IEEE

Brno University of Technology
Student Branch

Proceedings of IEEE Student Branch Conference

Mikulov 2017

28. - 30. 8.

Název: Proceedings of IEEE Student Branch Conference Mikulov 2017

Editor: Ondřej Zach

Vydavatel: Vysoké učení technické v Brně
Fakulta elektrotechniky a komunikačních technologií

Rok vydání: 2017

Vydání: první

Organizační výbor konference:

Ing. Miroslav Cupal
Ing. Martin Hrabina
Ing. Ondřej Zach

ISBN 978-80-214-5526-9

Conference was supported by

Gold sponsors

Bronze sponsor

Other partners

Preface

Dear friends,

it is already for the 13th time when you can read the Proceedings of a conference organized by IEEE Student Branch, Brno University of Technology. Although 13 is often considered as an unlucky number, let us hope it will bring only the good to this year's conference. In its reborn tradition, we are happy to welcome you at the conference for recent and future PhD students. For this year, we have chosen the beautiful south-moravian city of Mikulov as the venue of the conference.

The aim of the conference was always to bring together young scientist to share their ideas and to have the possibility of familiarizing with the research at other departments and universities and this year will not be an exception.

Organizing a conference is never an easy task. Therefore, we would like to thank all the companies and organizations that supported us, namely Rohde & Schwarz, NXP and Era. Big thank you also belongs to IEEE Czechoslovakia Section and Department of Radio Electronics, Brno University of Technology.

To conclude, allow me to thank you, the participants of this year conference. By your participation, you support the very idea of what IEEE represents. We wish you a pleasant time in Mikulov.

On behalf of the organizers

Ondřej Zach

Table of Content

Comparative Study of Technology Appropriation Concepts

Anna Maria Al Zubaidi-Polli
Department of Telecooperation
Johannes Kepler University
Linz, Austria
anna.polli@fh-hagenberg.at

***Abstract*—The stated aim of this conference is to debate new ideas from different research fields. This paper seeks to contribute to this discussion by exploring concepts of technology appropriation from a number of different socio-cultural studies. When a technology is appropriated, it is likely to create new sales and even possible new markets for the technologies. Therefore, designers and programmers should encourage this form of appropriation and be prepared to learn from it when developing and designing new technologies in the future. In an attempt to seek common ground among existing theories of technology appropriation, ten research articles on appropriation have been selected, which range from the root discussions about appropriation to the state-of-the-art publications within Human Computer Interaction. An analysis has been conducted by raising the same critical questions to each of the selected papers, to gain an enhanced foundation of these various understandings of the term technology appropriation. This paper will conclude with an overview of the various understandings and provide guidelines for researchers on how to design technologies for successful appropriation.**

Keywords—technology appropriation; concepts; definition; comparison; HCI; design for appropriation

I. INTRODUCTION

The concept of appropriation has appeared in Computer Supported Collaborative Work (CSCW) and Human Computer Interaction (HCI) research on and off for the last 15 years. Since then, research has broadened our understanding about how users give technology a meaning, as well as how people fit technology into the routines of their everyday life [1, 2, 3, 4]. Appropriation may occur because existing features of a specific designed technology do not meet the intended user's needs. People domesticate the technology and use it in their own ways. At this stage, we know the technology has been appropriated and *has become the users' own*, and not simply what we as designers provided them with [2]. When looking closer at research, the term 'appropriation' is differently defined, and the concepts differ from each other. Researchers are encouraged to investigate the process of appropriation [5]. Dourish claims appropriation is a key problem when it comes to the development of interactive systems, as it illustrates the success of the technological deployment. This highlights the importance of appropriation, as it lies in the intersection of design and the place of studies [6]. Therefore it is important to provide an overview of the various concepts of technology appropriation used in Computer Science for researchers to better situate their research, to be aware of these various understandings, and to find one common ground to all these various concepts and definitions.

II. DIVERSE CONCEPTS OF TECHNOLOGY APPROPRIATION

A. Allan MacLean et al. – User-Tailorable Systems

Allan MacLean et al. mention it is impossible to design systems, which are appropriate for all users and all situations. Instead, they argue about the most useful techniques to empower users to tailor systems. The authors point out a connection between a person's skills and a ***person's tailoring power***, meaning that the amount of required skills has an impact, regardless of whether a person has a low or high tailoring power. The authors differentiate between workers, tinkerers and programmers. Programmers and tinkers enjoy exploring the computer system and therefore have a high tailoring power. Workers have a low interest in computer systems and are the ones with the least tailoring power. Therefore the authors' advice is to give the workers a sense of ownership when using a system, so they feel more in control when changing the system. Furthermore, the authors hint that systems should be easily changeable so the users develop their own personalized environment in the system. Hence they can influence the design of the system and adopt, adapt and shape it to their needs. The authors identified the *handyman* who eases the interaction for the workers, so they can appropriate the environment in their own personal day-to-day activities through making the functionality accessible [7].

B. Bonnie A. Nardi & James R. Miller – Twinkling light and nested loops

Nardi and Miller talk about the ***developmental process of use*** of spreadsheets and practices. The authors don't emphasize on the individual user, but on how spreadsheets are collaboratively used. Indeed, one of their findings was that people learnt more when they cooperatively constructed spreadsheets. With the help of *capable-peer-users*, they learned how to use certain spreadsheets products, which they wouldn't have on their own. When users felt like they were approaching the limits of their understanding, they reached out to more experienced users. This paper explores ideas that are described today as characteristics of appropriation, such as

users asking other colleagues, or more capable peers, for support [8], which is a highly discussed characteristic of appropriation today [2].

C. *Michelle Gantt & Bonnie A. Nardi – Gardeners and Gurus: Patterns of cooperation among CAD users*

The authors identified ***different stages of customization***, which depend on the varying levels of people's computer expertise and interest to customize their environment. Gantt and Nardi differentiate between end users, local developers (called gardeners) and electrical engineers (the gurus). Based on the knowledge and skills in programming, the users customize the system differently. The authors identified reasons why users have not customized the studied system. One user, for instance, was not allowed by his manager to customize, while the others preferred not to specialize in the system as they got support from gardeners and gurus. The gardener "is a cross between both worlds" and is required to speak the language of the end users and the language of the administrators. A similar concept is described in section *A* [7]. Software designers need to find out, for a given product, where the continuum of computer expertise starts and ends, and then incorporate into the product a range of capabilities that take advantage of the range of users arrayed along the continuum [9].

D. *Randall H. Trigg & Susanne Bødker – From Implementation to Design: Tailoring and the Emergence of Systematization in CSCW*

The authors talk about tailoring and describe the term tailoring as an organizational process of adapting and appropriating technology. Tailoring is only a part of the larger effort of adapting technology done by particular users. They refer to it as a ***'phenomenon of use'***. Tailors adapt software to the workplace in which they are embedded to modify and extend their practices. First, these changes are completed by the individual, then distributed to other colleagues and are perhaps further customized on their own. Tailoring is seen as an ever-increasing variety of use patterns. The authors describe ***tailoring as a cooperative process***. The authors' advise to aim to flexible systems that can run on different platforms so the product can be integrated with other software. Thus the customizability of the systems by users will have no programming skills required [10].

E. *Jennie Carroll et. al. – Just what do the youth of today want? Technology appropriation by young people*

In this paper the authors are trying to gain insight into the early stages of technology use, stages they call 'appropriation'. The authors describe ***'appropriation' as a process by which people adopt, shape, use the technology and integrate it into their everyday life***. A model of technology appropriation was created, which illustrates the transformation of technology envisioned by the designer and used by the user in real life. These two levels indicate three states of appropriation, where technology has either never seriously been deliberated or reflected upon (non-appropriation), or where the technology is understood, accepted and adapted by the user (appropriation), which might exceed then to the third and last state of the model called 'dis-appropriation', namely, the state when appropriation is reached, but still could get rejected after a while. One criterion of appropriation for the users is whether the technology adds value to their lifestyle. A further criterion for appropriation is that a mobile technology may first need to satisfy the critical mass in a social group, before the technology may become appropriated. [3].

F. *Wendy March et al. – Designing Technology for Community Appropriation*

The authors are stating that all new devices around us are opening up for ***user adoption and adaptation***. As the title of their paper already hints, they discuss ***appropriation within a community*** and do not focus on the individual. Designing for appropriation seems even more challenging when looking at technologies specifically aiming to a whole community, than to an individual. In their workshop at CHI, the authors were investigating methods and strategies for designing systems that allow appropriation and modifications. One method they suggest is re-directing the user from the design process to the re-design process. They suggest inviting people from the early stage in the design process as a tool for reflection, to create open systems for user appropriation. The focus is placed on openness, transparency and adaptability. The authors suggest a way to capture people's imagination when using a system, namely, to create an "alternative universe" without limitations of here and now. So users have new channels where free expressions and public communication can occur. This will give an increased awareness and control over the energy in the things people use [11].

G. *Alan Dix – Designing for Appropriation*

Based on his frequent observations of ethnographies and field studies, Alan Dix explains that people do not play by the rules. Instead, they ***adapt and adopt the technology in ways the designer hasn't anticipated***, which makes it very hard to design for appropriation. Designing for appropriation is often seen as an oxymoron, as it appears to be impossible to design for the unexpected. The paper talks about appropriation as an important and positive phenomenon, since it is an essential ***sign of user's acceptance of technology***. He argues that one might not be able to design for unanticipated use, but one can design to allow the unexpected. At the point when the user domesticates a technology, then they feel comfortable to use it in their own ways and don't limit themselves with what the designer gave to them. This is what Alan Dix means by *appropriation*. When this ***stage is reached in the process of use***, the technology becomes the user's own. This stage implies that appropriation might occur when no existing tool is available for a required task, but appropriation is simply easier than this. Appropriating a technology has its advantages, such as the feeling of being in control and winning a sense of ownership. Alan Dix concludes that when designing for appropriation, one needs to be aware of the environment, as specific needs may change through a deployment of a particular technology, so design for use must be design for change [2].

H. *Gunnar Stevens et. al. – Appropriation Infrastructure:*

Appropriation needs to be understood as a ***phenomenon of collaborative and creative activities,*** the authors explain. Appropriation work needs to be understood as a core concept

in the field of End User Development. Appropriation work may lead to software usages that go beyond what has been envisioned by the designers of the software, it is a specific part of IT artifact's usage. The authors propose a new approach to support appropriation by integrating a communication infrastructure into software applications that follows an "easy-to-collaborate"-principle. Such an appropriation infrastructure stimulates the experience of sharing a heterogeneous product community and supports the situated development of usages. The tailorability of the artifact should be enhanced by means of communicative appropriation. They propose a design for appropriation infrastructure where the software users and administrators (i.e., different users, power users or system administrators, who work with tailorable software) can easily respond to the activities of the product community (software designers and developers). By means of this easy to collaborate principle, the authors have embedded a discussion software feature to lower the burden [12].

I. Susanne Bødker & Ellen Christiansen – Poetry in motion

The authors define appropriation as ***"the way that users evaluate and adopt, adapt and integrate a technology into their everyday practices"***. They explicitly explore the process for appropriation and dis-appropriation (see section *E*). They have investigated in specific criteria for appropriation, so a long-term change and use of an application is required. Furthermore, adoption relies on reaching the critical mass (taken from Roger's framework of adoption and innovation [13]). Then they experienced the process of taking something that belongs to others and making it one's own, by using Wertsch's segmentation of appropriation [14]. This is described through four stages (1) anticipation, (2) initial familiarity, (3) development of repertoires of routines and the (4) development of new forms of use. Their findings conclude the need to address both individual and social appropriation, as appropriation takes place at two levels: the social and the individual level [15].

J. Phil Turner – How we cope with digital technology

Turner suggests that users frequently personalize and configure digital technology and, in doing so, IT becomes better suited to the ways in which they want to use it. In short, Phil Turner calls this process' appropriation'. He defines appropriation as something users do to digital technology, whereas when other people (i.e., designers) change technology to better suit users, it is the consequence of user modeling and results in customization rather than appropriation. ***Appropriation is described as an umbrella term, encompassing user configuration, 'do-it-yourself' design, re-use, sustainability and personalization***. Digital technologies have a malleability, pliability, configurability, compliance, and flexibility, which invite people to make it their own. These acts of appropriation can range from slight changes (e.g., change of wallpaper) to actively modify or even re-design the technology itself. Phil Turner presents ***three dimensions of appropriation: user configuration, ensoulment and personalization***. Within the dimension of user configuration, the author talks about modifying toolbars and populating the smartphones with 'apps'. The second dimension of appropriation, referred to as 'ensoulment' is about designing (re-designing) artifacts that are adaptable to the user, empowering and reflecting the user's personal identity. The idea behind the third dimension of personalization is a process by which users are not just personalizing their phones with specific ring-tones, but they can also add functionality. While recognizing that personalization increases the sense of ownership of the artifact, it undoubtedly makes digital technology both more fun and easier to use [16].

III. Results and Discussion

A. Various definitions of appropriation

After a critical review of the literature on appropriation, various definitions were identified. The definitions varied from appropriation as being a phenomenon, a concept, a process, a development of use, an every-increasing use of patterns or an early stage of use.

[3, 7, 8, 10, 16] define appropriation as a process. [7] describes tailoring as a process where users evolve gradually along using a system with their own changing expertise and requirements. [8] talks about it as a developmental process of the use of technology. [10] describes it as an organizational process of adapting and appropriating technology. [3] talks about it as a process, how people adopt, shape, use the technology, and integrate it into everyday life. [16] defines appropriation as something users do to digital technology, this he calls 'process of appropriation'. These five research papers talk about appropriation as a process, but when looking closer at the contexts in which they use the term 'process', it varies quite extensively from paper to paper. Could it be the process of users, the process of use of the technology, or the process of people adopting, adapting or integrating technology into life? Although the term process was repeatedly used in the listed references above, their perspectives differ from one to the other. One common understanding though is that the user and the technology are involved in the process of appropriation, and both evolve over the process of use. Research papers [2, 3, 9] elaborate on appropriation as a stage of technology use: [3] gains insight into the early stages of technology use, stages they call 'appropriation'. [2] implies that this stage is reached in the process when users domesticate the technology and don't limit themselves with what the designer gave them. [15] refers to Wertsch's segmentation of appropriation. These segmentations are an even more detailed approach. The stage of process is divided into four stages (1) anticipation, (2) initial familiarity, (3) development of repertoires of routines and the (4) development of new forms of use.

Another frequent definition of appropriation is the term 'appropriation as a phenomenon' (see in [2, 10, 17]). [10] refers to it as as a process: they at the same time refer to researchers studying the phenomenon of use [7, 9]. [2] talks about appropriation as an important and positive phenomenon, since it is an essential sign of user's acceptance of technology. [17] approaches appropriation more from a collaborative angle: The author suggests that appropriation needs to be understood as a phenomenon of collaborative and creative activities. These three references look into different kinds of research fields, where as [17] looks into the collaborative aspect of appropriation: [2] concentrates on the interplay and

change after the phenomenon of appropriation has taken place, and [10] talks about it more from a general approach.

One further definition of appropriation was found in the ten selected research articles. [17] talks about appropriation as a core concept, where appropriation work needs to be understood as a core concept in the field of End User Development. Moreover, [15] talks about using the concept of 'appropriation' to analyze the interview data and to look into new opportunities of use by app technologies. Out of the critical review of literature, four profoundly different understandings were analyzed. One talks about appropriation as a process, a stage of a process, a phenomenon or as a concept (shown in Figure 1).

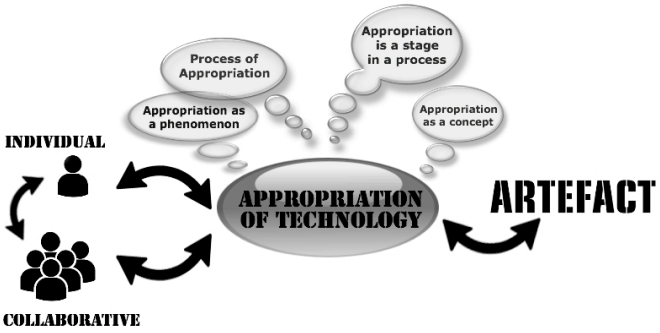


Fig. 1. Model of Technology Appropriation – Illustrates the four diverse understandings of the term appropriation

B. *Which Technology supports Appropriation?*

[2] talks about the idea that designing for appropriation is an oxymoron: "plan for the unexpected". But the author argues that whilst you cannot design for the unexpected, you can design so that people will be more likely to be able to use what you produce for the unexpected – so they do the 'final' design when the need arises. Therefore as [7, 10, 17] elaborate on, technologies need to be flexibly designed. They should be able to run on different platforms to start with, and little or no programming skills should be required. As [2] points out, the needs of the users change with the use over time. Therefore, flexibility to support and not control the user is needed to provide visibility, make the functioning obvious and allow interpretation by including elements with which users can add their own meanings. [11] talks about designing the technology to be open for appropriation, which is in line with the guidelines that [2] introduced. These authors propose that the user be involved in the design process, and that the user be taken step by step through the activities. These presented design guidelines in this paper may lead to successful appropriation of technologies. Designers and programmers are advised to apply these guidelines for future technologies to create new sales and to open new markets for technologies.

IV. Conclusion

When it comes to talking about appropriation, there exist several different understandings. Not only when it comes to the granularity, the role, the qualification of appropriation, but even when analyzing all these various definition are these diverse understandings for the term 'technology appropriation'. In this paper diverse understandings were compared. The take-away-message from this analysis is that there is no common understanding of appropriation, but identifying all these various understandings opens a basis to explore a working definition of appropriation for researchers. Moreover, the presented design guidelines will support designers when creating future technologies that will allow successful appropriation.

References

[1] S. Bødker and A. M. Polli, "Between initial familiarity and future use: A case of collocated collaborative writing," in *COOP 2014-Proceedings of the 11th International Conference on the Design of Cooperative Systems, 27-30 May 2014, Nice (France)*, 2014, pp. 137–154.

[2] A. Dix, "Designing for appropriation," in *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI... but not as we know it-Volume 2*, 2007, pp. 27–30.

[3] J. Carroll, S. Howard, F. Vetere, J. Peck, and J. Murphy, "Just what do the youth of today want? Technology appropriation by young people," in *System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on*, 2002, pp. 1777–1785.

[4] M. Muller, K. Neureiter, N. Verdezoto, A. Krischkowsky, A. M. Al Zubaidi-Polli, and M. Tscheligi, "Collaborative Appropriation: How Couples, Teams, Groups and Communities Adapt and Adopt Technologies," in *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion*, New York, NY, USA, 2016, pp. 473–480.

[5] E. Balka and I. Wagner, "Making Things Work: Dimensions of Configurability As Appropriation Work," in *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work*, New York, NY, USA, 2006, pp. 229–238.

[6] P. Dourish, "Process Descriptions As Organisational Accounting Devices: The Dual Use of Workflow Technologies," in *Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work*, New York, NY, USA, 2001, pp. 52–60.

[7] A. MacLean, K. Carter, L. Lövstrand, and T. Moran, "User-tailorable systems: pressing the issues with buttons," in *Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people*, 1990, pp. 175–182.

[8] B. A. Nardi and J. R. Miller, "Twinkling Lights and Nested Loops: Distributed Problem Solving and Spreadsheet Development," *Int J Man-Mach Stud*, vol. 34, no. 2, pp. 161–184, Feb. 1991.

[9] M. Gantt and B. A. Nardi, "Gardeners and Gurus: Patterns of Cooperation Among CAD Users," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 1992, pp. 107–117.

[10] R. H. Trigg and S. Bødker, "From implementation to design: tailoring and the emergence of systematization in CSCW," in *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, 1994, pp. 45–54.

[11] W. March, M. Jacobs, and T. Salvador, "Designing technology for community appropriation," in *CHI'05 extended abstracts on Human factors in computing systems*, 2005, pp. 2126–2127.

[12] G. Stevens, V. Pipek, and V. Wulf, "Appropriation infrastructure: mediating appropriation and production work," *J. Organ. End User Comput. JOEUC*, vol. 22, no. 2, pp. 58–81, 2010.

[13] E. M. Rogers, *Diffusion of Innovations, 5th Edition*. Simon and Schuster, 1962.

[14] J. V. Wertsch, *Mind as Action*. Oxford University Press, 1998.

[15] S. Bødker and E. Christiansen, "Poetry in motion: appropriation of the world of apps," in *Proceedings of the 30th European Conference on Cognitive Ergonomics*, 2012, pp. 78–84.

[16] P. Turner, "How We Cope with Digital Technology," *Synth. Lect. Hum.-Centered Inform.*, vol. 6, no. 3, pp. 1–109, Jul. 2013.

[17] Stevens, G., "Appropriation infrastructure: Mediating appropriation and production work," *J. Organ. End User Comput. JOEUC*, vol. 22, no. Nr. 2, Special issue, pp. 58–81, 2010.

The Automatic Product Pick and Place of the Robotic 3D Printer Cell as a part of Digital Factory

Jakub Arm
Dep. of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
Email: xarmja00@stud.feec.vutbr.cz

Zdenek Bradac
Dep. of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
Email: bradac@feec.vutbr.cz

Vaclav Kaczmarczyk
Dep. of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
Email: kaczmarczyk@feec.vutbr.cz

***Abstract*—This paper deals with terms Industrie 4.0, Cyber-physical system, Virtual factory, Digital Twin, manufacturing cell, and automatic object pick and place. As a demonstration project, manufacturing cell containing 3D printer machine, robotic arm, and hypothetical output conveyor is presented. The control system architecture of this system as a cyber-physical system is approached to comply with Industrie 4.0 requirements. So the manufacturing cell behaves as an I4.0-component. The cyber twin of this system is created using Process Simulate software. The Digital Twin of the created product is also described as much as their integration into the hypothetical Digital Factory. The automatic pick and place method of the created product is approached based on the simulation using the created cyber twin.**

I. Introduction

Industrie 4.0 is the term number one in Europe. It is a project of German government and it should bring money to research and many companies for their intelligent modernization. It should also popularise technician and other disciplines to bring up educated people who will run the growing informational society. After all, the main idea is to digitalize the society and prepare to live in the cyber-physical world. This concept is complex across many disciplines like industry, energy, health, environment, mobility, or even society.

In the industry area, the Smart Manufacturing, what is also the domain in the USA, covers the concept of intelligent modernization and consists of areas like autonomous robots, simulation, system integration, big data, augmented reality, additive manufacturing, cloud computing, and cyber-security. The main part for researching is to define an architecture of the SF (Smart Factory). This should be the modern factory where products communicates with machines according to their virtual model. This concept uses parts like IIOT (Industrial Internet of Things), IoS (Internet of Services), and CPS (Cyber-Physical Systems). To accomplish the definition, the researchers has to solve many problems regarding inter-communication, security, standards and norms, decentralization, reference architecture, and engineering [1].

The Cyber-Physical System (CPS) integrate computing, networking, and physical components. The challenge is about the intersection of the physical and the cyber. This intersection combines engineering models and methods from mechanical, environmental, civil, electrical, biomedical, chemical, aeronautical and industrial engineering with the models and methods of computer science and engineering. The origin meaning of CPS term is more like embedded system that is characterized by tight integration and coordination between computation and physical process [2].

The CPS status in the concept of Industrie 4.0 is called I4.0-component describing an object (machine component or controller) plus an administration shell connected with other shells using Industrie 4.0 compliant communication. The communication is also used by cloud services like ERP (Enterprise Resource Planning) using SOA (Service-Oriented Architecture). The core of the I4.0-component is defined as an object such as electronic axis or robot. So it can be also a group of more CPSs [3]. I4.0-components are interconnected via information bus that delivers the unicast messages and transports broadcast messages for subscribers. The I4.0 component has the service oriented interface providing identification, diagnostics, process data, command functions, and safety functions [1].

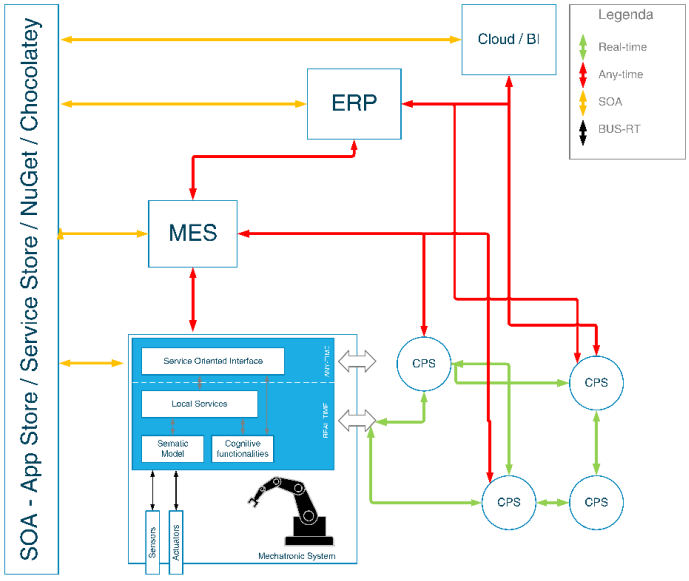


Fig. 1: The integration of I4.0-component (CPS) [1]

The Digital Manufacturing is the use of an integrated computer-based systems to create product and manufacturing process definitions simultaneously. Using virtual environment and information association, it allows optimization, commissioning costs reducing, product life-cycle management, or

virtual process automation [4]. The preferred standard tool for data association is OPC UA (OLE for Process Control Unified Architecture) as a Process control level in ISA 95 hierarchy [5].

The Digital Factory as the Industrie 4.0 platform uses the concept of the Digital Manufacturing for the cellular manufacturing. So every cell has its virtual model like the CPS called Cyber Twin. Therefore the production can be simulated and optimized using software tools, e.g. PLM (Product Life cycle Management) by Siemens. PLM tries to cover all the product phases such a specification and design, manufacturing plan, manufacturing preparation, manufacturing, and services. At the engineering level, factories, cells, and even machines can be made in virtual reality. Therefore some tools like collision detection, CNC tool path trajectory optimization and generation (NX CAM software), assembly line optimization (Tecnomatix Process Simulate) can be used.

The association of the virtual model of a product with the real one is called Digital Twin. Digital Twin is created by the specification phase of a product. Then it is growing as the product is going through the phases. Digital Twin can be used as a believable virtual clone for simulation, optimization, or documentation. The simplest implementation is based only on reading sensor data. Industrial twin present information and data in an asset/device model graph based on the physics. Oracle company even uses a special strategy containing Virtual Twin, Predictive Twin, and Twin Projections component [6].

Industrie 4.0 relies also on additive manufacturing technology. It is more economical than classical milling technology. Nowadays, many kind of plastic is used or even some iron alloys can be used. This technology is also preferred by rapid manufacturing. 3D printing technology is based on composing the final product layer by layer using thin plastic filament that agglutinates together.

II. The system architecture as an I4.0-component

In our system, there is a 3D printer and a robotic arm. The 3D printer is the classical box plastic filament printer where XY-core style is used with bed on Z axis. The robotic arm is the classical 5DOF robot where the last joint is a grasping grip. The task of the robotic arm is to pick the printed products and be able to put them on a output conveyor.

Each component has its own specific motion controller that can be integrated into a system. So each component can be considered as a separated CPS that communicates real-time with other CPSs and provides an interface to communicate with a superior system in the form of G-commands. As the I4.0-component object, we consider the whole system because a superior system can not command 3D printer movement and robotic arm movement separately.

The system architecture depicted on Fig. 2 shows the real-time communication among subsystems and the interface for the information bus that is service oriented. The administration shell functions are implemented in this interface so the system acts as an I4.0-component that can be integrated in the virtual factory system. The service is based on REST (Representational State Transfer) service hosted on the central unit using *Flask* library in Python language.

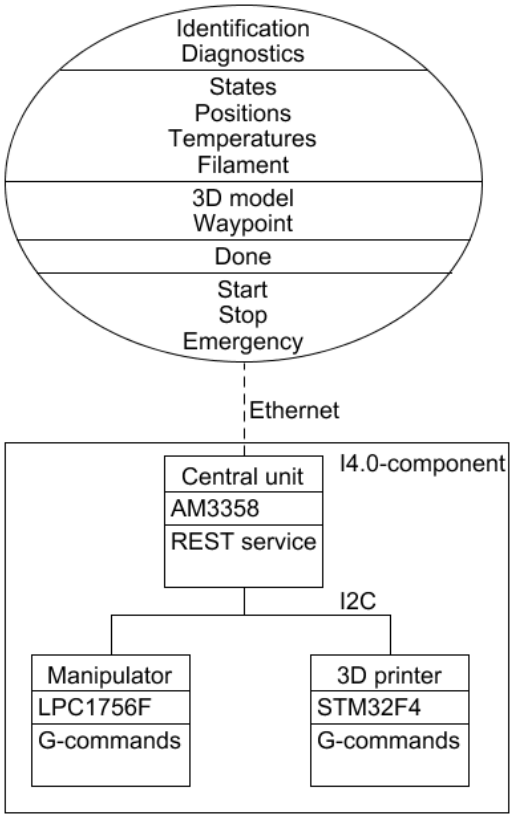


Fig. 2: I4.0-component architecture

In the virtual factory system, there has to be an execution service responsible for virtual process execution that uses cyber twins of machines and digital twins. In our case, the work cell cyber twin is described in Chapter II-A and the product digital twin is described in Chapter II-B. The execution service initiates data transmission through the information bus where state of the real system is copied as the state of the cyber twin and digital twin data are filled. Also other parts of the digital factory are connected like a conveyor for the created products. This simulation is not part of this project and planned for the continuing project.

A. Cyber Twin

The created cyber twin of the real work cell is cut down only to the virtual model for now to check geometry collision. It is implemented in Tecnomatix Process Simulate software that can execute simulation according to the specified CNC or PLC program.

The function of the cyber twin (see Fig. 3) in this project is to check non-colliding robotic manipulator path when the created product represented using the Digital Twin method is picked up and the successful action is performed (see Chapter III-A). The improvement lies in the better physics simulation such a product friction or rigidity.

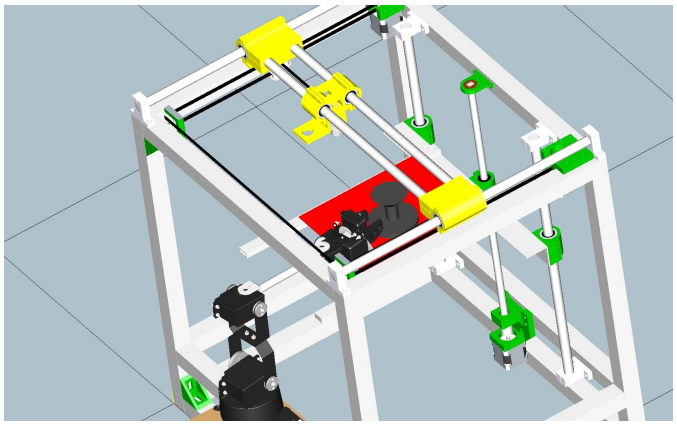


Fig. 3: The created cyber twin of the cell

B. Digital Twin

The specific Digital Twin associated with the specific product is growing informationally through the product life cycle phases. It covers data from the product 3D model and material to the machine where it was created, needed resources, and information concerning final quality. This data format is summarized in table I.

The proposed picking algorithm (see Chapter III-A) should be executed between specification and planning phase because the result of algorithm feasibility can be taken into account in the planning phase. In that case, the product will be sweeped away by the manipulator to the output box to be sorted out.

III. Automatic picking of the created product

Automatic picking of the created product from the bed of a 3D printer is a complex problem that should be solved rather at the mechanical level because of the product sticking and then at the software level using meta-heuristic methods such a genetic algorithm or neural network. The proposed method combines computer vision, processing, simulation, and optimization algorithms to get the optimal picking point or non-feasible solution. The followed assumptions are taken:

- The sticking is minimal because of suitable bed material (smooth glass) or the product is quickly warmed up and then cooled down to unstick the product. So the product can be without any effort picked by some manipulator.
- The precise product picking is needed for the precise placing, e.g. on the conveyor. In other cases, a plow can be used to put the product off. Or the robot can sweep away the product to a box that has to be manually or using another robot sort out.
- The product can be approximated using primitives such a box, sphere, or cylinder without any impact because of difference between the approximation and a real product causing a feasible picking solution does not work. This is accomplished in case of CAD-like models.

A. The proposal of picking by the robotic arm

The proposal is based on the method pipeline using the simulation as the main arbiter. Firstly, the product model is decomposed to primitives. Then the each primitive is processed to get possible picking points. Then the possible solutions are simulated in the virtual environment to check reachability and collision free. Therefore only the feasible solutions remain if any. Then an optimization is performed to get the optimal picking point solution.

The primitive fitting method decompose a mesh to the primitives such a plane, sphere, or cylinder. It is based on point clustering by the best fit to the primitive. This method can be tweaked be the sensitive parameter that corresponds to the number of the approximated clusters. This method is suitable for decomposition of CAD models [7].

The trickiest part of the algorithm is the generator of possible solutions (picking points). The possible solution is stored if the each primitive can be grabbed by the specified grip. Every primitive such as sphere, box, and cylinder can be grabbed using infinite grabbing points. So the points are trimmed using a grid that can be tweaked. The future improvement is to take into account the dynamics of the object to trim out solutions when a object would move while is being picked.

The most important and innovative part of the picking point optimization problem is the simulation in the virtual reality. This is achieved using Tecnomatix Process Simulate software that can be controlled externally via .NET API, e.g. in C# environment. This simulation returns the information of non-colliding robot orientation with the printer or with other part of the product when it is exactly on the picking point (solution of the previous algorithm). The collisions are detected by Collision Detection module using the algorithm presented in [8]. The simulation can also throw out all possible solution therefore there will be no feasible solution.

The optimization chooses the optimal solution from the feasible solutions if any. This is the unconstrained problem over the definite set of solutions. So the optimal solution is a solution that minimizes the penalty function (see Chapter III-B). That is written as follows:

$$x^* = argmin_{x \in S} f(x) \tag{1}$$

where x^* is the optimal solution, S is the finite set of the possible solutions, and $f(x)$ is the penalty function.

So the proposal is to combine all mentioned methods to check feasibility of picking and to find the optimal picking point. The pipeline of the proposed algorithm is depicted on 4. This algorithm can also produce the non-feasible state so it is recommended to use this algorithm also in the product design phase instead only in the manufacture preparing phase.

B. The penalty function

The penalty function will find the best solution from a finite set of solutions given the criterion. Criterion is based on the printer and robotic manipulator disposition. The penalty function uses the calculation of the distance that is expressed using the squared Euclidean metric.

To define the penalty function, some assumptions are taken. The possible solutions are feasible in reality. The solution as

Specification	Planning	Prepare	Manufacture	Services
3D model [mesh]	Filament length	Change filament	Program flow	Quality check
Color	Filament type	Extruder temperature	Measure time	Documentation
•	Estimated time	Bed temperature	Real-time data	Identification
•	Schedule	Check state	Machine	•
•	Position	Is grabable	•	•

TABLE I: Life cycle of the 3D printer product

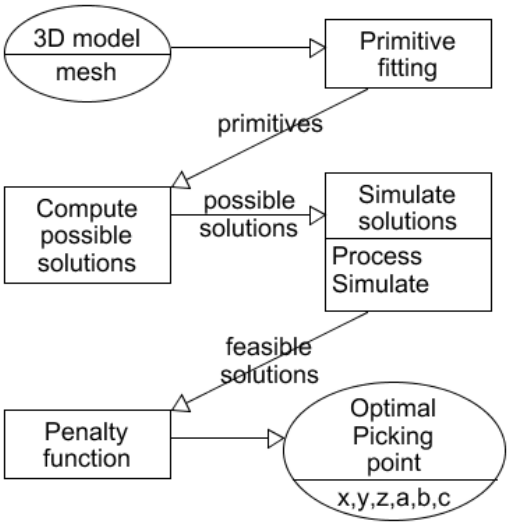


Fig. 4: Pipeline of the proposed picking algorithm

the picking point is defined in 3D space as $X = [x, y, z, a, b, c]$ where a, b, c are angles around x, y, z axis in radians. Robotic manipulator is placed from the printer origin in positive y axis. The picked object is rigid so every picking point leads to pick the object all right. The penalty functions tries to reduce an impact of object dynamics.

