, the VISA standard is used. As explained in Section 11.2.2, several implementations of VISA are proposed by the different manufacturers [14-16]. In addition, the VISA libraries are operating systems (OS) dependent, which can quickly turn the creation of a generic call method into something hard to handle.

Fortunately, PyVISA [27] is available. It is a Python package that works as a front end capable of calling multiple back ends implementing the VISA standard. This package starts to be recognized by instrumentation engineers and companies as a reliable way to control instruments, as proved by the donation of Keysight to the PyVISA's developers for test purpose.

PyVISA offers a hidden mechanism that makes the connections transparent from the software's perspective. The user has to provide the instrument's address, which will be parsed in order to identify the communication protocol. Then, the corresponding Resource object will be created, allowing the user to send commands to the instrument. The process is managed by a ResourceManager, responsible for calling the VISA libraries and opening the communication channels to the Resources.

Of course, the future of PyVISA cannot be predicted. The library could be discontinued or not be maintained anymore. In such cases, R-testbench should be able to find another solution, almost transparent for the user. To enable this, R-testbench works in parallel with PyVISA, as shown in Fig. 11.2. Using the object-oriented paradigm, an encapsulated instance of ResourceManager is created and used to build a communication channel with an instrument. Therefore, no direct call to PyVISA is performed by the user.

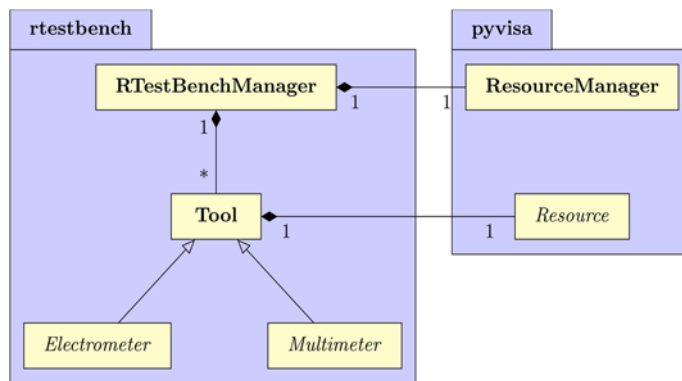


Fig. 11.2. R-testbench Unified Modeling Language (UML) class diagram: manager, tool interfaces mechanism, and interactions with PyVISA [28].

#### 11.4.2.2. Automated Instrument Recognition

The major difficulties in software instrument control are the many ways used by the manufacturers to design their instruments and how they operate. It forces the user to spend a lot of time in reading the documentation to understand the internal behavior of the tools, which is not necessarily relevant to carry out the experiments.

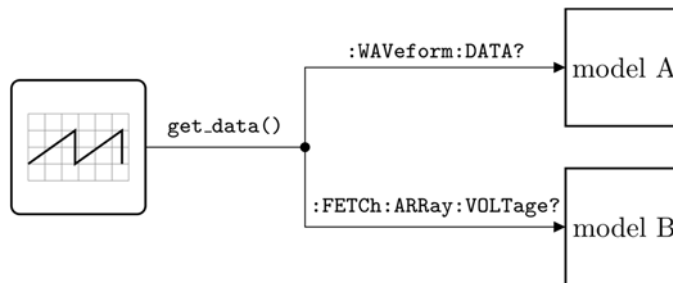
A first attempt of generalization was performed with the Standard Commands for Programmable Instruments (SCPI) [17]. This standard offers a common syntax to program all instruments in a unified style. However, it does not prevent the user from dealing with the specific features of each instrument, such as the data format.



To fix this issue and enhance the user experience, the R-testbench software architecture takes advantage of the object-oriented paradigm to encapsulate the SCPI commands and specific features of a given model of instrument into a dedicated class. These classes inherit from generic abstract classes that will define the interface common to a family of instruments, e.g., oscilloscopes, multimeters, or electrometers. These abstract classes also inherit from an even more generic class, named Tool, which defines the attributes and methods common to all tools that can be controlled by software using the VISA and SCPI standards (Fig. 11.2).

The proposed architecture offers three main advantages. Firstly, the development related to one specific model of instrument is fast, as it consists in embedding the SCPI commands into class methods and defining the invariants as class attributes. Secondly, a factory mechanism can be used to automatically identify the instrument and build the related object, by sending the “\*IDN?” request. This method also ensures that the instrument answers properly. Finally, polymorphism can be used to make the real type of object transparent to the user and offer a unique interface per family of instrument, as depicted in Fig. 11.3. Of course, the specific object can always be directly used, if necessary.

In addition to the Tool inheritance system, a class named RTestBenchManager has been created. Its role consists in calling a VISA resource manager, which is currently provided by the PyVISA library, and creating all resources necessary to provide interactions and other functionalities to the user, such as a logger. All instrument instances are attached to the Manager object, for proper life cycle management. Finally, this class provides facilities for saving data.

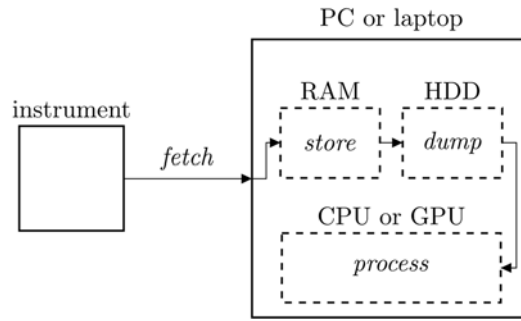


**Fig. 11.3.** Example of unique interface taking advantage of the polymorphism and encapsulation mechanisms to hide the specific SCPI commands into a single Python function [29].

## 11.5. Data Management Considerations

Scientific research experiments and characterization for industrial purposes share the same problem: the number of data that are retrieved and must effectively be managed. Efficiency is mandatory for two main reasons: the data quality and the computational load. This is monitored using three parameters: the data integrity, the memory usage, and the execution speed.

For this purpose, the data management has been divided in four steps: (1) fetch from instruments, (2) local temporary storage, at execution time, on the machine running R-testbench (e.g., a laptop), (3) permanent saving, i.e., dump data into a file, and (4) processing, for example to extract statistical information. The whole process is illustrated in Fig. 11.4.



**Fig. 11.4.** Illustration of the four-step data management in the case of data acquired with one instrument controlled from a laptop or a personal computer.

This section reviews those four steps to highlight the strategies used for efficient data management.

### 11.5.1. From the Instruments to the Computer: the Fetch Operation

As data fetch is controlled by the PyVISA front end and the VISA libraries, there is no possibility to optimize the related data management. The only possible action consists in choosing the interface used to communicate with the instruments.

The chosen interface must be compatible with VISA to be controlled by R-testbench. Another solution consists in choosing any standard suited for instrumentation [30], and, if not VISA-compatible, to use a gateway whose communication protocol is supported, such as TCP/IP.

### 11.5.2. From Acquisition to Temporary Saving: the Local Storage

Right after the fetch action, when data arrive to the hardware running R-testbench, a local storage in the RAM is necessary in order to process further the data or to save them in the permanent memory. According to the adaptability specification, the library offers the choice to store data in Python built-in containers, such as lists or dictionaries, as well as specialist objects provided by libraries optimized for scientific computations, e.g., NumPy or pandas.

Nevertheless, despite the flexibility given to the user, the default data container used for temporary local storage is the NumPy's N-dimensional array, *ndarray* [21]. It currently

reaches the best compromise between memory management, execution speed, and user-friendly manipulation of huge arrays.

### **11.5.3. From Temporary to Permanent Saving: the Dump Action**

At this step of the process, all acquired data are temporarily stored locally, for example in the RAM of a laptop. Afterwards, two actions are jointly performed: data aggregation and permanent saving.

On the one hand, it is generally required to aggregate related data for saving. For instance, the vector of timestamps corresponding to the measurement of a time-dependent parameter, such as the voltage on an oscilloscope.

On the other hand, data must be permanently saved into files, for later processing or reading. Depending on the applications, the expected-file formats can be different. The common compromise comes on either the possibility to read immediately the data values by opening the file with a text editor, or an efficient storage that needs the minimum memory space for the maximum number of data, without jeopardizing the accuracy.

Regarding both requirements, the pandas API offers a rich and easy-to-use interface that fits the needs [31]. Currently, R-testbench takes benefits of pandas to offer the possibility to save data in one text format, CSV, and in three binary formats, pickle, Feather, and HDF5. CSV (Comma Separated Values) has the advantage to be human readable and to be supported by a plethora of data processing tools. Pickle is a Python-specific module for object serialization that can convert an object into a byte stream to write the corresponding data into a binary file [32]. Feather is a format based on the Apache Arrow columnar memory specification for data representation that is very performant [33, 34]. HDF5 (Hierarchical Data Format) is available in several programming languages, and very popular for data science and artificial intelligence [34].

### **11.5.4. From Raw Data to Refined Information: the Data (Post-) Processing**

Data processing can be performed in two ways: in real time during the acquisition, or in post-acquisition. The former has the advantage of freeing memory as it is not necessary to store raw data permanently, the interesting information being already extracted. However, a simultaneous processing requires either a fast and efficient computation tool (hardware or software) or a huge buffer to accumulate the acquired data while they are being processed. The latter method, post-processing, does not constrain the computation capabilities, but requires the storage of the raw data, which can be a problem from the memory perspective.

Currently, R-testbench implements the post-processing mechanism, in order to allow the user to run the Python library without hard constraints on the hardware. Moreover, because of the various data files formats available, many data processing tools can be used, including open-source ones. This is compliant with the adaptability specification.

## 11.6. Validation: Tests, Results, and Discussion

Throughout the design and implementation process, the software has to be validated. Four techniques are presented in this section: continuous integration (CI), performance characterization, alpha test phase and public release on an open-source platform.

### 11.6.1. Continuous Integration and Alpha Tests

Continuous integration (CI) is a software development methodology that consists in (1) committing, i.e., sending the modified code to a unique repository, (2) versioning, which means tracking the modification of the code, and (3) building and testing the code whenever it is modified to ensure that no anterior functionalities have been broken by the most recent update [36, 37].

Among popular versioning tools, one can find Git, Subversion (SVN), Mercurial, and CVS. The selection of a specific tool is generally a personal matter, even though their internal behavior varies. The most important feature regarding the implementation of a version control tool is either the local copy of the versioning database on each station which has a copy of the project or a unique copy of the versioning database on the server that hosts the project. In the case of R-testbench, the tool selected for modification tracking is Git, because it is fully compatible with the GitHub hosting platform (cf. Section 11.6.3 for further details).

Automated testing has been implemented by using PyTest, to develop the unit tests, and GitHub Actions and Travis CI, to create the test scenarios including different Python versions (3.6, 3.7, and 3.8) and operating systems, as presented in Table 11.1.

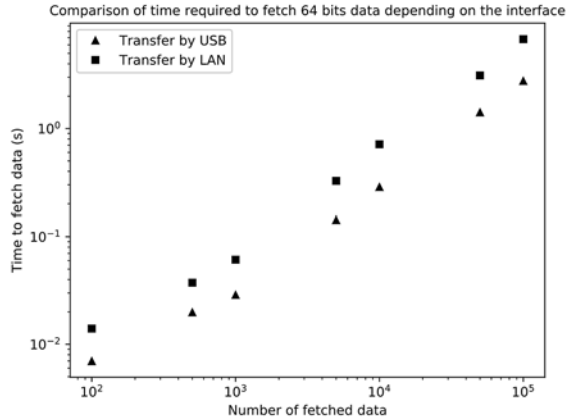
The R-testbench Python package has also been successfully tested by the developer in the so-called alpha tests. They were realized in the frame of experiments involving the characterization of semiconductors devices irradiated with ionizing radiation, e.g., beta rays [36]. This use case requires remote instrumentation, for safety reasons.

**Table 11.1.** List of the verified versions of operating systems distributions that are supported, and test methods that used to perform the verification. CI stands for continuous integration.

Distributions	Versions	Test methods
macOS	10.15 (Catalina)	CI (GitHub)
Ubuntu	16.04 (Xenial Xerus)	CI (GitHub, Travis)
	18.04 (Bionic Beaver)	Alpha tests, CI (GitHub)
Windows 10	N/A	Alpha tests
Windows Server	2019	CI (GitHub)

### 11.6.2. Performance Characterization

All performance measurements were realized with the Keysight B2985A electrometer [38], the instrument which was involved in alpha tests. All results were achieved by running the acquisitions and computing the properties seven times and averaging the values. This has been determined empirically to minimize the standard deviation. The time required to fetch data from an instrument and store locally in a NumPy ndarray object evolves linearly with the number of data, as demonstrated in the case of a transfer by USB and LAN (TCP/IP) in Fig. 11.5.



**Fig. 11.5.** Comparison of time required to fetch data from a Keysight B2985A electrometer to the computer [28]. Data are stored locally in a numpy.ndarray container.

To evaluate the overhead time due to the software latency and the local storage, the theoretical transfer time,  $t_{\text{transf}}$ , must be computed. It depends on the theoretical transfer speed of the related interface,  $v_{\text{transf}}$ , the number of transmitted data,  $N$ , and the number of bits used to code one datum,  $n_{\text{bit}}$ , as shown in eq. (1):

$$t_{\text{transf}} = \frac{N \cdot n_{\text{bit}}}{v_{\text{transf}}}. \quad (1)$$

The instrument used for the test is equipped with a USB 2.0 interface, which means that the theoretical speed limit is 480 Mbit/s. If the protocol overhead for high-speed bulk transactions is taken into account, it leads to a speed limit of useful data of 425.984 Mbit/s [39]. Hence, using (1), a thousand 64-bit data should require

$$t_{\text{transf}} = \frac{1000 \cdot 64}{425.984 \times 10^6}, \quad (2)$$

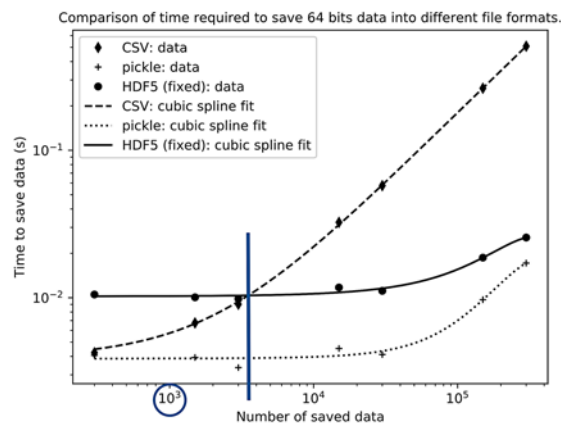
leading to a minimum transfer time of 150.24  $\mu\text{s}$ .

Compared with the results shown in Fig. 11.5, the minimum transfer time due to the interface is two orders of magnitude lower. Consequently, the overhead time necessary to

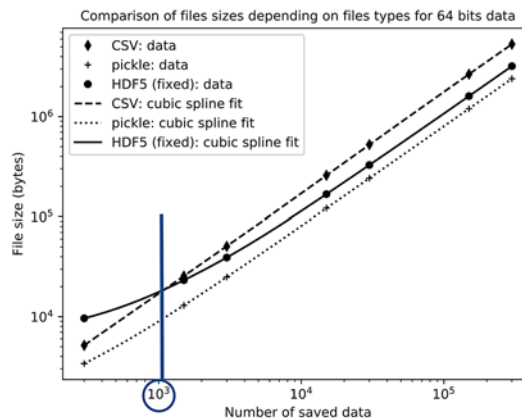
call the VISA libraries that allow to communicate with the instruments, and to store the data in a NumPy ndarray object, is a hundred times longer than the transfer time.

Further measurements are necessary to determine the bottleneck of the fetch operation. This requires the possibility to monitor the low-level mechanisms of the VISA libraries, and to test several implementations. Another problem could arise from the data storage with NumPy, which is managed by the PyVISA package.

The speed and memory requirements for the permanent storage have been characterized, for three different files formats: one text format, CSV, as well as two binary formats, pickle and HDF5. The corresponding results are presented in Fig. 11.6 and Fig. 11.7, for the saving time and the size of the generated data files, respectively.



**Fig. 11.6.** Comparison of the evolution of the time required to save data into CSV, pickle and HDF5 files formats with the number of data [28]. Data are saved to files using the pandas I/O API.



**Fig. 11.7.** Comparison of the evolution of sizes of CSV, pickle and HDF5 files with the number of data [28]. Data are saved to files using the pandas I/O API.

The best performance is expected to come from binary formats, for both size and speed aspects. However, for a small amount of data, the CSV format generates smaller files, and generate files faster, than the HDF5 format. This can be explained by the fact that the latter has been designed for Big Data, which assumes the manipulation of huge datasets. On the studied range, which corresponds to a maximum of 300,000 data, the pickle format exhibits the best behavior. This is confirmed by other studies [40], [41]. Nevertheless, the range of data should be extended to draw definitive conclusions, as observed HDF5 and pickle formats tends to have the same performance for bigger datasets, as observed in Fig. 11.6.

### **11.6.3. Public Release**

R-testbench has been made publicly available on a GitHub repository [6], an open-source platform widely used by researchers and programmers, from both industry and academia, to share and contribute to open-source programming projects. The project is licensed under the OSL-3.0.

There is currently one contributor and maintainer, A. Quenon, as well as one pull request to extend the Python package to instruments that are not originally supported. As the package has been made publicly available for the SEIA' 2020 conference [28], and because of the contributions to PyVISA due to the strong dependency, the interest for this project is expected to grow up.

### **11.6.4. Results Discussion**

The proposed library has been clearly characterized for the execution speed and the memory space required to save results. To the best of the authors' knowledge, there is no available studies that offer the same type of characterization for the other available software solutions for remote instruments control.

Nevertheless, R-testbench has several advantages, as shown in Table 11.2. Specifically, it is open-source, offers high-level programming capabilities, performs automatic instrument recognition and supports several types of files formats to save results. Currently, the main drawbacks are the few numbers of specific instruments that have been implemented and the small size of the development team. Of course, this is a matter of time.

## **11.7. Conclusions**

This chapter presents the design of R-testbench, an open-source Python library devoted to remote instrumentation and test bench automation, as a guide through the development of a reliable and open-source software solution.

**Table 11.2.** Comparison of the main characteristics of software solutions for measurement instruments remote control and automation [6]–[11], [27], [42], [43].

Tool	License	Programming			VISA compatible protocols	Automatic recognition	Supported results files types
		Level	Language	Graphical			
LabVIEW	Proprietary	high or low	MathScript	yes	yes, and more	no	many
MATLAB	Proprietary	low	Matlab	partial	yes, and more	no	many
Octave	GPL	low	Octave	no	no (PXI)	no	bin, CSV, HDF5, mat, txt
Scilab	GPL	low	Scilab	no	yes	no	SOD (HDF5)
PyMeasure	MIT	high	Python	no	yes	no	CSV
<i>R-testbench</i>	<i>OSL</i>	<i>high</i>	<i>Python</i>	<i>no</i>	<i>yes</i>	<i>yes</i>	<i>CSV, feather, HDF5, pkl</i>

The R-testbench library has been programmed using the object-oriented paradigm, to encapsulate implementation details and offer a user-friendly high-level programming interface. It has been designed to be reliable, which is achieved by using the VISA standard for the communications with the controlled instruments. The solution is adaptable to several types of communication interfaces, e.g., GPIB or TCP/IP, as well as to several types of data processing tools, because of the support of different files formats for saving results.

The library has been validated by implementing continuous integration, including code versioning and automated build and tests. It is proved to be cross-platform. Alpha tests were successfully carried out by controlling remotely an experiment involving the characterization of semiconductors devices irradiated with ionizing radiation. The source code is publicly available on GitHub and is becoming mature.

Future developments include the implementation of specific models of instruments, as well as the support of other files formats, such as the increasingly popular JSON text format. It would also be interesting to create bindings with database management tools, e.g., Mongo DB or MySQL. Finally, the possibility to run R-testbench on portable platforms, such as a Raspberry Pi or an FPGA, together with hardware acceleration, can be explored. The latter is a clue for efficient real-time data processing.

Tips to remember from this chapter:

- Identify the standards that rule the area that is targeted by your application to increase the reliability of the software;
- If possible, use encapsulation to hide specific, technical details to the user and provide an intuitive and clean programming interface;
- Provide generic and flexible solutions that can adapt to common cases;



- Think about the validation of the tool from the beginning of the design process;
- Use a version control tool;
- Automate the installation and unit tests by using continuous integration tools.

## Acknowledgements

The authors are grateful for D. Binon, for his advice on instruments and test benches, and his help with proprietary instrumentation software. They also thank J. Hanton, for his help for solving IT issues, as well as S. Devouge and F. Defontaines for their support and kindness.

This work was supported by the Fonds de la Recherche Scientifique — FNRS under Grant n° 33678493.

## References

- [1]. A. A. Michelson and E. W. Morley, On the relative motion of the Earth and the luminiferous ether, *Am J Sci*, Vol. 34, No. 203, November, 1887, pp. 333–345.
- [2]. The CMS Collaboration *et al.*, The CMS experiment at the CERN LHC, *JINST*, Vol. 3, No. 08, August 2008.
- [3]. Z. J. Deng, N. Yoshikawa, S. R. Whiteley, and T. Van Duzer, Data-driven self-timed RSFQ high-speed test system, *IEEE TAS*, Vol. 7, No. 4, December 1997, pp. 3830–3833.
- [4]. V. de Miguel Soto, J. Jason, D. Kurtoglu, M. Lopez-Amo, and M. Wuilpart, Spectral shadowing suppression technique in phase-OTDR sensing based on weak fiber Bragg grating array, *Opt. Lett.*, Vol. 44, No. 3, February 2019, pp. 526–529.
- [5]. L. E. Hecht, L. Clark, and the Linux Foundation, Survey: Open Source Programs Are a Best Practice Among Large Companies, *The New Stack*, August 30, 2018. <https://thenewstack.io/survey-open-source-programs-are-a-best-practice-among-large-companies/> (accessed Oct. 13, 2020).
- [6]. A. Quenon, *R-testbench*, 2020. [Online]. Available: <https://github.com/Arkh42/rtestbench>
- [7]. National Instruments, What is LabVIEW ? <https://www.ni.com/en-us/shop/labview.html> (accessed April 24, 2020).
- [8]. Octave Wiki contributors, Instrument control package, January 12, 2020. [https://wiki.octave.org/Instrument\\_control\\_package](https://wiki.octave.org/Instrument_control_package) (accessed April 24, 2020).
- [9]. J. W. Eaton and Octave developers, Simple File I/O (GNU Octave (version 5.2.0)), January 31, 2020. [https://octave.org/doc/v5.2.0/Simple-File-I\\_002fO.html#Simple-File-I\\_002fO](https://octave.org/doc/v5.2.0/Simple-File-I_002fO.html#Simple-File-I_002fO) (accessed October 14, 2020).
- [10]. Scilab team, Signal acquisition & instrument control. <https://www.scilab.org/software/atoms/signal-acquisition-and-instrument-control> (accessed May 06, 2020).
- [11]. Scilab team, save format - Format of files produced by ‘save.’ [https://help.scilab.org/docs/6.1.0/en\\_US/save\\_format.html](https://help.scilab.org/docs/6.1.0/en_US/save_format.html) (accessed October 14, 2020).
- [12]. PyMeasure Developers, PyMeasure scientific package — PyMeasure 0.8.0 documentation, Mar. 29, 2020. <https://pymasure.readthedocs.io/en/latest/index.html> (accessed October 14, 2020).
- [13]. IVI Foundation, VPP-4.3: The VISA Library. October 19, 2018.

- [14]. National Instruments, NI-VISA Overview, March 20, 2020. <https://www.ni.com/en-us/support/documentation/supplemental/06/ni-visa-overview.html> (accessed Oct. 09, 2020).
- [15]. Rohde & Schwarz, R&S@VISA, 2020. [https://www.rohde-schwarz.com/dk/applications/r-s-visa-application-note\\_56280-148812.html](https://www.rohde-schwarz.com/dk/applications/r-s-visa-application-note_56280-148812.html) (accessed October 09, 2020).
- [16]. Keysight Technologies, Software I/O Layers - VISA, VISA COM, SICL, Keysight 488 - Technical Overview | Keysight, Sep. 01, 2009. <https://www.keysight.com/main/editorial.jsp?ckey=1461160&id=1461160&nid=-33002.0.00&lc=dut&cc=BE> (accessed October 09, 2020).
- [17]. SCPI Consortium, Standard Commands for Programmable Instruments (SCPI), May 1999, [Online]. Available: <https://www.ivifoundation.org/specifications/default.aspx>
- [18]. International Organization for Standardization, ISO/IEC 14882:2017, *ISO*, December 2017. <https://www.iso.org/standard/68564.html> (accessed October 09, 2020).
- [19]. Python Software Foundation, PEP 1 -- PEP Purpose and Guidelines, *Python.org*, April 07, 2013, <https://www.python.org/dev/peps/pep-0001/> (accessed October 09, 2020).
- [20]. B. Schäling, *The Boost C++ Libraries*, 2<sup>nd</sup> ed. XML Press, 2014.
- [21]. S. van der Walt, S. C. Colbert, and G. Varoquaux, The NumPy Array: A Structure for Efficient Numerical Computation, *Computing in Science & Engineering*, Vol. 13, No. 2, April 2011, pp. 22–30.
- [22]. B. Stroustrup, *The C++ Programming Language*, 4<sup>th</sup> ed. Addison-Wesley, 2013.
- [23]. Open Source Initiative, The Open Software License 3.0 (OSL-3.0). <https://opensource.org/licenses/OSL-3.0> (accessed October 08, 2020).
- [24]. Python Software Foundation, History and License, September 03, 2020. <https://docs.python.org/3/license.html> (accessed October 09, 2020).
- [25]. Open Source Initiative, The MIT License. <https://opensource.org/licenses/MIT> (accessed June 22, 2021).
- [26]. Open Source Initiative, The 3-Clause BSD License. <https://opensource.org/licenses/BSD-3-Clause> (accessed June 22, 2021).
- [27]. G. Thalhammer, T. Bronger, F. Bauer, H. E. Grecco, and M. Dartiailh, *PyVISA: Control Your Instruments with Python*, 2020, Accessed: February 24, 2020. [Online]. Available: <https://pyvisa.readthedocs.io/en/1.11.1/index.html>
- [28]. A. Quenon, E. Daubie, V. Moeyaert, and F. C. Dualibe, R-testbench: a Python library for instruments remote control and electronic test bench automation, in *Proceedings of the 6th International Conference on Sensors and Electronic Instrumentation Advances (SEIA' 2020) and the 2nd IFSA Frequency & Time Conference (IFTC' 2020)*, Porto, Portugal, September 2020, pp. 47–50.
- [29]. A. Quenon, E. Daubie, V. Moeyaert, and F. Dualibe, R-testbench: a Journey in Open Source Programming for Remote Instrumentation with Python, *Sensors & Transducers*, Vol. 245, No. 6, pp. 90–98, October 2020.
- [30]. J. Park, S. Mackay, and E. Wright, *Practical Data Communications for Instrumentation and Control*. Newnes, 2003.
- [31]. W. McKinney, Data Structures for Statistical Computing in Python, in *Proceedings of the 9<sup>th</sup> Python in Science Conference*, Austin, Texas, June 2010, pp. 56–61.
- [32]. Python Software Foundation, pickle — Python object serialization, May 02, 2020. <https://docs.python.org/3/library/pickle.html> (accessed October 16, 2020).
- [33]. W. McKinney, *Feather: Fast, Interoperable Data Frame Storage*, 2019, [Online]. Available: <https://github.com/wesm/feather> (Accessed: October 16, 2020).
- [34]. The Apache Software Foundation, Apache Arrow Overview, *Apache Arrow*. <https://arrow.apache.org/overview/> (accessed October 16, 2020).
- [35]. The HDF Group, HDF5 File Format Specification, July 2019.
- [36]. T. Durieux, R. Abreu, M. Monperrus, T. F. Bissyandé, and L. Cruz, An Analysis of 35+ Million Jobs of Travis CI, in *Proceedings of the IEEE International Conference on Software*

- 
- Maintenance and Evolution (ICSME' 2019)*, Cleveland, OH, USA, October 2019, pp. 291–295.
- [37]. E. Laukkanen, J. Itkonen, and C. Lassenius, Problems, causes and solutions when adopting continuous delivery—A systematic literature review, *Information and Software Technology*, Vol. 82, February 2017, pp. 55–79.
- [38]. Keysight Technologies, Keysight B2980A Series Data Sheet, February 12, 2019. [Online]. Available: <https://www.keysight.com/en/pc-2444652/b2980a-series-femto-picoammeter-and-electrometer-high-resistance-meter>
- [39]. Compaq Computer Corporation *et al.*, Universal Serial Bus Specification, April 2000.
- [40]. pandas development team, IO tools (text, CSV, HDF5, ...) — pandas 1.1.3 documentation, October 05, 2020. [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/io.html](https://pandas.pydata.org/pandas-docs/stable/user_guide/io.html) (accessed October 16, 2020).
- [41]. I. Zaitsev, The Best Format to Save Pandas Data, *Towards Data Science*, March 29, 2019, <https://towardsdatascience.com/the-best-format-to-save-pandas-data-414dca023e0d> (accessed October 16, 2020).
- [42]. MathWorks, Instrument Control Toolbox, <https://www.mathworks.com/products/instrument.html> (accessed April 24, 2020).
- [43]. MathWorks, Supported File Formats for Import and Export - MATLAB & Simulink, 2020, [https://www.mathworks.com/help/matlab/import\\_export/supported-file-formats.html](https://www.mathworks.com/help/matlab/import_export/supported-file-formats.html) (accessed October 14, 2020).