The penalty function is therefore defined as follows:

$$f(X) = (x-x_c)^2+(y-y_c)^2+(z-z_c)^2+a^2+b^2+(c-\frac{\pi}{2})^2 \quad (2)$$

where $[x_c, y_c, z_c]$ is the center mass point of the object.

IV. Conclusion

We have proposed the interface of the manufacturing cell containing 3d printer and robotic manipulator according to the I4.0-component concept. We have also described the created printer product life cycle as the Digital Twin. We have created the manufacturing cell in the Tecnomatix Process Simulate virtual environment as a part of the cyber twin. We have proposed the pick and place method based on 3D primitive fitting, optimization, and simulation using the created cyber twin to get the optimal pick point of the object under the mentioned circumstances. In generally, object pick and place is very complex problem that should be solved for use in a modern factory. This project is aimed to the extension of the practical knowledge about implementing Industrie 4.0 concept as the Digital Factory.

Acknowledgment

The completion of this paper was made possible by the grant No. FEKT-S-17-4234 - 'Industry 4.0 in automation and cybernetics' financially supported by the Internal science fund of Brno University of Technology.

References

[1] J. Tichy, "Systemy mes/mom v prostredi industry 4.0," Brno University of Technology, Tech. Rep., 2017.

[2] P. Leitão, A. W. Colombo, and S. Karnouskos, "Industrial automation based on cyber-physical systems technologies: Prototype implementations and challenges," *Computers in Industry*, vol. 81, pp. 11 – 25, 2016, emerging ICT concepts for smart, safe and sustainable industrial systems. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0166361515300348

[3] L. Monostori, B. Kádár, T. Bauernhansl, S. Kondoh, S. Kumara, G. Reinhart, O. Sauer, G. Schuh, W. Sihn, and K. Ueda, "Cyber-physical systems in manufacturing," *CIRP Annals - Manufacturing Technology*, vol. 65, no. 2, pp. 621 – 641, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0007850616301974

[4] Siemens PLM Software, "Digital manufacturing," Siemens, Tech. Rep., 2017.

[5] ISA Committee, "Isa-95 standard," Tech. Rep., 2014.

[6] Oracle, "Digital twins for iot applications," Oracle Corporation, Tech. Rep., 2017. [Online]. Available: http://www.oracle.com/us/solutions/internetofthings/digital-twins-for-iot-apps-wp-3491953.pdf

[7] M. Attene, B. Falcidieno, and M. Spagnuolo, "Hierarchical mesh segmentation based on fitting primitives," *Vis. Comput.*, vol. 22, no. 3, pp. 181–193, Mar. 2006. [Online]. Available: http://dx.doi.org/10.1007/s00371-006-0375-x

[8] M. S. U. Qadeer, "Automatic coordination and collision avoidance - using sequence planner and process simulate," Chalmers University of Technology, Tech. Rep., 2011.

Localization of Current Fields in the Vicinity of a Railway Platform

Jiří Cigánek
VŠB - Technical University of Ostrava
Ostrava, Czech Republic
ciganek@vsb.cz

***Abstract* — The article deals with an examination of current fields in the vicinity of a railway platform of the SUDOP type. It compares a fault-free state in which the rail is connected with a gabion through a lightning arrester and a case in which there occurs a breakdown of the lightning arrester. The solution is reached with the use of the Comsol Multiphysics program, which solves partial differential equations by means of the finite element method.**

Keywords: platform; current density; rail; gabion

I. INTRODUCTION

The occurrence of the so-called stray currents represents a major problem in the vicinity of a railway track. Stray current is a current which passes from a rail during the crossing of a railway vehicle into the earth and returns back to the power supply station. These currents disrupt concrete foundations in the vicinity of the railway substructure. Station platforms are some of the places through which these currents pass. This contribution deals with an examination of these fields in the railway platform of the SUDOP type. It compares a fault-free state in which a rail is connected with a gabion through a lightning arrester by means of a conductor and a case in which there occurs a breakdown of the lightning arrester. In this case there therefore occurs a significant voltage surge in the gabion and thus also passage of stray currents.

The term "platform" refers to a part of the railway substructure (a transport area and a pathway) intended for boarding or alighting of passengers and for handling of minor consignments. A gabion or a gabion structure is an element of a cube or cuboid shape made of a hexagonal steel wire mesh, welded steel nets or possibly high-strength polymer geogrids filled with natural aggregate, quarried rock, earth, recycled material, etc.

II. MATHEMATICAL MODEL

The basic field quantity describing current fields is a current density vector $\boldsymbol{J}$. It is a differential quantity and it therefore determines the quantity of the current flowing through the elementary area for every place of the solved space. The current density J depends on the resistivity ρ, or rather on the conductivity σ according to the differential form of the Ohm's law.

$$\boldsymbol{J} = \sigma.\boldsymbol{E} \tag{1}$$

In every place where it is possible to determine the intensity of the electric field $\boldsymbol{E}$ and where the conductivity is known, it is also possible to determine the current density vector. Its direction in the isotropic medium is the same as the direction of the vector of the electric field intensity.

The way to determine the current density vector is more complicated in the anisotropic medium where conductivity is in the form of a tensor. In most cases a tensor has only diagonal components, this means there are various conductivities in the direction of various coordinates, but the intensity of the electric field in the direction of one coordinate will not affect the current density in the direction of another coordinate, i.e. for example

$$J_x = \sigma_{xx}.E_x, \quad J_y = \sigma_{yy}.E_y, \quad J_z = \sigma_{zz}.E_z, \tag{2}$$

The differential form of the Ohm's law is then

$$J = \sigma.E = \sigma.(-\operatorname{grad}\varphi) \tag{3}$$

where

$$\sigma = \begin{bmatrix} \sigma_{ii} & 0 & 0 \\ 0 & \sigma_{jj} & 0 \\ 0 & 0 & \sigma_{kk} \end{bmatrix} \tag{4}$$

Electromagnetic problems associated with solving current fields in a railway superstructure and a substructure work generally with a resulting current, or current density, which is connected with intensity, or potential. Forced quantities are potentials or the primary (forced) intensity of the electric field $\boldsymbol{E}$v. The primary intensity determines the distribution of charges and currents and it is used for specifying this distribution. For the purposes of an analysis it is possible to accept that the primary intensity determines the secondary distribution of charges and currents and thus the secondary intensity $\boldsymbol{E}$sec, i.e.:

$$\nabla \times \boldsymbol{E} = -\boldsymbol{\partial} B/\boldsymbol{\partial} t \tag{5}$$

$$\nabla.\boldsymbol{D} = \rho_c \tag{6}$$

$$\nabla \times \boldsymbol{H} = \boldsymbol{J} \tag{7}$$

$$\nabla.\boldsymbol{B} = 0 \tag{8}$$

where the symbol ρ_c does not stand for resistivity, but for the volume density of the total charge. Similarly to the intensity $\boldsymbol{E}$, the current density $\boldsymbol{J}$ has also two components, i.e. forced

current density J_v and eddy current density $\sigma.E$. Material relationships must be added to these equations on the boundary

$$D = \varepsilon.E, \quad B = \mu.H, \quad J = \sigma.E \tag{9}$$

A dynamic solution of Maxwell's equations includes also the influence of eddy currents and it is advisable to use here a continuous vector potential

$$\nabla \times \boldsymbol{A} = \boldsymbol{B} \tag{10}$$

$$\nabla \times \boldsymbol{H} = Jv + \sigma E = Jv - \sigma.\partial A/\partial \tau \tag{11}$$

therefore $$\nabla \times (\nabla \times A) = \mu(Jv - \sigma.\partial A/\partial \tau) \tag{12}$$

It is often advisable to perform the solution of direct-current fields for the scalar potential of the electric field. The advantage of using the scalar potential for large ground fields resides in the replacement of the vector field *E*, which is given by 3 components, with only a single scalar variable. The boundary value problem is therefore described with one scalar variable φ and after its solution all components of the electric field intensity and subsequently also all components of the current density are obtained. The potential has an important role in connection with converting a spatial model into a circuit model described by a node voltage method.

The Laplace's equation in the Cartesian system applies to the scalar electric potential in homogeneous environments without sources:

$$\nabla^2\varphi = \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2} = 0 \tag{13}$$

The duality of the stationary current field with the electrostatic or magnetostatic field was used in modelling 3D direct-current fields of the railway electrification system.

III. Creating a Model

The Comsol Multiphysics program was chosen for the solution; this program makes it possible to use the finite element method to solve physical problems described by partial differential equations. These problems are generally solved on the basis of environment definitions which are described by partial differential equations and on boundary conditions entered in points, on edges or surfaces of a given model. The results can be obtained in several steps.

A. Geometric Layout

The first step is to choose and draw a geometric form of the solution. Its layout is shown in Figure 1.

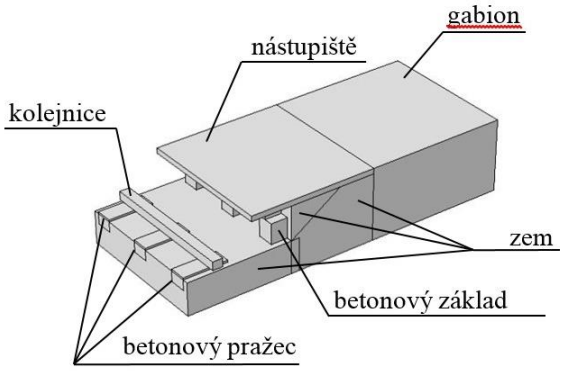


Figure 1. A geometric layout of the model

kolejnice = rail nástupiště = platform

gabion = gabion betonový pražec = concrete sleeper

betonový základ = concrete foundations zem = earth

The geometric layout is subject to several simplifications. The rail was drawn as a rectangle whose height and width correspond to the actual dimensions of a rail. A further simplification was made on the gabion. Although it is normally a mesh of wires, as was stated above, in the model it was simplified because it was modelled as a full surface. This simplification was necessary because of entering material properties. Other components and their dimensions correspond to the reality.

Two models were made for the purposes of the calculation; the first model corresponds to the fault-free state, where the rail is connected with the gabion through a lightning arrester. The model was prepared in a way that the rail is not connected with the gabion. The second case represents a situation in which there is a breakdown of the lightning arrester and the rail in the model is connected with the gabion by means of an insulated steel conductor.

B. Defining the Parameters

The material parameters were defined by means of the electric conductivity σ and the relative permeability μ_r. These are the primary parameters for calculating the model.

The conductor connecting the rail with the gabion was defined by means of the relative permeability μ_r = 4 000 and the electric conductivity σ = 1.12e7 S/m. The insulation of the conductor was defined by μ_r = 1 and σ = 1e-12 S/m. The material properties of the rail were defined by means of the measured BH characteristics. Table values of μ_r = 1 and electric conductivity σ = 10 S/m were assigned to the concrete sleepers under the rail, the concrete footings under the platform and the platform itself. The gabion was defined with the values of relative permeability μ_r = 1 and σ = 4.1e6 S/m.

The most difficult task was to define various types of earth. It is problematic to determine precisely the material properties of earth because its precise composition is unknown. The resistivity of earth depends on its composition; in the case of soil it can range between $\rho = 10^{-1} \div 10^{2}$ Ωm; resistivity of clay is $\rho = 10 \div 10^{2}$ Ωm and resistivity of sand is $\rho = 10^{2} \div 10^{5}$ Ωm. A gravel subgrade will also have higher resistivity – in the range of $\rho = 10^{2} \div 10^{5}$ Ωm. The Comsol program requires entering of data by means of electric conductivity; resistivity

was therefore converted to conductivity. In the present case the values for the earth were selected in the following way: compacted antifreeze material $\mu_r = 1$, $\sigma = 0.1$ S/m, hard surface $\mu_r = 1$, $\sigma = 0.001$ S/m.

C. Entering the Conditions

The conditions were entered in the following way: the condition of the electric potential $\varphi = 1\ 000$ V was set for the upper surface of the rail. The electric potential value is the highest presumed value which can occur on the rail. The condition of a zero electric potential $\varphi = 0$ V was entered on the bottom side of the model. Other external boundaries were assigned the condition of electric insulation $\boldsymbol{n} \times \boldsymbol{J} = 0$.

D. Covering the Model with a Network

After the boundary conditions have been set, the model is prepared for generating a network in whose nodes the needed partial differential equations were calculated. The parameters for dividing the network can be set manually on the individual volumes, surfaces or lines. Furthermore, it is also possible to set the density of the network. The denser the network is, the more nodal points there are and the higher the number of the partial differential equations is. However, the calculation becomes more difficult at the same time. In our case a fine division corresponding to the number of approx. 750 000 nodal points was sufficient. As Figure 2 shows, the program makes the network intuitively denser in the smaller areas.

A stationary linear solver was used for the solution.

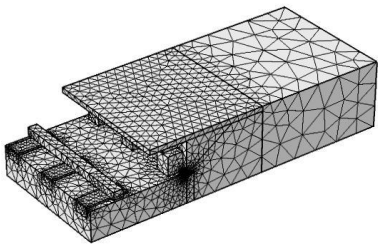


Figure 2. The model covered with a network

E. Simulation Results

As was already stated above, the results are calculated in two models.

First, the model which simulates a fault-free state will be considered. This means the state in which the rail is connected with the gabion through the lightning arrester.

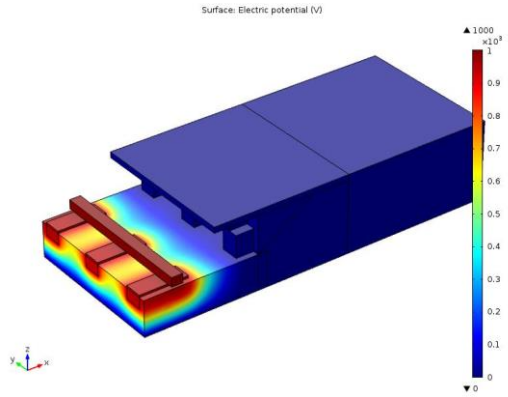


Figure 3. Distribution of the electric potential, a fault-free state

Figure 3 shows distribution of the electric potential in the model. The figure shows that the value of the electric potential decreases in the direction from the rail to the earth. The value of the electric potential in the platform is approx. 2 V, the value of the potential in the concrete foundations under the platform is 3 V and the value of the potential in the gabion is around 1 V.

The distribution of the current density has a higher expressive ability. (Fig. 4) The maximum value of the current density is in the rail, 5.5e4 A/m^2. In the concrete foundations under the platform, in the part which is covered by the earth, the current density is about 150 A/m^2, in the part above the earth the current density value is about five times lower, i.e. 30 A/m^2. In the area of the platform the current density value is very low, around 10 A/m^2. In the gabion the current density value can be considered to be 0 A/m^2, because the calculated value is in the order of magnitude 10^{-5} A/m^2.

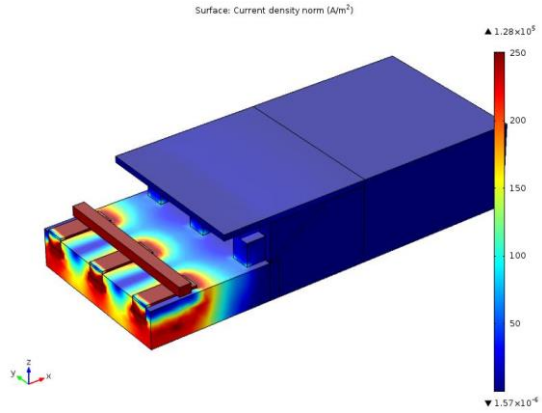


Figure 4. Distribution of the current density, a fault-free state

The second model simulates a breakdown of the lightning arrester which protects the connection between the rail and the gabion. In the model the conductor is anchored to the bottom side of the rail between the second and the third sleepers and it is connected with the gabion in the depth of 0.5 m under the ground. The results have been made transparent so that the effect of the conductor on the concrete foundations can be seen.

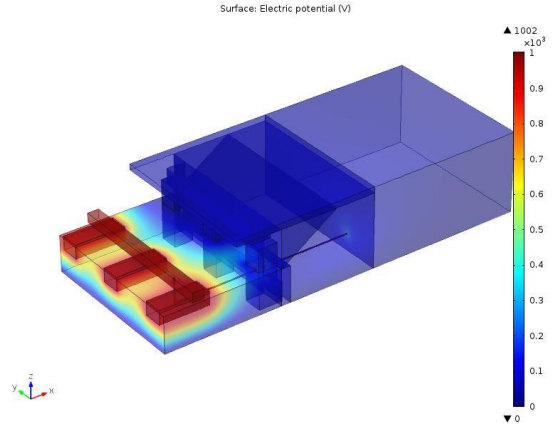


Figure 5. Distribution of the electric potential, a faulty state

The maximum value of the electric potential is again 1,000 V in the rail. (Fig. 5) The other values, however, are now significantly higher than in the case of the fault-free state. In the platform the value of the electric potential is approx. 40 V. In the concrete foundations the potential reaches the values of approx. 60 V and the value of the electric potential in the gabion is about 20 V.

The current density reaches significantly higher values, which is apparent in Figure 6.

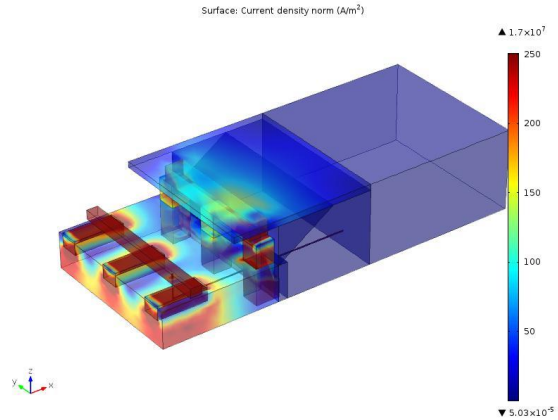


Figure 6. Distribution of the current density, a faulty state

The maximum value of the current density is again present in the rail. In the concrete foundations, in the part which is under the level of the ground, the current density is dependent on the proximity of the conductor which passes between the second and third concrete foundations. (Fig. 6) Its value is, therefore, in the range from approx. 500 A/m^2, in the case of the most distant concrete foundation, up to 2 000 A/m^2, which is the value in the third concrete foundation. In the level above the ground the current density reaches values from 100 A/m^2 to 450 A/m^2. The current density value in the platform is the highest in the places between the second and the third concrete foundations; there the value is 145 A/m^2. In contrast to the fault-free state the current density in the gabion also increases, to the value of approx. 1.5 A/m^2.

IV. Conclusion

The aim was to prove if current density and consequently a stray current pass through concrete foundations. And if yes, what values they reach. This simulation was calculated with the maximum value of the electric potential on the rail, for example, during the passage of a train. The resulting values give an idea about the distribution of the current density and therefore about the passage of the current inside the concrete foundations and also in the earth under the platform. The simulated results confirm the supposition that the higher the value of the electric potential is, the higher the value of the normal current density is.

As the models given above show, a stray current passes from the rail to the concrete sleeper and enters the earth. Then it passes through the vertical concrete foundations. A part of the current flows directly in the earth, while another part of the current flows through the surface of the platform and the gabion into the earth.

In the case of a fault-free state the current density values are very low and in some parts even negligible. However, despite this fact the current density value in the concrete foundations under the surface of the ground is approximately 150 A/m^2. This certainly does not have a positive effect on the ageing of the concrete and there may occur degradation of the concrete in these places. Nevertheless, a much worse case occurs if there is a breakdown of the lightning arrester. In comparison with the fault-free state the current density values are very high in this case. The current density value in the platform increases by about 5 times. If the worst case is considered, the current density value in the concrete foundations increases by approximately 13 times, which represents a really very high value.

As can be seen, this fact has an adverse effect on the concrete sleepers under the rail and also on the concrete foundations of the platform. One of the possible consequences of these stray currents is that their occurrence may cause degradation of the concrete and subsequently its cracking.

The illustrative figure above shows how a stray current disrupts the concrete foundations in the vicinity of the railway substructure.

V. Bibliography

[1] J.Paleček, Selected Chapters from Electrical Engineering Problems in Transportation Engineering university textbook, VŠB-TU Ostrava 1996

[2] L. Ivánek, Models and Transfer Parameters for Propagation of Reverse Currents in Electric Traction post-doctoral thesis, Ostrava 1998

[3] ČSN EN 50122-2 Railway applications – Fixed installations - Part 2: Protective provisions against the effects of stray currents caused by DC traction systems as amended

[4] Promotion material from the company Humusoft

Conductive Threads for Microwave Applications

Miroslav Cupal
Department of radio electronic
Brno university of technology
Brno, Czech Republic
cupalm@phd.feec.vutbr.cz

***Abstract*—This paper deals with parameters of conductive threads for using in microwave and millimeter waves field. The resistance and test application are presented in the paper. The thirteen threads of different manufacturers were measured and their properties are compared. The conductive materials of threads are carbon, silver and stainless steal.**

Keywords—conductive threads, resistivity measurement method, microwave, waveguide

I. Introduction

Wearable electronics are in the front of interest in past years. Many types of electronics devices are integrated into clothes for communication or control of other devices like smart phones or other wearable electronics [1]. The conductive threads are used as low voltage supply wires or as low-frequency signal wires. Some threads are used for manufacturing of protective clothes for EMC test workers and etc [2].

There are many types of the conductive threads that are used in wearable applications in smart clothes. The threads which have been measured are of two majority manufacturers. The first manufacturer is company Amann which is a leader in conductive threads field, the second is Bakaert other manufacturers are Statex, Imbut, Shakespeare and R-stat. The conductivity threads measurement methods are described in chapter II.

I measured 13 conductive threads of three different conductive materials of different thickens and non-conductive part. The threads are in Table 1. Manufacturer, type, conductive material, non-conductive material and thickness of all threads are in Table 1. The thickness of the threads is given in unit tex or dtex. The unit tex is defined by follows:

$$T(tex)=\frac{M(g)}{L(km)}, \qquad (1)$$

where T is linear mass density, M is mass in (g) and L is the length in km. The unit tex is used for natural fibers, for synthetic fibers is used unit dtex defined as:

$$1\,dtex=\frac{1\,g}{10\,km}. \qquad (2)$$

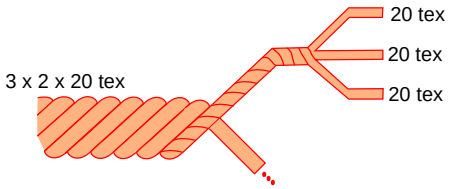


Fig. 1. Conductive thread diagram 120 tex.

TABLE 1. CONDUCTIVE THREADS OVERVIEW

Name	Manufacturer	Materials	MND [tex/dtex]	Resistivity [Ω/m]
F109	Shakespeare	PAD/ carbon	144 dtex	NA
Bekonix VN 14.1.9.200Z	Bekaert	steel	110 tex	70
silveR-stat	R-stat	PAD/ silver	240 dtex	NA
X-static	NFT	PAD/ silver	160 dtex	NA
Elitex	Imbut	PAD/ silver	220 dtex	70
Elitex	Imbut	PAD/ silver	440 dtex	20
Bekonix VN 12.1.2.100Z	Bekaert	steel	235 tex	30
Bekonix VN 12.3.2.175S	Bekaert	steel	760 tex	9.3
Bekonix VN BK 50/3	Bekaert	PES/ steel	50/3 Nm	NA
I-texh 20	Amann	PES/ carbon	111 x 3 dtex	NA
C-tech 80	Amann	PES/PAD/ silver	93 x 3 dtex	NA
Schieldtex 235f	Statex	silver	560 dtex	< 100
Silver-tech 120	Amann	PAD/ silver	93 x 3 dtex	< 3000

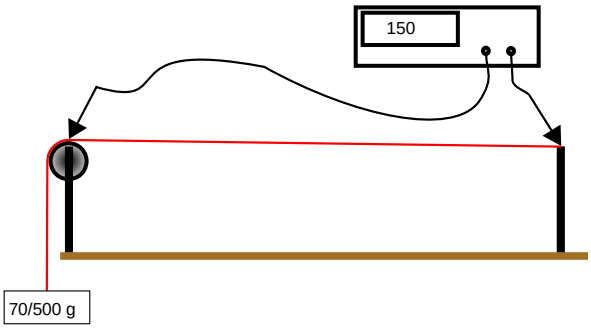


Fig. 2. Measuring conductive treads.

Threads are usually manufactured from two or more fibers. Conductive threads can combine conductive and non-conductive fibers. The thickness of the thread from some fibers can be determined as 3 x 2 x 20 tex (dtex). It means 6 fibers (20 tex) plied together as shows Fig. 1.

There are three groups of thread divided by the conductive material. The threads can be divided to carbon, silver, and stainless steel. The resistivity in Table 1. is resistivity given by the manufacturer, if it is known.

II. Measurement Methods

The threads are fabricated of the different combination of conductive and non-conductive parts. The conductance is given by the type of conductive material and its amount in a thread, a structure of thread and non-conductive material. Different threads can be used in different applications (DC, EMC shielding, signal wires).

The resistivity of the threads has been measured with RLC meter Instek LCR-819 and common multimeter. The one meter of each thread has been measured under two loads and two different tensions. The thread has been placed between two fix point and one end has been loaded. The work place diagram is in Fig. 2. The loads were 70 g and 500 g. Some threads are weak and they were measured only with 70 g load. Measured results for all threads are in Table 2. The resistivity was measured at DC, 1 kHz, 10 kHz and 100 kHz for both loads.

The threads with resistance lowest than 100 Ohm were chosen for next testing. The substrate integrated waveguide has been designed and side walls were sewed by conductive threads. The substrate integrated waveguide was designed on common substrate FR4 with relative permittivity 4.2 and $tg\delta = 0.02$. The cutoff frequency has been chosen as 5 GHz and operation frequency 8 GHz. The dimensions of metallized vias created sidewalls were chosen as diameter 1 mm and the distance between two vias are 1 mm. Metalized vias of one testing example were created by common PCB technology and this example is as a standard for comparison.

III. Simulated and Measured Results

The resistivity was measured for all frequencies and both loads and it is shown in Table 2. Carbon threads have resistance around hundred mega Ohms and are able to electromagnetic for low power. Two stainless steel threads (Bekonix VN 12.3.2.175S and Bekonix BK50/3) have resistance around kilo–ohms. This higher resistance is given by structure of threads. The other two threads have resistivity around tenth ohms. These two threads are really good conductive for DC and low frequency signals but their MND is relatively height and it is difficult to sew with them. The last part are silver threads. There are big differences between resistivity of these threads but the resistivity is relatively low for all threads. The resistivity is from a hundred ohms to the unit of ohms. The threads Elitex 220 dtex and Elitex 440 dtex have the best resistance and their MND is relatively low and the threads are able to hand and machine sewing.

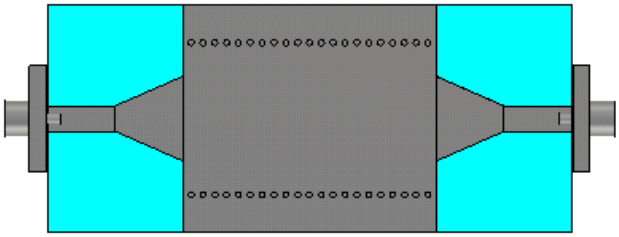


Fig. 3. Substrate integrated waveguide for conductive thread test.

Table 2. Measured resistivity

	70 g				500 g			
T. no.	DC	1 kHz	10 kHz	100 kHz	DC	1 kHz	10 kHz	100 kHz
1.	7.70 M	2.54 M	2.50 M	697.70k	NA	NA	NA	NA
2.	73.33	68.92	68.9	69.72	69.61	68.94	68.96	69.42
3.	226.74	228.8	228.6	229.4	234.05	234	232.2	233.2
4.	232.25	255.4	254.6	254.6	NA	NA	NA	NA
5.	75.26	72.4.2	74.92	63.6	NA	NA	NA	NA
6.	14.08	14.22	14.25	14.30	15.52	15.56	15.70	15.68
7.	50.19	36.20	37.46	31.80	38.35	29.88	29.80	29.86
8.	11.01	9.23	9.18	9.32	10.82	9.18	9.19	9.25
9.	3.02k	4.04k	2.23k	13.34 k	1.60 k	1.90k	1.67k	1.53k
10.	7 k	3.26 k	3.82k	2.39k	2.5k	2.4k	2.34k	2.63
11.	>200 M	NA	75.78 M	0.5M	>200 M	NA	NA	NA
12.	15.00	13.88	13.95	13.95	14.11	13.32	13.64	13.67
13.	758	848	823.2	743	643.02	594.32	633	652

The threads no. 2, 5, 6, 7, 8, 12 have been chosen to creating sidewalls of substrate integrated waveguide. The substrate integrated waveguide has been simulated in CST Microwave Studio. The waveguide is depicted in Fig. 3. The

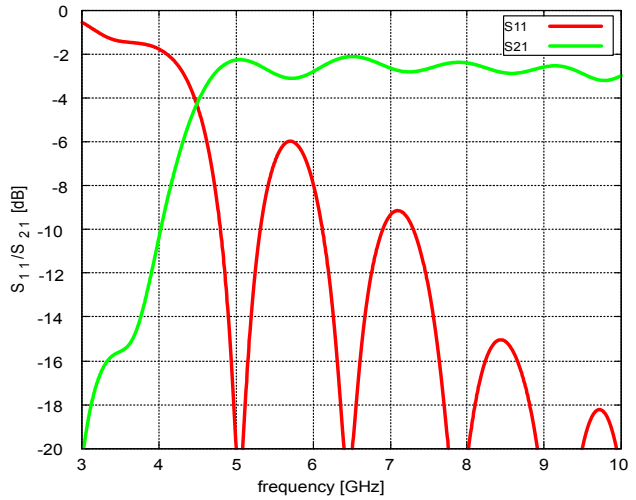


Fig. 4. Simulated transmission and reflection coefficient of SIW structure.

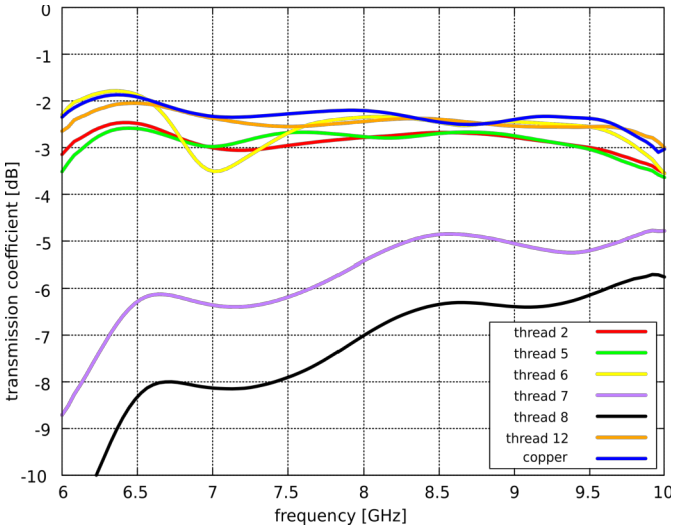


Fig. 5. Measured transmission coefficient of SIW structure.

structure did not optimize for best S_{21} but it is only as a testing example. The cutoff frequency was set at 5 GHz. The simulated results are shown in Fig. 4. The transmission coefficient is between -2 and -3 dB.

The simulated substrate integrated waveguide has been manufactured in two examples. The waveguide with metallized vias as conductive side walls and the second one has vias only. Through these vias were sewed by conductive threads create sidewalls. Measured results for metal vias waveguide and for conductive threads waveguide are in Fig. 5. The transmission coefficient S_{21} for copper vias is in good agreement with simulation results.

The conductive threads 2, 5, 6 and 12 are silver threads with different DC conductivity and very similar result as side walls of FR4 substrate integrated waveguide. The thread no. 12 is in very good agreement with copper vias waveguide. The thread no. 6 has very good result too, but there is a gap around 7 GHz. The transition coefficient is around -2.5 dB. The threads no. 2 and 5 are a little worse than previous. The transmission coefficient is between -2.5 dB and -3 dB. All

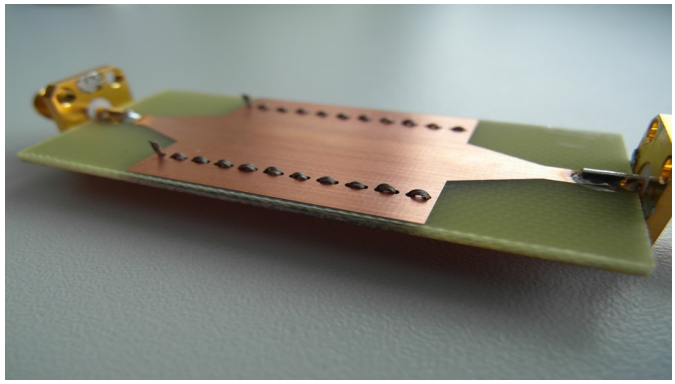


Fig. 6. SIW structure with conductive thread.

these threads are suitable for conductive walls or other MW components. The threads no. 7 and 8 are stainless steel threads. The transmission coefficient of the waveguide rapidly changes with frequency. S_{21} is -6.4 dB for the 7th thread and -8.1 dB for 8th threads at frequency 7 GHz. The SIW and thread no. 7 is in Fig. 6. The threads 7 and 8 are harder than other and the conductive connection between copper part of waveguide and threads is not ideal.

IV. Conclusion

Thirteen conductive threads from three manufacturers were measured. These threads were different types of different conductive materials. The DC and low frequency resistance have been measured. The thread Elitex 440 dtex, Bekonix VN 12.1.2.100Z, and Statex have the resistance between 10 and 20 Ohms per meter. The threads have been tested as substrate integrated waveguide sidewalls, and compared with SIW with copper vias. The thread no. 12 achieved results very similar to copper one. The threads no 6, 2 and 5 achieved very good results too. The transmission coefficient S21 was higher than -3 dB. The threads no. 7 and 8, have unsatisfactory results as conductive walls.

Acknowledgment

The research was supported by the MPO TRIO project FV10087. For simulations and experiments, equipment of the SIX Center was used (the grant LO14001). A partial support of Brno University of Technology is acknowledged (the grant FEKT-S-17-4713).

References

[I] C. Zysset, K. Cherenack, T. Kinkeldei and G. Tröster, "Weaving integrated circuits into textiles," International Symposium on Wearable Computers (ISWC) 2010, Seoul, 2010, pp. 1-8.

[II] L. O. Hoeft and E. W. Tokarsky, "Measured emc shielding characteristics of fabric made from metal clad aramid yarn and wire," IEEE International Symposium on Electromagnetic Compatibility. Symposium Record, Washington, DC, 2000, pp. 883-886 vol.2.

Buildings Control for Energy Saving

Kateřina Gecová, David Vala
Dep. Cybernetics and BMI
VSB - Technical University of Ostrava
Ostrava, Czech Republic
Katerina.Gecova.st@vsb.cz, David.Vala@vsb.cz

Abstract*—*The article deals with measurement and control of indoor building parameters. In the field of modern buildings, attention is now increasingly focused on the issue of low-energy buildings, which is linked to efforts to eliminate energy losses. One of the solutions to this problem is to improve the quality of door and window seals, which prevents heat leakage. However, the negative effect of this measure lies in the limited ventilation of the interior, which reduces the air quality and the increased humidity in the building can cause mold. A more significant risk, however, is the impact on human health. Limited ventilation causes a gradual increase in temperature, humidity and, above all, carbon dioxide produced by breathing. The result may be an excessive load on the thermoregulatory system, fatigue and decreased concentration ability. For these reasons, it is important to monitor the relevant environmental parameters and then regulate them to meet the required limits

Keywords—building control; measurement; energy saving; internal environment; carbon dioxide

I. INTRODUCTION

This paper is based on the thesis of the author [1] and focuses on the issue of monitoring the state of the environment in the building, which provides the necessary information for the effective management of the relevant parameters. The reason is both to minimize energy losses and to create a health-conscious environment.

For air monitoring, it is possible to use a range of commercially available products that can be used to obtain air quality data with individual air control systems in the building. However, it is always necessary to select such sensors that will suit the control unit communication interface used for controlling the building. One of the most widespread bus systems for building control is the KNX bus. The main objective of this work is to design and test prototype devices that enable the KNX to provide the control system with information about the air condition of individual building components. For example, information of the concentration of carbon dioxide, carbon monoxide, dust particles, temperature, humidity or pressure.

There are commercially available devices that measure and evaluate some of these parameters, but in some cases may not be entirely appropriate, for example because of a cost. The aim of the thesis is to design a device that would allow measuring with sufficient accuracy all the available qualitative parameters that affect the comfort of a person while staying inside the buildings. The device is designed modularly so that it can only be installed by those sensors that are needed for a particular use.

II. ENVIRONMENTAL IMPACT ON HUMANS

The microclimate, which is the sum of the climatic conditions of the indoor environment, can have a significant effect on the function and health of the human organism. Main parameters considered include temperature and relative humidity or air velocity. In addition, various chemicals and dusts can be present in the air.

The assessment of microclimatic conditions can be conceived in two ways, namely objectively and subjectively. An objective assessment is based on measured environmental parameter values. Subjective assessment then corresponds to the degree of well-being of individuals in the given environment. From this perspective, these subjective views are categorized into a scale. In this review, all other factors, such as gender, age, health and mental status, the degree of acclimatization to the environment, the nature of the work, the clothing, and so on, are always included in the result.

For this work, basic parameters were selected, which are mostly measured and evaluated in the interior. These are temperature, relative humidity, air pressure, carbon dioxide and carbon monoxide concentration, as well as dustiness. [2], [3]

A. Temperature

The thermal comfort is important when assessing the heat load. This term can be described as the subjective feeling of a person staying in certain environment. In a state of thermal comfort, the temperature of the body with ambient temperature is so balanced that the human thermoregulation system is not loaded. The total heat production by humans is therefore offset by the heat flow from the body.

Since the feeling of cold or excessive heat is not desirable for the ideal degree of thermal comfort, the temperature range of the environment needs to be adapted. The production of heat from the human body depends, among other things, on whether a person is at rest or doing some physical activity at a given moment. To avoid overheating the body, it is necessary to remove heat from the body by cooling the environment. Although there are several factors involved in the thermal

comfort of a person, there are studies that have found that a person doing light work performs 100% at 22 °C.

High temperatures are manifested by vasodilatation, as well as increased blood flow, sweating, increased fatigue, and decreased ability to concentrate, which can lead to various injuries. If heat shock occurs, body temperature increases to over 41 °C, perspiration stops and coma becomes. At this stage there is a high probability of irreversible brain damage or death.

When exposed to a cold environment, blood flow is reduced (vasoconstriction), blood pressure and heart rate increase, and oxygen consumption is also increased. The defensive function of the organism in this case is thermogenesis, a muscle tension (tremor) that can produce up to a tenfold increase in body heat production. The most important thing is to keep the body temperature constant at 37 °C. Extremely low temperatures can cause frostbite or even death, due to a blood circulation failure. The low temperature combined with low humidity leads to paralysis of the blood capillaries, resulting in blood congestion and swelling of the limbs.

B. Relative air humidity

The indoor air humidity value is manifested by outdoor humidity, the number of people and also various technological or other sources. The relative humidity in the room should be around 30-65%, since the ideal value for the human body is between 40% and 45%. By heating the interior, however, the humidity decreases considerably, down to around 20%. This condition may cause more intensive drying of upper respiratory tract mucous membranes, resulting in greater susceptibility to disease, particularly due to decreased protective function and easier penetration of harmful substances into the lower respiratory tract. It is therefore advisable to increase the humidity in the room in these cases, but not more than about 40%. If, on the other hand, there is a high temperature and humidity, the natural reaction of the body in the form of sweating is not sufficient as the sweat from the body evaporates worse. [4]

C. Air pressure

Changes in weather are also of considerable significance for the human organism. These are related to bioclimatology. Although to some extent a person is adapted to weather changes, more severe fluctuations may lead to headaches, joints, or muscles in more sensitive people. Other symptoms of increased sensitivity to weather changes (meteorosensitivity) may also include difficulty breathing, drowsiness and malaise.

At low air pressure the blood pressure rises, which can be dangerous for people with circulatory disorders. A higher degree of hypersensitivity to these changes (meteorotropism) may increase heart rate and even heart attack.

D. Carbon dioxide

Carbon dioxide (CO_2) is a common component of air, and fresh outdoor air has about 400 parts per million (ppm) of this gas. CO_2 is formed by the oxidation of organic matter and it comes into the interior of buildings as a product of breathing of people, animals and plants. The effects of individual concentrations of carbon dioxide on humans are shown in the table below.

TABLE I.

Concentration of CO_2	**Effect on the human organism**
350–400 ppm	outdoor environment
< 1000 ppm	recommended indoor CO_2 level
1200–1500 ppm	maximum recommended indoor CO_2 level
1000–2000 ppm	symptoms of fatigue and decreased concentration
2000–5000 ppm	headaches
5000 ppm	maximum safe concentration without health risks
> 5000 ppm	nausea and increased heartbeat
> 15000 ppm	breathing difficulties
> 40000 ppm	possible loss of consciousness

With limited ventilation, it increases its concentration. In an hour, one is able to produce between 13 and 77 liters of CO_2, depending on the activity being performed.

Carbon dioxide makes it possible to assess the level of ventilation as its concentration can be easily measured. The human organism ceases to perceive the concentration of odors at some point and the olfactory organs adapt to the environment in which it occurs. However, at a certain concentration, the effect of carbon dioxide will be reflected by decreased ability to concentrate, fatigue and so on. [5]

E. Carbon monoxide

Carbon monoxide (CO) is a colorless gas without taste and odor, irritating and lighter than air. Natural concentrations in the air range from 50-230 $\mu g/m^3$, in short-haul transport areas or for example in road tunnels, short-term concentrations may exceed tens of mg/m^3.

Carbon monoxide practically does not penetrate the skin, so the predominant exposure route is inhalation. CO toxicity is due to the fact that it prevents/restricts the transfer of oxygen from the lung to the tissues. By diffusion, carbon monoxide quickly gets from the lungs to the blood where it binds to iron in the hemoglobin blood stain to form carboxyhemoglobin (COHb), thus limiting blood oxygen transfer capacity (the affinity of hemoglobin to carbon monoxide is 200-250 times higher than that of oxygen).

During the exposure to a stable CO concentration in the air, the COHb percentage first increases rapidly, after 3 hours it begins to equalize and at equilibrium after 6-8 hours of exposure. Excretion of CO from the organism takes place according to the same rules as the intake, the half-life is in the range 2-8 hours.

From a health point of view, it is recommended that the blood COHb level does not exceed 2,5%, a value that does not

have a negative effect on the sensitive population (e.g. people with heart disease, developing fetus). [6]

F. Dust

Dust particles are referred to as PM (particulate matter) and are divided into two fractions, PM2,5 and PM10. The names of these fractions result from the particle size, so as PM2,5 they are referred to as those that reach a maximum of 2,5 μm. Because they are very small particles, they are dangerous to humans by allowing them to get deep into the lungs by inhalation. Often they contain heavy metals and toxic organic compounds. In terms of both composition and particle size, this fraction threatens human health more than PM10, the particle size of which ranges from 2,5 μm to 10 μm. The probability of penetration into deeper lung structures is lower than the previous one.

Among the most common health problems and diseases caused by dust particles are respiratory problems (cough, difficulty in breathing), asthma impairment, impaired lung function, lung cancer, cardiovascular disease, cardiac arrhythmia, a mild heart attack, or death.

III. AUTOMATION OF BUILDINGS

Building management functions can be divided into several sub-groups. The first one is heating, ventilation and air conditioning (HVAC). Room temperature can be set, for example, by number of people or depending on outside temperature. Other environmental control options may also include air quality monitoring for optimal ventilation settings or automatic cooling or heating shut down when windows are open. Lighting can then be controlled depending on the presence of people, the brightness of the room, or the brightness of the ambient light. Optimum room environments also use blinds. In the summer, they can prevent room overheating, and in the winter, on the other hand, will allow maximum use of sunlight. Automated buildings can also have security features, such as opening electrically adjustable windows for smoke detection or personal data entry checks according to biometric data. These features are just examples of how building automation can be used, but there are significantly more end-applications. [7]

KNX/EIB bus is the worldwide standardized communication system, which serves in the buildings as a network information connection of devices, such as sensors, actuators, metering devices, etc. Buses, which ensure a mutual interconnection of particular elements, as well as a dig-ital transition of data, can be realized in a form of a Twisted Pair (TP), or with help of a power cable, an optic cable or via the remote radio connection.

Manufacturers focusing on KNX/EIB products offer a series of productions from sensors, through actuators up to the sources. The total quantity of KNX compatible devices of different brands reaches the number of several thousands. These are mainly power supplies, amplifiers, sensors, actuators and so on.

IV. DEVELOPMENT/DESIGN OF SENSORS UNIT

To meter the required parameters of the inner environment, the following sensors were selected.

- For temperature and relative humidity metering, there was selected a sensor SHT21 produced by Sensirion Company based on the capacitive principle of humidity measurement. In this case, I^2C interface is the interface. The metering range for temperature is -40 to 125 °C, with the accuracy of measurement of ±0.3 °C and the resolution of 0,01 °C. The metering range of the sensors for humidity corresponds to the values from 0 to 100 %. The accuracy of measurement is ±2 % RH, and the resolution is 0.04 % RH.

- BMP280 produced by Bosch Company is a sensor for measurement of absolute barometric pressure, which disposes by a digital output in a form of the I^2C interface, and its function is based on the piezo-resistive effect. The metering range of the sensor is 300 - 1100 hPa, and the absolute accuracy corresponds to ±1 hPa at the pressure values from 950 to 1050 hPa and the temperature from 0 to 40 °C. Temperature can be also measured by means of BMP280 with the range from -40 to 85 °C.

- A sensor Telaire T6713 was selected for measurement of CO_2 concentration in the atmosphere; its metering range is from 0 to 5000 ppm. For the range of concentration from 400 to 5000 ppm, the accuracy corresponds to the value of ±30 ppm from the measured value. Accordingly, the accuracy at the concentrations from 400 to 2 000 ppm is ±25 ppm. The sensor works on the principle of the non-dispersion infrared spectrometry (NDIR).

- The sensor for carbon monoxide measurement TGS5042 has the metering range from 0 to 10000 ppm, and the accuracy less than ±5 % (for the range from 0 to 500 ppm). The principle of its function is an electrochemical element that works as a battery, and the resulting value is the current between the working electrode and the counter-electrode.

- The optical dust sensor GP2Y1010AU0F produced by Sharp Company works on the basis of an infrared LED diode and a photo-transistor. This sensor is suitable essentially for dusts with very fine particles, such as cigarette smoke. The accuracy of measurement is from 0.35 to 0.65 V/(0.1 mg/m^3). The density of dust particles depends on the measured voltage value.

For the realization of all necessary settings and possibility of reading of the measured values, it was be necessary to choose a certain communication medium. A microcontroller serves for these purposes, by means of which it is possible to set parameters of separate sensors according to the particular requirements of the module, as well as performance of collection of the measured data in the required format. Though, for now, in this state, there is no possibility to send data between sensors and the KNX bus itself. The mediator of this communication is the KNX chip NCN5120.

V. EXPERIMENTAL METERING

For demonstration of the changes of separate parameters of the environment depending on the conditions of the given

room, several short-term and long-term measurements were performed. The mentioned activities were carried-out in one of the classrooms of VŠB-TU during learning. Figure with graphs from this metering is shown in the Fig. 1 below.

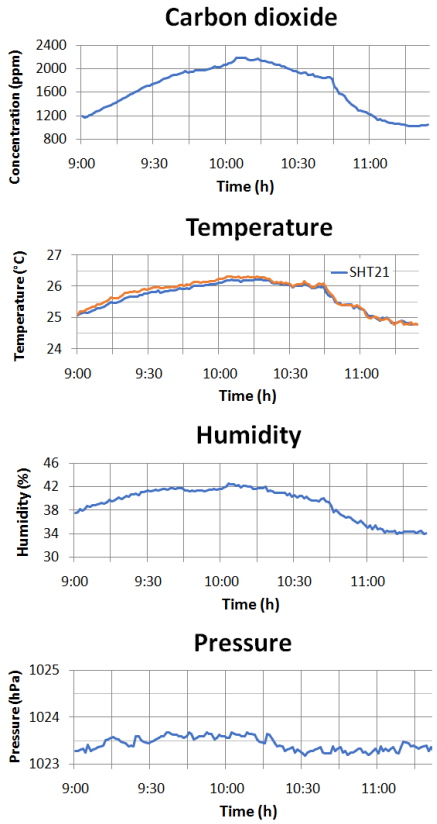


Fig. 1: Curves of individual parameters from classroom measurements

During staying of 18 persons in a closed, non-vented room, CO_2 concentration during the first hour of measurement reaches the values up to almost 2200 ppm, which very significantly exceeds the recommended hygienic limits. Concentration of carbon dioxide should get around at most up to 1500 ppm. In case of values above this boundary, signs of tiredness or non-concentration can occur, and above 2000 ppm also headaches.

To verify correctness of the solution, the resulting module was connected to the KNX bus. The measurement was realized in the KNX EB312 laboratory on the Electrotechnics and Informatics Faculty. The ranges of physical and group addresses within possibilities of the laboratory were used for the experiment.

For performance of connection of all parts of the module, the panel of KNX elements was switched-on, and afterwards the ETS5 program was initiated for the work with the KNX devices. Consequently, a hierarchy of the individual addresses for particular sensors was created, and a corresponding type is assigned to each of them according to the unit. Each data type has its specific code in ETS5. Consequently, gradual measurement of all parameters was initiated, which were written into the table in ETS5 in corresponding sequence.

VI. Conclusion

The realized module of the environment sensors was proposed as a universal unit for measurement of the monitored parameters in the inner building premises. Beside of that, the device also offers an additional connection of other sensors by means of both via remaining analogue inputs or via the I^2C microcontroller interface. However, the scalability of the device can be also used in the opposite direction. These are the cases where it is not necessary to monitor the state of the environment here in detail. So, several from the already connected sensors were not be used. By this reason, the module is configured so that any from the sensors can be omitted and the device can be fitted for any individual requirements of the user.

Other differences from the commercial device are undoubtedly variability and setting of the limit values of the sensors. The sensors of the usual devices are usually set to certain boundary values of the given measured parameter, which remain practically constant. It can be, for example, a value of 1500 ppm for the carbon dioxide concentration, whereas an alarm is sent in case of this limit exceeding. The realized device is however able to change these values dynamically, evaluate data even during the course of the measurement and fit itself on the basis of these statistics to the given environment.

Acknowledgment

This paper has been elaborated in the framework of the project „Control of technological systems with renewable and alternate energy sources 2017“ SP2017/150 and „A practical training program for staff developing a low-carbon economy in the border region“ reg. no. CZ.11.3.1190.0/0.0/15_005/0000048, supported by European Union in framework program INTERREG V-A Czech republic-Poland.

References

[1] GECOVÁ, Kateřina. Návrh jednotky senzorů prostředí kompatibilní s KNX. Ostrava, 2017. Diplomová práce. VŠB – TUO.

[2] Vliv změny klimatu na zdraví lidí, zvířat a rostlin. In: PUBLIC HEALTH [online]. Brusel: KOMISE EVROPSKÝCH SPOLEČENSTVÍ, 2009 [cit. 2017-01-18]. Dostupné z: http://ec.europa.eu/health/ph_threats/climate/docs/com_2009-147_cs.pdf

[3] Mikroklimatické podmínky vnitřního prostředí pracovišť. In: STÁTNÍ ZDRAVOTNÍ ÚSTAV [online]. Praha: Kolektiv pracovníků SZÚ, 2007 [cit. 2017-01-18]. Dostupné z: http://www.szu.cz/tema/pracovni-prostredi/mikroklimaticke-podminky-vnitrniho-prostredi-pracovist

[4] Technika a způsoby měření parametrů vnitřního prostředí. In: TZB-info [online]. Praha: Topinfo, 2012 [cit. 2017-01-18]. Dostupné z: http://vetrani.tzb-info.cz/vnitrni-prostredi/9217-technika-a-zpusoby-mereni-parametru-vnitrniho-prostredi

[5] Měření koncentrací škodlivých plynů uvnitř staveb. AZ-Radon s.r.o. [online]. České Budějovice: AZ - Radon, 2016 [cit. 2017-01-18]. Dostupné z: http://www.az-radon.cz/mereni-koncentraci-skodlivych-plynu-uvnitr-staveb/

[6] Oxid uhelnatý » Celý článek. STÁTNÍ ZDRAVOTNÍ ÚSTAV [online]. Praha: Kolektiv pracovníků SZÚ, 2016 [cit. 2017-01-18]. Dostupné z: http://www.szu.cz/tema/zivotni-prostredi/cely-clanek-1

[7] MERZ, Hermann, Thomas HANSEMANN a Christof HÜBNER. Automatizované systémy budov: sdělovací systémy KNX/EIB, LON a BACnet. Praha: Grada, 2008. Stavitel. ISBN 978-80-247-2367-9

Chipless RFID Tag Using U-Folded Dipoles in a Metal Frame

Jaroslav Havlíček
Department of Electromagnetic field
Czech Technical University in Prague
Prague, Czech Republic
havlij18@fel.cvut.cz

***Abstract*—This paper presents experimental 5-bit chipless RFID tag based on frequency domain information coding. The tag is composed of U-folded dipoles and a metal frame. Its main feature is frequency notched RCS response which provides high read range of the tag. Proposed simulation results of the tag explain a usage of an enhanced arrangement of the U-folded dipoles in the frame and demonstrate performance of the tag. Information encoding ability of the tag was investigated. Effects recognized during investigation of the tag will be employed for design of a tag with higher bit capacity.**

Keywords—chipless RFID, tag, scatterer, radar cross-section, spectral response coding

I. INTRODUCTION

The radiofrequency identification (RFID) is a widespread technology, which is applied mainly in industry (production process monitoring), logistics (acceleration of storing procedures), commerce (antitheft feature), health care (elimination of patient confusions) and inventorying (e.g. books in libraries). RFID is expected to have an important role in the internet of things concept [1].

There is a perfect opportunity for further spreading in commerce, provided that the RFID transponders (tags) are suitable for optical barcode replacement [2]. The advantage of the tag consists in the possibility to be read without necessity of clear line of sight. Low price of the tag is essential for reaching this goal. Unfortunately conventional tags which consists a chip are about hundred times more expensive than barcode [3]. Perspective approach to reduce the tag cost is chipless RFID which provides different methods of storing information without chip.

Promising examples of the chipless RFID technology are tags based on spectral signature [4], [5], which in most cases consist of resonating scatterers. Presence or non-presence of each scatterer's peak in tag's spectral signature represents logic one or zero, respectively. This method of the chipless tag design is employed in this paper. Second part discusses required properties of the scatterer and describes U-folded dipole. Third part presents a 5-bit chipless RFID tag which is using U-folded dipoles coupled with a metal frame.

II. U-FOLDED DIPOLE

There are many different types of planar resonators which can be used as a scatterer for the chipless RFID tag. [6] Each of these resonators can be described via three important quantities. First of them is radar cross-section (RCS) which determines the reading range of the tag. The higher RCS results in the longer reading range.

Second quantity is bandwidth (BW) of resonance peak which depends on quality factor of resonance. It can be evaluated from frequency characteristic of the resonator as frequency range of three decibel resonance peak drop. This value is important from spectral bit capacity point of view. For example when bandwidth is 200 MHz there cannot be stored more than only five bits for each GHz in the spectrum.

Last quantity is product ka where k is wave number and a is radius of circle circumscribed the resonator. This quantity represents rate of electric reduction of the resonator. Resonators with ka below one are suitable candidates because small scale is essential for purposes of chipless RFID.

Theory of electrically small antennas says that resonance structures with small ka have high quality factor Q which implicates low bandwidth which is beneficial for high spectral bit capacity and low RCS which is inappropriate for reading range. Therefore there is always a trade-off between spectral bit capacity and reading range of the tag.

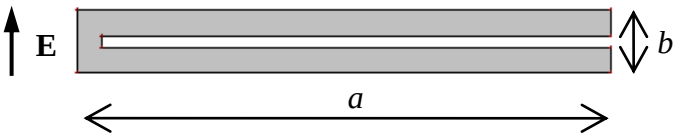


Fig. 1. The layout of the U-folded dipole

U-folded dipole represents quite advantageous choice from the point of view of this trade-off [6] but its biggest advantage is simple geometry. This scatterer consists of two parallel arms, which are connected to each other at one end; see Fig. 1. The width of all lines equals 1 mm and the distance between arms accounts for 0.5 mm, which gives the scatterer a width of b = 2.5 mm. The length of arms, which determines the required resonance frequency (the arm length is similar to one quarter of resonant wavelength) vary from a = 21.5 mm to

a = 25.5 mm (with 1 mm step). The scatterer shows a sufficient RCS equal to -34.3 dBsm and a narrow bandwidth BW_{3dB} = 18.1 MHz. However, its disadvantage is embodied in a mediocre factor ka = 0.791 [6]. The performance of the single scatterer as well as all variants of the tag was simulated on the low-loss dielectric substrate (ε_r = 3.38, tan δ = 0.002) with the thickness of 0.2 mm using MoM IE3D software.
U–folded dipole design is not a new contribution of the author. It has been already published [6, 7].

III. Design of 5-bit Chipless Tag

Five U-folded dipoles were coupled with a metal frame; see Fig. 2. Size of the frame is m = 29.5 mm by n = 28.5 mm. The width of all its lines equals 1 mm and the distance between U-folded dipoles and inner frame edge is also 1 mm (left side of the frame is an exception, this distance depends on a scatterer length). Order of scatterers in the frame is '4 3 1 2 5'. Numbers represents order of scatterer's resonant frequencies, for example number 5 represents the shortest scatterer with highest resonant frequency. Advantage of this enhanced arrangement in comparison with basic length ascending arrangement (see Fig. 5) consists in more uniform RCS response.

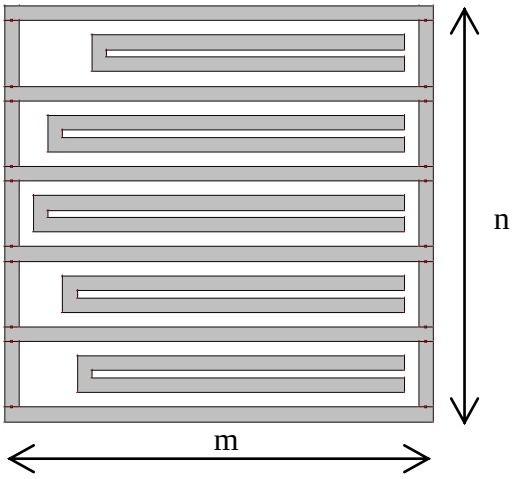


Fig. 2. The layout of the tag with enhanced arrangement

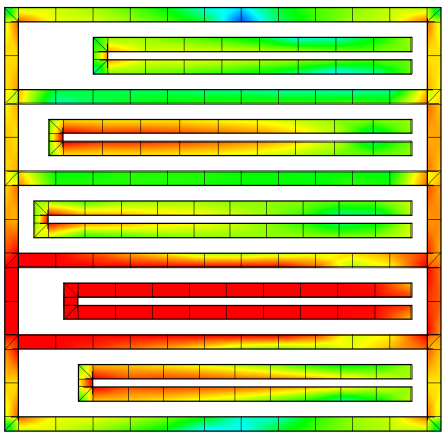


Fig. 3. The current distribution of the tag with enhanced arrangement

RCS response of the tag is notched, which means that each scatterer causes a narrow dip at its resonant frequency in wide resonant peak of the frame. The dip is a consequence of coupling between a U-folded dipole and the frame, which can be observed from current distribution at frequency point f = 2.56 GHz in Fig. 3.

RCS response in Fig. 4 illustrates performance of the tag with enhanced arrangement. Maximum of the frame's resonant peak is approximately 18 dBsm. All information coding dips are clearly recognizable. Depth of the dips is between 4 and 9 dB. Moreover coding a zero bit by removing particular scatterer doesn't affect dips of other U-folded dipoles which is essential for readability of the tag.

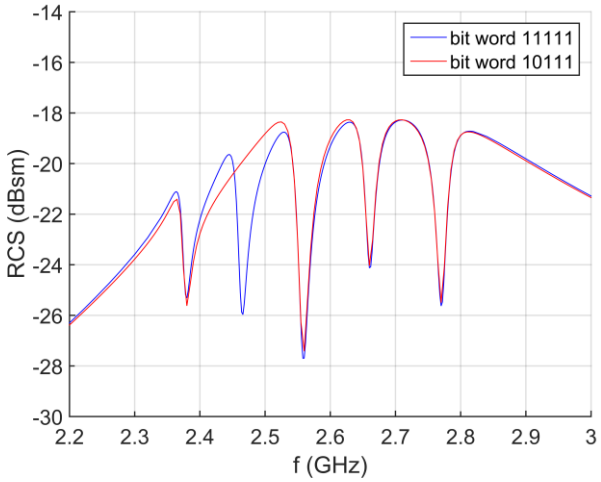
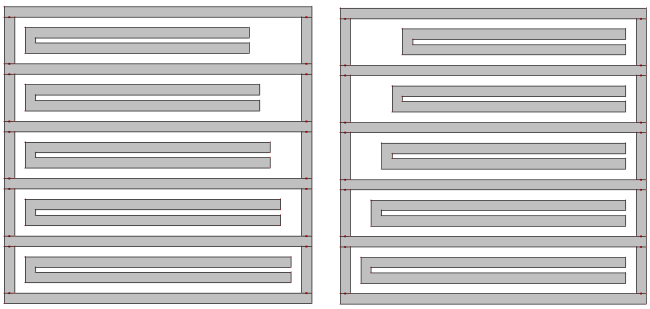


Fig. 4. Simulated performance of the tag with enhanced arrang. while coding

Two variants of basic arrangement using an ascending order of the scatterers according to their length were designed and investigated before the enhanced arrangement above; see Fig. 5. It is obvious that difference between variants consists only in alignment of the U-folded dipoles in the frame.



(a) (b)

Fig. 5. The tags with basic arrangement: aligned left (a), aligned right (b)

RCS responses of both tags are in the Fig. 6. Especially from the left aligned variant it can be observed that dips belonging to the side scatterers, which have only one neighbor are significantly weaker than rest of the dips. The first resonant dip is even beyond the limit of reliable reading. Comparison of the both responses demonstrates that greater distance between

inner frame edge and the shorted end of the scatterer's arms are beneficial for depth of the dips.

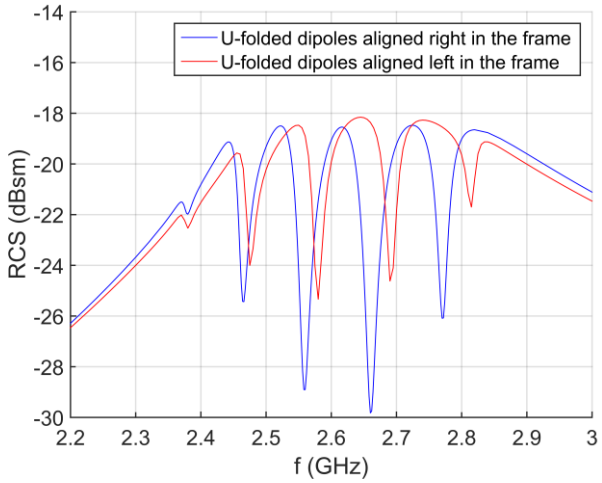


Fig. 6. Simulated comparison between variants of U-folded dipoles alignments in the frame

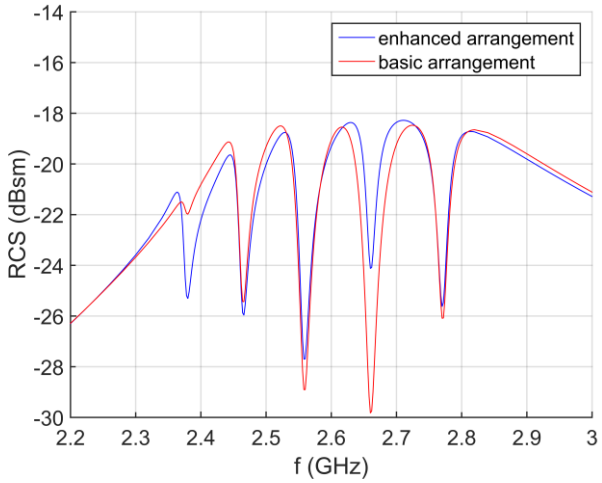


Fig. 7. Simulated comparison between basic and ehanced arrangement

The idea of enhanced arrangement is based on these effects. The longest scatterer was moved to the middle of the tag in order to increase its dip depth. The two shortest scatterers were chosen for the side positions because low dip depth which is essential for this position can be compensated by high distance between the shorted ends of the scatterers and the left side of the frame. The tag using enhanced arrangement has thanks to these changes more uniform RCS response than the tag with basic arrangement; see Fig. 7.

IV. Conclusion

The paper discusses design and performance of 5-bit chipless RFID tag, which is composed of the U-folded dipoles and the metal frame and it is working in the frequency domain mode. Employing a metal frame in the tag design allows achieving higher level of RCS response. It is shown that basic arrangement of the scatterers in the frame is inappropriate due to low RCS response uniformity. This problem was solved by proposed enhanced arrangement. Described tag has suitable properties for chipless RFID, but in current stage of development can be used only for experimental purposes. Practical usage is limited by its insufficient bit capacity. Further work on this project will be focused on a design of chipless RFID tag with significantly higher bit capacity based on the same principle of coupling a scatterer with a metal frame.

Acknowledgment

This work was supported by the Czech Technical University in Prague under the project SGS16/226/OHK3/3T/13.

References

[1] S. Dey, J. K. Saha and N. C. Karmakar, N. Smart Sensing: Chipless RFID Solutions for the Internet of Everything. IEEE Microwave Magazine, 2015, vol.16, no.10, pp.26-39. DOI: 10.1109/MMM.2015.2465711

[2] P. Harrop, R.Das, *Printed and Chipless RFID Forecasts, Technologies & Players 2011-2021*. [Online] Cited 2016-04-25. Available at: http://www.idtechex.com/research/reports/printed-and-chipless-rfid-forecasts-technologies-and-players-2011-2021-000254.asp.

[3] S. Preradovic, and N.C. Karmakar, "Chipless RFID: Bar Code of the Future," *IEEE Microw. Mag.*, vol. 11, no. 7, pp. 87-97, Dec. 2010.

[4] R. Rezaiesarlak, and M. Manteghi, "Complex-Natural-Resonance-Based Design of Chipless RFID Tag for High-Density Data", *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 2, pp. 898-904, 2014.

[5] A. Vena, E. Perret, and S. Tedjini, "A Fully Printable Chipless RFID Tag With Detuning Correction Technique", *IEEE Microwave and Wireless Component Letters.*, vol. 22, no. 4, pp. 219-229, 2016.

[6] J. Havlicek, M. Svanda, J. Machac, M. Polivka, "Improvement of Reading Performance of Frequency-Domain Chipless RFID Transponders", Radioengineering, vol. 25, no. 2, pp. 209-211, 2012.

[7] M. Polivka, M. Svanda, J. Machac, Chipless RFID Tag with an Improved RCS Response, *Proc. of 44th European Microw. Conf.* (EuMC), 2014, pp.770-773, ISBN: 9781479955985, DOI: 10.1109/EuMC.2014.6986548

Vessels Model for Laser Speckle Contrast Imaging Method

Branislav Hesko
Faculty of electronics and communication
Brno University of Technology
Brno, Czech republic
Email: xhesko00@stud.feec.vutbr.cz

Vratislav Harabiš
Faculty of electronics and communication
Brno University of Technology
Brno, Czech republic
Email: harabis@feec.vutbr.cz

Radim Kolář
Faculty of electronics and communication
Brno University of Technology
Brno, Czech republic
Email: kolarr@feec.vutbr.cz

***Abstract*—This conference paper deals with LSCI (laser speckle contrast imaging) method, which is used to visualise blood flow rate. A simple experimental setup has been designed for studying method properties. This setup uses microparticles as a model of red blood cells and glass capillary as vessels model. After processing sequences, it is evident, that final image intensity is a function of fluid flow. Moreover, we have shown the imaging process being stable, dependent on exposition time.**

I. Introduction

Blood flow rate is an important property of the human physiology. Usually, methods based on ultrasound Doppler are widely used. With the invention of laser in the 60s, laser based methods could be invented, for example, laser Doppler imaging, which examines frequency shifts or laser contrast speckle imaging, which uses speckle noise to estimate relative blood flow. [1], [2]

Laser speckle contrast imaging (LSCI) is a set of methods, that are using speckle pattern to estimate blood flow. The speckle pattern is typical for systems with coherent and monochromatic waves, such as ultrasound or laser systems. They are noninvasive and unharmful, the only risk remains in laser intensity, as for retina application, the laser power has to be under hygienic limits [2], [3].

Physically, speckle noise is arising based on interference. The incident wave is scattered on objects with dimension comparable to the wavelength. These objects are behaving like individual wave sources, due to the Huygens principle. Scattered waves are interfering and strike detector, where it is possible to observe local speckle pattern [4]. Mathematically we model speckle noise as multiplicative noise with Rayleigh probability distribution defined as: [5], [3]

$$p(A) = \frac{A}{\Psi} \cdot e^{\frac{A^2}{2\Psi}} \tag{1}$$

where A is amplitude and Ψ is distribution parameter. The principle of LSCI method is simple. When scatterers responsible for interference, like erythrocytes, are moving, the speckle pattern becomes blurred. This is caused by finite camera exposition time. It is possible to describe scatterers (erythrocytes) speed distribution using contrast values. Contrast value is a function of velocity and expressing it using image properties, we obtain equation: [6], [7]

$$K(x,y) = \frac{\sigma_{\mathrm{N}}}{\mu_{\mathrm{N}}}. \tag{2}$$

$K(x,y)$ is contrast value at position (x,y), σ_{N} is standard deviation and μ_{N} is mean in a pixel neighborhood of size N. Generally speaking, a neighborhood could be defined in various shapes, but the isotropic square neighborhood is used. In the case of an ideal light source with respect to coherence and monochromaticity, the value of standard deviation is equal to mean value in a neighborhood. This introduces condition for contrast values [8], [3]:

$$K(x,y) \leq 1. \tag{3}$$

This condition is not usually fulfilled. Dependency of contrast values on red blood cells velocity is not linear, therefore usually more respecting relation is used. It assumes Lorentz velocity distribution of moving scatterers, which is connected with Brownian motion assumption, thus independent movement of red blood cells. Then, we define decorrelation time τ_{c} as function of contrast value as: [8], [3]

$$K^2 = \beta \left(\frac{\tau_{\mathrm{c}}}{T} + \frac{\tau_{\mathrm{c}}^2}{2T^2} \cdot \left(e^{\frac{-2T}{\tau_{\mathrm{c}}}} \right) \right) \tag{4}$$

In this equation, β is system dependent constant and T is exposition time. Then, $\frac{1}{\tau_{\mathrm{c}}} \sim v_{\mathrm{s}}$, so the inverse of the decorrelation time is a function of scatterers velocity v_{s}. The decorrelation time from the equation (4) better respects nonlinearities in velocities then contrast values from the equation (2). Finally, by substituing $x = \frac{T}{\tau_{\mathrm{c}}}$ and by modifying equation (4) one obtains [6], [9]:

$$K(T, \tau_{\mathrm{c}}) = \sqrt{\beta \cdot \frac{e^{-2x} - 1 + 2x}{2x^2}} \tag{5}$$

This equation is easier to implement, but the exponential function under root square hardens analytical solution, therefore we have to solve it numerically. Then, we can convert contrast images, that are easily computable, into inverse decorrelation time images [8].