# Chapter 12

## Gas Sensing Technologies Application in Combustion

**Pavel Shuk**

*In memory of my lovely wife Galina Shuk (Galochka)*

### 12. Introduction

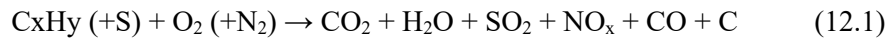
Gas sensing technologies are critical in many applications, i.e., indoor/outdoor air quality control, automotive industry, greenhouse gas monitoring, medical applications and industrial process control, especially combustion. Unlike some other sensing technologies like temperature, pressure, flow or level measurements, detecting a single variable, gas sensing technologies are much more complex with many variables affecting gas sensor performance. Gas sensor encounters many challenges related to the measurement environment variation, e.g., temperature, pressure or flow constant changes, cross interference to some other gases (low selectivity), poisoning. Besides, gas sensors designed for the process should provide reliable and stable output to be used in process control. Gas sensor should fulfill a “SSS” requirement, be Sensitive, Selective and Stable in process environment under regulation and control.

Combustion process regulation is mostly based on oxygen measurement and an improved combustion efficiency and stability can be achieved with the concurrent measurement of CO at ~100...200 ppm CO level. Monitoring of CO in the combustion process offers a new unique opportunity for better efficiency and safe flame operation with CO serving as a marker, indicating incomplete combustion. At combustion startup or in the event of a burner malfunction, the combustor could be filled with an explosive mixture faster than the response time of normal flame safety sensors installed in a combustor. An extra methane (CH<sub>4</sub>) or hydrocarbons (C<sub>x</sub>H<sub>y</sub>) measurement in a combustion analyzer would provide safety feature for combustion diagnostics and process control.

In this Chapter different gas sensing technologies will be discussed separately for oxygen (O<sub>2</sub>), carbon monoxide (CO) and methane (CH<sub>4</sub>) measurements, very important and sometimes critical for combustion process regulation and efficiency/safety control. With many gas sensing technologies for O<sub>2</sub>, CO and CH<sub>4</sub> measurements available on the market for environmental control only gas sensing technologies applied for combustion process regulation like high temperature electrochemical with zirconia solid electrolyte and optical using different infra-red lasers, proofed to be highly reliable in combustion process, will be discussed in details including application options.

## 12.1. Combustion Process

Combustion takes place when fuel reacts with the oxygen from the air to produce the heat used in the operation of the boilers, furnaces, kilns, and engines. The common combustion process products are carbon dioxide (CO<sub>2</sub>) and water (H<sub>2</sub>O) with the traces of other gases like sulfur dioxide (SO<sub>2</sub>) and nitrogen oxides (NO<sub>x</sub>) coming from fuel impurities and air nitrogen oxidation and carbon monoxide (CO) and soot (C) as products of incomplete combustion:



In the post combustion process or secondary combustion, unburned fuel (C<sub>x</sub>H<sub>y</sub>), carbon monoxide (CO) and soot (C) will react with oxygen to complete combustion:



The point of perfect combustion, referred to a stoichiometric point, in real combustion systems (Fig. 12.1, red solid line) can't be achieved because of non perfect gases mixing, fuel/air flow rates changes, and many other variables in combustion.

Oxidizing combustion process with oxygen excess causes heat losses and pollution, resulting in a shorter burner life. Reducing combustion with excess fuel or lean oxygen produces a sooty emission, greatly reducing the burner life span. For every 1 % of excess oxygen in combustion process, the fuel needed to produce the same heat should be increased by 1-3 % (Fig. 12.2). Monitoring and regulating some of the gases in flue gas or exhaust would highly contribute to combustion efficiency improvement.

Combustion control is mostly accomplished with oxygen measurement alone [1-2] and an improved combustion efficiency and stability can be achieved with the concurrent measurement of carbon monoxide at ~100...200 ppm CO levels (Fig. 12.3) and ~1...6 % of oxygen regulation, depending on the fuel type. CO monitoring in the combustion process offers a new unique opportunity for safe flame operation with CO serving as a marker, indicating incomplete or fuel-rich malfunction [2-4].

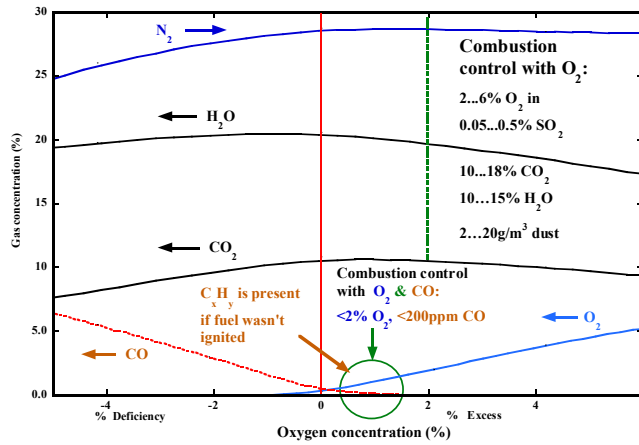


Fig. 12.1. Combustion flue gas diagram with  $O_2/CO$  control and  $C_xH_y$  detection for safety.

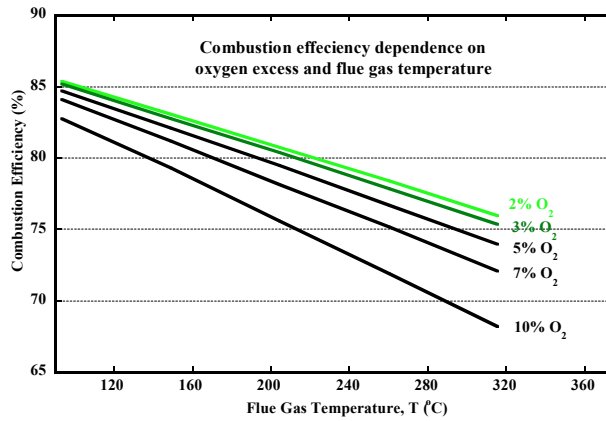


Fig. 12.2. Combustion efficiency dependence on oxygen excess and flue gas temperature.

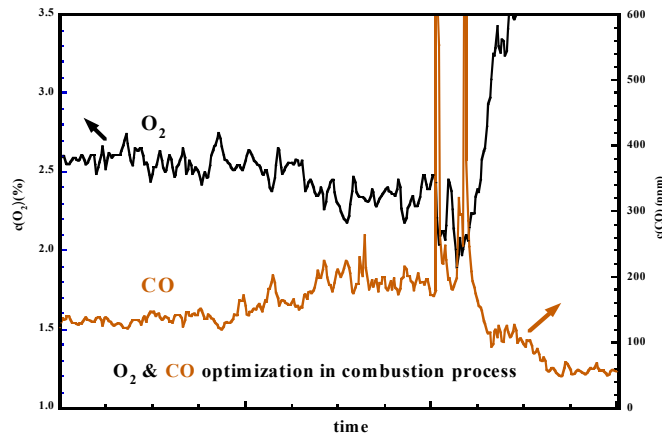
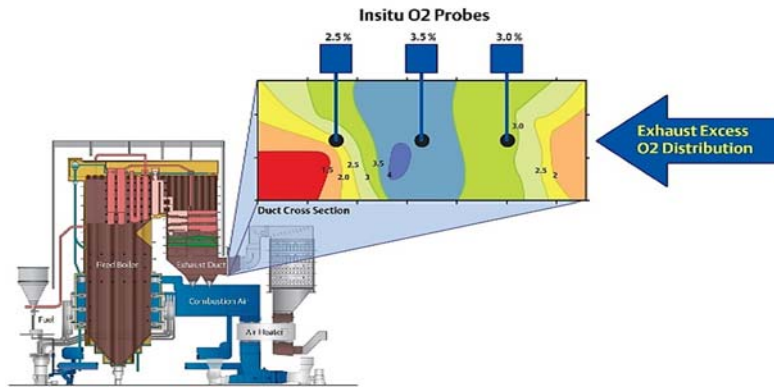


Fig. 12.3. Oxygen and carbon monoxide optimization in combustion control at coal fired Power Plant [4].

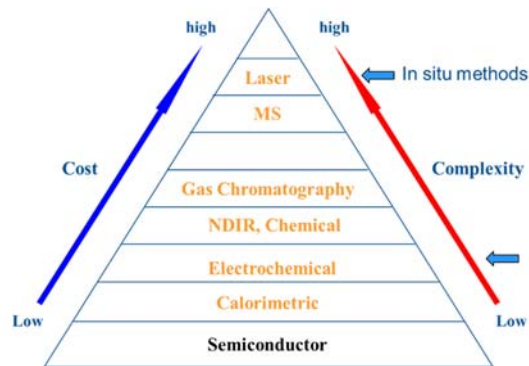


**Fig. 12.4.** Combustion process control based on oxygen measurement and flue gas stratification [3].

## 12.2. Gas Sensing Technologies for Combustion Process Control

There are many gas sensing technologies available for combustion process monitoring (Fig. 12.5) and these technologies complexity, reliability, sensitivity, selectivity, stability in combustion environment were carefully investigated before application in analytical instrumentation for combustion regulation.

Only a few of these gas technologies, such as electrochemical high temperature sensors, calorimetric sensors and laser techniques, have proved to be reliable for extractive, in situ, or across stack flue gas application.



**Fig. 12.5.** Gas sensing technologies for combustion control [3].

## 12.3. Oxygen Gas Sensing Technologies in Combustion

The oxygen sensors employed under ambient conditions include Clark type cells [5], paramagnetic [6], and optical sensors [7]. The high temperature application limits oxygen

gas sensing technologies to electrochemical (potentiometric or amperometric) sensing technologies [8] and resistive semiconductor sensors [9]. Potentiometric sensors measure the equilibrium voltage of the cells according to the Nernst equation and would require the reference. The amperometric oxygen sensor with the oxygen diffusion barrier limiting the current and fixed applied voltage can be only applied in very clean environment and was discussed in detail elsewhere [10-11]. The resistive oxygen sensor is based on semiconductors like doped  $\text{TiO}_2$  with the defect concentrations varying exponentially with the oxygen concentration [12].

There are three major oxygen gas sensing technologies high temperature zirconia potentiometric [13-15], tunable diode laser spectroscopy, TDLS [16-19], and paramagnetic [20], applied in-situ or in extractive mode for combustion process.

### 12.3.1. Zirconia Oxygen Potentiometric Technology

Zirconia potentiometric oxygen gas sensing technology, widely used nowadays in oxygen analyzers for different industrial combustion applications and in automotive Lambda sensors, was invented in 1961 by Peters and Möbius [21] and Weissbart and Ruka [22]. This technology was accepted in the power industry, traditionally more conservative regarding the new technology applications. The first industrial zirconia oxygen analyser for the process gas application was developed in the early 70s by Westinghouse Electric Co (Rosemount Analytical Inc.) and was based on the advanced platinum cermet electrode and zirconia solid electrolyte technologies developed initially for the application in solid oxide fuel cells. All industrial zirconia oxygen sensors are based on an electrochemical cell with yttrium stabilized zirconia (YSZ) solid electrolyte and two platinum based electrodes printed and sintered on the opposite sides of the zirconia ceramic and exposed to the process and reference gases [23]. The process side of the cell is gas tight separated from the reference side with the special high temperature ceramic/metal/glass seals, and zirconia ceramic is used in the form of disc (Fig. 12.6), tube, or thimble.

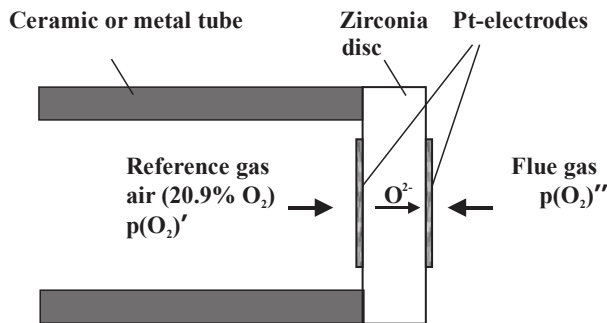


Fig. 12.6. Zirconia potentiometric oxygen gas sensing cell [15].

Differential oxygen chemical potentials on the oxygen cell electrodes would develop electromotive force (EMF),  $E$ , according to the Nernstian equation:



$$E = \frac{RT}{4F} \ln \frac{p(O_2)_{\text{Process}}}{p(O_2)_{\text{Ref}}}, \quad (12.5)$$

with R universal gas constant, T the process temperature in Kelvin (K) and F the Faraday number.

By using fixed oxygen partial pressure on the reference electrode (e.g., air with 20.9 % O<sub>2</sub> and  $p(O_2) = 2.1 \times 10^4$  Pa), the sensor signal of the thermally balanced oxygen sensor will be only depending on the sensor temperature (Fig. 12.7). Oxygen potentiometric gas analysis was established by H.-H. Möbius [24] permits measurements of chemically bonded oxygen in very small concentrations <1ppm (part per million) or < 1 ppb (parts per billion).

Industrial zirconia oxygen sensors are operating at elevated temperatures with oxygen ion migration in zirconia ceramics establishing oxygen equilibrium at the process and reference electrodes and zirconia electrolyte interfaces at >300 °C depending on the electrode composition and microstructure. Over 1400 °C oxygen electrolytic permeability of the zirconia solid electrolyte ceramic might compromise oxygen measurements. Higher application temperature will favor oxygen sensor performance but would be highly limiting the high temperature ceramic/metallic materials applied in the zirconia oxygen sensor packaging. Most industrial zirconia oxygen sensors are operating at elevated temperatures around 700...800 °C. Disc-shape zirconia ceramics brazed into special metal alloy supporting tube initially developed by Rosemount Analytical Inc. offer an advantage of the much less differential temperature profile across the cell on the reference and process sides compared to the tube or thimble ceramic shape and less thermal stress considering a good thermal expansion coefficients (TEC) optimization of all involved in the oxygen sensor packaging metallic and ceramic materials [15].

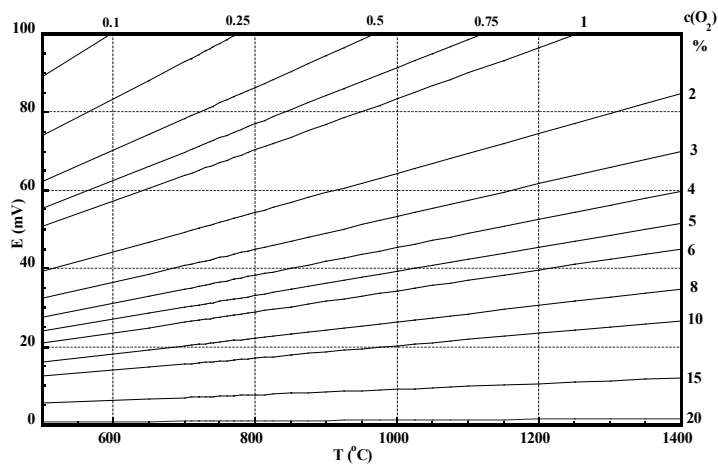


Fig. 12.7. Potentiometric Zirconia oxygen sensor signal-concentration-temperature diagram [15].