II. Experimental setup

Comparing to other blood flow measurement methods, LSCI requires simpler and cheaper setup. A camera and laser system are only needed. Block scheme of our measurement setup is shown on figure 1.

We are using laser with maximum power $100\,\mathrm{mW}$ and wavelength $\lambda = 658\,\mathrm{nm}$. Laser diode 658P040 is controlled by laser controller LDC 205C. As vessels model, we have chosen

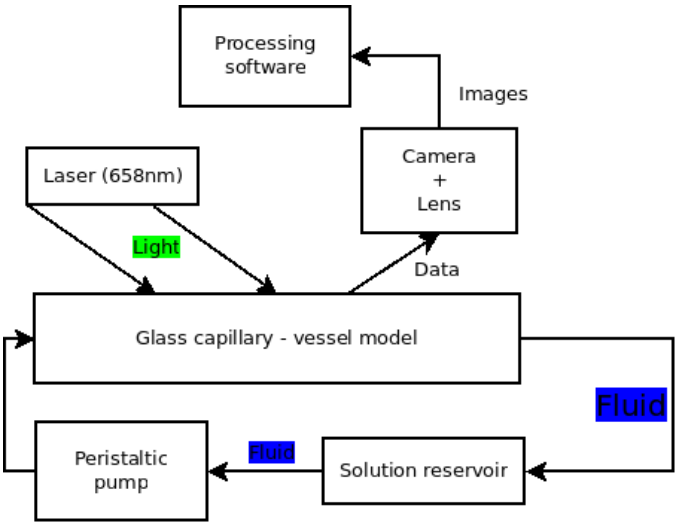


Fig. 1: Main components of designed experimantal setup. Peristaltic pump secures fluid flow, laser and camera are main parts of LSCI system.

glass capillary 5 mm wide and 0.5 mm deep. Blood modeling fluid is water with latex microparticles. Two different diameters of microparticles have been used, 0.1 μm and 1.1 μm made by SIGMA, but no significant difference between them have not been observed. Due to the high microparticles clustering, the glass capillary has to be often changed.

III. Experiment and Methods

Data acquired using our experimental setup have been processed offline. We are using programming language Python with scientific libraries NumPy and OpenCV. Each acquired video has approximately 200 images, with frame rate 30 fps and size 1280×1024 pixels. Vessel model is situated approximately in the middle and positioned vertically. Images do not need preprocessing, only resorting in the case of multi-exposition measurements. Firstly, images are transformed into contrast images.

A. Transformation into contrast images

Scatterers, i.e. red blood cells or microparticles movement cause visible blurring in the capillary. Using equation (2), it is possible to visualize this blurring. Neighborhood size is the only parameter and its size is chosen as 7×7 or 9×9 [6]. To speed up the computation process, local mean values are computed by convolution with the averaging mask of neighborhood size. The standard deviation in each point is computed from a definition of variance as:

$$\sigma_{M,N}(x,y) = \sqrt{\langle X^2(x,y)\rangle_{M,N} - \langle X(x,y)\rangle^2_{M,N}} \qquad (6)$$

where the first term is computed using convolution with the averaging mask of neighborhood size and squared image. The second term is computed from local mean values by squaring them. Finally dividing both terms, one obtains contrast image. The effect of contrast transformation is shown on the image 3.

B. Image filtration

As blurring and thus contrast information is a random variable, the resulting image suffers from random contrast fluctuations in both spatial and time domain. As the spatial video resolution is sufficient, spatial averaging filtering with mask of size 31 was used. Similarly, with heart frequency at 1 Hz, temporal averaging filter with size 7 could have been used, without any impact on temporal resolution. Filter length is chosen based on experience, sufficiently small to not influence spatial and temporal resolution. Filtered images are then transformed into inverse decorrelation time, which enhances blurring caused by moving particles over a static background.

C. Decorrelation time transformation

Contrast images are transformed into inverse decorrelation time images using (5). This equation is solved numerically, with an assumption, that contrast lies in the interval $[0, 1]$. This interval is divided into 1000 values and for each value, equation (5) is solved using modified Powell method. The vector of calculated inverse decorrelation time values is used as a look-up table and each contrast value $K(x, y)$ is transformed. The transformation curve for different exposition times is shown in figure 2. Note, that interval is divided linearly, but the dependency of inverse decorrelation time is not linear.

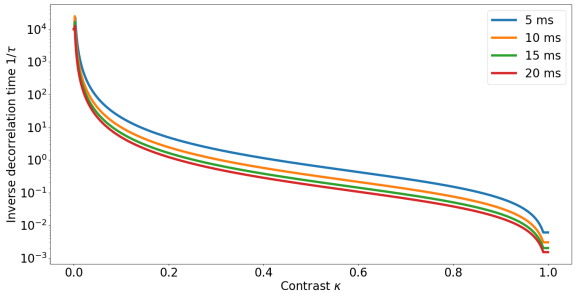
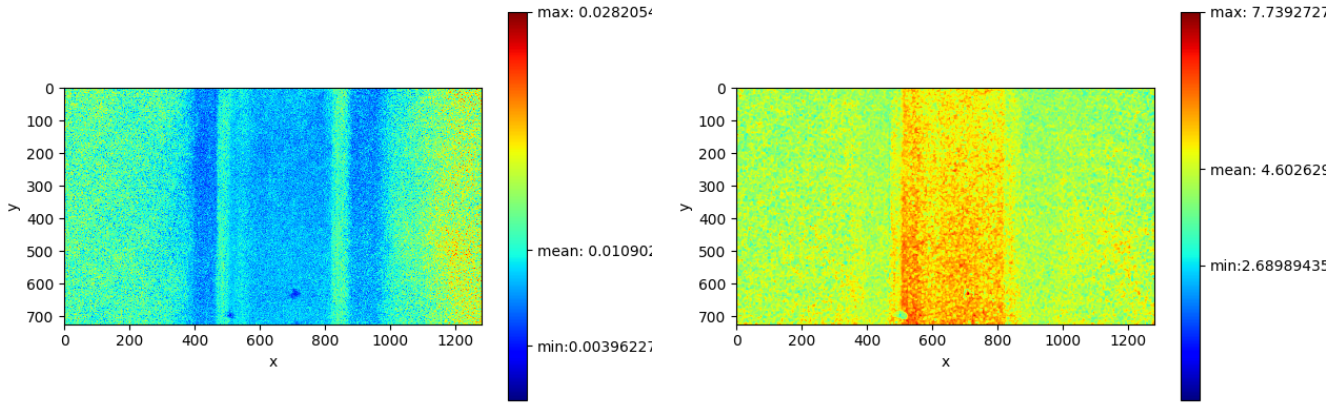


Fig. 2: Semilogaritmic plot of inverse decorrelation time as dependency on contrast value for different exposition times. Note that curves seems to be shifted only.

IV. Results

For different exposition times and velocities, we have acquired videos in single-exposition mode. In single exposition mode, all the images from a sequence are acquired with the same exposition time. To secure same conditions, sequences are recorded in a short time.

Each sequence is first transformed into contrast images, then filtered and finally transformed into decorrelation time images. Contrast values can be considered as a random variable and filtration prevents contrast to have extreme values. If extreme values are present, due to the nonlinearity of inverse decorrelation time values transform, extremity will be enhanced. Mean contrast image from a sequence and corresponding inverse decorrelation time image are shown in figure 6.



(a) Original image

(b) Inverse contrast image

Fig. 3: Comparison between original and inverse contrast image. Transformation into contrast values identifies fluid flow.

A. *Stability of the process*

First of all, one has to ensure, that the process will be stable in time. This condition means, that contrast values and derived inverse decorrelation time will be ergodic, ie the energy will not change in time. If this condition is not fulfilled, images cannot be cross compared, thus they have to be dealt independently. Ergodicity of the process is proved in figure 4, where the mean value from each frame and mean capillary value are calculated and plotted. Both mean values oscillate at the beginning of the experiment, which is probably a transitional state to equilibrium. After 50 frames, the system remains stable and only small oscillations are present. This is the side effect of randomness of the process. Moreover, mean values in the capillary are lower than entire image means. This is expected because velocity is inversely dependent on contrast: $v_c \sim 1/K$.

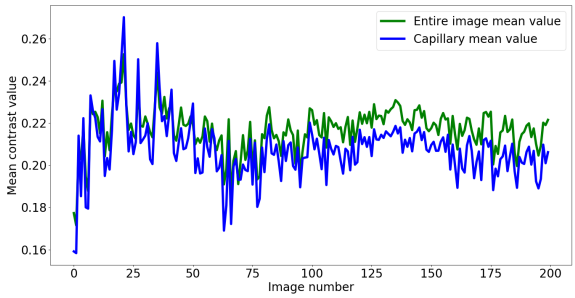


Fig. 4: Stability of measurement process. Mean contrast value from entire image and mean contrast value inside capillary are plotted.

Every single exposition sequence have been processed and within each sequence following values have been calculated:

- Mean value in the sequence.
- Mean capillary value in the sequence.

These values have been calculated for inverse contrast images, filtered inverse contrast images and inverse decorrelation time images. We have acquired sequences for exposition times $[5, 10, 15, 20]$ ms and fluid speed $[0.075, 0.75, 7.5]$ ml·s^{-1}.

B. *Experimental results evaluation*

First of all, dependency of inverse filtered contrast on fluid flow is shown in figure 5. As inverse contrast value should be a function of speed, we expect, that with increasing flow, the inverse contrast will increase too. This is the case for exposition time of 15 ms, where both mean value in capillary and mean value from the whole sequence is increasing with the fluid flow. For exposition time of 5 ms, mean value increase first, then decrease or remains approximately constant.

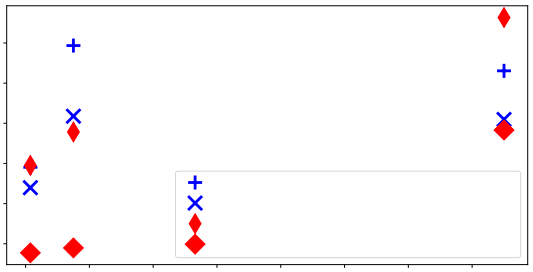
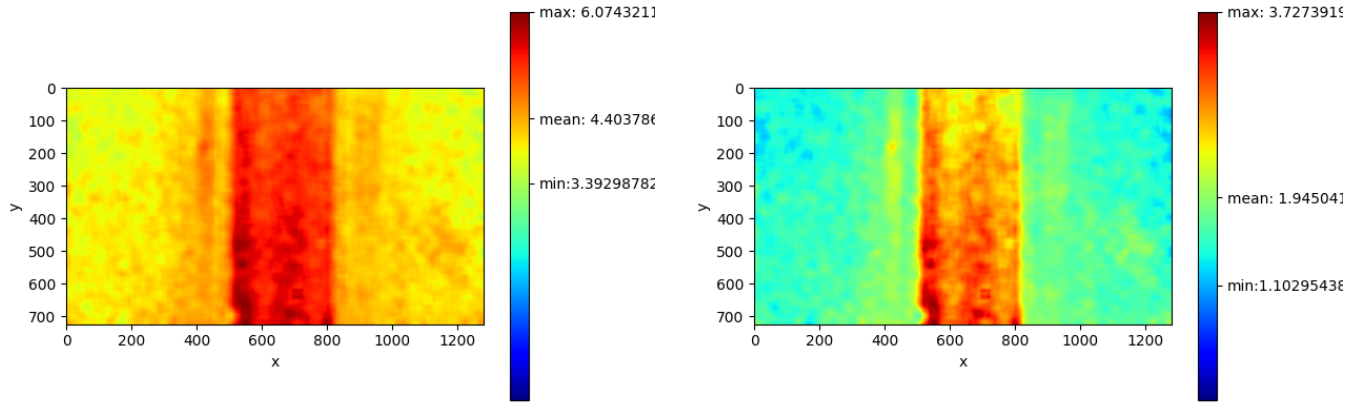


Fig. 5: Dependence of inverse filtered contrast on fluid flow for exposition times 5 ms and 15 ms.

Finally, we have plotted dependency of decorrelation time on exposition time. From the figure 7, we can deduce, that for flow $I = 0.75$ ml/min decorrelation time is decreasing with increasing exposition time and for flow $I = 7.5$ ml/min, we have found similar conclusion, with difference at 30 ms.



(a) Filtered inverse contrast image.

(b) Inverse decorrelation time image.

Fig. 6: Comparison between filtered inverse contrast image and inverse decorrelation image. Note, that transformation into decorrelation time enhances blood flow with respect to background

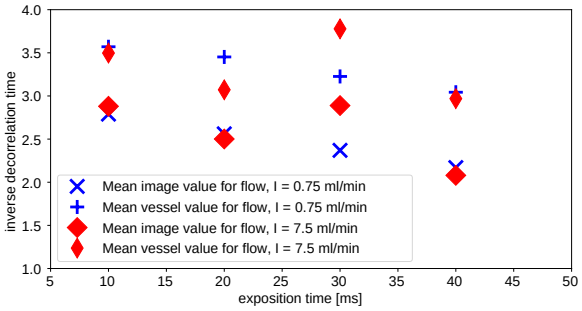


Fig. 7: Dependence of decorrelation time value on exposition times, for flow 0.75 ml/min and 7.5 ml/min.

V. Conclusion

LSCI method is an interesting, unharming and cheap way to determine blood flow. We have presented a new, experimental setup, which allows method properties examination.

Transforming sequences into contrast values allows blood flow visualisation and random noise suppression. By filtering contrast images, we have obtained smoother blood flow visualisation, because we suppose, that axial velocity is constant. Transformation into inverse decorrelation time results in enhancing contrast of brighter structures, while suppressing low contrast values, ie background, see figure 2.

The whole experiment is relatively stable, with an oscillation at the beginning. Stability was determined using mean value for each frame in the sequence.

Based on the figure 5 we can conclude, that the dependency of the inverse contrast value is not strictly linear and is also a function of exposition time. From the figure 7 we can conclude, that with increasing exposition time, inverse decorrelation time is decreasing. The explanation is that contrast is higher with increasing exposition time, as images are more blurred and more enlightened and according to the figure 2, contrast and inverse decorrelation time are inversely proportional.

We have shown some properties of our experimental LSCI setup, most of them are in accordance with theory expectations. For better evaluation and absolute flow measurement in the future, we have to measure more exposition times and more fluid velocities.

Acknowledgment

The authors would like to thank to Michaela Jahůdková for a valuable work on video acquisition and experimental setup.

References

[1] J. D. Briers, "Laser doppler, speckle and related techniques for blood perfusion mapping and imaging," *Physiological Measurement*, vol. 22, no. 4, p. R35, 2001.

[2] M. Draijer, E. Hondebrink, T. van Leeuwen, and W. Steenbergen, "Review of laser speckle contrast techniques for visualizing tissue perfusion," *Lasers in medical science*, vol. 24, no. 4, p. 639, 2009.

[3] J. Senarathna, A. Rege, N. Li, and N. V. Thakor, "Laser speckle contrast imaging: Theory, instrumentation and applications," *IEEE Reviews in Biomedical Engineering*, vol. 6, pp. 99–110, 2013.

[4] J. W. Goodman, *Statistical optics*. John Wiley & Sons, 1985.

[5] C. B. Burckhardt, "Speckle in ultrasound b-mode scans," *IEEE Transactions on Sonics and Ultrasonics*, vol. 25, no. 1, pp. 1–6, jan 1978.

[6] L. M. Richards, S. M. S. Kazmi, J. L. Davis, K. E. Olin, and A. K. Dunn, "Low-cost laser speckle contrast imaging of blood flow using a webcam," *Biomedical Optics Express*, vol. 4, no. 10, p. 2269, sep 2013.

[7] J. C. Dainty, *Laser speckle and related phenomena*. Springer Science & Business Media, 1975, vol. 9.

[8] D. Briers, D. D. Duncan, E. Hirst, S. J. Kirkpatrick, M. Larsson, W. Steenbergen, T. Stromberg, and O. B. Thompson, "Laser speckle contrast imaging: theoretical and practical limitations," *Journal of biomedical optics*, vol. 18, no. 6, pp. 066 018–066 018, 2013.

[9] D. D. Duncan and S. J. Kirkpatrick, "Can laser speckle flowmetry be made a quantitative tool?" *JOSA A*, vol. 25, no. 8, pp. 2088–2094, 2008.

Feature Comparison under Different Noise Conditions for Gunshot Detection Task

Martin Hrabina, Milan Sigmund
Brno University of Technology
Department of Radio Electronics
Technicka 12, 61600 Brno, Czech Republic
hrabina@phd.feec.vutbr.cz, sigmund@feec.vutbr.cz

***Abstract*— Presented work investigates performance of three feature sets (LPC, LPCC and MFCC) in distinguishing gunshots from non-gunshot, mostly urban sounds under white noise conditions ranging from 30 dB to 0 dB with 10 dB step, including clean signal. Results show, that LPC coefficients are best at 30 dB with comparable results achieved by LPCC. MFCC are significantly better than others at 20 dB and 10 dB. Performance at 0 dB was balanced between LPC and MFCC – LPC had more true detections and MFCC achieved better score for false alarms.**

Keywords*— *gunshot detection, feature analysis, linear predictive coding coefficients, cepstrum, noise

I. Introduction

This work is a part of our effort of developing reliable gunshot detection system. It follows after first attempts, where we developed simple algorithm to distinguish gunshots from non-gunshot sounds (mostly impulsive, such as branch cracking) on relatively small dataset [1] and subsequently implemented it on signal processor [2]. In recently finished, but not yet published paper [3], we were investigating influence of frame size on extracted features, where we concluded that for some non-Fourier Transform based features, 11 ms is sufficient. In this publication, we also investigated influence of position of signal within frame on feature variability.

The aim of this paper is to test how features perform in task of distinguishing gunshots from other signals under different noise conditions. To keep consistency, we are using similar features as in previous publications, but we would like to expand feature list and include some newer, such as MPEG-7 descriptors.

Other publications that deal with feature evaluation use for example MRMR metric [4],[5] or use other techniques, like wrapper and filter feature selection [6].

II. Sound Dataset

Sound data come mainly from two sources. Gunshots are from free Firearm Sounds dataset [7] where a dataset of 1533 gunshots is available. The recordings proceed from various weapons (24 types), different angles (originally labeled as “front”, “behind”, “right” and “left”) and different distances (originally labeled only as “short”, “mid” and “long”). Non-gunshot signals consist of mostly urban noises, such as drilling, recordings of crowded places, jackhammer etc. from Urban Sounds Dataset [8]. The last and smallest part of used signals was recorded directly by us. All recordings used are in WAV format, mono, 44.1 kHz and 16 bit sampling (downsampled where needed, especially with sounds from [7]).

Different noise conditions are established adding white Gaussian noise with different sound-to-noise ratios (SNR). This noise was chosen due to its uniform power distribution in spectrum. In the future, use of real-life noises (such as from [8]) is planned instead of white noise. Noise was added using Matlab awgn(‘measured’) function, with ratios 30 dB, 20 dB, 10 dB and 0 dB.

All signals were segmented into frames of 11 milliseconds (485 samples at 44.1 kHz). This sample size can be applied because we do not use Fourier Transform (thus a power of two is not preferred). Additionally, in previous tests [3] 11 ms was chosen as the most suitable from candidates of 23 ms, 11 ms, 8 ms, 5 ms and 3 ms, with 11 ms being comparable with 23 ms (1024 samples, which is frequently used in similar applications) and outperformed lower resolutions in terms of false alarms / true detections (although only slightly for 8 ms).

Figures 1 and 2 show examples of gunshots from AK-47 assault rifle, with the first showing frontal recording with muzzle blast and shockwave and the second showing recording from behind of the shooter.

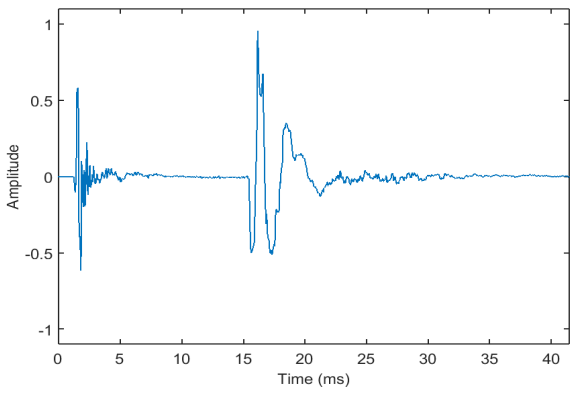


Fig. 1. Frontal recording of AK-47

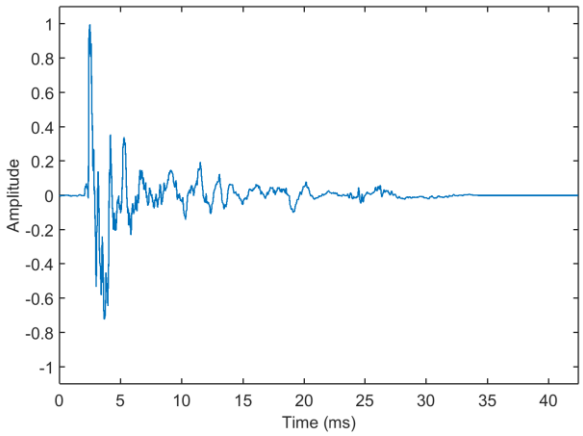


Fig. 2. Rear recording of AK-47

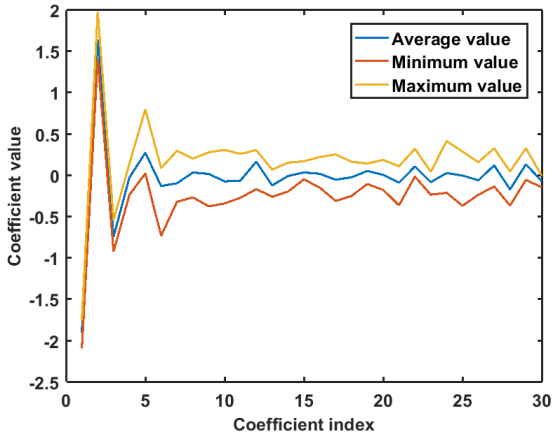


Fig. 3. LPC coefficients

III. Features and Methods

A. *Features used*

Investigated features consist of linear predictive coding coefficients (LPC), linear predictive cepstral coefficients (LPCC) and mel-frequency cepstral coefficients (MFCC) of even orders from 8 to 30. These features were selected because of our previous experience with them and their frequent use in similar applications [9], [10].

Figures 3 to 5 illustrate values of extracted features (minimal, average and maximal) of individual coefficients corresponding to the order 30 in each feature set. In these experiments, 7 recordings of AK-47 gunshots recorded from 100 m (frontal recording) were taken into account.

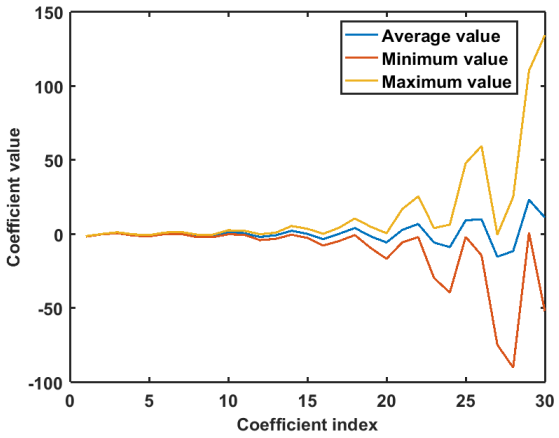


Fig. 4. LPCC coefficients

B. *Methods of Evaluation*

Performance of individual features and ranks was evaluated by observing false alarms and true detections ratio. As a detector, neural networks pattern recognition Matlab tool was used. These networks has one hidden layer, default value of 10 neurons in hidden layer was kept. Also, default value of division into non-overlapping training (70%), validation (15%) and testing (15%) sets was kept. We find these values as in range of common settings to be used [11]. Data division was based on random permutation for each training (i.e. each variation of feature and/or order).

Since neural networks use optimization algorithms such as gradient descent, which may yield different results each time, all results published in this paper are average of results of five training rounds.

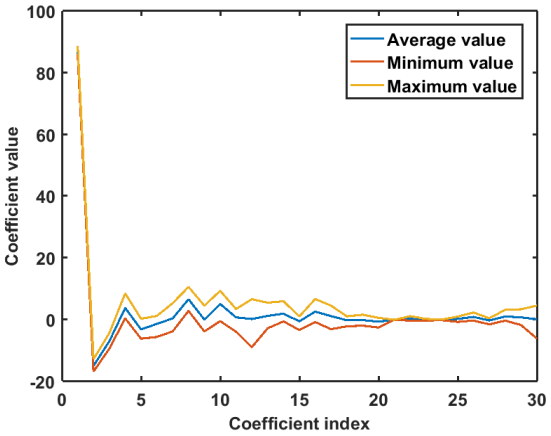


Fig. 5. MFCC coefficients

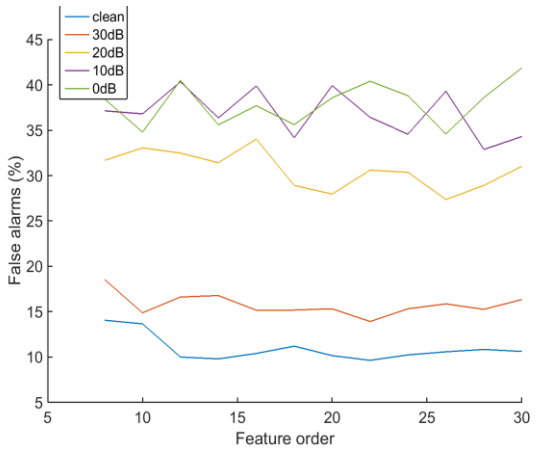


Fig. 6. LPC false alarms

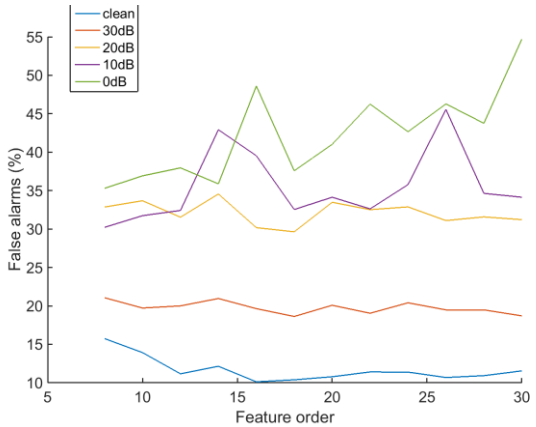


Fig. 7. LPCC false alarms

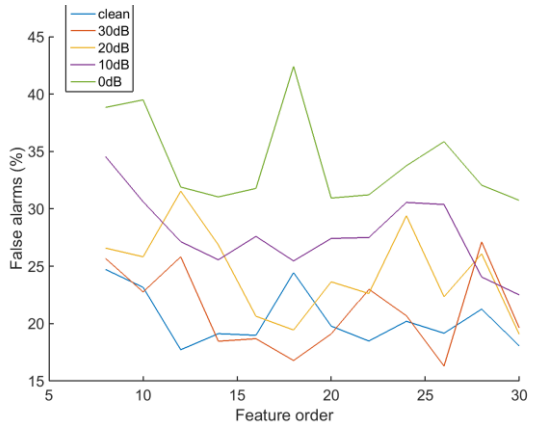


Fig. 8. MFCC false alarms

IV. Results

Results show, that performance beyond certain order (consequently, number of coefficients) varies only slightly, which is probably due to high correlation between features. Best results for no noise or low noise conditions were achieved by LPC coefficients with similar, but slightly worse results achieved by LPCC. MFCC performed by far best for 20 dB and 10 dB. At 0 dB LPC and MFCC coefficients were comparable, with LPC achieving more true detections and MFCC less false alarms. Figures 6 to 8 show false alarms for each feature depending on feature order and SNR.

Table 1 presents best results achieved for each SNR by different features, showing True Detections (TD), False Alarms (FA) and feature order for achieved results.

TABLE I. True Detections and False Alarms for different features and SNR

SNR	LPC			LPCC			MFCC		
	TD [%]	*FA [%]*	*order*	*TD [%]*	*FA [%]*	*order*	*TD [%]*	*FA [%]*	*order*
clear	86,7	9,6	22	86,3	10,1	16	75,1	17,7	12
30 dB	82,7	13,9	22	80,5	18,7	30	77,5	16,8	18
20 dB	62,1	27,4	26	60	29,6	18	74,8	19,4	18
10 dB	43,8	32,9	28	54,4	30,2	8	68,5	22,5	30
0 dB	49,8	34,6	26	49,7	35,3	8	44,6	30,9	20

V. Conclusion

In this paper, we have compared for gunshots detection some of the features usually used in sound recognition under different noise conditions. Results show, that although higher orders achieved best results, order itself has minor role on performance. With increasing noise levels, preference shifts from LPC to MFCC, but at 0 dB SNR, LPC and MFCC are comparable. In next research, we would like to investigate influence of various noise types on performance, e.g. urban noises instead of white noise, and also extended feature list including some of newer features (e.g. MPEG-7 descriptors).

Acknowledgment

Research described in this paper was financed by Czech Ministry of Education in frame of National Sustainability Program under grant LO1401. For research, infrastructure of the SIX Center was used. The research was also financially supported by the Brno University of Technology Internal Grant Agency under project no. FEKT-S-17-4707.

References

[1] M. Hrabina and M. Sigmund, "Acoustical detection of gunshots," in Proc. of 25th Int. Conf. Radioelektronika, Pardubice, 2015, pp. 150-153.

[2] M. Hrabina and M. Sigmund, "Implementation of developed gunshot detection algorithm on TMS320C6713 processor," in Proc. of SAI Computing Conference, London, 2016, pp. 902-905.

[3] M. Hrabina and M. Sigmund, "Analysis of linear predictive coefficients for gunshots detection based on neural network," unpublished.

[4] J. Palanichamy and K. Ramasamy, "An improved feature selection algorithm with conditional mutual information for classification

problems", in 2013 International Conference on Human Computer Interactions (ICHCI), Chennai, 2013, pp. 1–5.

[5] E. Vozarikova, J. Juhar, and A. Cizmar, "Acoustic Event Detection Based on MRMR Selected Feature Vectors", Journal of Electrical and Electronics Engineering, vol. 5, no. 1, pp. 277-282, 2012.

[6] J. Suto, S. Oniga, and P. P. Sitar, "Comparison of wrapper and filter feature selection algorithms on human activity recognition", in 2016 6th International Conference on Computers Communications and Control (ICCCC), Oradea, 2016, pp. 124–129.

[7] B. Jaszczak and B. Nelson, (2014). Free Firearm Sound Effects Library. [online] Available at: http://www.airbornesound.com/downloads/free-firearm-sound-effects-library-expanded-edition [Accessed March 2017].

[8] J. Salamon, C. Jacoby, and J. P. Bello, "A dataset and taxonomy for urban sound research", in Proc. of 22nd ACM International Conference on Multimedia, Orlando, USA, 2014, pp. 1041-1044.

[9] M. Sigmund, "Search for keywords and vocal elements in audio recordings," Elektronika ir Elektrotechnika, 2013, vol. 19, no. 9, pp. 71-74.

[10] S. Yang, J. Cao, and J. Wang, "Acoustics recognition of construction equipments based on LPCC features and SVM", in Proc. of 34th Chinese Control Conference (CCC), 2015, pp. 3987-3991.

[11] T.-E. Chen et al., "S1 and S2 Heart Sound Recognition Using Deep Neural Networks", IEEE Transactions on Biomedical Engineering, vol. 64, no. 2, pp. 372-380, Feb. 2017.

DNA Reads Feature Extraction Analysis

Robin Jugas
Department of Biomedical Engineering
Brno University of Technology
Brno, Czech Republic
e-mail: jugas@feec.vutbr.cz

***Abstract*—As the price of genome sequencing decreased, the complete genome sequencing is the new standard for exploring biological abilities of organisms. The commonly known contaminants like proteins and RNA are efficiently ousted but there is another contaminant which is diverse DNA. DNA contaminants are common in sequencing output and until recently nobody reckons with it, they are also part of already published genomes. The paper deals with issues by using alternative approach of genomic signal processing and other computational methods. The main objective is to detect contaminants in sequencing output and evaluate the analyses of usable features.**

Keywords— sequencing; genomic signal processing; features; contaminants; clustering; PCA; bioinformatics

I. INTRODUCTION

Since Sanger sequencing method was introduced in 1977 the development of sequencing DNA molecules has changed in various ways. Scientists are currently developing so called third generation sequencing methods and the price per sequenced human genome decreased from $ 100M to $ 1000 in 2015. The most used methods of the second generation also called new generation sequencing (NGS), generally, produce a huge amount of data. It is caused by a need to shear a DNA into many fragments and then sequence these fragments as no sequencer is able to sequence the whole DNA molecule at once. The original sequence is then obtained by process called assembly, which searches for overlaps and alignments between the reads and joins them together.

The paper focuses on the partial issue of high-throughput NGS sequencers which is contaminant DNA. There are some contaminants during DNA isolation like RNA and proteins, which are chemically washed out, but some studies already discovered that detectable levels of contaminants are present in sequencers output. Those contaminants are DNA of diverse species. [1,2,3] E.g., human DNA parts was observed in *E. coli* sequencing output [1].

I performed analyses of detecting contaminant DNA reads in NGS output by using genomic signal processing, features extraction, principal component analysis and machine learning.

Genomic signal processing (GSP) is generally alternative sight of studying DNA. The DNA is composed of four alternating letters representing nucleotides (A, C, T, G) forming basic unit of bioinformatics – base pair [bp]. GSP transform a letter based DNA to numerical based DNA using several methods which suit different utilization [4]. Numbers can show what letters cannot, e.g. hidden patterns, repetitions, similarities, etc. Some basic letter based approaches can be accomplished by using GSP, e.g. multiple alignments [5], taxonomy classification [6], etc. in less demanding manner.