A special oxygen cell with improved thermal balance and a ceramic or metallic diffuser for the process gas particulate filtration was designed to improve the O<sub>2</sub> analyzer, delivering an excellent stability (< 0.1 % O<sub>2</sub>) and reproducibility ( $\pm 0.01$  % O<sub>2</sub>) in 2...5 % oxygen combustion concentration range (Fig. 12.8).

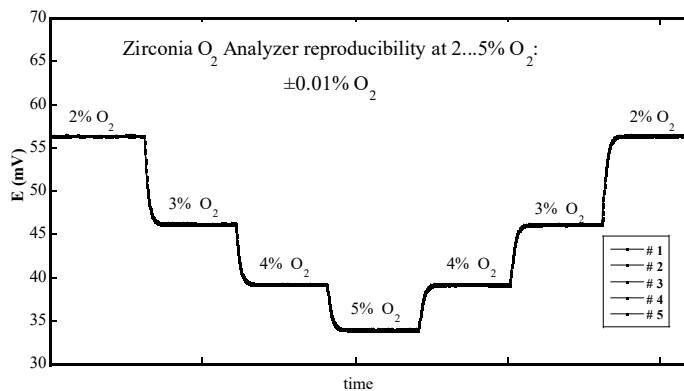


Fig. 12.8. Zirconia O<sub>2</sub>-Analyzer reproducibility.

There have been numerous attempts to miniaturize oxygen sensor with innovative sealed chamber reference and oxygen pump [25], glass sealed metal/metal oxide reference electrode [26-27], but the evaluation couldn't confirm these design reliabilities in traditionally difficult stack gas or waste combustion gas analysis. After 60 years, zirconia potentiometric oxygen gas sensing technology is still dominating the combustion market because of the high reliability, accuracy in the most of applications and a relative low price. Ametek, Rosemount (Emerson Electric Corp.), Yokogawa, and ABB are the major players on combustion in-situ oxygen measurements market. The performance of Zirconia O<sub>2</sub>-Analyzers on the market is summarized in Table 12.1.

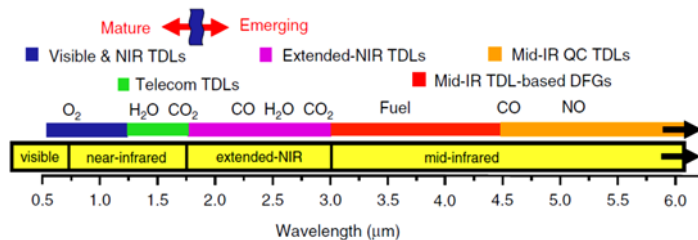
### 12.3.2. Tunable Diode Laser Spectroscopy (TDLS)

Tunable diode laser spectroscopy (TDLS) is an innovative optical measurement technique utilizing semiconductor lasers to detect a variety of gases (Fig. 12.9), including oxygen in the near infrared (IR) range to deliver part per millions (ppm) level measurements [16-18]. Over the last ~50 years, tunable diode laser (TDL) absorption sensing has become an established method for non-intrusive gas measurements in combustion environment due to the robust, compact, low-cost TDL sources, and the convenient overlap of TDL wavelengths with absorption bands of O<sub>2</sub>, CO, NO<sub>x</sub>, H<sub>2</sub>O, CO<sub>2</sub> combustion gases [19]. TDLS with the ability to interface directly with the process gas and eliminating the need for the very high cost and high maintenance sample handling systems would be ideal for the real time dynamic measurement of the process conditions. With federal, state and private companies multi-billion investments, TDL analyzers have now comprehensively established themselves as a core measurement technology within the portfolio of the gas analysis techniques available on the market [18-19].

**Table 12.1.** In-situ Zirconia O<sub>2</sub>-Analyzers performance and materials [15]

Company	In-situ O <sub>2</sub> -probe	Accuracy	Operation temperature	Materials
Rosemount (USA)	6888 O <sub>2</sub>	±0.75 % or ±0.05 % O <sub>2</sub>	704 or 825 °C	316L SS E-Brite
Yokogawa (Japan)	ZR202G	±1.0 % error	up to 700 °C	316 SS Hastelloy
ABB (Swiss)	Endura Z20	±0.75 % or ±0.05 % O <sub>2</sub>	up to 800 °C	316L SS
Ametek (USA)	WDG In-situ flue	±1.0 % or ±0.05 % O <sub>2</sub>	up to 677...800 °C	310 SS
Enotec (Germany)	Oxitec 5000	±0.2 % of MV*	up to 800 °C 1400°C (with cooling tube)	316 SS
Fuji (Japan)	ZKF8	±0.5 % of FS**	up to 800 °C	316 SS 304 SS
GE-Panametric (USA)	FGA311	±3 % or ±0.1 % O <sub>2</sub>	up to 650 °C up to 1050 °C (HT)	316 SS Inconel
Energy Support (Japan)	TF-10	±1.0 % of FS**	up to 300 °C	305 SS 308 SS
LamTec (Germany)	LS2	±2.0 % or ±0.2 % O <sub>2</sub>	up to 300 °C	305 SS 308 SS
Zirox (Germany)	SS28	±5.0 %	up to 800 °C	1.4841 (314 SS)
ETC (UK)	ETC 2800	±1 % of MV*	up to 540 °C	305 SS 308 SS
Ceramic Oxide Fabricators (Australia)	JS probe	±0.1 % O <sub>2</sub>	up to 600 °C	316 SS
Novatech (Australia)	1231	±1.0 %	up to 800 °C	316 SS

\* MV - Measured Value, \*\* FS - Full Scale, SS – Stainless Steel

**Fig. 12.9.** Diode laser types and combustion gases vs. wavelength [19].

The infrared absorption spectrum of the oxygen is like a fingerprint, providing oxygen identification at ~760 nm (Fig. 12.10) and measurements in a wide concentration range down to ~80...100 ppm as a detection limit. With the longer optical path length, a higher

absorption and better sensitivity will be achieved, laser wavelength can be changed over a wide range by adjusting laser operation temperature and over a narrow range by changing the applied current to the laser.

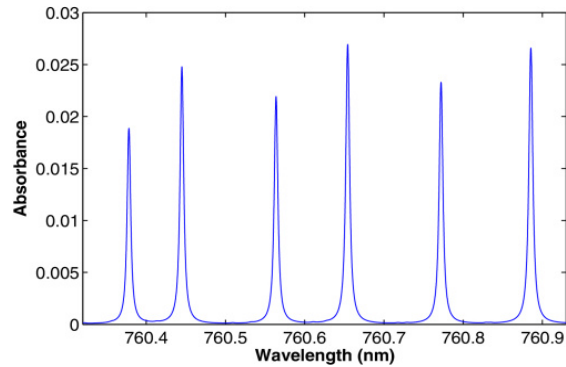


Fig. 12.10. Partial spectrum of oxygen at 760 nm [28].

TDL analyzers have been shown to be capable of the effective and accurate oxygen measurement even losing up to 90-95 % of the original signal in the dusty environment. TDL spectroscopy is highly distinguishable from the conventional process photometry by the laser ability to be scanned across the narrow oxygen absorption peaks many times per second. While the conventional optical filter bandwidth is  $>5$  nm, a laser with 0.2...0.3 nm typical range scan would provide much better resolution and selectivity [29]. More details on TDLS theoretical background and applications can be found in the latest reviews [19, 30-33]. Performance and application options of some of TDL  $O_2$  analyzers on the market are summarized in Table 12.2.

Table 12.2. TDL  $O_2$  analyzers performance and application options [15].

Company	Product	Detection limit	Accuracy	Application
Yokogawa (Japan)	TDLS8000	Application dependent	$\pm 1.0$ % or $\pm 0.01$ % $O_2$	Across the duct
Siemens (Germany)	SITRANS SL	0.02 % $O_2$	$\pm 1.0$ %	Across the duct
ABB (Swiss)	LS4000	0.05 % $O_2$	$\pm 0.2$ %	Across the duct
Ametek (USA)	<b>TDLAS 5100</b>	0.01 % $O_2$	$\pm 1.0$ % or $\pm 0.03$ % $O_2$	Extractive
Servomex (UK)	SERVOTOUGH MiniLaser Oxy	0.01 % $O_2$	$\pm 1.0$ % or $\pm 0.01$ % $O_2$	Across the duct
Sick Maihak (Germany)	TRANSIC151LP $O_2$	0.01 % $O_2$	$\pm 0.2$ % $O_2$	In-situ probe (<80 °C)
Mettler-Toledo (USA-Swiss)	TDLS GPro 500	0.01 % $O_2$	$\pm 2.0$ % or $\pm 0.01$ % $O_2$	In-situ probe

Table 12.2. Continued.

Company	Product	Detection limit	Accuracy	Application
Fuji Electric (Japan)	ZSS6 O <sub>2</sub> /CO Analyzer	N/A	±2.0 % of FS	Across the duct
NEO Monitors (Norway)	LaserGas III SP	0.01 % O <sub>2</sub>	±1 %	Across the duct
Focused Photonics Inc. (China)	LGA-4100 (in-situ) LGA-4500 (extractive)	0.01 % O <sub>2</sub>	±1 %	Across the duct & extractive
TDL Sensors Ltd (UK)	LGM-CD	0.05 % 20 ppm O <sub>2</sub>	N/A	Across the duct
Zolo Technologies (USA)	ZoloBoss	0.01 % O <sub>2</sub>	±0.5 %	Across the duct
Oxigraf Inc. (USA)	O2iL, O2iC, O2E, O2E2, O2C, O2D, O2L	0.01 % O <sub>2</sub>	±0.2 %	Extractive

### 12.3.3. Paramagnetic Oxygen Gas Sensing Technology

Oxygen with two unpaired electrons is a well-known paramagnetic gas and would be attracted into a strong magnetic field. These paramagnetic properties were implemented for relatively selective oxygen measurements method used by industries for over fifty years and applied in three different methodologies: magnetodynamic, thermomagnetic, or magnetic wind and non-contact Quinke method [20].

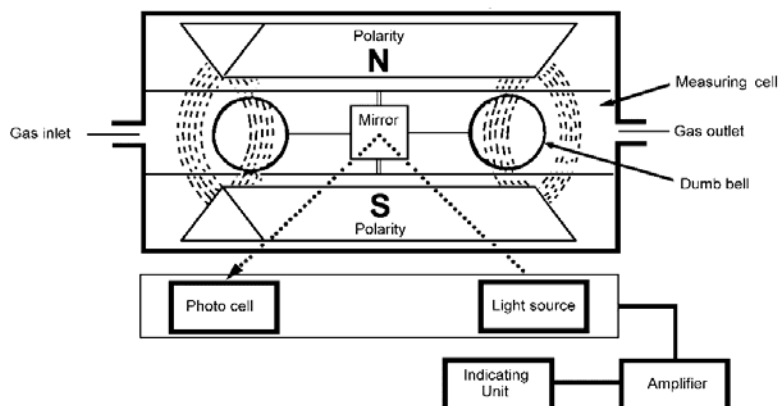


Fig. 12.11. Paramagnetic O<sub>2</sub>-sensor schematically.

The magnetodynamic oxygen sensors are dumbbell or of the magnetic pressure type. In the dumbbell type, the O<sub>2</sub>-sensor incorporates two nitrogen-filled glass spheres mounted

on the rotating suspension (Fig. 12.11). This assembly is in a strong magnetic field and with the gas sample containing oxygen introduced into the magnetic field, the oxygen molecules will tend to approach the strongest part of the magnetic field pushing the dumbbell aside and causing the suspending wire to be twisted. This twist is detected as the movement of the light from the reflective mirror located at the suspending wire center. The output from the photo-sensor would be fed back to a coil around the suspension assembly restoring torque to the original “zero” position, and current would be directly proportional to the magnetic susceptibility of the sample oxygen gas and oxygen partial pressure. The applied current will be proportional to the oxygen concentration in the gas mixture and would be used for the precise oxygen calculation. This method with the high linearity and good response is however highly sensitive to the different mechanical shocks like vibration.

In the magnetic pressure type, the O<sub>2</sub>-sensor has a small amount of reference gas like pure nitrogen or air is externally supplied through a small hole made on one of the magnetic poles. When the sample gas is introduced into an uneven magnetic field, paramagnetic oxygen gas would be attracted toward the strongest part of the magnetic field, in which the magnetic poles have come closer to each other. The increase in pressure resulting from the attraction of oxygen can be detected with a capacitor microphone type detector or a mass flow sensor. In order to ensure the stable signals detection, the two magnetic poles are alternately excited to amplify the alternating current. The selectivity of this method is pretty high, but a reference gas would be required for these measurements.

Thermomagnetic O<sub>2</sub>-sensor (magnetic wind method) has two chambers with a heating wire element located at the centre. A magnetic field would be provided only on the measurement chamber. After the analysing oxygen gas enters the measurement chamber, oxygen would be attracted by the magnetic field and then heated using the heating wire element. The magnetic oxygen flow would cool the heating wire element as its intensity varies proportional to the oxygen concentration. The resistance changes during this process would be picked up as an unbalanced voltage at the Wheatstone bridge to measure the oxygen concentration. The magnetic wind type analyzer has no movable parts and is more stable in the vibration environment. Carbon dioxide with some positive interference for oxygen by the small variation in the concentration can be compensated.

Quincke O<sub>2</sub>-detector utilizes a pneumatic Wheatstone bridge for the measurements using a differential pressure or flow to determine the oxygen concentration. A reference gas is used so the sample does not come in contact with the sensor. The reference gas is divided into two flow paths recombining at the reference gas outlet, where the sample gas is also introduced. A magnetic field is in one arm of the reference gas outlet creating a back pressure related to the presence of the oxygen in the gas sample.

Measurement range for the paramagnetic O<sub>2</sub>-sensor is typically between 0.05 % and 100 % O<sub>2</sub> and sensor can be very precise under certain conditions [20] but would require quite expensive gas sample conditioning system to stabilize the process gas flow, pressure, and temperature and to remove hazardous gases and dust. The major disadvantages of the paramagnetic oxygen sensors are the price, low stability in the vibration environment, and relative low accuracy ( $\pm 0.10 \dots 0.25$  % O<sub>2</sub>). Paramagnetic analyzers have been historically

limited to the extractive application with general use or intrinsically safe instrumentation version, and maximum external operation temperature up to 120-145 °C.