II. METHODS

A. Numerical representations

Numerical representations are general methods for transforming DNA string compounded by four chars (A, C, G, T) representing the four nucleotides in DNA sequence. Thus, we acquire numerical sequence (signal). There are more published methods for the transformation [7], but I chose representations based on phase representation [2, 3]. To every nucleotide is assigned corresponding value (A = $\pi/4$; T = $-\pi/4$; C =$-3\pi/4$; G = $3\pi/4$) – the phase of a complex number. From basic phase representation are two others derived – cumulated phase and unwrapped phase representation. The cumulated phase is cumulative sum of the phase values of the signal from first to the last element. The unwrapped phase is corrected phase, the absolute value between the two values is kept less than π by adding or subtracting multiples of 2π. The unwrapped phase denotes the difference between positive and negative transitions, e.g. A→G, G→C, C→T, T→A are positive transitions (anti-clockwise). [8]

B. Features

From every read were extracted some chosen features. The features should ideally represent the read and be distinct enough to discriminate the reads. Several features were chosen: Hjorth parameters, energy, standard deviation, the average and maximum value of both first and second derivation. The energy was defined as $E = \sum_{-\infty}^{\infty} |x(n)|^2$. Hjorth parameters – activity, mobility and complexity was defined as $A = var(y(t))$, $M = \sqrt{var(\partial y(t))/var(y(t))}$, $C = M(\partial y(t))/M(y(t))$, where y(t) is signal. [10] Their origin is in EEG processing as slope descriptors. The activity represents mean power, mobility is the measure of mean frequency and complexity quantifies any deviation from the sine shape. [10]

C. Correlation

For acquiring information about relations between the features the Pearson correlation coefficient was computed using equation (1), where A and B denote two variables, μ_A and σ_A are the mean and standard deviation of A, respectively, and μ_B and σ_B are the mean and standard deviation of B [11]:

$$\rho(A,B) = \frac{1}{N-1}\sum_{i=1}^{N}\left(\frac{\overline{A_i - \mu_A}}{\sigma_A}\right)\left(\frac{B_i - \mu_B}{\sigma_B}\right) \quad (1)$$

D. Principal component analysis

The principal component analysis serves as one of the features extraction methods. It allows to describe features redundancy and reduce it to lower dimensionality. PCA method computes the linear combinations of variables and chooses the one with the highest variance. It creates the new set of variables called principal components, each principal component is a linear variation of original variables. Principal components are orthogonal. The number of principal components is the same as the original variables, but usually, first few components cover the majority of the total variance of the original data. [12]

E. Clustering methods

Clustering methods serve to divide the dataset into categories which should correspond with some real features of the data. K-means is heuristic unsupervised machine learning algorithm. It divides data into the preselected number of clusters. The algorithm consists of repeating two steps: assignment step and update step. In assignment step each observation is assigned to cluster according to the squared Euclidean distance criterium or another metric [13] :

$$S_i^{(t)} = \left\{ x_p : \left\| x_p - m_i^{(t)} \right\|^2 \le \left\| x_p - m_j^{(t)} \right\|^2 \right\} \quad (1)$$
$$\forall j, 1 \le j \le k$$

where $m_i^{(1)}, \ldots, m_k^{(1)}$ is the initial set of k means, x_p is assigned into one cluster $S^{(t)}$. In update step, the new means are calculated [13]:

$$m_i^{(t+1)} = \frac{1}{\left|S_i^{(t)}\right|}\sum_{x_j \in S_i^{(t)}} x_j \quad (3)$$

The algorithm converges when no assignments are changed. It does not guarantee to provide a global optimum in all cases. [13]

F. Dataset

The analysis was tested on synthetic datasets. The datasets are based on original *Escherichia coli* genome (GenBank NC_000913) and *Homo sapiens* 21q chromosome (GenBank BA000005) DNA sequences. For simulating NGS reads the wgsim [8, 9] tool was used. For mapping reads to reference genome the BWA [16] tool was used. The BWA output is SAM file (Sequence Alignment Map), which contains reads together with their position in the reference genome. It is based on Burrows-wheeler transform and provides high-quality output. [17] The total output is 5000 *E. coli* reads mixed with another 300 reads from human genome representing contaminants. The reads are simulating Illumina sequencing platform reads and have 300 bp in length.

III. Results and Discussion

The algorithm was programmed in MATLAB environment. The output of BWA tool, the SAM file, contains both mapped and unmapped DNA reads as letter strings. Reads were converted into the DNA signals using the cumulative sum of phase. From every read the nine features were acquired – three Hjorth parameters, signal energy, standard deviation, the average and maximum value of first derivation and average and the maximum value of the second derivation.

Before carrying principal component analysis is needed to preprocess the dataset. Thus, the features were normalized by subtracting mean and divided by standard deviation of the variable: $y_{norm} = (y - \bar{y})/\sigma(y)$.

A. Correlation

The Pearson correlation coefficient was computed between variables. This should provide information about features dependence. The significant correlation coefficient occurred between several features: Activity, Complexity, SD, and Energy was correlating together. In Table 1 there are correlation coefficients (>0,8) between several features.

Table 1 – Correlation between features (>0,8)

	A	M	C	EN.	SD
A	1	0	0,87	0,87	0,86
M		1	-0,84	0	-0,84
C			1	0,80	0,99

B. Principal component analysis

The principal component analysis was computed using SVD algorithm with internal MATLAB function pca. The 3D space was chosen for further analysis and clustering, so three principal components were requested. All three components exceed 78 % of total variability. In figure 1 is a PCA biplot, which also denotes dependency between energy, activity, SD, and complexity.

The PCA provided a new set of variables with reduced dimensionality to 3D (originally 9). That allows features to be easily displayed as the graph.

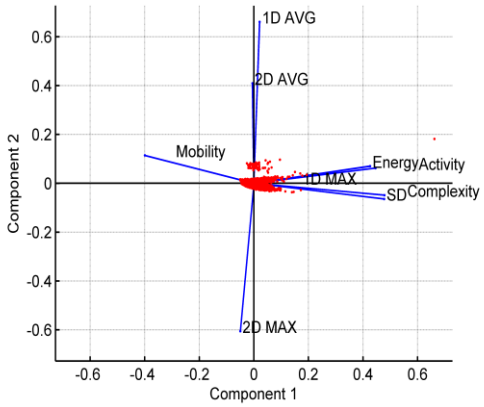


Fig. 1. Biplot of principal components

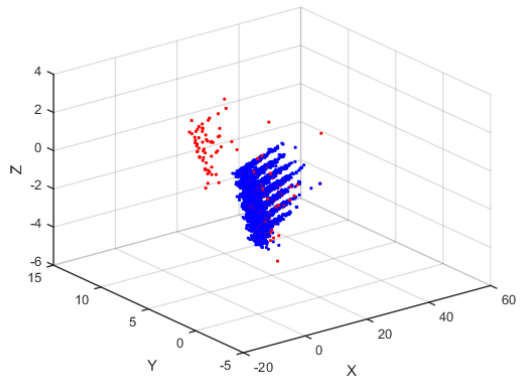


Fig. 2. 3D graph of reads features, red - contaminants

In figure 2 are displayed reads represented by reduced features, the red dots are contaminants (human origin) whereas blue ones are *E. coli* reads. There is clearly visible that many reads have similar features as original bacterial reads. Those will be demanding to distinguish as contaminants.

C. K-means clustering

The k-means algorithm was performed at reduced features data. As distance metric was chosen city-block distance (sum of absolute differences): $d(x,c) = \sum_{j=1}^{p} |x_j - c_j|$, where x is observation and c is centroid coordinates. The algorithm was started with given initial centroid positions, as algorithm tends to give inaccurate solutions with random initial centroids. The number of clusters was set to two, as I wanted to differentiate the contaminants from original reads.

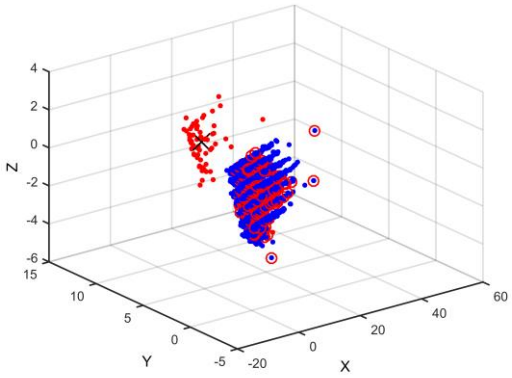


Fig. 3. k-means clusters. Red-contaminants clustered correctly, red circle clustered incorrectlv. Crosses - centroids

In figure 3 there are contaminant reads colored red, the original reads are blue, and those contaminants clustered wrongly as original colored by a red circle. The different view from the side is in figure 4.

D. Evaluation

The method was tested on ten different datasets. The dataset contains 5000 original reads of E. coli together with 300 reads of a human genome. The evaluation consists of computing confusion matrix and accuracy.

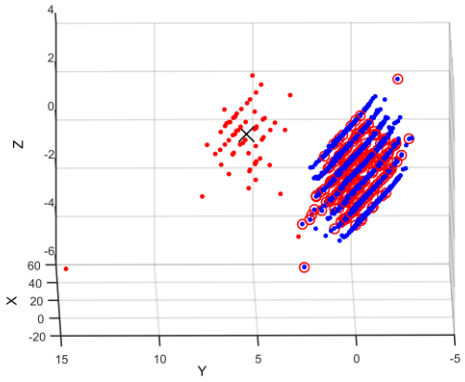


Fig. 4. 3D graph of reads features, side view

The analysis of using feature extraction from reads shows overall usability of this approach. The features extraction is a simple and undemanding process. Principal component analysis is also suitable since e.g. Hjorth descriptors alone are not discriminative enough. So, combining more features and reduced them into principal components is an effective approach.

The k-means clustering in the case of this dataset reaches its limit because many contaminants share similar coordinates in space with original data. It makes any division very demanding and sophisticated.

Table 2 – Confusion matrix (1. dataset)

	Predicted positive	Predicted negative
Actual positive	63	237
Actual negative	0	5000

In Table 2 there is confusion matrix for one of the datasets. The original reads create a solid cluster in graph space, so all of them was clustered correctly. The clustering accuracy reached 96 % but sensitivity was only 21 %. As it was stated, it is caused by similar features with original reads. For other datasets the results were similar, reaching in average 25 % of sensitivity and 96 % of accuracy.

IV. CONCLUSION

The state-of-the-art in DNA processing field called bioinformatics is standard letter based string methods. The string methods are already developed for many tasks and produce quality results but they cannot provide further information about the sequence. The genomic signal processing can provide such information. By transforming DNA into the numerical signal, we can apply different methods which were available only to common signal processing and apply them to DNA signals. In the paper was presented the possibility of using GSP for disposing of contaminants in sequencing data, which was only recently seriously considered.

The features were extracted from reads and were analyzed using Pearson correlation coefficient between the variables. The features dimensionality was reduced by principal component analysis and plotted in 3D space. Then the k-means clustering was applied in order to discriminate the contaminant DNA from original DNA. The average rate of discrimination was 25%, the rest of reads share similar features with original data. So, k-

means clustering reaches its limit in this case. A further improvement would be possible by using supervised machine learning and other methods which are available for genomic signal processing domain.

ACKNOWLEDGMENT

This work was supported by grant project GACR 17-01821S. Computational resources were partially provided by the CESNET LM2015042 and the CERIT Scientific Cloud LM2015085, provided under the programme "Projects of Large Research, Development, and Innovations Infrastructures".

REFERENCES

[1] R. W. Lusk, “Diverse and widespread contamination evident in the unmapped depths of high throughput sequencing data,” *PLoS One*, vol. 9, no. 10, 2014.

[2] S. Mukherjee, M. Huntemann, N. Ivanova, N. C. Kyrpides, and A. Pati, “Large-scale contamination of microbial isolate genomes by Illumina PhiX control,” Stand. Genomic Sci., vol. 10, no. 1, p. 18, 2015.

[3] M. Laurence, C. Hatzis, and D. E. Brash, “Common contaminants in next-generation sequencing that hinder discovery of low-abundance microbes,” PLoS One, vol. 9, no. 5, pp. 1–8, 2014.

[4] D. Anastassiou, “Genomic signal processing,” IEEE Signal Process. Mag., vol. 18, no. 4, pp. 8–20, 2001.

[5] H. Skutkova, M. Vitek, K. Sedlar, and I. Provaznik, “Progressive alignment of genomic signals by multiple dynamic time warping,” J. Theor. Biol., vol. 385, pp. 20–30, 2015.

[6] H. Skutkova, M. Vitek, P. Babula, R. Kizek, and I. Provaznik, “Classification of genomic signals using dynamic time warping.,” BMC Bioinformatics, vol. 14, no. Suppl 10, p. S1, 2013.

[7] M. Abo-Zahhad, S. M. Ahmed, and S. a. Abd-Elrahman, “Genomic Analysis and Classification of Exon and Intron Sequences Using DNA Numerical Mapping Techniques,” Int. J. Inf. Technol. Comput. Sci., vol. 4, no. 8, pp. 22–36, 2012.

[8] P. D. Cristea, “Phase analysis of DNA genomic signals,” Proc. 2003 Int. Symp. Circuits Syst. 2003. ISCAS ’03., vol. 5, p. V-25-V-28, 2003.

[9] P. D. Cristea, “Large scale features in DNA genomic signals,” Signal Processing, vol. 83, no. 4, pp. 871–888, 2003.

[10] B. Hjorth, “EEG analysis based on time domain properties,” Electroencephalogr. Clin. Neurophysiol., vol. 29, no. 3, pp. 306–310, 1970.

[11] I. The MathWorks, “Correlation coefficients (corrcoef),” 2017. [Online]. Available: https://www.mathworks.com/help/matlab/ref/corrcoef.html. [Accessed: 31-Jul-2017].

[12] L. I. Smith, “A tutorial on Principal Components Analysis Introduction,” Statistics (Ber)., vol. 51, p. 52, 2002.

[13] D. J. C. MacKay, Information theory, inference, and learning algorithms. Cambridge University Press, 2003.

[14] H. Li, “wgsim,” 2017. [Online]. Available: https://github.com/lh3/wgsim. [Accessed: 30-Jul-2016].

[15] M. Escalona, S. Rocha, and D. Posada, “A comparison of tools for the simulation of genomic next-generation sequencing data,” Nat Rev Genet, vol. advance on, no. 8, pp. 459–469, 2016.

[16] H. Li and R. Durbin, “Fast and accurate short read alignment with Burrows-Wheeler transform,” Bioinformatics, vol. 25, no. 14, pp. 1754–1760, 2009.

[17] A. Hatem and D. Bozdag, “Benchmarking short sequence mapping tools,” Proc. - 2011 IEEE Int. Conf. Bioinforma. Biomed. BIBM 2011, pp. 109–113, 2011.

Neural Network Localization Passengers Inside the Car

Martin Kotol
Department of Radio Electronics, Brno University of Technology,
Brno, Czech Republic, xkotol00@stud.feec.vutbr.cz

***Abstract*—In the paper, we introduce a neural estimator of seats occupied by passengers in a vehicle. The estimator is built from the radial basis function (RBF) neural network. Occupied seats in a vehicle are estimated from transmissions between antennas operating at 60 GHz bandwidth.**

I. INTRODUCTION

In the paper, an artificial neural network (ANN) is applied to estimate occupied seats in a vehicle. The localization operates in the unlicensed frequency band from 55 GHz to 65 GHz to separate measuring wireless links by high attenuation of atmosphere.

Frequency responses of transmission among measuring antennas are influenced by the number and the position of passengers in a vehicle. ANN is trained such a way to identify requested parameters from measured transmissions. Such a neural estimator has not been published in the open literature yet.

In Section II, radial basis function (RBF) neural networks are reviewed. Section III describes in-car measurements performed for collecting training data and testing ones. Section IV discusses processing of measured data, and Section V concludes the paper.

II. ARTIFICIAL NEURAL NETWORKS

An artificial neural network (ANN) can be defined as a network consisting from a large number of simple non-linear units (neurons) which are mutually connected.

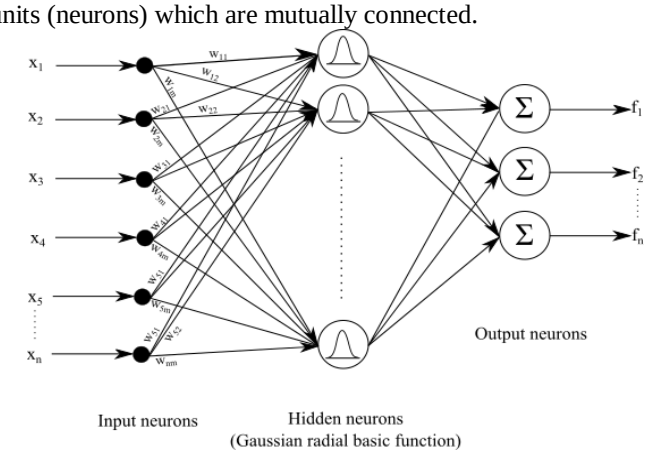


Fig. 1. RBF neural network.

Figure 1 shows the general structure of a RBF network. Here, x_1 to x_n are inputs, w_{11} to w_{nm} are weights between neurons, and f_1 to f_n are outputs. During training, weights are adaptively changed so that given input patterns are associated with respective output targets.

III. DESCRIPTION OF MEASUREMENT

In order to estimate occupied seats in a vehicle, transmissions among measuring antennas were processed in the band from 55 GHz to 65 GHz. For measurement, two configurations of transmit antennas (Tx) and receive ones (RX) were used (see Figure 2):

- First, the transmit antenna was located in the back left corner of the car roof and receiving antenna was placed into the opposite front right corner.
- Second, antennas were situated in the front center and the rear center of the roof.

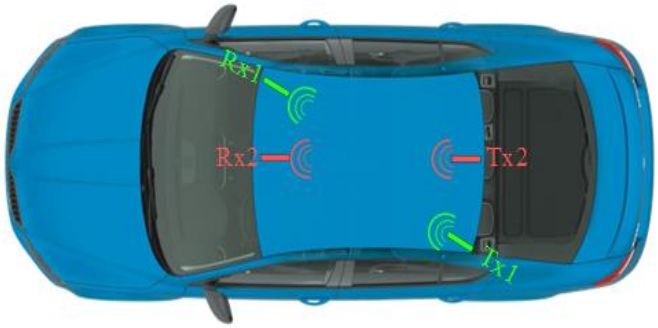


Fig. 2. Location of antennas.

The transmit antenna and the receive antenna were built from an open end of the waveguide WR15. In Figure 3, the receive antenna is depicted, and in Figure 4, the transmit antenna is shown.

Transmissions between red antennas and green ones (Figure 2) were measured for all the possible configurations of passengers in a car. Each configuration was measured five times to suppress of random variations.

Outputs of described measurements were used to train an ANN. The trained ANN was exploited to estimate occupied seats from measurements in a different car occupied by different passengers sitting in slightly shifted positions (to demonstrate robustness of the trained neural estimator).

Fig. 3. Receive antenna.

Fig. 4. Transmit antenna.

IV. PROCESSING MEASUREMENT VALUES

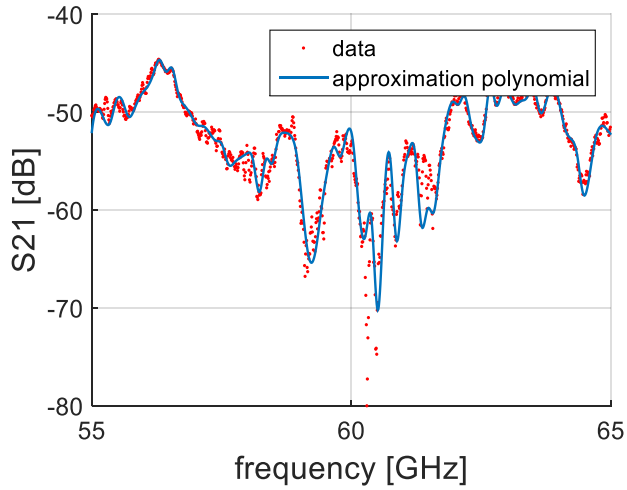


Fig. 5. Measured frequency response of the transmission between the antennas (red). Polynomial aproximation of frequency response (blue).

Figure 5 shows measured values of the transmission between the transmit antenna Tx1 and the received antenna Rx1 (red points). The measured values were approximated by a polynomial function to partially smooth data for training neural networks (blue line).

As an input training pattern, we created a matrix of operational frequencies and corresponding transmission coefficients for all the measured configurations of passengers. As an output training pattern, we used a matrix of corresponding configurations of passengers (see Table 1). The number 1 indicates the occupied seat in the car and the number 0 is related to an unoccupied seat. All the configurations of passengers were considered. The training process was stopped when reaching the relative training error lower than 10^{-6}.

For verification of the trained ANN, we created a new input matrix. Elements of the validation matrix were related to different frequencies than used for training.

In order to demonstrate robustness, the trained ANN, transmissions between antennas were measured for different passengers sitting in a different car in different positions. Examples are given in Table 2.

TABLE I. SELLECTED TRAINING CONFIGURATIONS OF PASSENGERS *(1 - OCCUPIED, 0 - EMPTY)*

Number configuration crew	**Occupation passengers in the car**				
	driver	***passenger 1***	***passenger 2***	***passenger 3***	***passenger 4***
1	1	1	1	1	1
4	1	1	1	0	1
9	1	1	0	0	1
11	0	0	1	1	1
13	1	1	0	1	0
21	1	0	0	0	1
28	1	0	0	0	0

TABLE II. SELECTED TESTING CONFIGURATIONS *(D-DRIVER, DJ-DRIVER WITH JACKET, P1-PASSANGER 1, P1J- PASSANGER 1 WITH JACKET, P2J- PASSANGER 2 WITH JACKET, P2BJ- PASSANGER 2 WITH JACKET AND SIT ON THE POSITION BETWEEN PASSANGER 2 AND PASSANGER 3, P3J- PASSANGER 3 WITH JACKET, P4J- PASSANGER 4 WITH JACKET)*

Number configuration crew	**Occupation passengers in the car**				
	driver	***passenger 1***	***passenger 2***	***passenger 3***	***passenger 4***
1	DJ	P1J	P2J	P3J	P4J
4	D	P1J	0	P3J	0
5	DJ	P1J	P2J	0	0
6	D	P1J	P2J	0	P4J
10	DJ	0	P2BJ	0	0
16	D	P1	P2J	0	P4J
22	DJ	P1	0	0	0

Figures 6 and 8 show training configurations. Yellow circles indicate occupied position in a car. Figures 7 and 9, show testing configurations. Green circles indicate correct estimates of occupied positions, and the red circle corresponds with the wrong estimate of the occupied position.

Fig. 6. Training configuration no. 9
(occupied positions – yellow)

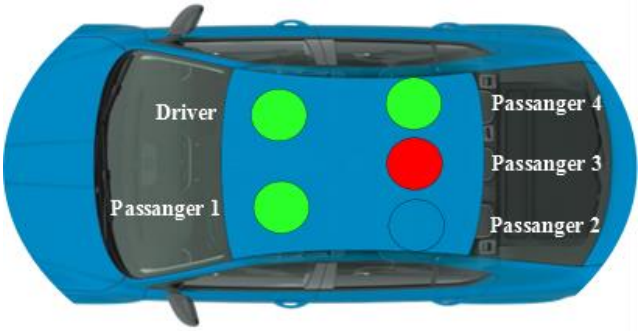


Fig. 7. Testing configuration no. 9. *(Correct identification – green, wrong estimation – red)*

Fig. 8. Training configuration no. 11
(occupied positions – yellow)

Fig. 9. Testing configuration no. 11 *(Correct identification – green, wrong estimation – red)*

Figure 10 shows the number of incorrect estimations for all measured configurations of passengers (the number of configuration of passengers follows the axis x).

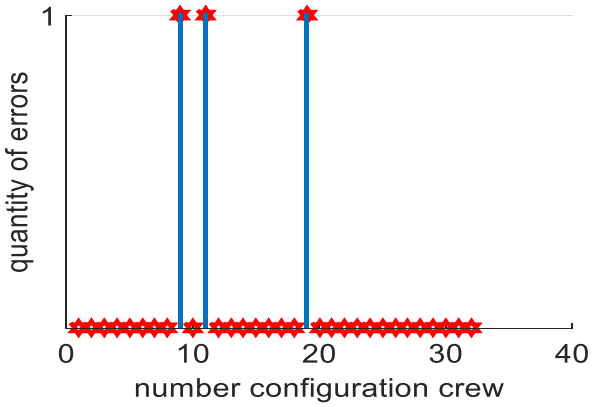


Fig. 10. The number of wrong estimations versus measured configurations.

Figure 11 shows the number of incorrect estimations for all measured non-standard configurations of passengers (is the number of configuration follows the axis x).

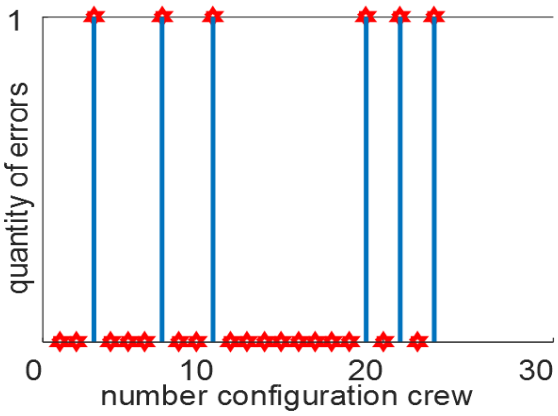


Fig. 11. The number of wrong estimations versus measured independent configurations

V. CONCLUSION

In the paper, we introduce a neural estimator of passengers in a vehicle.

The estimator uses radial basis function ANN. For the ANN, we created an input training pattern and an output testing pattern. From measured frequency responses of transmission between the transmit antenna and the receive antenna, the neural estimator identifies the corresponding configuration of passengers in a car.

Next, we measured non-standard configurations of passengers for verification Results clearly show that a real configuration of passengers is in a good agreement with the results provided by ANN.

ACKNOWLEDGMENT

Research described in this paper was financially supported by the Czech Science Foundation under grant no. P102/12/1274. Measurements were performed in laboratories of the SIX Research Center.

[1] Blumenstein, J.; Mikulasek, T.; Marsalek, R.; Prokes, A.; Zemen, T.; Mecklenbräuker, C., "In-Vehicle mm-Wave Channel Model and Measurement," in Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th , vol., no., pp.1-5, 14-17 Sept. 2014

[2] Specht, Donald F., "A general regression neural network," in Neural Networks, IEEE Transactions on , vol.2, no.6, pp.568-576, Nov 1991

[3] D. Kriesel, "A Brief Introduction to Neural Networks," avail. at http://www.dkriesel.com, n 2007.

[4] Jingwen Tian, Meijuan Gao, Fan Zhang, "Network Intrusion Detection Method Based on Radial Basic Function Neural Network," In Proceedings of E-Business and Information System Security, 2009, Wuhan, China.

[5] Z. Raida, "Modeling EM structures in the Neural Network Toolbox of MATLAB," IEEE Antennas and Propagation Magazine, 2003, vol. 44, no. 6, p. 46-56.

Measured Properties of Anisotropic Materials Used in Aerospace Applications

David Krutilek
Department of Radio Electronics
Brno University of Technology
Brno, Czech Republic
xkruti01@stud.feec.vutbr.cz

***Abstract*—This paper deals with the theoretical aspects of the measurement of shielding effectiveness. An alternative method of these measurements is described in this contribution. The theoretical design and parameters of measurement the speciments are discussed in other parts of the paper. Carbon fibers are electrically conductive and can be added to cement to enhance the shielding effectiveness. The paper describes possible methods for measuring shielding effectiveness of anisotropic materials with carbon fiber.**

Keywords—Carbon composite; Shielding effectivenes; Measuring; Aviation industry; Coaxial flange.

I. INTRODUCTION

The use of various materials, such as magnesium, plastic, fabric or wood in the field of aircraft construction tends to disappear. Reduction of these materials started in the mid-1950s. Even aluminum is notably being diminished in use these days. Its use has decreased from 80 percent of airframes produced in 1950 to only around 15 percent aluminum and aluminum alloys used for airframe construction. The materials mentioned above are replaced with nonmetallic aircraft materials, especially with reinforced plastics and advanced composites [1].

In order to enhance aircraft design, the aircraft industry began to produce synthetic fibers during the 1940s. From that time on, composite materials have been used much more frequently. When speaking about composites, the majority of people recall fiberglass and eventually also graphite or aramids (Kevlar). Even though the use of composites began in aviation and aircraft industry for the first time, many other industries have included composites into their production, such as auto racing, sport equipment, boating as well as defense industry.

One of the definitions says that a "composite" material represents a mixture of different materials or things. The meaning of this definition is too general – it may refer to metal alloys which are made from various metals to enhance diverse characteristics, such as strength, ductility, conductivity or other desired properties. Similarly, the composition of composite materials can be seen as a combination of reinforcement components, e.g. a fiber, whisker, or particle. These components form a structure, as they are surrounded and held in place by a resin. When put apart, reinforcement and resin are totally different than when combined. Although these two materials are in their combined state, they can be still identified and each component mechanically separated. One composite, concrete, is composed of cement (resin) and gravel or reinforcement rods for the reinforcement to create the concrete.

Compared to other traditional materials used aircraft industry, composites are superior not only because of their increased strength but also because of their ability to accommodate the performance needs. Since the use of composites is increasing, the costs, design, inspection ease as well as information about strength to weight properties definitely help composites become the material of choice for aircraft construction [2].

This article is focused on determination of selected electric parameters of composite materials used in aircraft industry. Verification of gained results will be implemented with the aid of numeric models. Functions as well as required parameters of this template are verified with simulations. The obtained results are discussed below.

II. THE SPECIMENT HOLDER

This test method provides a procedure for measuring the electromagnetic (EM) shielding effectiveness (SE) of a planar material for a plane, far-field EM wave. From the measured data, near-field SE values may be calculated for magnetic (H) sources for electrically thin specimens [3], [4]. Electric (E) field SE values may also be calculated from this same far-field data, but their validity and applicability have not been established.

The measurement method is valid over a frequency range of 30 MHz to 1.5 GHz. These limits are not exact, but are based on decreasing displacement current as a result of decreased capacitive coupling at lower frequencies and on overmoding (excitation of modes other than the transverse electromagnetic mode (TEM)) at higher frequencies for the size of specimen holder described in this test method. Any number of discrete frequencies may be selected within this range. For electrically thin, isotropic materials with frequency independent electrical properties of conductivity, permittivity, and permeability, measurements may be needed at only a few frequencies as the far-field SE values will be independent of frequency. If the material is not electrically thin or if any of the

parameters vary with frequency, measurements should be made at many frequencies within the band of interest.

A. Apparatus

A basic equipment setup is shown in Fig. 1. Physical dimensions of a specimen holder are given in D4935 – 10 [5]. Fig. 2 shows a half of template spread into particular structural elements. The specimen holder is an enlarged, coaxial transmission line [6] with special taper sections and notched matching grooves to maintain a characteristic impedance of 50 Ω throughout the entire length of the holder. This impedance is checked and any variations greater than ±0.5 Ω are corrected.

Fig. 1 General test setup.

There are three important aspects to this design. First, a pair of flanges in the middle of the structure holds the specimen. This allows capacitive coupling of energy into insulating materials through displacement current. Second, a reference specimen of the same thickness and electrical properties as the load specimen causes the same discontinuity in the transmission line as is caused by the load specimen. Third, nonconductive (nylon) screws are used to connect the two sections of the holder together during tests. This prevents conduction currents from dominating the desired displacement currents necessary for the correct operation of this specimen holder

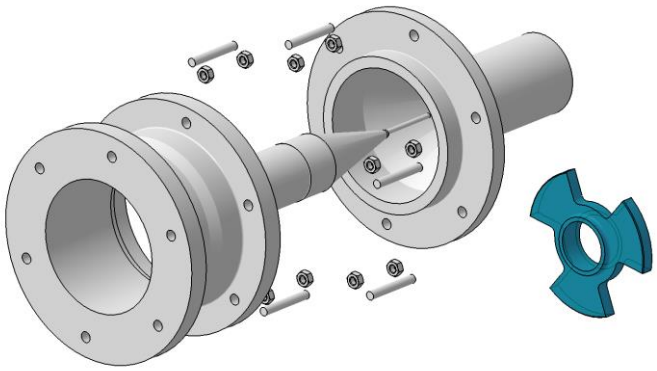


Fig. 2 Displayed half of measuring apparatus is structured as follows: grey parts are from brass, blue parts represent plexiglass.

B. Test Specimens

The reference and load specimens shall be of the same material and thickness. Both are shown in Fig. 3. Dimensions are shown in [5] – annulus of referential sample proved to have diameter 33 and 76.2 mm. The load specimen can be larger than the outer diameter of the flange on the holder but keeping them to the recommended dimensions will expedite handling.

Specimen thickness is a critical dimension. For the best repeatability of SE measurements, reference specimen and load specimen shall be identical in thickness. For this test method, two specimens are considered to have identical thickness if the difference in the average thicknesses is less than 25 μm and the thickness variation within and between specimens is less than 5 % of the average.

Materials of the specimens may be either homogeneous or inhomogeneous, single or multiple layered, and conducting or insulating. Measured SE values of inhomogeneous materials are dependent on geometry and orientation, and results are less repeatable than for homogeneous materials.

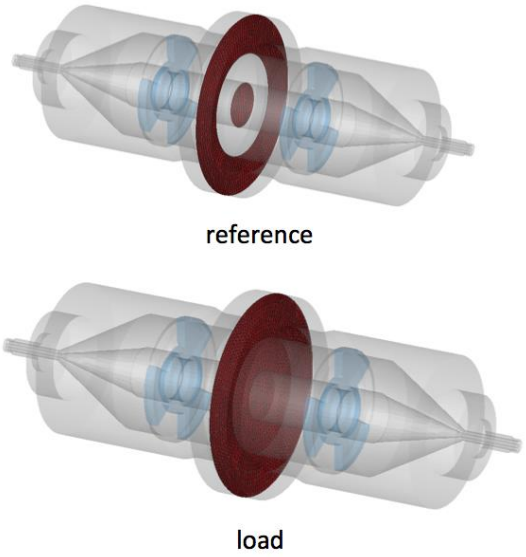


Fig. 3 Tested specimens in holder – reference and load.

III. VERIFICATION OF THE HOLDER

An initial check of the specimen holder should be performed with a time-domain reflectometer or other suitable instrument to ensure that a characteristic impedance of 50 ± 0.5 Ω has been achieved during construction and that this impedance has not been degraded during shipment or handling. A time-domain system can give location of a mismatch in addition to its magnitude. Each time the ancillary equipment is connected to the specimen holder, good practice requires measurement of a reference specimen to ensure the measurement system is in proper working order.

The dynamic range (DR) of the system can be checked by comparing the maximum signal level obtained with a reference specimen to the minimum signal level obtained when using a metallic load specimen. The lower limit of the measurement system sensitivity is a function of the sensitivity and bandwidth of the receiver. Narrowing the bandwidth of the receiver lowers the detectable level but increases the measurement time. Leakage caused by connectors or cables may reduce the DR of the system by providing a parallel signal path that does not pass through the specimen. If a step attenuator placed in series with the specimen holder causes a change in the minimum signal detected that corresponds to a change in attenuator setting, and if the step attenuator itself does not cause a leakage path, leakage is negligible and the DR measured above is correct. If

the levels do not correspond, the attenuation should be increased until a one-to-one correspondence is achieved to determine the DR. Since leakage from a coaxial connector is determined not only by the quality of the connector, but also by the amount of torque used in tightening the connector, connections should be rechecked.

Fig. 4 shows transmitting parameters of template without specimen. It is visible that pass-through signal occurring in required frequency domain is not notably decreasing and resonances are located out of measured area.

The norm also highlights the following fact: If a standard reference specimen such as gold film deposited on mylar is available, measurement of its SE value can provide assurance that the entire system is working properly. A specimen with the surface resistivity of 5 Ω commonly possess SE = –32 ±3 dB [5]. This was also proved by gained results, which is illustrated in Fig. 5. Any other known specimen may be used to check setup-to-setup repeatability.

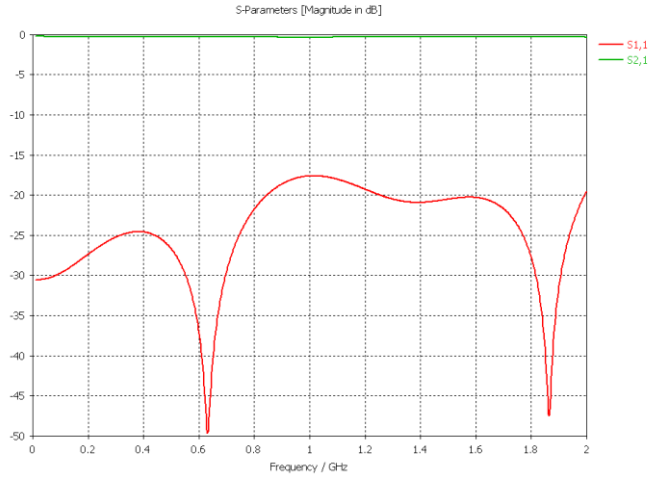


Fig. 4 Transmitting parameters of template without specimen.

SE is usually expressed in decibels (dB) by the following equation:

$$SE = 10log\frac{P_1}{P_2}(dB), \quad (1)$$

where P_1 is received power with the material present, and P_2 is received power without the material present.

If the receiver readout is in units of voltage, use the following equation:

$$SE = 20log\frac{V_1}{V_2}(dB), \quad (2)$$

where V_1 and V_2 are respective voltage levels with and without a material present.

According to these equations, SE will have a negative value if less power is received with the material present than when it is absent

IV. Uncertainty Estimation of SE Measured Values

This part gives an estimate of uncertainty in the measured values of SE. The sources of error we considered are operator errors, specimen-caused errors, and measurement system errors.

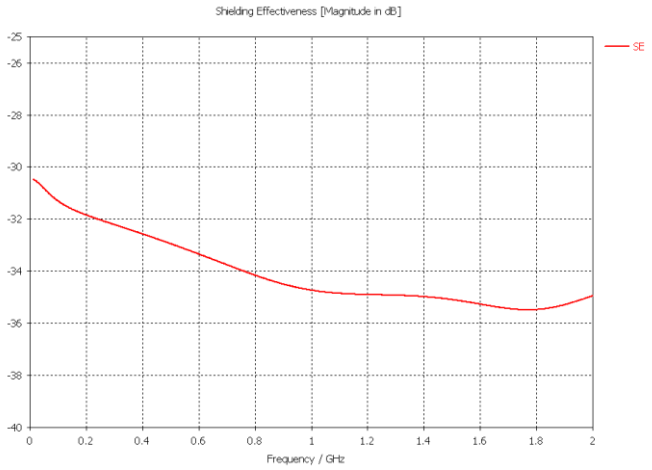


Fig. 5 A specimen with the surface resistivity of 5 Ω.

Operator errors may be caused by either carelessness or lack of experience and training. No limits can be placed on such errors, but the deviation from any norm may be large enough that an experienced observer will be able to determine that the results are indeed erroneous.

Specimen-caused errors are affected by irregularities in specimens, either in preparation or in inherent structure. Isotropic, homogeneous specimens with smooth surfaces will give the most repeatable results. If the reference specimen and load specimen have different thickness, a bias error will be introduced. If both specimens have the same thickness, but the thickness is irregular over each specimen, random errors will be introduced. Inhomogeneities or anisotropicities in specimens cause various effects depending on size, distribution, and geometric arrangement. Experience with measurements on many types of specimens indicates that repeatability of measured data may be expected except if the surface is rough. A round robin of measurements made on different types of specimens bears this out.

Source	Systematic	Random
Mismatch	±0.5 dB	±0.5 dB
Power instability in signal generator	±0.4 dB	±0.4 dB
Receiver calibration	±3.3 dB	±0.1 dB
Total	±1.2 dB	±1.0 dB

Tab.1 Summary of estimated uncertainties [5].

Measurement-system errors are caused by impedance mismatches, generator instabilities, leakage paths, limited dynamic range, limited frequency range, and receiver errors. To the extent that an experienced, well-trained operator can make measurements over the appropriate frequency range, within the dynamic range of the system, avoid leakage paths

within the measurement system's dynamic range, use suitable attenuators to avoid mismatches, monitor, and adjust input power to keep it constant, then measurement system errors may be reduced to a very modest part of the total error.

In Table 1, a summary of estimated errors under favorable conditions by a skilled operator is given. The systematic error in the receiver is probably irrelevant since the SE values are based on difference measurements. The random error that relates to drift over a few-minute time period is very relevant. If no attenuators were used, the mismatch error on the generator side would be excessive since in one case, with the reference specimen installed in the holder, the impedance seen by the signal generator is determined almost entirely by the receiver, and in the other case, with the load specimen installed, the impedance seen by the signal generator is almost a short circuit for conductive specimens. The actual change in impedance is greatly reduced by the attenuator between the signal generator and the specimen holder. The actual change of impedance level seen by the signal generator may also load the signal generator and cause the output power to vary from one condition to the other. These changes can be monitored by use of a bidirectional coupler, and corrections can be made to compensate for them. This couplers not shown as part of the setup shown in Fig. 1, so the error given for generator instability is based on no compensation. The size of the corrections measured with a coupler was the basis to determine the magnitude of this effect if no compensation is used.

Results from a round robin of measurements performed by five different individuals at five different organizations on five sets of specimens of three different materials indicate what level of agreement may be expected. These results reflect variations as a result of all causes. Results from all five sets are not included since a few data points are obvious outliers, probably caused by operator error. One material, stainless steel deposited on an ABS base, had a very smooth surface. The standard deviation for this specimen based on measurements by five individuals on five specimens was 0.9dB. A second material, carbon coated on one side of an ABS base, was not quite as smooth, and the standard deviation for this specimen based on measurements by five individuals on five specimens was 2.0 dB. The third material, nickel coated on an ABS base, showed substantial surface roughness. The standard deviation for this specimen based on measurements by five individuals on five specimens was 6.0 dB. For all three materials, the same specimen holder and numbered specimens were used. Each of the five individuals used different ancillary equipment.

The systematic uncertainty, or bias error, is ±1.2 dB. The random error, based on a standard deviation of 2.5 dB from the round robin results on reasonably smooth specimens, is within ±5 dB.

V. Conclusions

The aim of this article was to verify the function and design of measuring device with the aid of numeric methods.

The calculations show that required frequency extent was from 30MHz to 1.5 GHz. The simulation also included a sample with surface drag 5 Ohm. SE with value –32 ±3 dB was set according to norm recommendations.

The template is currently being produced. Physical prototype will serve for measurement of composite material parameters. The obtained data will be used for creation of numeric models. These models will be employed while simulating the effects of lightning affecting composite aircraft or the effects of HIRF.

Acknowledgment

Research described in this paper was financed by Czech Ministry of Education in frame of National Sustainability Program under grant LO1401.This work also receives funding through specific research grant project FEKT-S-14-2281. Special thanks belong to EVEKTOR Company for providing and manufacturing the specimen holder.

References

[1] ASTM F3114-15, Standard Specification for Structures, ASTM International, West Conshohocken, PA, 2015, www.astm.org

[2] Daniel, I. M., Ishai, O.,"Engineering Mechanics of Composite Materials (2nd Edition)" Oxford, New York 2006, ISBN 978-0-19-515097-1

[3] Wilson, P. F., and Ma, M. T., "A Study of Techniques for Measuring the Electromagnetic Shielding Effectiveness of Materials," NBS Technical Note 1095, May 1986.

[4] Adams, J. W., and Vanzura, E. J., "Shielding Effectiveness Measurements of Plastics," NBSIR 85-3035, January 1986.

[5] ASTM D4935-10, Standard Test Method for Measuring the Electromagnetic Shielding Effectiveness of Planar Materials, ASTM International, West Conshohocken, PA, 2010, www.astm.org

[6] Černohorsky, D., Novaček, Z., Raida, Z., Electromagnetic waves and lines.. Brno: Vutium 1999. 136 p. ISBN 80-214-1261-5.

Bicycle Charger for Mobile Phone

Jan Kufa
Department of Radio Electronics
SIX Research Center, Brno University of Technology
Brno, Czech Republic
xkufaj00@stud.feec.vutbr.cz

***Abstract*— Nowadays, we cannot imagine life without a mobile phone. It does not matter what we do, we just want to have it and use it. However, modern battery-powered devices have low battery life, whether it is a smartphone or outdoor navigation. In spite of the large network coverage of power supply, there are still places where we cannot connect a device to the power supply. One example can be a long bike ride. The propose of a suitable method of charging mobile devices while riding a bicycle is the aim of this work.**

Keywords— Dynamo; Battery Charger; LiFePO4

I. INTRODUCTION

Today's mobile devices, such as mobile phones, smartphones, GPS navigation and various other multimedia devices, have very low stamina. Unfortunately, there is no significant increase in the capacity of used batteries in portable devices. This is a major drawback if people want to travel somewhere where it is not possible to charge the device. The simplest solution is to make an extension power supply with an external battery and a voltage stabilizer. With such a device, the time of power on navigation can be extended approximately twice. This is still inadequate in some situations, for example, if we take a day trip or a longer sightseeing tour on bicycles. Therefore, it is necessary to choose another solution. The aim of this paper is to suggest a suitable way of charging mobile devices by using a bicycle. A Mobile phone that is used on a bicycle has large power consumption. This is because a mobile must have turn on the backlight of the display all the time. It is also necessary to continuously receive the GPS signal and have a connection to the internet because it is necessary to route the way.

Photovoltaic panels could solve the problem with lack of energy. Nevertheless, the panel must have a large size to be able to efficiently charge our equipment, and this source is unreliable in our climatic conditions. Therefore, dynamo-wheel charging seems to be optimal solution, which also charges an additional battery, so that the device can be powered even when the wheel does not move.

In the first part, there are discussed various options for charging the battery on a bicycle and described requirements for this device. The second part describes in detail used blocks that were used in the engagement and the reasons of their choice. In the last section, there are described the final parameters of the charger.

II. REQUIREMENTS FOR DEVICE FUNCTIONALITY

It was crucial for propose of the device to meet the following requirements.

A. Automatically switch on charging at defined speed

It is desirable that the device turns on charging according to a certain speed. Ideally automatically after exceeding the set speed. This is useful while driving uphill, when is not suitable, that the dynamo brakes the wheel.

B. Possibility to supply from other energy sources

It was required for equipment that device can be charged by using the bicycle or using other energy sources. For example, a car outlet, AC (Alternating Current) and DC (Direct Current) wall charger. To ease of use, the polarity of the source cannot be conditioned.

C. Protection against destructions

From the measured parameters of the dynamo, it was seen if it is unloaded, then voltage linearly increases with increasing driving speed. Therefore, it was necessary to develop an OVP (Over-Voltage Protection) to prevent damage the circuit under unusual use. Dynamo parameters are described in detail in Chapter 4 A.

III. VARIANTS OF SYSTEM SOLUTIONS

We have several solutions of powering a mobile phone where we need unlimited energy.

A. Photovoltaic panels

If we wanted to use a photovoltaic panel, that would supply the phone by 5 W, it would have dimensions at least 500 cm^2. This is impossible to use on a bicycle. Moreover, the panel's nominal power would only be available at maximum sunshine.

B. Bottle and hub dynamo

In earlier times, a bottle dynamo was used to power light on a bicycle. That dynamo is mechanically rotated due to tire contact. This solution has a disadvantage, there is a high friction, which slows down the bicycle. Such dynamos also do not have high performance. For our purposes, it is better to use a dynamo that is integrated into the front wheel hub. See Figure 1. Such a dynamo has less friction and better power.

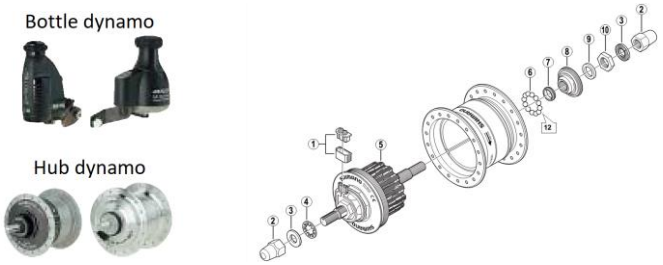


Figure 1 a) Types of dynamo b) Detail of hub dynamo

The hub dynamo parameters are described in detail in Chapter 2 A. The electronics circuits must be connected to the dynamo. In our case there are two main options:

1) Easy solution

Device with internal battery charger and 5V stabilizer. The easiest for engagement, size and price. From the measured characteristics of the dynamo, it is obvious that the dynamo with a full load does not exceed the voltage of 10 V. This is ideal for us. However, it would be necessary to constantly charge the cell phone with a high current (for example, when the battery in the phone is discharged). If the battery in the phone is charged as well as the internal battery, the voltage will increase linearly with the speed of the wheel. This would destroy the charging circuit. The simplified block diagram can be seen in Figure 2.

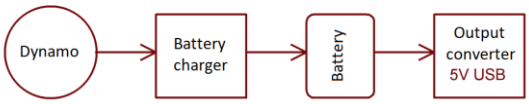


Figure 2 Block diagram of basic bicycle charger for phone.

2) More complex solution

This device contains: dynamo, overvoltage protection, input high voltage reducing converter, internal battery charger, a stabilizer to 5V. This is the best possible solution that is resistant to high speeds and to the failure to connect your mobile phone to the charger. If the batteries are already charged, the current taken from the dynamo is dropped and the dynamo is free from the load. Another advantage is the choice of the speed at which the charging is switched on. The disadvantage is that the entire input block must be designed for high voltage. In this project, a 100-volt threshold was selected, which corresponds to a speed of 80 km/h. An input rectifier, a filter capacitor and an input reducing circuit must be dimensioned to this voltage. In addition, an LED is provided to indicate that the 70 V has been exceeded for increased protection. Transil will interfere if the voltage continues in the rise to 85 V. Transil is also included in the circuit to eliminate voltage peaks. The simplified block diagram can be seen in Figure 3.

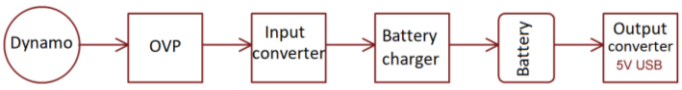


Figure 3 Block diagram of advanced bicycle charger for phone.

IV. Solution of Individual Parts

A. Power source

A hub dynamo is a small electrical generator built into the hub of a bicycle wheel. Let's note that the hub dynamo is not actually a dynamo, which creates DC, but a generator that creates AC. Proper labeling should be the alternator. Hub dynamo in the front wheel was selected due to good power efficiency. Specifically the Shimano DH-3N30. Consequently, it was necessary to measure the voltage characteristics. Under the load and without the load. This is in Figure 4 and Figure 5.

The equation (1) was calculated from the chart in Figure 5. By this equation can be determined, what the voltage will be at given speed, without load.

$$y = 0.8067x * 1.1004 \quad (1)$$

Where y represents the voltage and x represent the velocity.

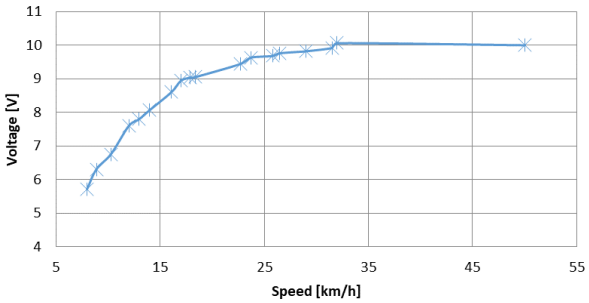


Figure 4 Voltage from dynamo under load.

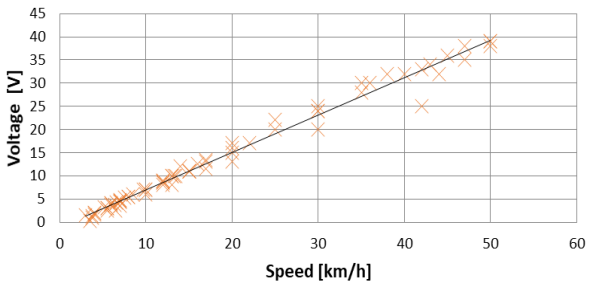


Figure 5 Voltage from dynamo without load.

The oscilloscope, which is connected to the computer, was used to measure the dynamo. Screenshot of an oscilloscope with the voltage at a speed of 10 km/h, without load, is shown in Figure 6.

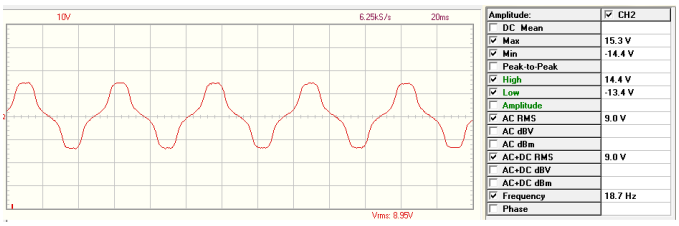


Figure 6 Output voltage at speed 10 km/h without load.

B. Input block

Electronic circuits that suitably adjust the voltage for the battery charger must be connected after dynamo.

1) Rectifier

Hub dynamo produces AC voltage, it is necessary to rectify this voltage. When choosing a rectifier, it is ideal to use two-way rectification by using a Graetz bridge. The most important parameter of the rectifier, in our case, is the voltage drop in the forward direction [2]. The basic problem occurs in slow speed. The low voltage from the generator is reduced by about 1.2 volts, due to the voltage drop on the rectifier. The decrease of voltage would represent approximately 20 % of the voltage loss. At higher speeds, the loss would not be so high. It was also necessary for the rectifier to handle a voltage of at least 100 V and a current of about 1 A.

Because of required low voltage drop it was necessary to select a rectifier that can manage much more current than we need. Rectifiers to larger current are characterized by a lower voltage drop at the same current load as the rectifiers to a smaller current. The most suitable rectifier is RS603, its voltage drop is 0.6 V.

2) Over voltage protection

To protect sensitive electronic components, an overvoltage protection has to be connected to the input. Most suitable protection is provided by the transil, which is mainly designed for large impulse currents. In the case of a high overload, it is not interrupted like the fuse, but is shorted. Thus protecting connected circuits from overvoltage. Transil with a peak power of 1.5 kW can absorb 6.5 watts in permanent mode, this value is sufficient to permanently protect the electronics from destruction. The transil is activated after exceeding the 85-volt limit. This is an enough reserve for the input converter, that has a maximum input voltage 100 V.

3) Voltage converter

A circuit that can operate at a voltage of 100 V has been selected as an input reducing converter. It would be ideal for the circuit that can operate from 6V with an adjustable low input voltage limit. Only National Semiconductor LM5116 meets required parameters. The LM5116 can work from 6 V to 100 V at the output voltage of 5.5 V, which is required as the supply voltage for the battery charger. A minimum switch on voltage can be changed by using the reference. Thanks to this, the circuit is disconnected at speeds up to 8 km/h and the charger does not burden the cyclist.

The theoretical efficiency of the converter is 93 %, if the circuit is loaded and the dynamo has a voltage about 10 V. With increasing voltage, the efficiency drops dramatically. With a 53 V input voltage, the theoretical efficiency is only 43 %. The real efficiency of the convertor is 80 % for the standard voltage usage. The main advantages of the circuit are robust 3.5 A peak output, programmable current protection, undervoltage shutdown and shutdown when overheating.

C. Battery + Battery charger

Nowadays most of batteries are used from 2 large families, namely NICD, NIMH and in the second group are Li-Ion, Li-Pol, LiFePO4. The NICD and NIMH batteries are used as a source of energy in several types of AA batteries, such as AAA, AA, D. Their main disadvantage is memory effect, low capacity and low cell voltage. Lithium-based batteries are used as built-in rechargeable batteries in mobile devices such as mobile phones, tablets, laptops. Due to the largest capacity per liter and almost unlimited number of cycles, a Lithium phosphate battery, which is referred to as LiFePO4, has been selected. These batteries are used by GM in their electric cars. The number of charging cycles is plotted in Figure 7. At room temperature and charging and discharging current 2 A, the battery capacity is still 95 % over 1000 cycles. The ANR26650 battery that has a capacity of 2600 mAh and an operating temperature range of -30 °C to 55 °C has been selected [3].

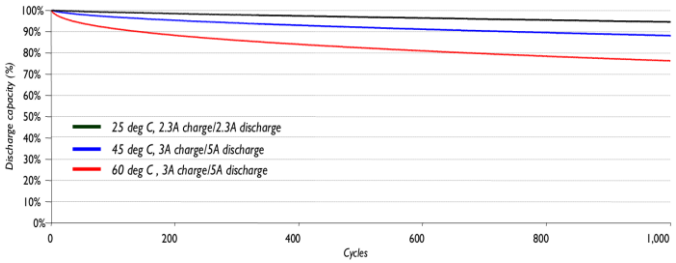


Figure 7 Cycle life performance [4].

The MCP73123 from Microchip was chosen as an integrated circuit for charging the internal battery. The maximum input voltage is 18 V and output current up to 1.1 A. Efficiency of the charging circuit is greater than 90 %.

D. Control panel

The device has a control panel that is attached to the handlebars. There is the option of manually switch off the charging, indicating a 60 km/h overrun and measuring the remaining capacity of the internal battery. A voltage comparator was used to measure capacity. Three LEDs realize the remaining energy status. These represent 80, 50 and 20 % of the remaining capacity.

E. Output block

LM3478 was chosen as the output boost converter circuit. It transforms the voltage of 3 V - 3.6 V from the battery to 5 V. This voltage is needed as the power supply for USB. The circuit is set to operate from 3V, which is a voltage when the battery is almost discharged. This ensures that the internal battery is never fully discharged and it will be turn off automatically. The maximum output current is 1.5 A. This current is enough to quickly charge mobile phone or even smaller tablets. The efficiency of the converter does not change throughout the range of current and voltage and ranges from 85 to 89 %.

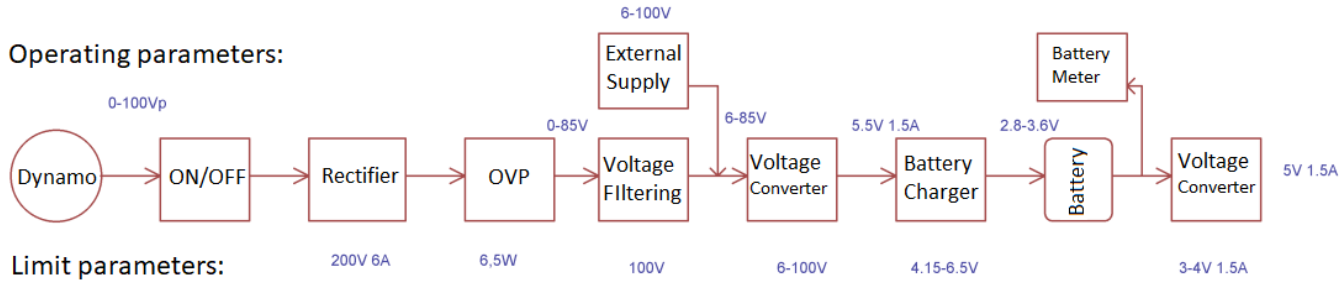


Figure 8 Block diagram of the bicycle charger for mobile phone including the voltage levels.

V. Conclusion

The aim of the work was to propose and produce a mobile phone charger by using a bicycle. The device automatically switches on when the bike goes faster than 8 km/h. So, it does not slow down the wheel when the cyclist goes uphill. The main power source is the hub dynamo in the wheel. The device can be also powered by AC and DC charger, regardless of polarity. The device has the OVP at the input. It does not destroy itself if cyclist drives faster than 80 km/h and the battery is not charging. The theoretical efficiency of the entire charger at maximum load is 73 %, with increasing input voltage the efficiency decreases. The maximum output current is 1.5 A. Measured efficiency at input voltage 10 V and output current 1 A is 60 %.

Acknowledgment

The research in this paper was financed by the Czech Ministry of Education in the frame of the National Sustainability Program under grant LO1401 and by the Brno University of Technology Internal Grant Agency under project no. FEKT-S-17-4426. The infrastructure of the SIX Center was used.

References

[1] Shimano, Datasheets DH-3N20 [online]. Available from: http://cycle.shimano-eu.com/media/techdocs/content/cycle/EV/bikecomponents/DH/EV-DH-3N20-2602A_v1_m56577569830677489.pdf

[2] S. Mahalingam and B. J. Baliga, "A low forward drop high voltage trench MOS barrier Schottky rectifier with linearly graded doping profile," Power Semiconductor Devices and ICs, 1998. ISPSD 98. Proceedings of the 10th International Symposium on, Kyoto, 1998, pp. 187-190. doi: 10.1109/ISPSD.1998.702665

[3] A. Wang, X. Jin, Y. Li and N. Li, "LiFePO4 battery modeling and SOC estimation algorithm," 2017 29th Chinese Control And Decision Conference (CCDC), Chongqing, 2017, pp. 7574-7578. doi: 10.1109/CCDC.2017.7978558

[4] A123, product page 26650 [online]. Available from: http://www.a123systems.com/lithium-ion-cells-26650-cylindrical-cell.html

Application of SWT for Heart Rate Monitoring Using Smartphone Camera

Andrea Němcová, Martin Vítek
Department of Biomedical Engineering
Brno University of Technology
Brno, Czech Republic
andrea.nemcova@phd.feec.vutbr.cz, vitek@feec.vutbr.cz

***Abstract*—Monitoring of heart rate using smartphone is very popular nowadays. There exist some applications for smartphone with usually unknown accuracy. But there is no certified application for this purpose. The article introduces new method for heart rate monitoring which uses smartphone´s built-in camera and LED light. The algorithm is based on extraction of photoplethysmogram and use of stationary wavelet transform. 52 various wavelets were tested in terms of their suitability for heart rate monitoring. The accuracy is compared with heart rate estimated from electrocardiogram. In the range of 55-111 bpm, the heart rate can be estimated with mean error of 0.4093 bpm and maximum error of 0.8458 bpm. The recommended settings of the algorithm is to use green photoplethysmogram (green color channel of the video), provide stationary wavelet transform with 6-level decomposition, rbio2.4 wavelet, and calculate heart rate using PPG peaks' distance. Data of lower and higher heart rate were tested as well.**

Keywords—heart rate, pulse rate, HR, wavelet transform, WT, smartphone, camera, Faros, ECG, photoplethysmogram, PPG

I. Introduction

It is estimated that in 2017 there are 2.32 billions users of smartphone worldwide [1] and this number grows every day. For this reason, the smartphone is a device with big potential in many areas from telecommunication to healthcare. The other reason for it´s wide use is the fact that people usually have the smartphone with them 24/7.

In healthcare area, the smartphone can monitor many health parameters from heart rate (HR) through oxygen saturation to blood pressure [2] without any additional device. HR can be estimated using three different built-in sensors: camera [3], [4], accelerometer [5], [6], and microphone [7], [8]. HR can be measured in a contact (usually from finger) [3] and non-contact (usually from face) [4] way.

There also exist mobile applications for heart rate monitoring using smartphone, which are accessible to the general public. Unfortunately, none of these applications is certified and majority of them does not offer any information about accuracy or used method. It depends on the users, whether or how much they believe the measured values. For nonexperts, moreover without any reference, it is very difficult decision and it can lead to misinterpretation. There also exist articles and conference papers, which usually describe the method in detail but do not offer the application for smartphone.

Physiological range of rest heart rate is between 60 beats per minute (bpm) and 100 bpm [9]. For well-trained people, HR can fall down to 35 bpm. During and after physical activity, HR is normally increasing above 100 bpm. At least 89% of people between age of 12 to 80 and over have resting HR between 56 bpm and 108 bpm [10].

II. Data and Methods

A. Data

For the testing purposes, we sensed the videos using the camera of the smartphone Honor 7 Lite. The camera was focused at infinity, the white balance was set to the tungsten mode and video resolution was set to 1280×720 px. The frame rate (FR) was set automatically on approximately 29.5 Hz (it fluctuates a little bit video to video). Measured subject lightly placed their finger on the rear camera to cover both the camera and the built-in LED. The LED was set on. The signal length is about 20 s. Altogether 10 signals with physiological HR were sensed. Also 4 signals with higher or lower HR were recorded.

As the golden standard for HR estimation is electrocardiogram (ECG), the reference ECG (Fig. 1 d)) was recorded simultaneously with videos. For this purpose, the mobile device for ECG monitoring eMotion Faros 180° was used. The sampling frequency was set to 250 Hz. In the ECG

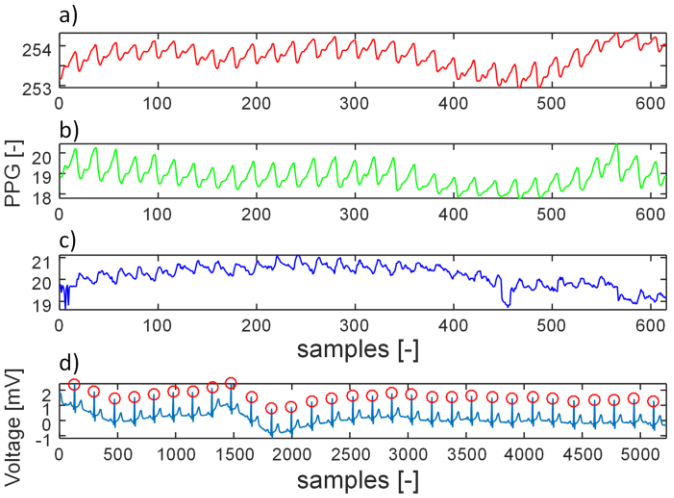


Fig. 1. PPG signals from red, green and blue video channels, and ECG signals with detected R waves (red circles) as a reference for HR.

signal the R waves (Fig. 1 d)) were detected using the QRS detector [11]. For HR estimation, also certified pulse oxymeters can be used. The accuracy of pulse oxymeter Nonin Onyx II 9560 is ±3 bpm.

B. *Preprocessing – photoplethysmogram*

The recorded video was divided into frames and each frame was divided in three RGB (red, green, blue) components. For each frame and each RGB component the mean value of the image was calculated. This way, we created three (red, green, blue) photoplethysmograms (PPG). Red, green and blue PPGs are shown in Fig. 1 a), b), c), respectively.

C. *Wavelet Transform*

For the HR estimation, the stationary wavelet transform (SWT) was used. The decomposition level was empirically set to 6. Altogether 52 wavelets from various wavelet families were tested in terms of their suitability for HR estimation. We tested Haar wavelet, Daubechies (db) wavelets, biorthogonal (bior) and reverse biorthogonal (rbio), coiflets (coif), symlets (sym), and Mayer wavelet (dmey).

D. *Heart rate estimation*

After the SWT decomposition, we further used the coefficients of the 4th highest frequency band (Fig. 4, blue) – in case of physiological HR. From the 4th band, HR in range of about 55 bpm to 111 bpm can be estimated. In this band, peaks were found and then HR was calculated in two ways. The first one (HR_1, equation (1)) counts the peaks in the time window and the second one (HR_2, equation (2)) calculates the distance between adjacent peaks.

$$HR_1(bpm) = \frac{NP*FR}{N} * 60, \quad (1)$$

$$HR_2(bpm) = \frac{FR}{\overline{PP}} * 60, \quad (2)$$

where NP is the number of peaks, N is the number of samples of the 4th frequency band, FR is the frame rate, and $\overline{PP}$ is the average distance of adjacent peaks.

In case of HR > 111 or HR < 55, different frequency band should be used to reach as high accuracy as possible. For higher HRs, the 3rd highest frequency band should be used, where HR between 111 bpm and 221 bpm can be estimated. For lower HRs, the 5th frequency band should be used and HR from 28 bpm to 55 bpm can be estimated in this band.

III. Results

A. *Physiological HR (55-111 bpm)*

HR estimated using particular wavelets was compared to the reference and then the percentage quality was calculated. This was performed for all RGB channels, both methods of HR estimation and all 10 signals altogether. The dependence of quality on the used wavelet is shown in Fig. 2. Two levels of tolerance were used for calculation of quality – 0 bpm (zero tolerance, blue color) and 1 bpm (green color). The highest quality was reached using the rbio2.4 wavelet for both levels of tolerance. The other suitable wavelets are from the reverse orthogonal wavelet family rbio2.2, rbio2.6, rbio2.8, rbio4.4, rbio6.8, and from the symlet wavelet family sym6 and sym8.

Then the methods and RGB channels were tested separately, which is pictured in Fig. 3 a) (tolerance 0 bpm) and b) (tolerance 1 bpm). Those 6 combinations (two methods and three colors) are placed on y-axis. The quality is expressed by colormap – the black color means zero quality and the brighter the color is, the better the quality is up to white color, which expresses the quality of 100%. From Fig. 3 a) it is obvious that the best wavelet rbio2.2 enables to reach up to 70% successfulness (with zero tolerance) of HR estimation (for 10 subjects). Fig. 3 b) shows that if the tolerance is set to 1 bpm, the quality reaches 100% for some wavelets. The most suitable wavelets seem to be rbio2.2, rbio2.4, rbio4.4, and rbio6.8, for which the quality is 100% for the second method of HR estimation and R and G channels. In case of B channel, the quality is never 100%.

TABLE I. Results of testing 2 methods of HR estimation and 3 (RGB) PPGs - mean difference between resulting HR and reference in bpm.

HR1 - R	HR2 - R	HR1 - G	HR2 - G	HR1 - B	HR2 - B
1.0877	0.6960	0.8103	0.4912	6.7478	6.8755

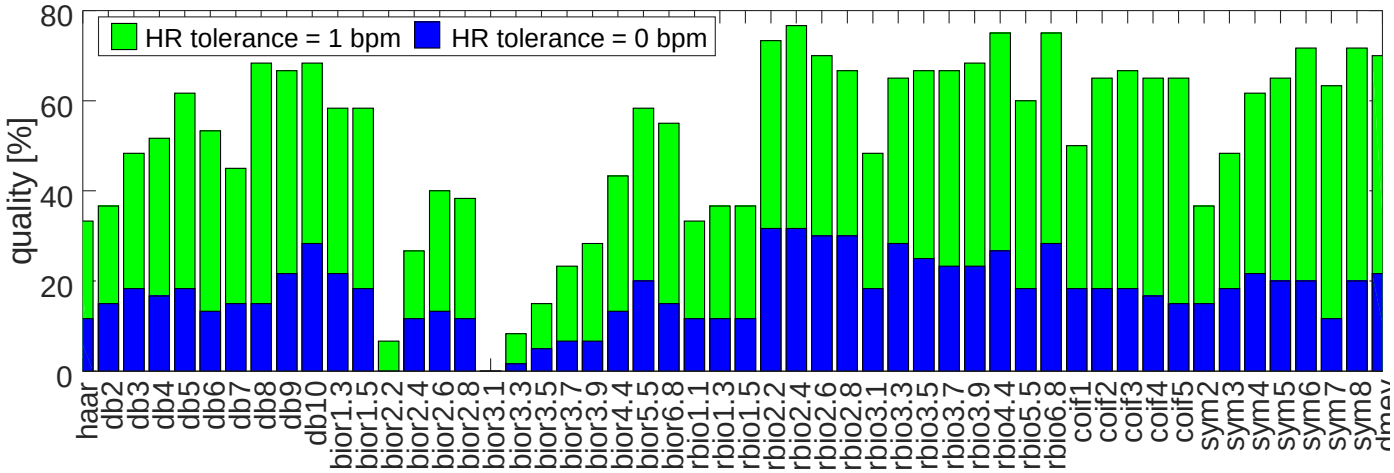


Fig. 2. Comparison of various wavelets in terms of HR estimation (percentage quality).

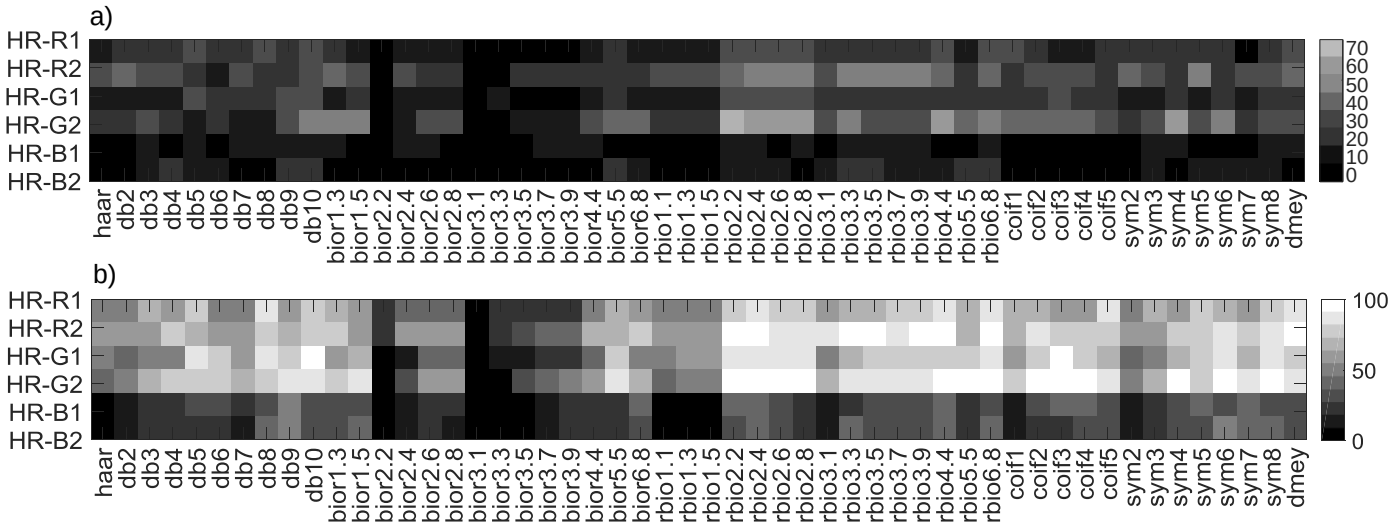


Fig. 3. The results (percentage successfulness of HR estimation) of testing two methods of HR estimation, three (RGB) channels, and 52 wavelets.