## 12.4. Carbon Monoxide (CO) Gas Sensing Technologies in Combustion

Combustion control is mostly accomplished with O<sub>2</sub> measurement alone and an improved combustion efficiency and stability can be achieved with the concurrent measurement of carbon monoxide, CO. Operation at ~100...200 ppm CO levels and ~1-6 % of oxygen depending on the fuel type would indicate conditions near the stoichiometric point with the highest efficiency in safe operation mode (Fig. 12.1).

There are three comprehensive carbon monoxide gas sensing technologies implemented so far in combustion analyzers on the market: calorimetric catalytic, mixed potential electrochemical and tunable diode laser spectroscopy (TDLS), or quantum cascade laser spectroscopy (QCLS) [2].

### 12.4.1. Calorimetric Catalytic CO Gas Sensing Technology

For the extractive type combustion analyzers, an advanced calorimetric CO-sensor was developed by several analytical corporations utilizing combustible reaction heat on a special selective catalyst film over the resistance temperature detector (RTD) implemented for the temperature measurements [2, 34-35]. The catalyst material in calorimetric CO-sensor was initially based on *Hopcalite*<sup>®</sup>, a commercial catalyst, mixture of copper and manganese oxides. Unfortunately, this catalyst showed a very short life in the severe combustion process environment and was replaced by a more reliable platinum group catalyst using a glass coating for protection in process containing sulfur oxides. The catalyst film was merged with RTD using highly thermally conductive packaging oxide materials to utilize the heat transfer of the carbon monoxide combustion reaction in the flue gas to RTD more efficiently. Catalyst film was printed on a metallic or ceramic substrate thermally connected to the RTD. The catalytic film oxidizes carbon monoxide selectively in the presence of other flue gas species. Developed calorimetric catalytic CO-sensors on the market are quite selective, and cross sensitivity of a typical CO-sensor is summarized in Table 12.3.

**Table 12.3.** Calorimetric catalytic CO-sensor typical cross-sensitivity [2].

Gas of interest	H <sub>2</sub>	O <sub>2</sub>	SO <sub>2</sub>	NO <sub>x</sub>	CH <sub>4</sub>	H <sub>2</sub> O
Selectivity	2 : 1	1 : >10,000	1 : >1000	1 : >1000	1 : 20,000	1 : >1000

Oxidation reaction occurs between combustible species and oxygen with defined heat released depending highly on the combustible concentration. Calorimetric catalytic CO-sensor can measure CO-concentration down to ppm level with a detection limit ~5...10 ppm CO. This technology is employed in Rosemount, Emerson Electric Corp.

(USA), Servomex (GB), Ametek (USA) and ABB (Switzerland) extractive combustion analyzers permitting fast and precise CO measurements (Fig. 12.12) with  $\pm 25$  ppm error or 5 % of the reading in a process flue gas environment up to  $\sim 1750$  °C temperature.

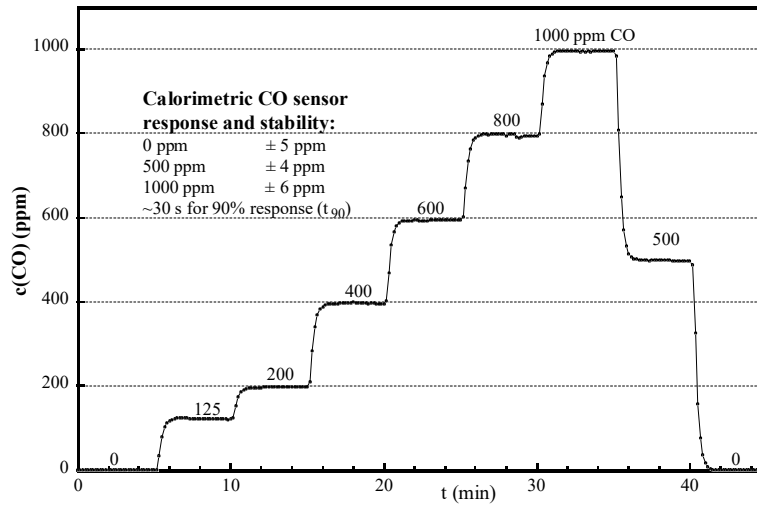


Fig. 12.12. Calorimetric catalytic CO-sensor response and stability [2].

Calorimetric CO-sensor has good,  $>99.9$  % linearity in a wide CO concentration range up to 2.5 % CO (Fig. 12.13) but to improve accuracy a calibration would be always recommended using the upper limit of CO concentration measurement range.

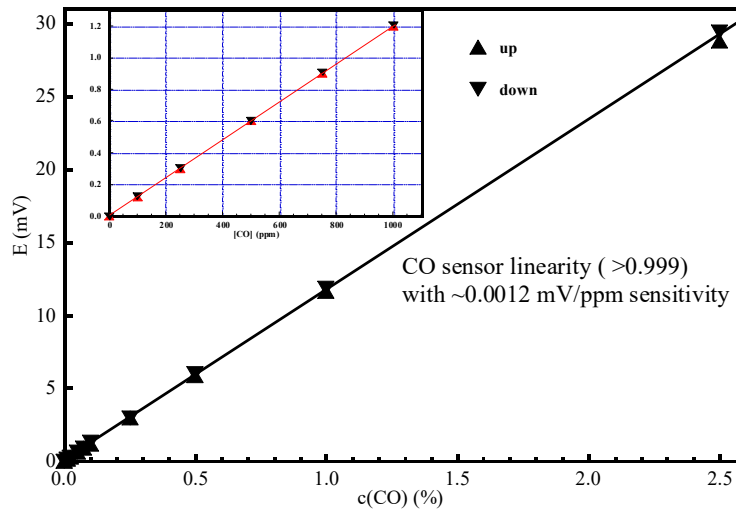
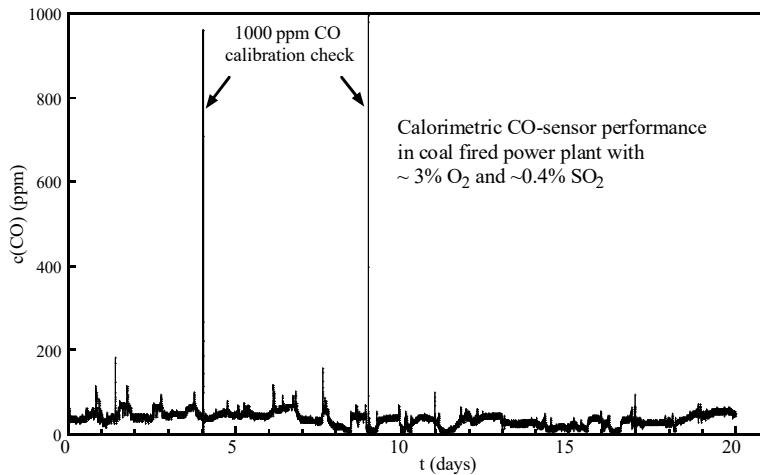


Fig. 12.13. Calorimetric catalytic CO-sensor linearity [2].





**Fig. 12.14.** Calorimetric CO-sensor evaluation in combustion analyzer at coal fired power plant [2].

Calorimetric catalytic CO-sensor is highly flow sensitive and gas flow stabilization and a flow sensor should be implemented to improve CO-sensor reliability in extractive combustion analyzer.

An industrial combustion analyzer is always measuring carbon monoxide equivalent,  $\text{CO}_e$  proven to be sufficient for the combustion process control (Fig. 12.14). Taking into consideration the flow and temperature sensitivity of the catalytic CO-sensor, gas temperature conditioning of incoming flue gas and precise internal flow and temperature control in the analyzer should be implemented. An in-situ process application for this type of sensor was proven to be much less precise and should be avoided because of the flow and temperature fluctuation in the process, affecting catalytic CO combustion reaction speed and compromising sensor reading.

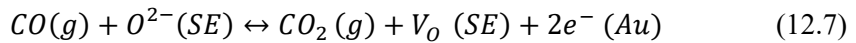
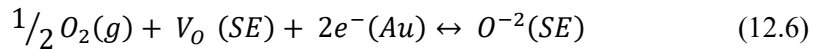
#### 12.4.2. Mixed Potential Electrochemical CO Gas Sensing Technology

Mixed-potential gas sensing technology employing oxygen ion conducting zirconia solid electrolyte and two different activity electrodes was invented by Sandler in 1976 [36] and was proposed for the detection of combustible gases. The presence of CO or any other combustible gas effects well known zirconia oxygen sensor (see Chapter 12.3.1.) with two competing reactions on electrodes. Oxygen reduction and CO or other combustible gas oxidation would establish a mixed potential highly dependent on CO concentration (Fig. 12.15). This developed mixed potential depends mostly on the electrodes electrochemical and catalytic activities, film morphology, and temperature. Various metals and metal oxide materials were investigated for sensitive electrodes in the mixed potential sensors in the past [37-38], and the gold metal was one of the most promising sensitive electrodes because of its low catalytic activity for the combustion reaction and chemical stability in hazardous chemicals environment.

	Catalytically Active Equilibrium Electrode (Pt)	Catalytically Inactive Mixed potential Electrode (Au)
<b>Models</b>		
<b>Reactions</b>	$CO(g) + \frac{1}{2} O_2(g) \leftrightarrow CO_2(g)$	$CO(g) + O^{2-}(YSZ) \leftrightarrow CO_2(g) + V_O^{2+}(YSZ) + 2e^-(Au)$ $\frac{1}{2} O_2(g) + V_O^{2+}(YSZ) + 2e^-(Au) \leftrightarrow O^{2-}(YSZ)$
<b>Equations</b>	$E_{eq} = \frac{RT}{4F} \ln \frac{p(O_2)_{process}}{p(O_2)_{ref}}$	$E_{mix} = E_{eq} - C * \ln(\varphi_{CO})$

Fig. 12.15. Mixed potential CO-sensor with two different catalytic activities electrodes [2].

The mixed potential value depends on two competing reduction/oxidation reactions on catalytically inactive Au-electrode:

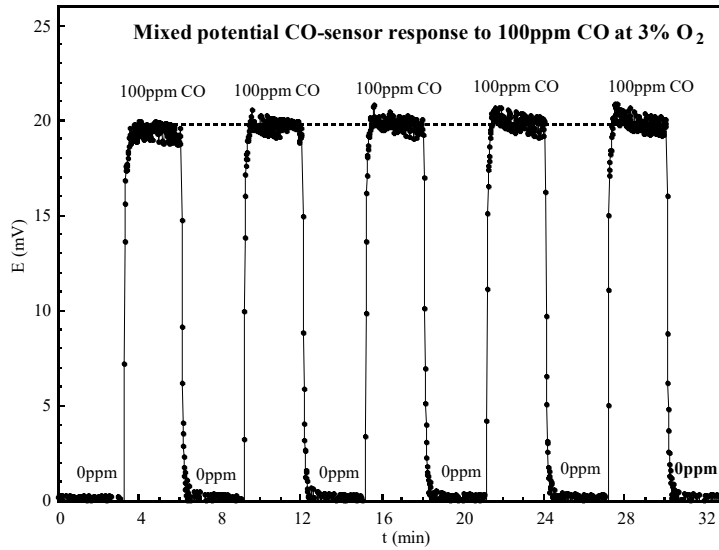


with  $V_O$  oxygen vacancy,  $O^{2-}$  oxygen ions in zirconia solid electrolyte (SE) and  $e^-$  electrons in Au- electrode. Mixed potential will be affected by changes in electrode triple phase boundary (TPB) area, limited diffusion in the gas phase or electrode, limited absorption and interference of other gases in combustion process, e.g.,  $H_2O$ ,  $CO_2$ ,  $SO_2$ ,  $NO_x$  and  $O_2$ . A mixed potential CO-sensor with catalytically inactive electrode based on Au-alloy produces a repeatable signal (Fig. 12.16) with significant  $\sim \pm 100$  ppm CO error impacted by oxygen concentration variation between 3 and 20 %  $O_2$  [2].

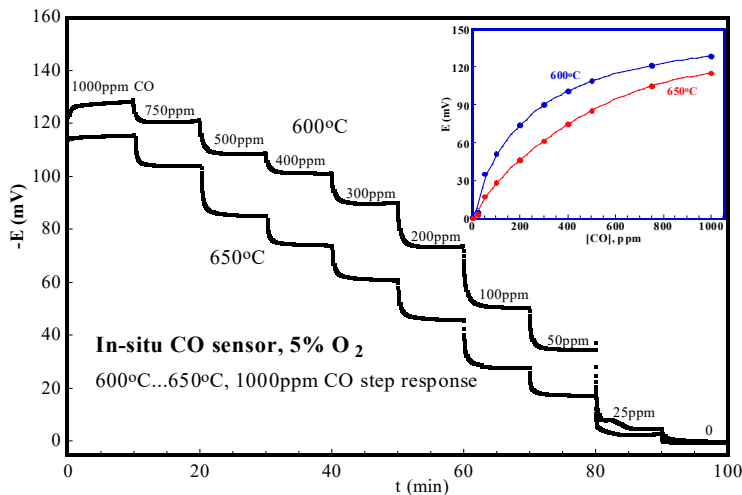
The gold film electrode re-crystallization and coarsening by temperatures over 500 °C will affect the electrochemical triple phase boundary and sensitivity of the sensor, limiting this electrode application at higher, over 500 °C temperature [38-39]. The CO-sensor with a gold composite electrode, especially with gallium oxide, exhibits a higher sensitivity to CO and long-term stability in a combustion process environment even at a higher temperature (Fig. 12.17).

Unfortunately, mixed potential CO sensor dependence on CO concentration is not linear and a special polynomial calibration algorithm should be developed based on evaluation in different combustion processes. An oxygen correction algorithm applied to mixed potential CO-sensor signal using an extra oxygen concentration measurement input would

improve CO measurement reliability in a challenging oxygen concentration environment [2]. Investigation with different catalytically-inactive electrodes based on Au-alloy or Au-composite has shown that sulphur dioxide ( $\text{SO}_2$ ) and methane ( $\text{CH}_4$ ) can highly compromise the mixed potential CO sensor reading (Figs. 12.18, 12.19).



**Fig. 12.16.** Mixed potential CO-sensor response and reproducibility [2].



**Fig. 12.17.** Mixed potential CO-sensor with Au-composite electrode response [2].

In many practical combustion applications, a removal of these interference gases using special chemicals, like lime for  $\text{SO}_2$ , would be required, or special gas filtering membranes or molecular sieves like zeolites [40-43] should be employed in sensor packaging to

reduce the interference. These membranes or molecular sieves would require an evaluation in combustion flue gas for their performance reliability and chemical stability in a highly challenging combustion temperature and chemical environment.