The brightest horizontal bands in Fig. 3 mean that these methods – HR-R2 and HR-G2 are more suitable for HR estimation than the other ones. This fact is moreover supported by the Table I (in the test, only the best wavelets, which are mentioned below, were used).

TABLE II. RESULTS OF TESTING 8 BEST WAVELETS – MEAN DIFFERENCE BETWEEN RESULTING HR AND REFERENCE IN BPM.

rbio2.2	rbio2.4	rbio2.6	rbio2.8	rbio4.4	rbio6.8	sym6	sym8
0.4377	0.4093	0.4203	0.4314	0.4908	0.5779	0.5727	0.5898

TABLE III. RESULTS OF TESTING THE BEST SETTINGS ON 10 SIGNALS.

wavelet	mean	median	max	min
rbio2.4	0.4093	0.3667	0.8458	0.0491

Further testing focused only on the best wavelets chosen previously: rbio2.2, rbio2.4, rbio2.6, rbio2.8, rbio4.4, rbio6.8, sym6, and sym8, and the second method of HR estimation from green channel. According to the Table II as well as Fig. 2 and Fig. 3 b), the most suitable wavelet is rbio2.4. For the best wavelet rbio2.4 mean, median, maximum and minimum differences were calculated as is shown in Table III.

B. Higher or lower HR (28-55 bpm and 111-221 bpm)

In case of higher HR, the algorithm is very similar to the one mention above (for physiological HR) except for use of 3rd (Fig. 4, red) frequency band of SWT instead of the 4th band. This variant was tested on 3 signals with mean error of 0.5155 bpm and maximum error of 0.8162 bpm. These results are comparable with the variant for physiological HR.

If HR is lower than 55 bpm, for HR estimation the 5th frequency band would be used, theoretically. It is quite difficult to find person with lower HR, therefore we tested the assumption on one signal with reference HR of 53 bpm, which is very close to the bound (55 bpm). The SWT decomposition is shown in Fig. 4 (black). In this picture we can see that the HR information manifests in the 4th and 5th frequency bands. But it is not clear

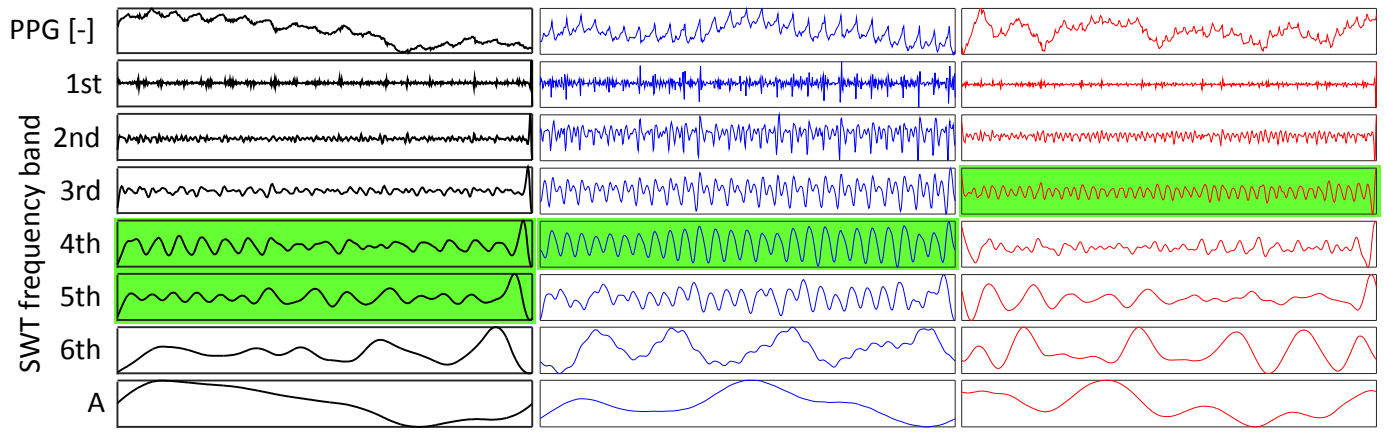


Fig. 4. Green PPG and SWT frequency bands for physiological HR (blue), lower HR (black), and higher HR (red). Frequency bands from which the HR is estimated is highlighted by green color.

whether to use the 4th or the 5th band for HR estimation. In both the 4th and the 5th frequency bands, the HR was calculated and the results are 58.4159 bpm and 43.3811 bpm, respectively.

C. Comparison with other methods

In this part, we compare our algorithm (for physiological range of HR) with other methods. The best method of HR estimation from [12] is based on green PPG filtration and detection of peaks. For HR estimation, equation (2) is used. This method reaches mean rounded error of 0 bpm and maximum rounded error of 2 bpm. If we round our results of HR, the mean error is 0 bpm and maximum error is 1 bpm. Microphone-based HR monitoring has mean error of 0.92 bpm [8]. The method of HR estimation using accelerometer has mean error of 0.17 bpm [6]. Both these methods used as a reference pulse oxymeter, which can be less accurate than HR estimation from ECG. The accuracy of method which estimates HR from camera [3] is 98.02%. The accuracy of our method calculated according to the equation published in [3] is 99.4253%. Our method´s accuracy also exceeds the accuracy of methods of HR estimation from user face, which are 98.92% and 98.53% [4].

IV. Discussion

The previous facts result in requirement to decide from which frequency band the HR will be estimated. For this purpose and to automatize the whole process of HR estimation we need some rule that precede the SWT. This rule can be based on some simple method for HR estimation, such as those mentioned in [12] and briefly described above. The easier variant of rule can be based on users' decision whether they are well-trained athletes, in normal condition, or after some physical performance (e.g. sport or stairs). One more variant may be possible: to directly quantify the suitability of each SWT frequency band for HR estimation.

The advantage of the method based on SWT is: robustness (it can estimate HR from signals of worse quality more accurately than simpler algorithms), easier setting of peak detection (due to the character of signal in relevant frequency band), and accuracy (at least in physiological range of HR excluding bound values of HR). Also the use of built-in camera has advantages: the measurement can be provided in any environment (it is not necessary to measure in quiet place as in case of use of microphone [8]), and in any position of person (the person does not have to lie as is recommended in method, which uses accelerometer [6]). The disadvantage is higher computational demand due to use of SWT. Another disadvantage seems to be lower accuracy in small segments around the borders (HR information is divided into two frequency bands).

V. Conclusion

HR was estimated using smartphone built-in camera and LED. For this purpose, extraction of RGB PPGs was done and the PPGs were decomposed using the SWT. Application of SWT on this problem is the innovation of this work. The level of decomposition was empirically set to 6. Altogether 52 wavelets were tested in terms of their suitability for HR estimation using two methods and three (RGB) PPGs. Each setting was tested on 10 signals with physiological value of HR. The error of the methods was calculated as the difference between the resultant HR and the reference HR estimated from ECG. The most suitable wavelet is rbio2.4 applied on the green PPG with the second method of HR estimation (using peak distance). In this case, mean error of HR is 0.4093 bpm and maximum error is 0.8458 bpm. The accuracy in physiological range of HR seems to be better than accuracy of pulse oxymeter. The physiological range of HR is sufficient for healthy teenagers and adults to measure their resting HR.

In case of higher (> 111 bpm) and lower (< 55) frequencies, the algorithm should be adjusted in terms of use of the suitable SWT frequency band. Estimation of higher HR has similar accuracy as for physiological HR. The accuracy of HR estimation in the border area is low, because the HR information is present in two SWT frequency bands. This should be adjusted as well in the future.

The results of application of SWT for HR monitoring using smartphone built-in camera are encouraging and the method has a great potential. Nevertheless, several issues should be solved to reach maximal accuracy.

Acknowledgment

This work was supported by grant project FEKT-S-17-4487.

References

[1] "Statista: The Statistics Portal", 2017. [Online]. Available: https://www.statista.com/. [Accessed: 17-Jul.-2017].

[2] F. Lamonaca, G. Polimeni, K. Barbé, and D. Grimaldi, "Health parameters monitoring by smartphone for quality of life improvement", *Measurement*, vol. 73, no. September 2015, pp. 82-94, 2015.

[3] S. A. Siddiqui, Y. Zhang, Z. Feng, and A. Kos, "A Pulse Rate Estimation Algorithm Using PPG and Smartphone Camera", *Journal of Medical Systems*, vol. 40, no. 5, 2016.

[4] S. Kwon, H. Kim, and K. S. Park, "Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone", in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012, pp. 2174-2177.

[5] C. Gavriel, K. H. Parker, and A. A. Faisal, "Smartphone as an ultra-low cost medical tricorder for real-time cardiological measurements via ballistocardiography", in *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 2015.

[6] J. Víteček and A. Němcová, "Heart rate measurement using smartphone accelerometer", in *Proceedings of the 23rd Conference STUDENT EEICT 2017*, 2017, pp. 21-23.

[7] A. Misra, R. Banerjee, and A. D. Choudhury, "Novel peak detection to estimate HRV using smartphone audio", in *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 2015.

[8] J. Víteček, O. Guryča, and A. Němcová, "Měření tepové frekvence s využitím chytrého telefonu", *Elektrorevue*, vol. 19, no. 2, pp. 66-72, 2017.

[9] M. S. Thaler, *The Only ECG Book You'll Ever Need*, Fifth Edition. USA: Lippincott Williams & Wilkins, 2007.

[10] Y. Ostchega, K. S. Porter, et al., "Resting Pulse Rate Reference Data for Children, Adolescents, and Adults: United States, 1999–2008", *National Health Statistics Reports*, no. 41, pp. 1-17, 2011.

[11] L. Maršánová, "Detection of P, QRS and T Components of ECG Using Phasor Transform", in *Sborník příspěvků studentské konference Blansko*, 2016, pp. 55-58.

[12] A. Němcová, R. Smíšek, et al., "Využití chytrého telefonu pro monitorování zdravotního stavu uživatele", in *Trendy v biomedicínském inženýrství*, 2017. Accepted for publication.

A Survey on Web Based User Interface Technologies

Marek Novák, Lukáš Janík
Department of Radio Electronics
Faculty of Electrical Engineering and Communication
Brno University of Technology
Technická 12
616 00 Brno (Czech Republic)
Email: xnovak0m@vutbr.cz, xjanik12@vutbr.cz

***Abstract*—This paper introduces web applications in the role of graphical user interface. Two particular protocols of communication between the server and the client are discussed - Common Gateway Interface (CGI) and Websocket. The communication principle is explained, with emphasis on system performance, memory footprint and data availability. The article lists few well-known web user interface frameworks with client side only, client side and server side and packaging architectures. The article mentions realized applications of web UI in embedded devices.**

I. Introduction

The Internet, a web of pages became common information source. The massive use of technologies, narrowly related to web pages - e.g. hypertext transfer protocol (HTTP), hypertext markup language (HTML), cascade style sheets (CSS), JavaScript (JS) caused, that these are being used even for purposes which they were originally not designated for. Thanks to their learning curve and fast development, they are often chosen as the base of standalone programs, running on the desktop e.g. Atom.io - Text editor (IDE), written in JS, GUI Layout in CSS. This approach also caused that many services (managed by desktop clients) gained their counterpart in cloud - e.g. Microsoft Office 365.

II. Advantages of Web UI

Web UI presents a way of implementation of cross platform applications (desktop, tablet, mobile phones). The look of the user interface (UI) is not limited by themes of upper graphical library (GTK, .Net Framework) and allows easy customization via cascade style sheet files (CSS).

Embedded devices are often used as a data collectors (sensors). Mostly it is desirable to send these data to remote device, process them and then graphically represent (print) them. The use of web UI allows easily sharing real-time data with more users thanks to off-loading of the embedded device - only data is transferred, the "expensive"computations are moved to the client. The device can also provide complex JavaScript-based user interface, even if it does not contain JavaScript (JS) interpreter. The web application can be self-contained package, without any external program required.

III. Server to Client Communication

A. Common Gateway Interface

The Common Gateway Interface (CGI) [1] is a standard protocol used by web browsers to communicate with console

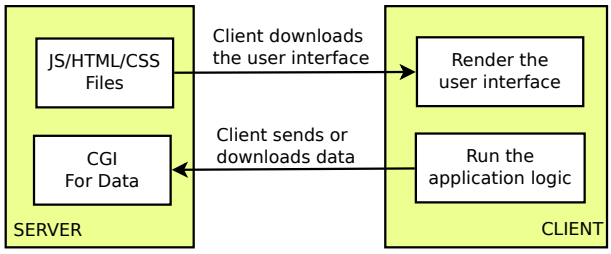


Fig. 1. Web user interface in embedded device

applications (scripts) running on the server. The data from web page are sent to standard input (stdin) of a CGI script; the CGI script responds via standard output back to the web application. CGI scripts thus can be written virtually in any scripting language (Python, Bash, ...) or even binary (compiled C code etc.).

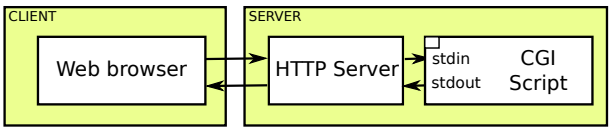


Fig. 2. Block scheme of the communication between web browser and CGI script.

1) CGI Communication sequence: The CGI Communication scheme is illustrated in Fig. 2. The web application sends a POST request on e.g. http://example.org/cgi-bin/clients. The web server on http://example.org will look for CGI script called *clients* and call a *fork*, creating a subprocess with CGI script. The POST data (if any) are passed to the standard input of the CGI script. The script creates a response, which is via standard output passed back via HTTP server to client. The CGI script has to assemble whole response, including the HTTP header.

Listing 1. An example of python CGI script sending a dictionary via JSON - formatted string

```
some_dictionary = {'some_string':'Alpha', \
                   'some_number':123}
# HTTP Header
print('Content-Type:application/json')
# Newline as a terminator
print()
# JSON Data
print(json.dumps(some_dictionary))
```

The CGI is available on most implementations of HTTP server (apache, nginx, lighttpd), even in the LwIP HTTP server, which is available for small microcontrollers.

B. FastCGI

Increasing number of requests causes system to frequently spawn new processes with CGI scripts, increasing total overhead. This led to invention of FastCGI. Instead of creating new process for each request, FastCGI uses persistent processes to handle those requests. This reduces the processing time of each request. The persistent process, however, causes consumption of certain amount of system memory, even if there is no communication.

Listing 2. AN EXAMPLE OF PYTHON FASTCGI SCRIPT

```
# Loop
while (FCGI.accept() >= 0):
  # HTTP Header
  print("Content-type:text/html\n")
  # Newline terminator
  print()
  # Parse request and perform appropriate action
  if (some test):
    do_something()
```

C. Websocket

Another way to communicate between the server and the client is called Websockets [2]. This is a more recent technique than CGI. The principle is that a client establishes an HTTP connection to the server and then asks the server to upgrade this connection to so called Websocket. After this operation, only raw application data can be sent via this newly created tunnel. The connection is kept established all the time, which is the main difference compared to CGI.

One of the protocols used on top of Websockets is Distributed Data Protocol (DPP). This protocol adopts a publish/-subscribe model, which means that a user asks for publication of certain data by sending a request to the server. The other side sends new data to the client anytime the server data are updated. This approach can be considered as a managed way to directly access the database on the server.

D. Comparison

Even if Websocket is a more recent approach for communication between client and server, it does not have only advantages. One of the disadvantages of it is a permanent memory allocation for any active user, which is connected to the server. This can be a serious issue for memory constrained embedded devices. On the other side, Websocket provides a real-time behavior, since the TCP connection does not have to be established again each time a request is done. Another even more important advantage of Websocket is that the server can actively begin application level communication, once the Websocket channel is established. With CGI, a method called long-polling has to be used, which consist in asking the server periodically about a status of an action. This is of course very inefficient. CGI has on the other side the advantage in its simplicity - it is very easy to implement it in embedded devices and it is stateless, since it does not keep any context from one to other user connection.

IV. WEB UI FRAMEWORKS

Frameworks used for Web UI development can be divided into three basic categories: packaging, client and server side and client side only frameworks.

A. Client Side Only

These frameworks are composed both of client-side Javascript and CSS. The CSS files offer some basic styling of the user interface, while the Javascript code is used to dynamically assign different style attributes to the generated user interface. This way, a very responsive UI is created, since no communication with the server is required - the elements on the web page are generated and altered directly by client-side Javascript based on user interactions. In this category, there are also libraries for more advanced data representation such as interactive plots and charts.

These are one of the well-known client-side UI frameworks:

- **Bootstrap, SematicUI** Provides simple way to create buttons, checkboxes, modal windows, drop-downs, drop-ups etc.
- **ChartJS, CanvasJS** Libraries providing responsive plotting and charting functions.

B. Client and Server Side

This category contains somewhat more complex frameworks and libraries. The aim is to provide a unified approach in creation of web based interface, which allows creation of ready-to-use components. Some of them require extensive computation power, which is one of the reasons why sometimes a portion of the component is running server-side. Another reason for server-side presence is the need to store, filter and access user application or user interface related data. Some frameworks are even created specifically for some type of database, such as MeteorJS, which is to be used with MongoDB database.

In this category, there are also high level frameworks allowing easy and fast creation of user interface, which can be accessed online by multiple users.

These are some examples of specified frameworks:

- **MeteorJS** Meteor enables creation of applications and user interfaces, which keep synchronized, when multiple users interact with it simultaneously. Both server and client side is written in Javascript and MongoDB database is used to store data in JSON (Javascript Object Notation) format. Each client is maintaining a websocket connection to so-called DPP (Distributed Data Protocol) server, which allows interactions on data used by the user, restriction of access or sharing data between multiple users etc.
- **Node-Red** Node-Red is a data flow based programming tool, which enables creation of server and client side application in a graphical way. This is achieved by interconnecting different GUI nodes by virtual wires.

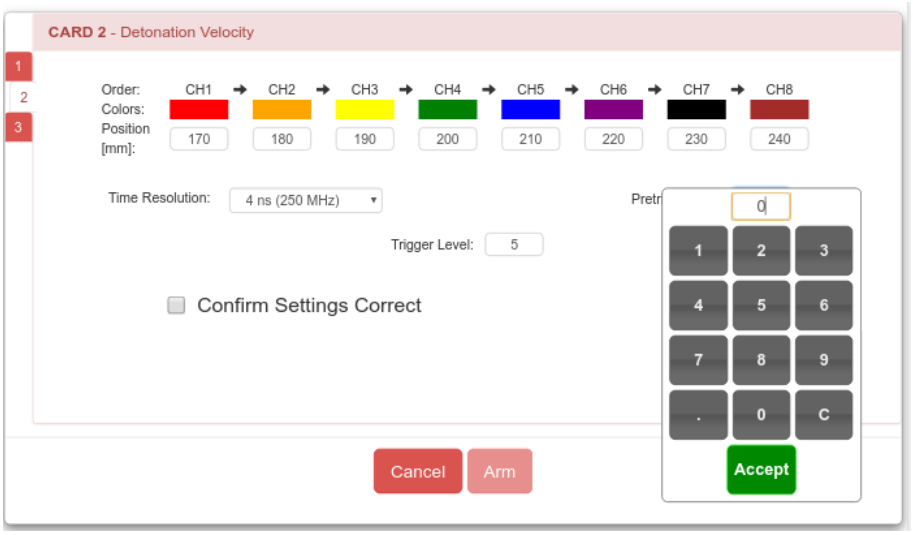


Fig. 3. An example of application of Web user interface in embedded detonation velocity measurement tool

- **ExpressJS, TotalJS** Alternatives to MeteorJS, which do not intergrate with a specific database and do not provide system for software components sharing.

C. Packaging Frameworks

In this section, there is currently only a few frameworks, which share a common history. The most known is called Cordova or PhoneGap. It enables to create an Android, iOS or Windows Mobile application by packaging an existing web page containing Javascript, HTML and CSS. No web browser is required to run these application, so they are completely standalone and self-contained. For accessing so called native functions of the target platform, a special application programming interface (API) is created. Thanks to it, the application can access device file system, camera, microphone etc. It can be therefore used to create a full-featured application with almost no limitations.

V. Conclusion

The implementation of standalone web user interface brings new view to the utilization of web technologies, with their specific benefits and disadvantages. There are few grown-up frameworks, powering either enterprise or community cloud services.

Whereas a standalone web application presents a possible way of implementation of multi-platform desktop and mobile applications, the client-server setup presents very suitable way of implementation of a remote user interface for embedded devices, allowing load balancing between the embedded device and the client device. Special care needs to be taken in cases of realtime applications. The common pitfall is the slightly different support of JS features across web browsers.

The web GUIs were successfully implemented e.g. in the administration of free space optical (FSO) transceiver (running on Xilinx Microblaze, Virtex-5) or various embedded measurement devices (running on Atmel Aria G25). The universal Web-based interface for remote access was patented in 2003 [3]. Later, in 2005, the client/server application for offline usage was patented [4].

Acknowledgment

Research described in this paper was financed by Czech Ministry of Education in frame of National Sustainability Program under grant LO1401. For research, infrastructure of the SIX Center was used.

References

[1] Robinson, D. and K. Coar, *The Common Gateway Interface (CGI) Version 1.1*, RFC 3875, DOI 10.17487/RFC3875, October 2004, https://www.rfc-editor.org/info/rfc3875.

[2] Fette, I. and A. Melnikov, *The WebSocket Protocol*, RFC 6455, DOI 10.17487/RFC6455, December 2011, https://www.rfc-editor.org/info/rfc6455.

[3] Davis, J, R. Umstattd and B. Frew. *Web-based interface for building management systems.* US 20030233432 A1.

[4] Davidson, J. G., R. J. Tompkins, N. R. Folkman and R. A. Landsman *Client/server web application architectures for offline usage, data structures, and related methods*. US 7594003 B2.

Energy Harvesting from Human Movement

Radka Pavelková, David Vala

Faculty of Electrical Engineering and Computer Science
VŠB - Technical University of Ostrava, Czech Republic
Radka.pavelkova.st@vsb.cz, david.vala@vsb.cz

***Abstract*—The article informs about the possibility of generating electrical energy from the human movement. Our body can power eg. mobile sensors for applications where the connection to the external sources is impossible. The observed energy comes from human walking when the bearer carries heavy objects on his back. A system includes the useful load, springs and other mainly electrical elements, which helps to get heavy load of backpack to resonance and thus an optimal yield of power generation.**

Keywords—backpack, energy harvesting, motion capture, kinematics

I. Introduction

The most of electrical equipment is powered by the energy from non-renewable sources. Nowadays requirements grow and it is important to invent the devices which are independent and work with low energy consumption. The technologies for getting small amount of the electric energy become more popular. Another point is that they are autonomous systems, where is the need to ensure a continual power supply source with help of batteries. Among the traditional energy sources can be energy extracted from a human body with the use of two methods: based on movement and heat. This article summarized in short, the content of the diploma work focused to the human walking and use of the energy from periodic movement for energy generation. In the beginning was necessary to analysis human movement. For this purpose, was used motion capture technology from XSens Company. The device works on the principle of sensors placed onto moving segments of the body, which read their accurate position in the time and calculates their speed and acceleration. The energy harvesting systems which works on base of human walking is special backpack with generator, which transfer the kinetic/potential energy to the electrical energy. The system must calculate also with spring forces, friction in the linear guidance and loses of energy by generator.

II. Analysis of human movements

As mentioned earlier in the introduction part, firstly we had to determine, how much kinematic energy may be transferred by the human body for the electric energy generation. At particular steps, the center of gravity of the human body copies the path with a typical shape of the circular contour sector. So, the center of gravity makes a movement in the vertical direction up and down, which corresponds with walking. Provided the mass of the rucksack must execute a corresponding movement, the body must hand-over the energy to the load to cover the path. The energy transfer from the body to the load can be calculated as an increment of the potential energy of the load in the change of the height. The energy corresponding, for example, with the use of a load of 20kg, which was tested in the most of previous experiments with this matter, the walking speed of 1.7m/s at the acceleration of 5m/s and the angle of deflection from the balanced position at the step of 30°, the approximate vertical deflection of 5cm, is approximately 11J. To confirm these presumptions, the results of the measurement taken over from the Motion Capture software were used.

III. Simulation in MATLAB Simulink

During designing of the rucksack system, the big role plays the force of the springs, absorption, friction, which rise because of the friction of the linear guidance and power take-off from the rotary generator. Also, the weight of the rucksack itself, respectively the load. It is necessary to consider these variables to reach the maximum electric power at the given frequency of the load movement along the linear guidance.

We could consider the general formula for calculation of the resonance frequency in case that no absorption of the system would occur. It occurs during this experiment, and that is by two ways. Firstly, due to the above mentioned friction of the bearings of the linear guidance, and secondly by taking-off of energy by the rotary generator. That is why there was made a mathematic mode, which also includes the following absorption coefficient b.

This mathematic model is described by the differential formula:

$$x'' = y'' - (b/m) * x' + (b/m) * y' - (k/m) * x + (k/m) * y$$

Where:
m – load mass,
x(t) – time flow of the load position,
y(t) – time flow of the rucksack mechanical structure position,
b – absorption coefficient,
k – springs force.

This differential formula was consequently re-written into the MATLAB Simulink environment and serves for setting of the rucksack parameters, which will cause the resonance movement and calculation with the maximization of the generated power.

IV. Rucksack structure

To get energy from the human body movement, the character of the movement must be also taken into consideration. One of the aspects of this movement is that it can evolve a sufficient speed in the vertical direction in the direct proportion to the walking speed. In case of the movement acceleration, the vertical speed will also be increased. The second important aspect is the fact that this movement cannot act with a big force directly. The most suitable is the use of transmission of the mechanic energy to the electric one by means of the electromagnetic generator. The third aspect is a relative big path of the movements and the space for the placement of the generator. These aspects induce a possibility of using of the linear generators, which does not require linear/rotary movement converters. However, due to its dimensions, it is more suitable to use a converter regarding a rotary generator.

The rucksack itself is a mechanic system - the moving mass suspended on the spring system. The linear movement is converted via the transmission by a toothed belt to the rotary movement. The electromagnetic rotary generator generates electric energy during its movement. The proposal of the mechanics must be sophisticated to ensure achievement of the maximum effectiveness with a variable weight of the load in the rucksack, and at the same time with the lightest structure to avoid loading of the carrier.

The key element of the system is the electronic circuits transforming the output from the generator to the useful voltage level. The circuits include a DC/DC converter with a non-inverted output with a possibility of the input current regulation. There is a profitably used feature that the input voltage can be higher and lower than the output one (it is influenced by the walking speed). The electric energy from the converter is accumulated by an ultra-capacitor. This energy is later used, for example, for the recharging of a battery or other electronic systems.

The electronic system maintains the voltage in the ultra-capacitor under the upper limit and consumes this excess energy, which cannot be more stored. Thanks to it, the circuit works in the automatic mode.

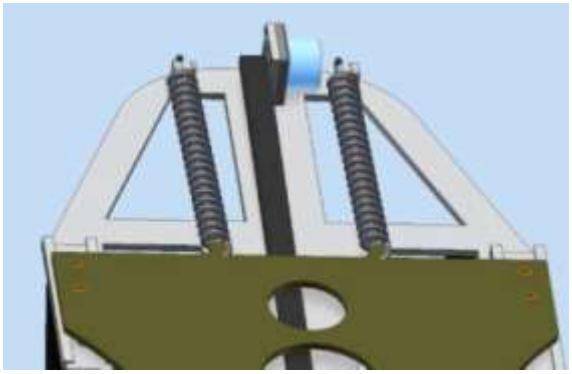


Fig. 1: The view of the 3D model of the springs and generator fixture with the transmission of movement by a toothed belt.

V. Development and realization

For better imagination, an electro-mechanic system was modeled in 3D in the Autodesk Inventor software and afterwards experimentally assembled from the accessible material.

The aim was that the weight of the structure does not exceed 3kg. This was achieved by choosing a suitable material and the cross-sections in the structure, thanks to which the weight of the system was significantly reduced. The mechanical components of the rucksack are made of wood and aluminum.

Fig. 2: The overall view to the electro-mechanic structure of the Energy rucksack.

VI. System energy gain metering

The verification of the features and the energy gain of the system was carried out by several methods. The device was modeled by means of the Motion Capture technology, consequently there was created a model in Matlab Simulink serving for the spring systems adjustment. The assembled real model was observed with simulated results.

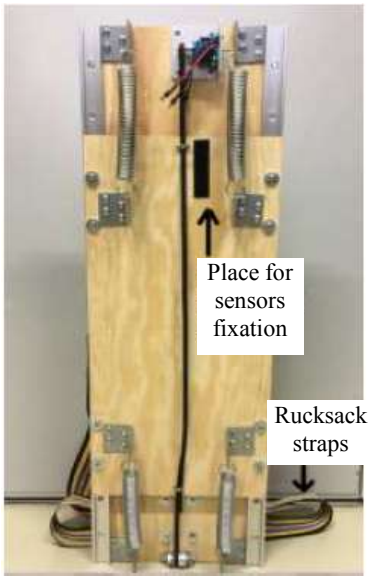


Fig. 3: Experimentally created structure of the EH rucksack.

During its movement, the rucksack copies the movements of the human body. Fig.4 represents the measurement on the electrohydraulic excitation system with the resonance frequency of 3Hz at the test load with the weight m=12kg, when the maximum energy gain of the system was achieved.

This frequency couldn't be achieved by human walking, and it will be necessary to correct the settings of springing and absorption of the system so that it corresponds with the ergonomic parameters of human walking.

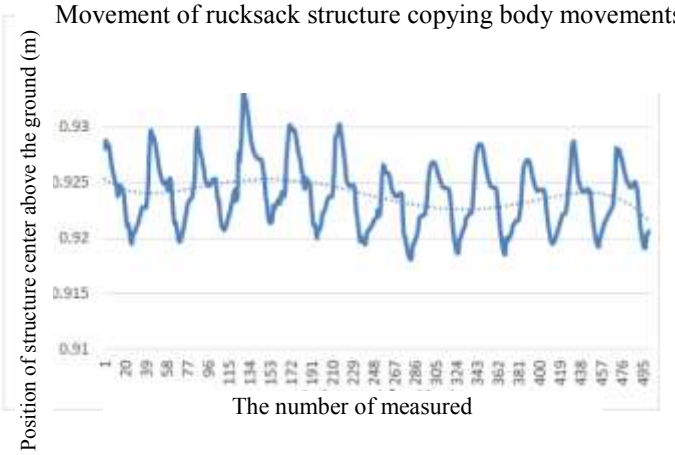


Fig. 4: The graph represents the position of the sensor situated on the stationary structure of the rucksack.

The decline of the amplitude of the weight movement along the linear guidance caused by the friction can be seen on the graph on fig.5. It occurs because of taking-off of the energy by the generator, as well as small friction in the guidance.

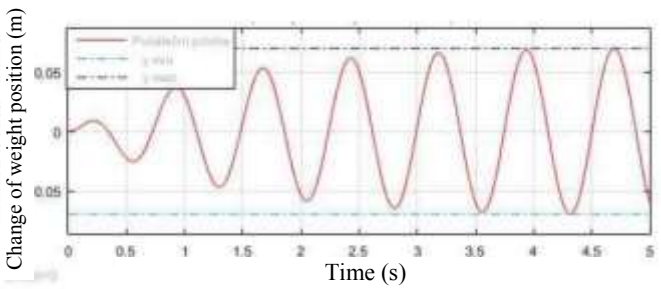


Fig. 5: The chance of the weight position on the springs at resonance.

Fig. 6: Anchorage of the cylinder and the Energy Harvesting device in the vertical position.

VII. Conclusion

The work proposed to create an experimental prototype of the rucksack with the Energy Harvesting system, which should generate electric energy at walking that can be used for powering mobile sensors. To enable designing of this device, I needed theoretical knowledge of the human movement kinematics, the Energy Harvesting technology, as well as possibilities of electric energy conversion.

Before work initiation, we explored already existing devices that use the movement of the human body to generate energy. As a suitable movement, human walking was selected, which is an enough big source for powering small electronics and mobile sensors.

With this knowledge, the human walking kinematics was analyzed by means of a mathematic simulation in the Simulink software and with help of the Motion Capture technology for observation of the human walking parameters by means of sensors, which monitor segments of the body. For this work, the shoulder movement data were used, which correspond with the movement of the human mass point and which are

also copied by the rucksack structure. The rucksack uses the human kinetic energy, and thanks to it shifts the weight along the linear guidance with the amplitude of 6 - 7 cm and the acceleration of 10 - 12 $m \cdot s^{-2}$ at the optimal laboratory conditions, and 4 cm and the acceleration up to 10 $m \cdot s^{-2}$ during a real test at walking.

The rucksack power is variable, as well as the human walking. That's why this experiment was measured on the hydraulic testing device, which can provide frequency and amplitude stability. An average power of the rucksack 2,76 W was measured. The power of the rucksack is limited by its structure, which has the bearing capacity up to 20 kg. During power metering, the rucksack was loaded by the weight just 7.5 kg. By this reason, the second prototype of the rucksack was created in the Inventor software, which preserves the weight less than 3 kg, but at the same time will have far higher bearing capacity about 40 kg. An adequate power increasing is expected.

The electric power of the rucksack depends on many factors. First, the frequency of the movement and the step length are important, also the weight of the selected load, as well as different asymmetry in the articular apparatus of the bearer. The walking speed or the weight of the load must be fit to it to enable comfortable achievement of the frequency near the resonance frequency, thereby the highest excitation and power of the rucksack. The structure of the rucksack must be always optimized for each bearer by reason of the different walking speed, the step length and the load weight. To adjust it, the springs force or the linear guidance length can be changed. There was also experimentally verified that during the lifting phase, this structure brings certain advantages for the bearer, especially a feeling of unloading and consequently decreasing of muscle tiredness.

ACKNOWLEDGMENT

This paper has been elaborated in the framework of the project „Control of technological systems with renewable and alternate energy sources 2017“ SP2017/150 and „A practical training program for staff developing a low-carbon economy in the border region“ reg. no. CZ.11.3.1190.0/0.0/15_005/0000048, supported by European Union in framework program INTERREG V-A Czech republic-Poland.