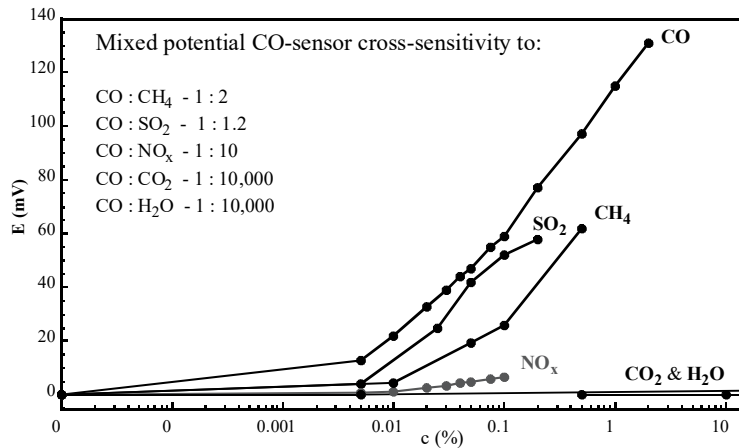


Fig. 12.18. Au-alloy based mixed potential CO-sensor typical cross-sensitivity [2].

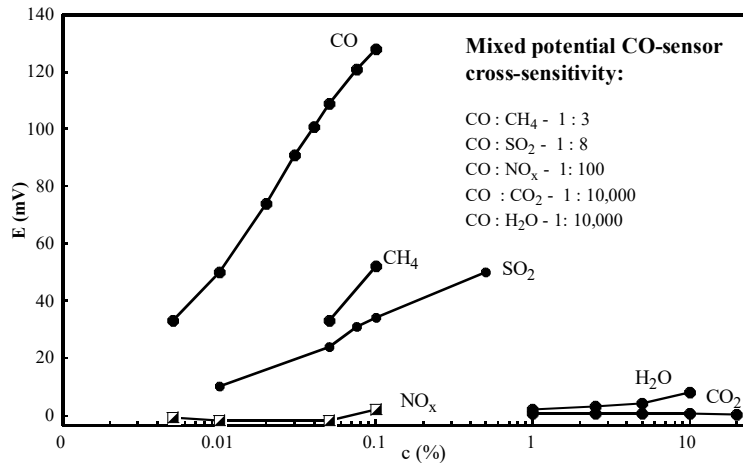
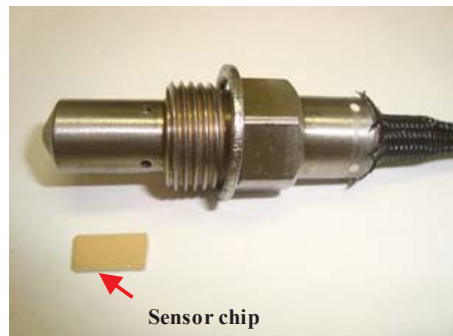


Fig. 12.19. Au-composite based mixed potential CO-sensor typical cross-sensitivity [2].

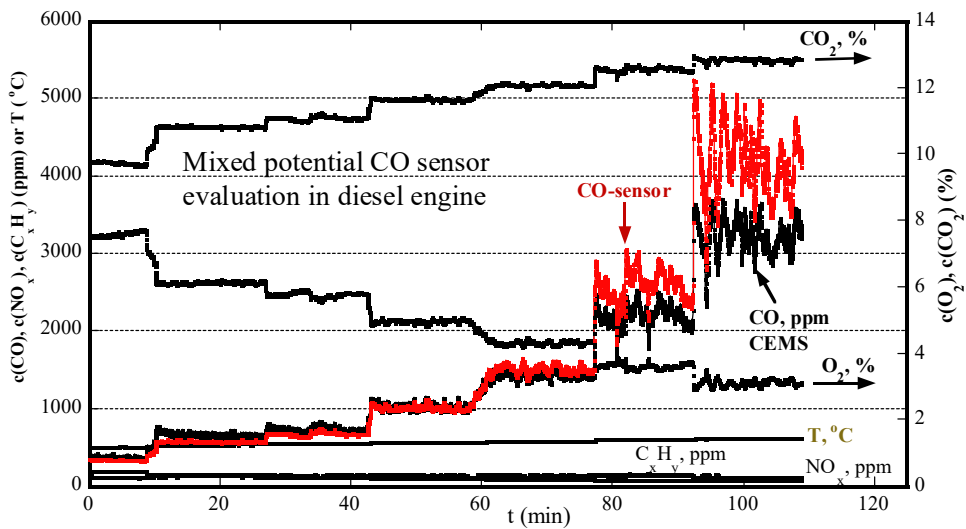
A mixed potential CO-sensor prototype with Au-composite electrode was packaged like a Lambda sensor (Fig. 12.20) using CO-sensor chip and mechanical electrical contacts. This CO-sensor was installed and evaluated in the Lambda sensor spot in the manifold of automotive engine Cummins ISF 109-167 PS with temperature recording and emission monitoring.

Extractive 5 channel Horiba MEXA 7500 Continuous Emission Monitoring System (CEMS) was used for O<sub>2</sub>, CO, CO<sub>2</sub>, NO<sub>x</sub>, C<sub>x</sub>H<sub>y</sub> measurements, and the mixed potential

in-situ CO-sensor was showing good agreement with an extractive CO measurement using CEMS (Fig. 12.21).



**Fig. 12.20.** Mixed potential CO-sensor prototype with Au-composite electrode [2].



**Fig. 12.21.** Au-composite based mixed potential CO-sensor evaluation in automotive engine [2].

Based on evaluation results, mixed potential CO-sensor has practically negligible cross-interference to  $\text{CO}_2$  or  $\text{H}_2\text{O}$  and quite low interference with  $\text{O}_2$  when considering combustion flue gas oxygen concentration range to be between 2 and 5 %  $\text{O}_2$ , depending on the fuel type used for combustion. Based on the  $\text{O}_2$ -sensor input signal, a correction related to the oxygen concentration using special algorithm could be implemented, improving the reliability of CO-sensor. Cross-sensitivity to  $\text{SO}_2$  and  $\text{CH}_4$  can be reduced by materials and microstructure optimization but this interference is still quite high to provide very reliable CO measurements in high concentration  $\text{SO}_2$  environment. Based on laboratory investigation, an algorithm related to  $\text{SO}_2$  and  $\text{H}_2\text{O}$  interference was developed

and with some success was evaluated at a coal fired power plant with up to 0.2 % SO<sub>2</sub> concentration [2].

There are two combustion analyzers on the market from Enotec GmbH (Germany) and Lamtec GmbH (Germany) with an implemented mixed potential CO-sensor providing  $\sim\pm 25$  % of the reading or  $\pm 5$  % of full range in-situ carbon monoxide equivalent, CO<sub>e</sub>, measurements in very clean gas or light oil combustion applications.

### 12.4.3. Tunable Diode Laser (TDL) and Quantum Cascade Laser (QCL) Spectroscopies

Tunable Diode Laser Spectroscopy (TDLS) and Quantum Cascade Laser Spectroscopy (QCLS) are innovative optical measurement techniques utilizing semiconductor lasers to detect a variety of gases in the near infrared (IR) range. The second overtone band of CO near  $\sim 1.57$   $\mu\text{m}$  has been utilized commonly for measurements using inexpensive telecommunication type diode lasers. However, given weak line strengths and interference from carbon dioxide (CO<sub>2</sub>) and moisture (H<sub>2</sub>O) always present in the combustion process, CO near  $\sim 2.3$   $\mu\text{m}$  would provide a better detection limit, free of interference.

Specialty TDLs and Quantum Cascade Lasers (QCLs) have been utilized for CO-measurements. The IR absorption spectrum of carbon monoxide (CO) is like a fingerprint, providing CO identification at 1.57  $\mu$ , 2.33  $\mu$  (TDL), or 4.8  $\mu$  (QCL) and measurements in a wide concentration range down to  $\sim 0.5$  ppm CO (TDL) or even  $\sim 10$  ppb CO (QCL) as a detection limit [2].

TDL CO analyzers could be used in different configurations, e.g., across the duct, extractive or an in-situ measurement using a probe. Depending on the application, different porosity filters can be employed at the probe end to increase analyzer sensitivity and reliability by reducing interference with particulates in the combustion process. Temperature and pressure variation in the process can introduce extra error in the measurement and requires compensation. There are some challenges with in-situ calibration for the systems employing an internal short reference cell. IR light reflection in a high temperature environment combined with wide background radiation from the fire box and process windows fouling might bring additional challenge in the application and might reduce the service life of the analyzer. There are many quite accurate TDL/QCL Analyzers on the market (Table 12.4), developed by Yokogawa (Japan), NEO Monitors (Norway), Focused Photonics Inc. (China), Rosemount (Emerson Electric Corp., USA), Mettler-Toledo (USA), Fuji Electric (Japan), Siemens (Germany), Servomex (GB) and Sick Maihak (Germany) with  $\sim\pm 1$  % (2 % for in-situ) error for CO measurements. These analyzers are providing often dual CO and O<sub>2</sub> measurements in a combustion environment across the duct or in-situ with CO detection limit being  $<1$  ppm CO.

However, TDL Analyzers are quite expensive and considering combustion process highly challenging temperature, pressure variation, and flue gas turbulent flow, an in-situ CO calibration or validation would be highly recommended for all these analyzers.

**Table 12.4.** TDL and QCL CO Analyzers on the market [2].

Company	Products	Detection limit	Application
Yokogawa (Japan)	TDLS8000	Application dependent	Across duct or extractive
Siemens (Germany)	SITRANS SL (O <sub>2</sub> & CO)	0.6 ppm	Across duct or extractive
Focused Photonics Inc. (China)	LGA-4100 LGA-4500 LGA-3500	0.6 ppm	Across duct or extractive
NEO Monitors (Norway)	LaserGas II SP LaserGas III SP	0.3 ppm	Across duct or extractive
Servomex (GB)	SERVOTOUGH LaserSP	0.3 ppm	Across duct or extractive
Sick Maihak (Germany)	GM901 CO	N/A	Across duct In-situ (<430 °C)
Rosemount (USA)	CT5800	0.05 ppm	Extractive

## 12.5. Methane (CH<sub>4</sub>) Gas Sensing Technologies in Combustion

At combustion startup or in the event of a burner malfunction or flame blowout, the combustor could be filled with an explosive mixture faster than the response time of normal flame safety sensors installed in a combustor. An extra methane (CH<sub>4</sub>) or hydrocarbons (C<sub>x</sub>H<sub>y</sub>) measurement in a combustion analyzer [4] would provide an extra safety feature for combustion control and diagnostics (see combustion flue gas diagram, Fig. 12.1), also increasing combustion efficiency (Fig. 12.3).

### 12.5.1. Mixed Potential Electrochemical CH<sub>4</sub> Gas Sensing Technology

Mixed-potential gas sensing technology with an oxygen ion conducting zirconia solid electrolyte and two different activity electrodes invented by Sandler [36] was proposed for the detection of combustible gases. The presence of CH<sub>4</sub> or any other combustible species will affect the electrochemical oxygen sensor signal with competing reactions of oxygen reduction and CH<sub>4</sub> or other combustible species oxidation, establishing mixed potential highly depending on the CH<sub>4</sub> concentration.

Mixed potential developed by this CH<sub>4</sub> sensor would depend mostly on kinetic factors and would be a strong function of the process electrode electrochemical and catalytic activities and morphology. The sensor signal will be affected by changes in the gas/electrolyte/electrode triple phase boundary (TPB) area, limited methane or oxygen diffusion in the gas phase, electrodes and TPB, limited adsorption, and other combustion gas interference. The fact that the response is controlled by reaction kinetics places a limit on a maximum operating temperature for the device in order to achieve a usable signal-to-noise in the application. Various metals and metal oxide materials were investigated for sensitive electrodes in the mixed potential sensors in the past [37], and a special signal

conditioning pulsed discharge technique was proposed to increase sensor sensitivity and selectivity [44-45].

Developed mixed potential  $\text{CH}_4$  sensor with tin-doped indium oxide sensitive electrode was very sensitive with a detection limit  $\sim 25$  ppm  $\text{CH}_4$ , [46]. Unfortunately, mixed potential sensors have not been especially selective with strong interference from other hydrocarbons ( $\text{C}_x\text{H}_y$ ), nitrogen oxides ( $\text{NO}_x$ ), carbon dioxide ( $\text{CO}_2$ ) and carbon monoxide ( $\text{CO}$ ) (Fig. 12.22).

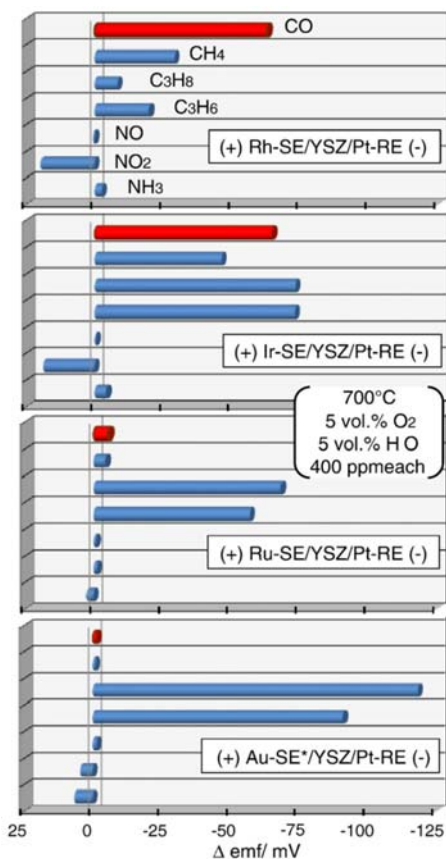


Fig. 12.22. Selectivity of mixed potential sensors with Rh, Ir, Ru and Au sensitive electrodes [4].

Recent efforts to improve mixed potential technology have focused on improving electrochemical interface morphology in order to permit fabrication of stable and reproducible devices.

### 12.5.2. Calorimetric Catalytic $\text{CH}_4$ Gas Sensing Technology

For extractive combustion analyzers, advanced calorimetric catalytic  $\text{CO}$  sensors were developed by several analytical corporations utilizing combustible reaction heat on the



special selective catalyst film over the resistance temperature detector (RTD) implemented in the sensor package for temperature measurements (Chapter 12.4.1). Conventional catalysts for methane combustion are based on noble metals (e.g, palladium or platinum), deposited on oxides such as cerium oxide ( $\text{CeO}_2$ ) or metals. Methane's complete combustion reaction is highly exothermic (-890 kJ/mol):

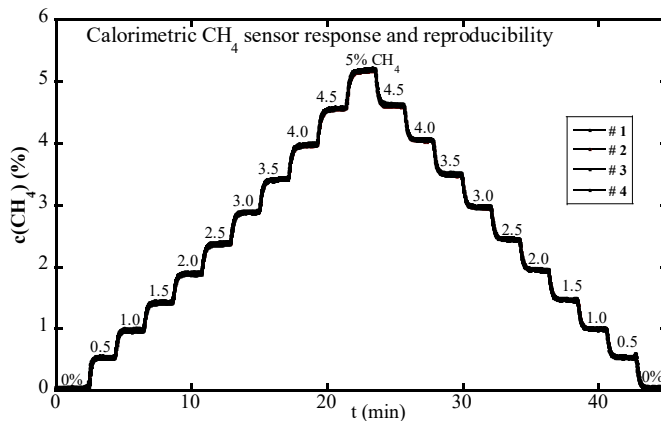


releasing  $\sim 25$  times more heat, compared to partial methane oxidation (conversion) to syngas:



Therefore, oxygen sufficiency or even oxygen excess would be critical for the calorimetric catalytic  $\text{CH}_4$  sensor's reliable performance. Optimized catalytic film would selectively oxidize methane in the presence of other flue gas species. Similar to the calorimetric catalytic CO sensor implemented on combustion market, the  $\text{CH}_4$  sensor would measure temperature and the associated sensor's RTD resistance change in the flue gas with the excess oxygen, depending on the methane concentration [47].

Catalyst microstructure and sensor temperature optimization would permit the calorimetric  $\text{CH}_4$  sensors reproducible response and good linearity (Figs. 12.23, 12.24).



**Fig. 12.23.** Calorimetric  $\text{CH}_4$  sensor response in 0...5 %  $\text{CH}_4$  mixture with 10 %  $\text{O}_2$  [4].

A calorimetric catalytic  $\text{CH}_4$  sensor has a response of  $\sim 30$  s for 90 % signal, and the flue gas flow control extractive environment would contribute to  $\text{CH}_4$  detection accuracy, reproducibility, and reliability in the very challenging temperature, pressure, flow, and chemical flue gas combustion environment.

The only calorimetric catalytic  $\text{CH}_4$  sensor available on the combustion market is quite selective and accurate with  $\sim \pm 5$  % of full scale or  $\pm 0.25$  %  $\text{CH}_4$  error in extractive methane measurements in a flow and temperature-controlled environment. Oxygen concentration

variation between 3 and 10 % might create up to a 0.4 % CH<sub>4</sub> error [4]. Some chemicals in combustion like silicon compounds, sulfur oxides, chlorine, and heavy metals might poison catalyst material, reducing sensor sensitivity and life.

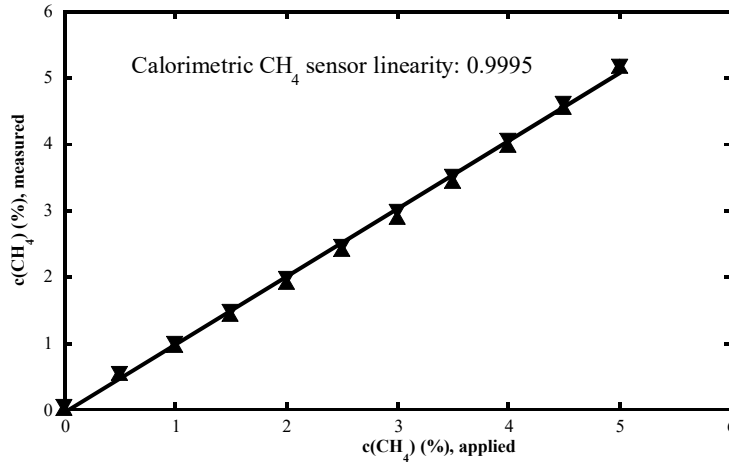


Fig. 12.24. Calorimetric catalytic CH<sub>4</sub> sensor linearity [4].

### 12.5.3. Tunable Diode Laser (TDL) and Quantum Cascade Laser (QCL) Spectroscopies

The IR absorption spectrum of methane (CH<sub>4</sub>) would provide reliable identification at 1.65 μ (TDL) or 7.9 μ (QCL) and measurements in a wide concentration range down to ~0.6 ppm CH<sub>4</sub> (TDL) or even ~10 ppb CH<sub>4</sub> (QCL) as a detection limit. The detection wavelengths and limits of some combustion gases are summarized in Table 12.5.

Table 12.5. Combustion gases absorption lines and detection limits.

Target gas	Wavelength (μ)	Detection limit (ppm)
Methane (CH <sub>4</sub> )	1.65 (TDL)	0.15
	7.9 (QCL)	0.003
Carbon monoxide (CO)	1.57	30
	2.33	0.5
	4.80	0.01
Oxygen	0.76	8.956
Carbon dioxide (CO <sub>2</sub> )	1.96	3
Nitric oxide (NO)	1.8	60
	2.65	1
Nitrogen dioxide (NO <sub>2</sub> )	0.68	0.3
Water (H <sub>2</sub> O)	1.39	0.06

A series of process adaption application tools were developed for in-situ, extractive, or in cross-pipe installations. The probe design would not require a special optical path alignment but would deliver single point measurement. In the probe design, the laser source and detector are located in the probe housing with the laser beam being reflected at the probe end back to the detector by quartz prism with max operation temperature of  $<250\text{ }^{\circ}\text{C}$  or by gold mirror with max operation temperature of  $<430\text{ }^{\circ}\text{C}$ . With a unique folded-path design, the probes can be installed in almost all pipes and stacks with no alignment needed.

Temperature and pressure variation in the process has to be compensated and might cause additional error in the measurements. IR light reflection at high temperatures, combined with wide background radiation from the fire box and process windows fouling, might bring additional challenges in many applications. There are quite accurate TDL analyzers with  $\sim\pm 4\%$  or  $\pm 0.02\%$   $\text{CH}_4$  error and sub-ppm detection limits in some applications. Considering the combustion process's highly-challenging flue gas temperature, pressure and flow variation, an in-situ multipoint temperature measurement and an in-situ  $\text{CH}_4$  validation would be highly recommended to improve these analyzers' reliability.

From all available major methane gas sensing technologies, only catalytic calorimetric and TDLS/QCLS have so far found practical application in instrumentation on the combustion market with quite reliable and accurate ( $\sim 2\text{--}5\%$  error) extractive or across the stack in-situ methane measurements.

## Acknowledgements

The author would like to thank Rosemount (Emerson Electric Corp., USA) and Chad McGuire for supporting his work on gas sensors projects, Professor Ulrich Guth (Kurt-Schwabe-Institute for Measuring and Sensor Technology Meinsberg e.V., Germany), Dr. Eric Brosha (Los Alamos National Laboratory, USA) for their very fruitful collaboration and Professor Ronald Hanson (Stanford University, USA) for sharing his knowledge on TDLS/QCLS application in combustion process.

## References

- [1]. N. Docquier and S. Candel, Process control and sensors: a review, *Progress in Energy and Combustion Science*, Vol. 28, 2002, pp. 107–150.
- [2]. P. Shuk, C. McGuire, Carbon monoxide gas sensing technologies in combustion process, *Sensors & Transducers*, Vol. 217, Issue 11, 2017, pp. 1-13.
- [3]. U. Guth, P. Shuk and C. McGuire, Gas sensing technologies in combustion: A comprehensive review, *J. Electrochem. Sci. Eng.*, Vol. 10, Issue 2, 2020, pp. 103-110.
- [4]. P. Shuk, C. McGuire and E. L. Brosha, Methane gas sensing technologies in combustion: A comprehensive review, *Sensor & Transducers*, Vol. 229, Issue 1, 2019, pp. 1-10.
- [5]. C. Wu, T. Yasukawa, H. Shiku and T. Matsue, Fabrication of miniature Clark oxygen sensor integrated with microstructure, *Sensors and Actuators B*, Vol. 110, Issue 2, 2005, pp. 342-349.

- [6]. A. C. Manning, R. F. Keeling and J. P. Severinghaus, Precise atmospheric oxygen measurements with a paramagnetic oxygen analyzer, *Global Biogeochem. Cycles*, Vol. 13, No. 4, 1999, pp. 1107-1115.
- [7]. C. S. Chu, Y. L. Lo, T. W. Sung, Review on recent developments of fluorescent oxygen and carbon dioxide optical fiber sensors, *Photonic Sensors*, Vol. 1, No 3, 2011, pp. 234-250.
- [8]. W. C. Maskell and B. C. H. Steele, Solid state potentiometric oxygen gas sensors, *J. Appl. Electrochem.*, Vol. 16, No. 4, 1986, pp. 475-489.
- [9]. R. Ramamoorthy, P. K. Dutta and S. A. Akbar, Oxygen sensors: materials, methods, designs and applications, *J. Mater. Sci.*, Vol. 38, No. 21, 2003, pp. 4271-4282.
- [10]. W. C. Maskell, Inorganic solid state chemically sensitive devices: electrochemical oxygen gas sensors, *J. Phys. E: Sci. Instrum.*, Vol. 20, No. 10, 1987, pp. 1156 -1168.
- [11]. W. Göpel, G. Reinhardt, M. Rösch, Trends in the development of solid state amperometric and potentiometric high temperature sensors, *Solid State Ionics*, Vol. 136-137, 2000, pp. 519-531.
- [12]. J. W. Fergus, Doping and defect association in oxides for use in oxygen sensors, *J. Mater. Sci.*, Vol. 38, No. 21, 2003, pp. 4259-4270.
- [13]. M. Kleitz, E. Siebert, P. Fabry and J. Fouletier, Solid state electrochemical sensors, in *Sensors, a comprehensive survey*, Vol. 2, Part 1, Chemical and biochemical sensors (Eds. W. Göpel, T. A. Jones, M. Kleitz, I. Lundström and T. Seiyama), *VCH Publisher*, Weinheim - New York, 1991, pp. 340-428.
- [14]. P. Shuk, Process Zirconia oxygen analyzer - state of art, *Technisches Messen*, Vol. 77, Issue 1, 2010, pp. 19-23.
- [15]. P. Shuk, Oxygen gas sensing technologies application: a comprehensive review, in *Sensors for everyday life, smart sensors, measurement and instrumentation*, Vol. 23 (Eds. S. C. Mukhopadhyay et al.), *Springer*, Heidelberg - New York, 2017, pp. 81-107.
- [16]. E. D. Hinkley and P. L. Kelley, Detection of air pollutants with tunable diode lasers, *Science*, Vol. 171, No. 3972, 1971, pp. 635-639.
- [17]. P. Werle, A review of recent advances in semiconductor laser based gas monitors, *Spectrochim. Acta*, Vol. 54, Issue 2, 1998, pp. 197-236.
- [18]. J. D. Tate, Advanced combustion diagnostics and control for furnaces, fired heaters and boilers, *DOE report DE-FG36-06GO16093*, 2010, pp. 1-40.
- [19]. R. K. Hanson, Applications of quantitative sensor laser to kinetics, propulsion and practical energy systems, *Proceedings of the Combustion Institute*, Vol. 33, Issue 1, 2011, pp. 1-40.
- [20]. R. P. Kovacich, N. A. Martin, M. G. Clift, C. Stocks, I. Gaskin and J. Hobby, Highly accurate measurement of oxygen using a paramagnetic gas sensor, *Meas. Sci. Technol.*, Vol. 17, Issue 6, 2006, pp. 1579-1585.
- [21]. H. Peters and H.-H. Möbius, Procedure for the gas analysis at elevated temperatures using galvanic solid electrolyte elements (Germ), *DD-Patent 21673*, 1961.
- [22]. J. Weissbart and R. Ruka, (1961) Oxygen gauge, *Rev. Sci. Instrum.*, Vol. 32, Issue 5, 1961, pp. 593-595.
- [23]. P. Shuk, E. Bailey and U. Guth, Zirconia oxygen sensor for the process application: state of art, *Sensors & Transducers*, Vol. 90, Issue 4, 2008, pp. 174-184.
- [24]. H.-H. Möbius, Basics of oxygen gas potentiometric analysis (Germ), *Z. phys. Chem.*, Vol. 230, Issue 5-6, 1965, pp. 396-416.
- [25]. D. M. Haaland, Internal-reference solid-electrolyte oxygen sensor, *Anal. Chem.*, Vol. 49, Issue 12, 1977, pp. 1813-1817.

- [26]. J. V. Spirig, R. Ramamoorthy, S. A. Akbar, J. L. Routbort, D. Singh and P. K. Dutta, High temperature zirconia oxygen sensor with sealed metal/metal oxide internal reference, *Sensors and Actuators B*, Vol. 124, Issue 1, 2007, pp. 192-201.
- [27]. Y. Tanaka, T. Sato, H. Ikeda and N. Miura, Cobalt-based solid reference-electrode usable in zirconia-based sensors for detection of oxygen or volatile organic compounds, *Sensors and Actuators B*, Vol. 114, Issue 11, 2014, pp. 899-903.
- [28]. Y. Krishna, S. O'Byrne and J. J. Kurtz, Baseline correction for stray light in log-ratio diode laser absorption measurements, *Applied Optics*, Vol. 53, Issue 19, 2014, pp. 4128-4135.
- [29]. S. Langridge, On line measurement of oxygen: review and new developments, in *Proceedings of the 55<sup>th</sup> ISA Analysis Division Symposium*, Vol. 481, 2010, pp. 113-127.
- [30]. P. A. Martin, Near-infrared diode laser spectroscopy in chemical process and environmental air monitoring, *Chem. Soc. Rev.*, Vol. 31, Issue 4, 2002, pp. 201-210.
- [31]. R. F. Curl and F. K. Tittel, Tunable infrared laser spectroscopy, *Ann. Rep. Prog. Chem. C*, Vol. 98, 2002, pp. 219-272.
- [32]. M. Lackner, Tunable diode laser absorption spectroscopy (TDLAS) in the process industries - a review, *Rev. Chem. Eng.*, Vol. 23, No 2, 2011, pp. 65-147.
- [33]. F. K. Tittel, R. Lewicki, R. Lascola and S. McWhorter, Emerging infrared laser absorption spectroscopic techniques for gas analysis, in Trace analysis of specialty and electronic gases (Eds. W. M. Geiger and M. W. Raynor), *John Wiley & Sons Inc*, New York, 2013, pp. 71-109.
- [34]. R. Kocache, D. Holman and J. Swan, Sensor for combustible gases, *US Patent 5,549,871*, 1996.
- [35]. P. Shuk, P. Murphy, New advanced CO-sensor for the process application, in *Processing of the 1<sup>st</sup> International Conference on Sensing Technology*, Palmerston North, New Zealand, 21-23 November 2005, pp. 226-227.
- [36]. Y. Sandler, Electrochemical sensor for reactive gas mixtures, *US Patent # 3,981,785*, 1976.
- [37]. N. Miura, T. Sato, S. Anggraini, H. Ikeda and S. Zhuiykov, A review of mixed-potential type zirconia-based gas sensors, *Ionics*, Vol. 20, Issue 7, 2014, pp. 901-925.
- [38]. R. Sorita, T. Kawano, A highly sensitive CO sensor: screening of electrode materials, *Sensors and Actuators B*, Vol. 35-36, Special Issue, 1996, pp. 274-277.
- [39]. P. Shuk, E. Bailey, J. Zosel and U. Guth, New advanced in-situ carbon monoxide sensor for the process application, *Ionics*, Vol. 15, Issue 2, 2009, pp. 131-138.
- [40]. A. Corma, From microporous to mesoporous molecular sieve materials and their use in catalysis, *Chem. Rev.*, Vol. 97, Issue 6, 1997, pp. 2373-2420.
- [41]. C. S. Cundy, P. A. Cox, The hydrothermal synthesis of zeolites: history and development from the earliest days to the present time, *Chem. Rev.*, Vol. 103, Issue 3, 2003, pp. 663-702.
- [42]. K. Na, C. Jo, J. Kim, K. Cho, J. Jung, Y. Seo, R. J. Messinger, B. F. Chmelka, and R. Ryoo, Directing zeolite structures into hierarchically nanoporous architectures, *Science*, Vol. 333, Issue 6040, 2011, pp. 328-332.
- [43]. Mesoporous zeolites: Preparation, characterization and applications, Eds. J. G. Martinez and K. Li, *John Wiley & Sons*, Weinheim, Germany, 2015, 608 p.
- [44]. B. Farber, Method of sensor conditioning for improving signal output stability for mixed gas measurements, *US Patent # 7,585,402*, 2008.
- [45]. B. Farber, Method of sensor conditioning for improving signal output stability for mixed gas measurements, *US Patent # 9,304,101*, 2016.

- [46]. P. K. Sekhar, J. Kysar, E. Broscha and C. R. Kreller, Development and testing of an electrochemical methane sensor, *Sensors and Actuators B*, Vol. 228, Issue 6, 2016, pp. 162-167.
- [47]. P. Shuk, C. McGuire and P. Ostby, Combustion analyzer with simultaneous carbon monoxide and methane measurements, *US Patent # 11,346,554*, 2022.



# Index

## 2

2D-FEM, 88, 90, 92, 101

## A

algorithm related to SO<sub>2</sub> and H<sub>2</sub>O  
interference, 268  
amperometric oxygen sensor, 255  
analytical instrumentation, 254  
Angular-movement sensor, 40  
Areal density, 155  
image, 157  
artificial neural networks, 174, 175  
Au-composite, 266, 267, 268  
AWG, 94, 96

## B

Basis material  
decomposition, 155  
image. *see Areal density image*  
BER - Bit Error Ratio, 179  
Big Data, 191, 192  
Blumlein bridge, 26  
Electronic analog, 28  
BMD. *see* Basis material decomposition  
boundary layer, 145  
Box proposer, 111

## C

Calibration, 109  
with Known Extrinsic Parameters, 121  
Metric Calibration, 115  
Calorimetric  
catalytic CO-sensor, 262-264  
sensors, 254  
carbon monoxide equivalent, CO<sub>e</sub>, 264, 269  
catalytic CH<sub>4</sub> sensor, 272, 273  
CH<sub>4</sub> Gas Sensing Technology, 270, 271  
checksum, 179

CIC decimation filter, 218  
circuit L298 integrated, 78  
Clark type cells, 254  
Classification, 109  
Classification of vehicles according to  
model, 129  
Classifier of vehicle colors, 126  
Height Estimation, 118  
cluster area, 151  
CMOS, 92  
CNN. *See* Convolutional neural network  
CoFeSiB wire, 222  
Collision Avoidance, 195  
Combustion  
control, 252, 262  
process regulation, 251, 252  
comfort, 135, 147, 150  
assessment, 148  
biomechanical, 135  
climate comfort, 137, 139, 150  
climatic, 135  
comfort correlations, 148  
comfort perception, 138  
comfort predictions, 141  
ergonomic, 137  
global thermal comfort, 138  
regional comfort, 139  
wearing comfort, 137  
communication protocols, 173, 174  
COMSOL, 88, 90, 92  
Convolutional neural network, 111  
Faster R-CNN, 111  
MobileNet, 113  
Region-based Fully Convolutional  
Network, 111  
ResNet, 113  
CoS - Class of Service, 200  
cross-sensitivity coefficients, 174  
Current Loops, 176  
curvature energy, 69, 70, 75, 77, 78, 81, 82  
CVD, 87, 98



**D**

Data container, 241  
 DCS - Distributed Control Systems, 181  
 Debris sensor, 43  
 decubitus, 149  
 deflects, 70, 71, 75  
 Design and Calibration, 61  
 DE-XRT. *see* *Dual energy X-ray transmission*  
 different gas sensing technologies, 252  
 Diffused Waveguide, 88  
 Digital off Diagonal GMI Sensor, 214  
 Digital-Down-Converter, 218  
 dilation, 69, 70, 74, 75, 78, 79, 81  
 Dirac function, 76  
 Direct Digital synthesizer, 215  
 Distortion, 112  
 Dual energy  
   detector, 154  
   X-ray transmission, 154  
     coal, 159  
     food safety, 160  
     laundry, 166  
     mining, 159  
     recycling, 168  
     sorting, 168  
 dynamic range, 227  
   ADC, 226

**E**

Eddy current probe, 19  
 Effective atomic number, 156  
 electrochemical, 252, 254, 255, 262, 264, 265, 270, 271  
   high temperature sensors, 254  
 energy  
   curvature, 69  
   spectra, 69  
 EPL - Ethernet Powerlink, 197  
 equivalent magnetic noise, 224  
 EtherCAT, 197, 203  
 evaporation, 138  
 Evaporative Capacity, 146  
 extractive combustion analyzers, 263, 271

**F**

Feature extractor, 110  
 femtosecond (fs) laser writing, 87  
 FF - Foundation Fieldbus, 176, 181, 182, 184-186, 188  
 field  
   observable, 69, 82  
   theory, 69  
 fieldbuses, 174  
 FIR filter, 220  
 Foreign body, 160  
   aluminum, 162  
   EPDM, 164  
   glass, 163  
   PTFE, 164  
   realistic, 165  
   steel, 163  
   synthetic bone, 165  
 Fourier  
   spectra, 76  
   transform, 76  
 FPGA, 216  
 FRT - Firm Real Time, 194  
 fundamental mode, 87

**G**

Gap sensor, 34  
 Gas sensing technologies, 251, 254  
 geometrical invariant, 70, 71  
 Giant Magneto Impedance Effect, 206  
 GitHub, 232, 237, 243, 246, 247  
 GMI  
   effect, Origin, 207  
   Sensor Conditioning Strategy, 210  
 ggravitational waves, 69, 70, 81-83  
 guard bands, 199

**H**

hadrons, 70  
 Hall sensor, 69, 71, 75-79  
 harsh measuring conditions, 174  
 HART – Highway Addressable Remote

Transducer, 174, 178-180  
H-Bridge, 78  
heat  
  balance, 136  
  flow  $H_c$ , 146  
  heat and humidity source, 141  
  Index, 137, 149  
  map, 143  
Heterogeneous video surveillance, 109  
homoiothermia, 139  
Hough transform, 116  
HRT - Hard Real Time, 194  
humidity map, 143  
Hybrid, 90, 99, 105

## I

I2C interface, 141  
Image region regressor, 111  
Impedance Modelling, 210  
Implementation, GMI sensor, 212  
improved combustion efficiency, 251, 252, 262  
Industrial  
  Ethernet, 191, 193, 196  
  Internet of the Things - IIoT, 191-193, 203  
  Measurements, 173, 174  
  zirconia oxygen analyser, 255  
influence variables, 173, 175  
interconnectivity, 193  
interoperability, 174, 178, 180, 181, 184  
interpolation algorithm, 143  
IP - Internet Protocol, 192, 193, 196, 197

## L

Lambert-Beer law, 155  
LAN - Local Area Network, 192, 195, 196  
Lane markings, 116  
laser techniques, 254  
Line detector, 154

## M

Magnetic sensors, 205  
mash material, 142  
Mass attenuation coefficient, 155  
Maxwell equation, 73

mean skin temperature, 140  
Measurement  
  Limitations, 51  
  resolution, 177, 178  
Meta-architectures, 111  
Methane ( $CH_4$ ) Gas Sensing, 270  
microclimate  
  skin near, 150  
microclimate, 136, 138-140, 145, 150  
  skin near, 151  
Mixed potential, 265, 266, 268, 270  
  CO sensor, 265  
  gas sensing, 264, 270  
moisture  
  flow  $H_e$ , 146  
  management, 149  
  Vapor Transmission Rate, 146  
Monitoring of CO in the combustion  
  process, 251  
MU - Manufacturing Units, 191-193, 203

## N

negligible cross-interference, 268  
Noise Performance, 213

## O

Object detection, 110  
off-diagonal, 213  
  GMI sensor, 223  
Ohm's law, 70, 71, 75  
One-stage detectors, 111  
open-source, 231-234, 237, 242, 243, 246  
optical sensors, 254  
Optimum Design, Measurement Accuracy  
  and Calibration of Optical Fiber Sensors  
  Based on Strain Transfer Theory, 49  
OSI - Open System Interconnection, 181,  
  182, 184, 192, 193, 202, 203  
Oxidizing combustion process, 252  
Oxygen  
  Gas Sensing Technologies, 254  
  identification at  $\sim 760$  nm, 258  
  measurement, 251, 252, 254, 259

## P

paramagnetic, 254, 255, 260, 261

pass to pass motors, 78  
 PCB printed circuit board, 142  
 Photonic bandgap, 91  
 PIC, 88, 92, 93, 94, 96, 100  
 Plasmonic, 90, 91  
 PLC - Programmable Logic Array, 192  
 pressure ulcer, 149  
 Processing Gain, 221  
 product testing, 150  
     non-destructive, 147  
 Profinet, 193, 196, 197, 203  
 Proposal classifier, 111  
 PTP - Precision Time Protocol, 196, 197,  
     200, 201  
 PVD, 98

## Q

QoS - Quality of Service, 194, 195  
 quantum cascade laser spectroscopy, 262

## R

rating scales, 148  
 real time, 174  
 Reducing combustion, 252  
 Re-identification of vehicles, 126  
 reliability, 231, 234, 235, 238, 247  
 resistive oxygen sensor, 255  
 rib, 87, 88, 93, 94, 103, 104  
 ribbon cable, 142  
 ridge, 87, 88, 91  
 RT - Real Time, 191, 192, 193, 194, 195,  
     196, 197, 198, 203  
 RTP, 88, 89

## S

scheduled, 198, 199  
 SCPI, 234, 236, 237, 239, 240  
 screw effect, 76  
 second curvature, 69, 74  
 secondary combustion, 252  
 Selectivity, 262, 271  
 Sensitive, Selective and Stable, 251  
 sensor  
     combi-sensor, 141  
     flash-mode, 142  
     mat, 142

packaging, 256, 266  
 selectivity, 173  
 SMD sensor chip, 142  
 wired sensor set, 143  
 SERCOS III, 193, 198  
 silica-titania platform, 100, 101, 103-105  
 simulations, 88, 90, 92  
 SiN, 92, 96, 97, 100  
 Single shot detectors, 111  
 Single-coil eddy current sensor, 22  
     Cluster composition, 37  
     Temperature effect, 31  
 skin effect, 209  
 Slot, 89, 90  
 SNR - Signal to Noise Ratio, 178  
 SOI, 88, 91-94, 96, 97, 100  
 sol-gel dip-coating, 99, 105  
 spacer material, 142  
 spectral  
     curvature, 69  
     density, 76, 79  
     torsion, 69, 79  
 spin oscillations, 69  
 SRT - Soft Real Time, 194  
 Standard Commands for Programmable  
     Instruments. *See* SCPI  
 steady-state, 141  
 stoichiometric point, 252, 262  
 Strain Transfer Theory, 53  
 subwavelength grating  
     SWG, 87, 89  
 sweat glands, 146  
 SWEATLOG, 151  
     sensor array, 147  
     sensor mat, 148  
     temperature and humidity distribution,  
         148  
 SWEATOR, 145-148, 150, 151  
     climate dummy, 144  
     heat and humidity source, 148  
     manikin, 144, 150  
     simulation system, 150  
     simulation systems, 144  
     sweating devices, 145  
     torso, 147  
 synchronization, 176, 184, 193, 195-198,  
     200, 201

## T

### TDL

- CO analyzers, 269
- spectroscopy, 259
- TDLS, 255, 257, 259, 262, 269, 274
- The study of the GMI effect, 208
- Thermal Resistance (Rct), 146
- thermophysiology, 138
- thermoregulation, 137-139, 144
- time-delay, 227
- Torque-measuring devices, 40
- torsion, 69-82
  - energy, 69, 70, 74, 76
- tunable diode laser spectroscopy, 255, 262

## U

- unconventional, 89, 104
- unscheduled, 192, 199
- user interface, 234, 236, 237, 238

## V

- Vanishing points, 116
- vapor phase, 146
- Virtual
  - concentration, 156

Instrument Software Architecture. *See*  
VISA

VISA, 233-241, 245, 247

## W

- water vapor
  - partial pressure, 146
  - permeable membrane, 145
  - Transmission Resistance (Ret), 146
- waveguide, 85, 93, 94, 96, 99, 104
- white noise, 225
- WHRT - Weakly Hard Real Time, 194, 195
- Wireless HART, 180
- work capacity, 138
- WR - White Rabbit, 200

## X

- X-ray, 153
  - projection, 157

## Z

- zirconia
  - oxygen sensor, 256, 264
  - potentiometric, 255, 257



# Sensors, Measurements and Networks

**Sergey Y. Yurish, Editor**

The 8<sup>th</sup> volume entitled '*Sensors, Measurements and Networks*' contains twelve chapters with the descriptions of latest advances in sensor related area written by 33 authors from academia and industry from 11 countries: Belgium, China, Czech Republic, Germany, México, Poland, Portugal, Russia, Senegal, Serbia, and the USA.

Coverage includes current developments in physical sensors, chemical sensors, measurements and sensor networks. With this unique combination of information in each volume, the Advances in Sensors book Series will be of value for scientists and engineers in industry and at universities, to sensors developers, distributors, and users.

Like the first seven volumes of this Book Series, the 8<sup>th</sup> volume also has been organized by topics of high interest. In order to offer a fast and easy reading of each topic, every chapter in this book is independent and self-contained. All chapters have the same structure: first an introduction to specific topic under study; second particular field description including sensing or/and measuring applications. Each of chapter is ending by well selected list of references with books, journals, conference proceedings and web sites.

This book ensures that our readers will stay at the cutting edge of the field and get the right and effective start point and road map for the further researches and developments. By this way, they will be able to save more time for productive research activity and eliminate routine work.



ISBN 978-84-09-44954-5