REFERENCES

1. Zboril, J. (2014) Alternativní zdroj napájení pro mobilní snímače polohy [online]. Ostrava, VSB-TUO Dostupné z: https://dspace.vsb.cz/bitstream/handle/10084/104550/ZBO0011_FMMI_B3922_3902R040_2014.pdf? sequence=1

2. Mastny, P. and coll. (2011) Obnovitelné zdroje elektrické energie. ČVUT Praha, ISBN 978-80-01-04937-2

3. Kaminsky, J., Vrtek, M. (1998) Obnovitelné zdroje energie VŠB-TUO, ISBN 80-7078-445-8

Comparison of Baseline Wandering Removal Algorithms for Automatic Classification of Electrocardiogram

Radovan Smíšek[1,2], Martin Vítek[1], Jana Kolářová[1]

[1]Department of Biomedical Engineering
Brno University of Technology
Brno, Czech Republic
xsmise00@stud.feec.vutbr.cz, vitek@feec.vutbr.cz,
kolarova@feec.vutbr.cz

[2]Institute of Scientific Instruments
The Czech Academy of Sciences
Brno, Czech Republic

***Abstract*—Baseline wandering (sometimes also baseline drift) is a noise that exacerbates the evaluation of ECG and reduces the success of automatic ECG classifiers. In previous works, many methods have been proposed to remove such noise. The aim of this article is to compare commonly used methods. Each signal is in parallel reprocessed by all analyzed methods and then enters an automatic classifier that is able to classify QRS complexes. According to the classification success, the proposed methods were compared. Based on the results of this work, the best method for removing baseline wandering in the ECG is wavelet filtration. The success of the classification is further improved by the combination of wavelet filtration and EEMD. The disadvantage of this combination is its very high computational complexity.**

Keywords—baseline wandering; baseline drift; ECG; automatic classification; EMD; EEMD; wavelet filtration

I. INTRODUCTION

Cardiovascular diseases nowadays represent the most common cause of death in Western countries [1]. Electrocardiography (ECG) is a common cardiological method for detecting cardiac disorders. ECG diagnosis is dependent on the use of computing technologies. One of the advantages of computing technologies is acceleration of diagnostics. This is necessary especially in Holter ECG, which are several-hours or even several-days ECG records of one subject. The cardiologist cannot view the entire record without using the classification software.

The ECG signal can be contaminated by different types of noise, which reduces the classification success. Signal preprocessing is used to increase the success of automatic classification. This work deals with removal of baseline wandering. This type of noise causes that the base axis (x-axis) of the signal moves up and down. Because of this phenomenon, the whole signal is shifted from its normal base.

Baseline wandering is caused by electrochemical processes at electrode-skin transition (up to 0.8 Hz), breathing (up to 0.5 Hz) and slow movements of the subject under investigation (up to 2 Hz) [2]. According to the American Health Association (AHA) the lowest frequency component in the ECG is around 0.05 Hz [3]. Baseline wandering thus affects the useful frequency components of the ECG. The aim is to find a method that eliminates baseline wandering but minimally affects the useful ECG components.

Various methods to remove baseline wandering have been described in the available literature. This work describes and compares 9 methods used to baseline wandering removal.

All methods for removing of baseline wandering are compared by the success of the subsequent automatic classification. The classifier used for classification (with slightly modified input parameters resulting in higher success rates) is described in [4]. Baseline wandering removal has a direct effect on the features used to enter the classifier, such as the area under the QRS and the magnitude of the R wave.

II. METHODS

A. Median filtration

The first tested method is median filtration. A floating window is moving across the signal. The floating window length is 1.3 seconds. The median value in the window is calculated at each shift. Number series of the individual median is the estimated signal trend. This trend is eventually subtracted from the original signal.

This method does not allow the trend estimate at the beginning and at the end of the signal, because it is always necessary to estimate 0.65 s of the signal before and after the analyzed signal segment. The advantage is that this method is very fast and it does not distort signal if the baseline wandering does not occur.

B. Chouchan's method

Chouchan's method was described in [5]. The overall median of the signal is subtracted from each sample in this method. This causes a rough suppression of the baseline

This work was supported by grant project IGA FEKT-S-17-4487.

wandering. Then a fifth order polynomial is fitted to the signal using the least squares method. The estimate of baseline produced this way is then subtracted from the signal. The final step is to subtract the median value in each RR interval from each sample at this RR interval [3], [5].

The advantage of this method is that the method is very fast and it does not distort the signal if the baseline wandering does not occur.

C. Shape-preserving piecewise cubic interpolation

The first step of this method is the detection of isoelectric points. First, the 20 ms section of the lowest standard sample deviation is searched in the interval from 20 to 80 ms before each QRS complex. This section is thus flattest section at a given interval. The isoelectric point is then a sample in this section having a median value of this flattest section.

These isoelectric points are then interpolated with shape-preserving piecewise cubic interpolation to produce a baseline estimate. This estimate is finally subtracted from the signal.

The disadvantage of this method is that if the isoelectric points are not correctly detected, the baseline wandering can be highlighted in comparison to the original signal.

D. Piecewise cubic spline interpolation using not-a-knot end conditions

The first step of this method is the detection of isoelectric points, which is the same as that described in chapter C. These isoelectric points are then interpolated with piecewise cubic spline interpolation using not-a-knot end conditions to produce a baseline estimate. This estimate is finally subtracted from the signal.

This method is one of the most commonly used to remove the baseline wandering from the ECG [3]. The disadvantage of this method is that if the isoelectric points are not correctly detected, the baseline wandering can be highlighted in comparison to the original signal.

E. Empirical mode decomposition

Empirical mode decomposition (EMD) is a method used to process non-stationary signals. EMD is based on interpolation of local maxima and local minima of the signal and subsequent subtraction of the average of interpolated points from the signal. This is how the signal is decomposed into the sum of so-called intrinsic mode functions (IMFs). The last further indivisible IMF characterizes the trend of signal.

In this work, the signal was divided into 10 IMFs and the sum of the last two IMFs was set as a trend of signal. The more recent IMFs are summed up, the more the trend is adapt to faster baseline changes.

IMFs are derived from a signal without the need to provide any signal information. This is an advantage compared to curve interpolation methods where it was necessary to detect specific points of the signal. The disadvantage is that this method does not give good results in the intermittent signal or in the signal smashed by intermittent noise [6]. Further, EMD distorts the trend at the beginning and at the end of the signal. Another disadvantage is high computational complexity.

F. Ensemble empirical mode decomposition

Problems of the EMD method for decomposition of intermittent or noisy signals and the problem of the signal edge analysis are solved by the ensemble empirical mode decomposition (EEMD). The principle of the method is to create a set of signals by adding different white noise realizations to the analyzed signal. The EMD method is then applied to each of the created signal realizations. The resulting IMFs are obtained by averaging the IMFs of each signal.

In this work, the signal was divided into 10 IMFs, and the sum of the last two IMFs was identified as a trend.

The disadvantage of this method is the high computational complexity, which is caused by the fact that the EMD has to be run as many times as the number of created realizations of the signal [7].

G. Filtration through the high-pass filter

Baseline wandering is a low-frequency noise, so it is possible to remove this noise by filtering through a high-pass filter. An IIR filter with a cut-off frequency of 0.5 Hz was used. Zero-phase forward-backward must be used to prevent phase distortion.

H. Zeroing spectral lines

Another type of filtration is filtering by zeroing spectral lines. Again, an upper pass filter is used. The method of zeroing the spectral lines has the advantage that only the required frequency components are removed. Phase distortion never occurs in this method.

I. Wavelet filtration

The last method to be compared is filtration using a discrete stationary wavelet transform. The decomposition was done in 8 bands using the db3 wavelet. The last band represents low frequencies and therefore the last band is thresholding after decomposition. Soft thresholding was used.

III. Results and Discussion

The methods described above can be divided into two groups. The first group consists of methods that result in a signal without baseline wandering without an estimate of the trend. Wavelet filtering, zeroing spectral lines, and filtering through the high-pass filter belong to the first group.

The second group are the methods which estimate the trend. This trend is subtracted from the original signal in the last step. Median filtering, shape-preserving piecewise cubic interpolation, piecewise cubic spline interpolation using non-a-knot end conditions, EMD and EEMD belong to the second group. The trend estimated by the individual methods of the second group can be seen in Fig. 1.

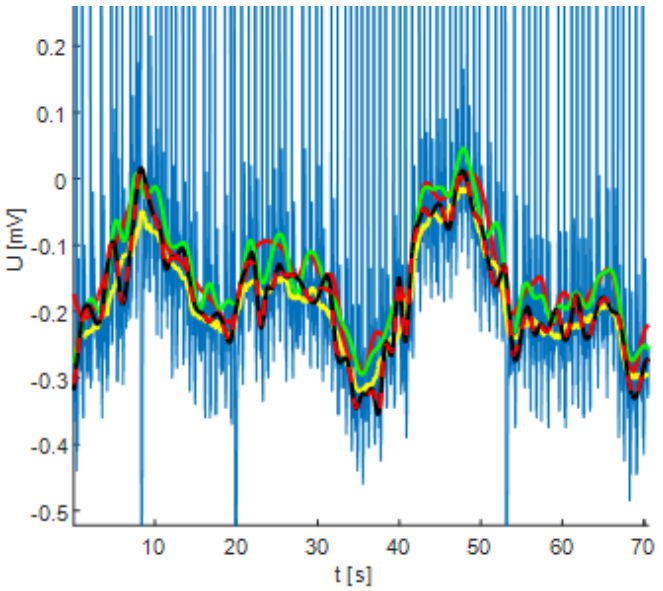


Fig. 1. Unprocessed ECG signal (blue) with estimated trend – median filtration (yellow), EMD (red), EEMD (green), shape-preserving piecewise cubic interpolation (red dotted line) and piecewise cubic spline interpolation using not-a-knot end conditions (black)

All of the above described methods have been tested on the entire MIT-BIH Arrhythmia Database [8], [9]. The database contains 48 records, which are about 30 minutes long. Each record has two leads. Sampling frequency is 360 Hz, the resolution is 11-bit within a 10 mV range. Each of these signals was pre-processed by all methods described above.

The QRS complexes of all pre-processed signals were classified by an automated classifier to compare the efficiency of the methods for removing baseline wandering. The comparison of the methods is based on the assumption, that the higher classification score achieved in the classification of QRS complexes indicates a better method for eliminating baseline wandering.

The classifier described in [4] was used for automatic classification. It is an automatic classifier of QRS complexes based on the SVM method. The SVM classifier was trained separately for each type of pre-processing. Training and testing of the classifier is realized by ten-fold cross-validation. All QRS complexes of all signals are randomly divided into ten equally sized subgroups. Nine subgroups are used for training and one for testing the classifier. Gradually, each of the subgroups is selected as the test (there is ten testing in total). [4]

The MIT-BIH Arrhythmia Database contains a total of 109820 QRS complexes. Each QRS complex was classified by cardiologists into one of the fifteen heartbeat types. These groups can be united into 5 groups according to the AAMI standard [10]: supraventricular ectopic beat (S), ventricular ectopic beat (V), fusion beat (F), unknown beat (Q) and any heartbeat not in the S, V, F or Q classes (N). All QRS complexes were classified into one of these five groups [4].

The success of the classification was evaluated using accuracy (Acc), sensitivity (Se) and specificity (Sp). These variables are calculated as follows [4]:

$$Acc = (TP_N + TP_S + TP_V + TPF + TPQ) \times 100 / Y, \quad (1)$$

$$Se_i = TP_i \times 100 / (TP_i + FN_i), \quad (2)$$

$$Sp_i = TN_i \times 100 / (TN_i + FP_i), \quad (3)$$

where TP (true positive) is the number of correctly classified complexes into the analyzed group, TN (true negative) is the number of correctly classified complexes into other groups than the analyzed group, FP (false positive) is the number of incorrectly classified complexes into the analyzed group, FN is the number of incorrectly classified complexes into other groups than the analyzed group, Y is the number of all classified QRS. Se and Sp are calculated for each group of QRS separately (subscript i). Resulting Se and Sp for each method are given by averaging partial values.

The resulting classification successes after pre-processing by individual methods for removing baseline wandering are given in Table 1.

TABLE I. CLASSIFICATION SUCCESS AFTER REMOVAL OF BASELINE VANDERING BY VARIOUS METHODS

Removal method	**Classification success**		
	Acc	***Sp***	***Se***
Without pre-processing	99,07	99,25	92,04
Median filtration	99,20	99,35	92,93
Chouchan's method	99,12	99,29	92,74
Shape-preserving piecewise cubic interpolation	99,16	99,34	92,73
Spline interpolation using not-a-knot end conditions	99,17	99,34	92,66
EMD	99,14	99,32	92,70
EEMD	99,18	99,35	92,96
Filtration through the high-pass filter	99,20	99,35	92,86
Zeroing spectral lines	99,24	99,39	93,19
Wavelet filtration	99,28	99,43	93,55
Zeroing spectral lines and EEMD	99,23	99,39	93,35
Wavelet filtration and EEMD	99,31	99,45	93,77
Median filtration and EEMD	99,19	99,36	93,05

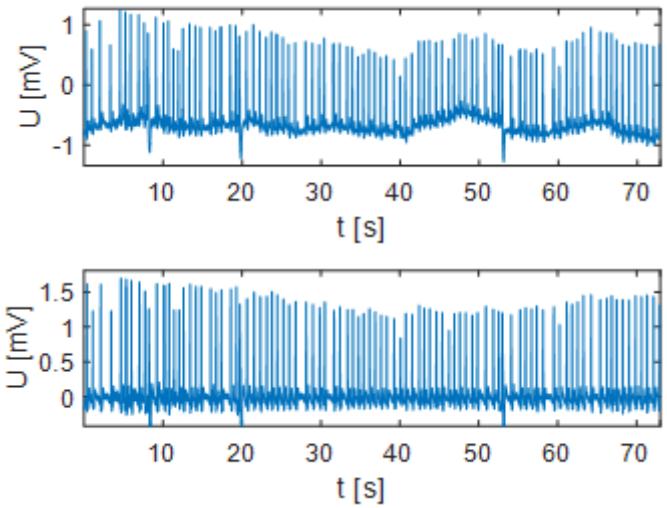


Fig. 2. The original unprocessed signal (above) and the pre-processed signal by wavelet filtration (below)

The best classification success after pre-processing with one method is achieved after wavelet filtration. The signal after this filtering is shown in Fig. 2. The second best method is zeroing the spectral lines. Furthermore, combinations of successful methods in combination with EEMD were tested. The success of the classification is further improved by the combination of wavelet filtration and EEMD. This combination of preprocessing is therefore the best to remove the baseline wandering before the automatic classification. The disadvantage of this method is its very high computational complexity.

All tested methods to eliminate baseline wandering increased the classification success rate compared to classification without prior elimination of baseline wandering. This proves, that removing the baseline wandering before the classification is useful.

Comparison of algorithms for the elimination of baseline wandering according to the following classification success was not carried out in the literature published so far. However, similar methods to this article were compared in [3]. They were compared depending on how much they affect the ST segment level. Five methods were used in this article as well as in [3]. These are the methods: wavelet filtering, EMD, Chouchan's method (this method was named Median filtering in [3]), Piecewise cubic spline interpolation using non-a-knot end conditions (this method was named Cubic Spline Curve Fitting in [3]) and filtration through the high-pass filter with the same cut-off frequency (it was used FIR filtration in [3], while IIR filtration was used in this article). The order of success of the methods that were used in the study [3] and in this work is very similar. Wave filtration is the most successful in both cases. High pass filtering is the second best in both articles from the methods that are in both articles. The worst results are given in both cases by Chouchan's method.

IV. Conclusion

A total of 9 methods were described and implemented to eliminate baseline wandering in the ECG. All methods were tested on the entire MIT-BIH Arrhythmia Database (48 records, 109820 QRS). The methods were compared according to the classification success of the automatic classification of the pre-processed signal.

The usefulness of removing baseline wandering was confirmed. All methods achieved better results than the original unprocessed signal. Wavelet filtration and then EEMD is the best for subsequent use for automatic ECG classification according to the results of this article. The disadvantage of this method is its very high computational complexity. Therefore, in some cases it may be better to use only wavelet filtration. This method is the best of all individual tested methods.

Acknowledgment

This work was supported by grant project IGA FEKT-S-17-4487.

References

[1] Cardiovascular diseases (CVDs). *World Health Organization* [online]. 2016 [cit. 2017-07-13].

[2] J. Kozumplík, "Biological signal analysis", "Analýza biologických signálů", Brno University of technology, Brno, 2012.

[3] F.A. Afsar, M.S. Riaz and M. Arif, "A comparison of baseline removal algorithms for electrocardiogram (ECG) based automated diagnosis of coronory heart disease," 3rd International Conference on Bioinformatics and Biomedical Engineering, iCBBE 2009.

[4] R. Smíšek and J. Kolářová, "ECG Signal Classification Based on SVM," STUDENT EEICT 2016, Brno, 2016, pp. 365–369.

[5] V. S. Chouhan, V. S. Mehta and S. S. Mehta, "Total removal of baseline drift from ECG signal," Proceedings - International Conference on Computing: Theory and Applications, ICCTA 2007, 2017, pp. 512–515.

[6] Y. Gao, G. Ge, Z. Sheng and E. Sang, "Analysis and Solution to the Mode Mixing Phenomenon in EMD," *2008 Congress on Image and Signal Processing*, Sanya, China, 2008, pp. 223-227.

[7] Y.-H. Wang, Ch.-H. Yeh, H.-W. V. Young, K. Hu, M.-T. Lo, "On the computational complexity of the empirical mode decomposition algorithm," Physica A: Statistical Mechanics and its Applications, vol. 400, 2014, pp. 159-167.

[8] GB Moody, RG Mark, "The impact of the MIT-BIH Arrhythmia Database," IEEE Eng in Med and Biol, vol. 20, 2001, pp. 45-50.

[9] AL Goldberger, LAN Amaral, L Glass, JM Hausdorff, PCh Ivanov, RG Mark et al., "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," Circulation, vol. 101, 2000, pp. 215-220.

[10] P. D. Chazal, M. O'Dwyer and R. B. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," IEEE Transactions on Biomedical Engineering, vol. 51, 2004, pp. 1196–1206

Symmetrical Double-Sided Two Way Ranging Using nanoPAN 5375 Modules

Jakub Streit
Faculty of Electrical Engineering
and Communication
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
Email: xstrei03@stud.feec.vutbr.cz

Václav Kaczmarczyk
Faculty of Electrical Engineering
and Communication
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
Email: kaczmarczyk@feec.vutbr.cz

***Abstract*—In the near future, rapid demand for technologies enabling location and navigation in indoor environment is expected. One of the basic principles used for positioning is to measure the distance between a moving node fixed and stations with known coordinates. The nanoPAN 5375 is an RF module that, in addition to wireless communication, can measure the distance others. It uses time of flight (TOF) and Symmetrical Double-Sided Two Way Ranging (SDS-TWR) to minimize clock drift. This paper proposes testing this module.**

I. Introduction

Precise indoor location is one of the challenges of modern technology. The solution has to be cheap, accurate and reliable. There are a wide variety of use such a technology: navigation in a large office building or shopping center, parcel, goods or material tracking, staff monitoring in dangerous area in industry and so on.

One of the most direct approach is based on measuring distance between a mobile node and a set of fixed anchors with known position. Distance is measured indirectly, by measuring a time of radio signal travels between the mobile node and the anchors. Actually, this is the same principle as the outdoor GNSS location systems (like GPS) are based on. Unlike GNSS systems, there are two major problems. The first one is an enormous jaggedness of the indoor environment. Direct line-of-sight is almost mandatory for precise results. This can be partially solved by increasing number of anchors, covering given area. And it is related to the second problem - highly accurate time measurement. GNSS systems solve this by using an atomic clock as time reference in the satellites [1]. This is very expensive and not suitable for a large number of cheap, small and energetically inexpensive anchors. Symmetrical Double-Sided Two Way Ranging is a one from numerous solutions of this problem.

II. Symmetrical Double-Sided Two Way Ranging

Symmetrical Double-Sided Two Way Ranging - SDS-TWR - is a method for minimizing impact of clock offset between two nodes on time of flight measurement precision. Because of this, usage of less precise clock sources, like a quartz oscillators, is possible, while maintaining sufficient precision of time measurement.

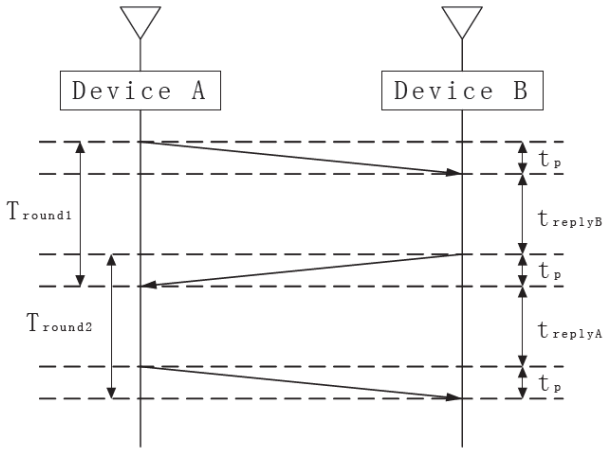


Fig. 1: Symmetrical Double-Sided Two Way Ranging [2]

Figure 1 shows principle of SDS-TWR. Device A starts measuring by sending a request to device B. Also starts measuring of time t_{round1}. When device B receives the request, sends ACK + their own request back to the device A and start measuring of time t_{round2}. Device A then receives ACK + request from device B, stops t_{round1} measurement and send ACK to device B. Device B stops measurement of t_{round2}, after receives device A ACK. Both measured values then must be collected at one device, usually device A. Consider fixed and known reply processing time in both devices, the propagation time t_p can be computed according equation 1.

$$t_p = \frac{1}{4}((t_{round1} - t_{replyA}) + (t_{round2} - t_{replyB})) \qquad (1)$$

There are also several modifications of this concept. Their improvement is often based on performing several measurements, which shares some part of communication. This approach gains better precision by simple averaging, while maintain shorter measurement time, than several complete cycles of SDS-TWR. Another advantage is energy saving. Some examples are in [3], [4] and [5].

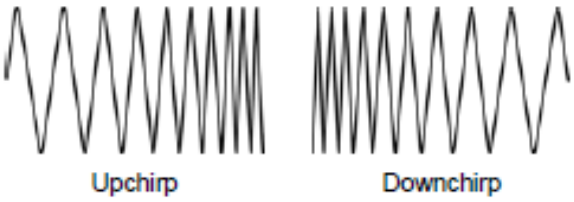


Fig. 2: Chirp Spread Spectrum [7]

TABLE I: Description of equation 2 fields

Field	Description
$TxRespTime$	nanoPAN register, 16 bit
$TxGateOff$	nanoPAN register, 4 bit
$RxGateOff$	nanoPAN register, 4 bit
$TxGateSum$	nanoPAN register, 14 bit
$RxGateSum$	nanoPAN register, 14 bit
$SPEED\_OF\_AIR$	299.792458 $[m/\mu s]$
CLK_4MHZ	4
CLK_32MHZ	32
CLK_LOD20	2000.0/244175
$PULSE\_DET\_UC\_MAX$	$2 * 24$
$rangingConst$	325.587060636

III. NANOPAN 5375

nanoPAN 5375 is a commercially available RF module, operating in 2.4 GHz ISM band, capable of measuring distance between two modules, using SDS-TWR. It is based on a nanoLOC TRX Transceiver NA5TR1 [6]. This chip uses Chirp Spread Spectrum (CSS) modulation technique, developed by Nanotron Technologies [7]. This technique uses monotonically increasing (Upchirp) or decreasing (Downchirp) frequency for coding symbols (see figure 2). It provides high reliability of communication because of better immunity against interference.

A. Distance measurement

nanoPAN 5375 uses SDS-TWR technique for distance ranging. Unfortunately, there is no specification of range, precision, or even resolution in the documentation ([8] [6] [9] [10] [7]). There is only an equation 2 in a sample code of a development kit [9]. But with little description of input data - see table I.

$$\begin{aligned} d = 0.5\ SPEED\_OF\_AIR \Bigg(& \frac{TxRespTime}{CLK\_4MHZ} \\ & - \frac{TxGateOff + RxGateOff}{CLK\_32MHZ} \\ & - \frac{CLK\_LOD20\ (TxUcSum + RxUcSum)}{PULSE\_DET\_UC\_MAX} \Bigg) \\ & - rangingConst \end{aligned} \tag{2}$$

Resolution of nanoPAN 5375, determined from equation 2 and table I is approximately 25.58 mm.

B. Shorter distance

Apart from classical SDS-TWR, there is an alternative way of calculating distance in the source code. Measurement is done in the same way as in SDS-TWR, but instead of averaging both distances, only the shorter one is returned. Assumption is, that some reflections in signal propagation path may cause greater error, than clock differences of modules. Therefore, the shorter measured distance is closer to the real one. Both of these methods are tested in chapter IV.

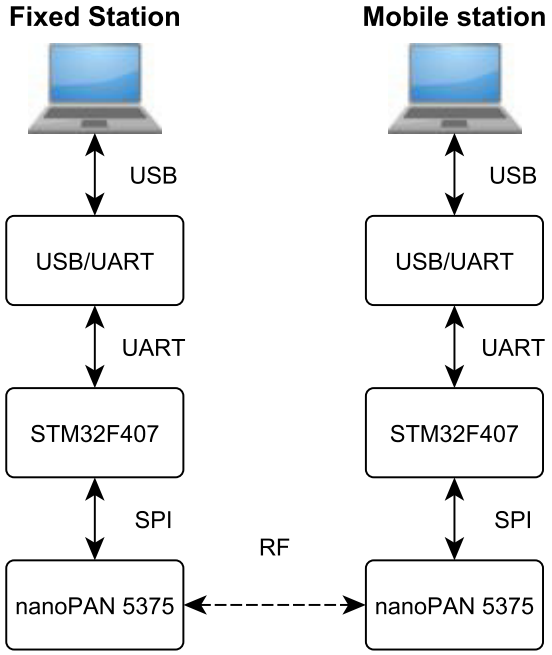


Fig. 3: HW setup

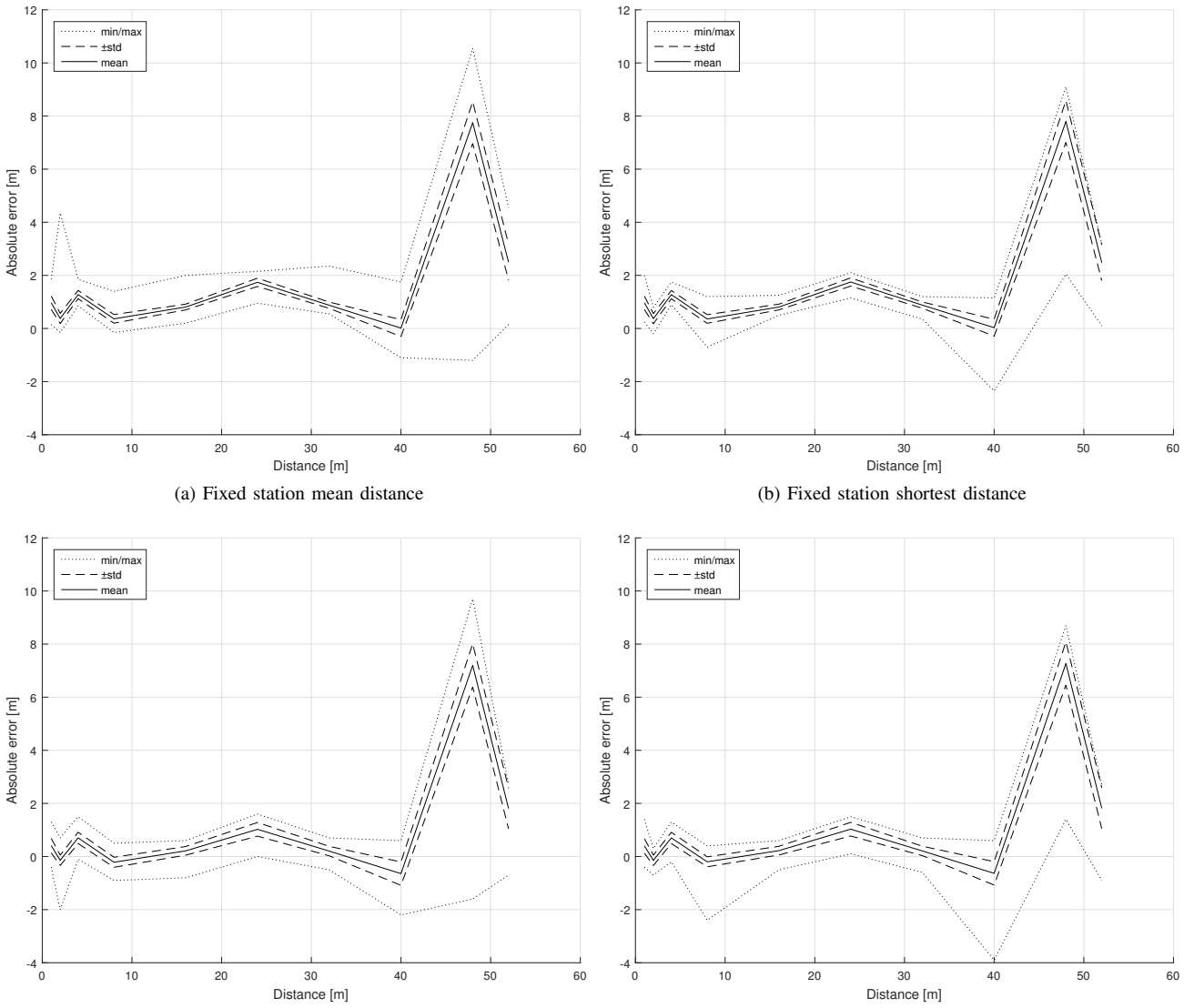
IV. EXPERIMENTAL MEASUREMENT

To test the nanoPAN 5375, several sets of measurement was performed. Figure 3 shows a block diagram of used hardware. Distance between two nanoPAN modules was measured, one is called Fixed station, other Mobile station. The measurement was done in a straight corridor of an office building. The Fixed station was on the one end, while the Mobile station has been placing to desired distances for each set. Each set consist of 1000 measurements with 500 ms period, performed by both stations. Minimum, maximum and mean absolute error was evaluated. Figure 4 shows results.

Until 40 m, mean absolute error is less than 2 m, while its standard deviation is less than 20 cm. Peak at 48 m is probably caused by some signal reflection and interference. It has to be closely studied.

V. CONCLUSION

The first tests of nanoPAN 5375 shows, that distance measurement resolution is sufficient for precise indoor location. Also repeatability looks well. But accuracy has to be improved. Maybe it could be reached by different settings of the module. Research continues.



(a) Fixed station mean distance

(b) Fixed station shortest distance

(c) Mobile station mean distance

(d) Mobile station shortest distance

Fig. 4: Direct line of sight measurement results

ACKNOWLEDGMENT

The completion of this paper was made possible by the grant No. FEKT-S-17-4234 - "Industry 4.0 in automation and cybernetics" financially supported by the Internal science fund of Brno University of Technology.

REFERENCES

[1] E. Kaplan and C. Hegarty, *Understanding GPS: principles and applications*. Artech house, 2005.

[2] Y. Jiang and V. C. M. Leung, "An asymmetric double sided two-way ranging for crystal offset," in *2007 International Symposium on Signals, Systems and Electronics*, July 2007, pp. 525–528. [Online]. Available: http://ieeexplore.ieee.org/document/4294528/

[3] H. Kim, "Double-sided two-way ranging algorithm to reduce ranging time," *IEEE Communications Letters*, vol. 13, no. 7, pp. 486–488, July 2009. [Online]. Available: http://ieeexplore.ieee.org/document/5166419/

[4] A. I. Baba and M. M. Atia, "Burst mode symmetric double sided two way ranging," in *2011 IFIP Wireless Days (WD)*, Oct 2011, pp. 1–3. [Online]. Available: http://ieeexplore.ieee.org/document/6098183/

[5] D. Neirynck, E. Luk, and M. McLaughlin, "An alternative double-sided two-way ranging method," in *2016 13th Workshop on Positioning, Navigation and Communications (WPNC)*, Oct 2016, pp. 1–4. [Online]. Available: http://ieeexplore.ieee.org/document/7822844/

[6] *nanoPAN 5375 Module Technical Specifications*, Nanotron Technologies GmbH, 05 2008, rev. 1.0.

[7] *nanoLOC TRX Transceiver (NA5TR1) Datasheet*, Nanotron Technologies GmbH, 03 2010, rev. 2.3.

[8] *nanoPAN 5375 RF Module*, Nanotron Technologies GmbH, 2008.

[9] *nanoPAN 5375 Development Kit User Guide*, Nanotron Technologies GmbH, Oct 2010, rev. 1.1.

[10] *nanoLOC TRX Transciever (NA5TR1) User Guide*, Nanotron Technologies GmbH, 12 2006, rev. 1.01.

Random Forests Pixel-wise Classification for Detection and Segmentation of Cells in the Images from Holographic Microscope

Tomas Vicar, Radim Kolar
Department of Biomedical Engineering
Brno University of Technology
Brno, Czech Republic
xvicar03@stud.feec.vutbr.cz, kolarr@feec.vutbr.cz

***Abstract*—Microscopic cell image analysis is widely used by biologists for cell behavior and cell morphology study. In dense cell cultures precise single-cell segmentation is challenging task and it is an important step for automatic cell analysis methods. This work introduces a novel method for robust single cell segmentation of images from holographic microscope. The method is based on pixel-wise classification with random forests for both background segmentation a cell detection, where cell detection image is refined with distance transform based detector. Final single cell segmentation combines both detection and background with seeded watershed. Proposed background segmentation part reaches results similar to other algorithms, but cell detection part of the algorithm is innovative and achieves significantly better result than commonly used detector.**

***Keywords*—cell segmentation, cell detection, random forests, distance transform**

I. Introduction

Nowadays, optical microscopy is widely used to quantify single-cell features such as cell size or shape. Especially for long-term live cell imaging, where up to thousands of images are recorded, manual processing of such amounts of data is impracticable. For this reason automatic cell segmentation is a subject of interest in a wide range of cell-based studies and it is an important step in the automation of analysis of cell based experiments. Holographic microscopy [1] is one of the methods for Quantitative Phase Imaging (QPI), very suitable for long-term label-free live cell imaging [2], but it is relatively new modality and no suitable segmentation approach exists.

Cell segmentation is challenging problem, because of heterogeneous shape, that dynamically change and have large variability across cell population. Furthermore, cell interactions and cell clustering during the long-term experiments create another challenging problem in the single cell segmentation task. There are typically tree main efforts in automatic cell image segmentation - cell detection (identification of single cells) for cell counting, background segmentation for the analysis of cell confluence (percent of covered area) and single-cell segmentation for single-cell features evaluation. In most of the published approaches (see review [3]) is single-cell segmentation composed of the cell detection and segmentation of the background.

The cell detection method can use different approaches, which are usually based on Distance Transform (DT) [4], morphology operations, Laplacian of Gaussian (LoG) filter [4], maximally stable extremal region, Hough transform or radial symmetry voting, etc. However, many of these methods are capable to detect only circular or elliptical shapes and have poor results for touching cells. Therefore, current methods used some supervised learning approach [3].

The most popular approaches for cell segmentation are based on thresholding, watershed, pixel clustering, graph methods or deformable models [3]. Nowadays, supervised learning methods for segmentation (e.g. WEKA [5] and Ilastik [6]) with pixel-wise classification are very popular thanks to their universality. These methods extract many features (e.g. local entropy, Gabor filters and gradient magnitude) for every pixel, which are used for pixel classification. These methods are capable to segmented cell images from different modalities like differential contrast microscopy, phase contrast microscopy and fluorescence microscopy. However they are not capable to achieve of complete single-cell segmentation (including the splitting of touching cells), but only background segmentation. The aim of this study is to use similar approach even for single-cell segmentation (including cell detection and the splitting of touching cells).

In this study, a novel approach for single-cell segmentation for QPI images is introduced. It's based on pixel-wise classification for both background segmentation and cell detection followed with seeded watershed for separation of single-cells.

II. Methods

A novel method for cell segmentation introduced in this paper, involve these main steps - feature extraction, pixel-wise classification for both cell detection and background segmentation and foreground division with seeded watershed (see Figure 1).

A. Random forests classifier

Random forests classifier is most commonly used classifier for image pixel-wise classification [6], whose main advantage is insensitivity to redundant features. Matlab implementation of random forest has been used, which implements the method described in [7].

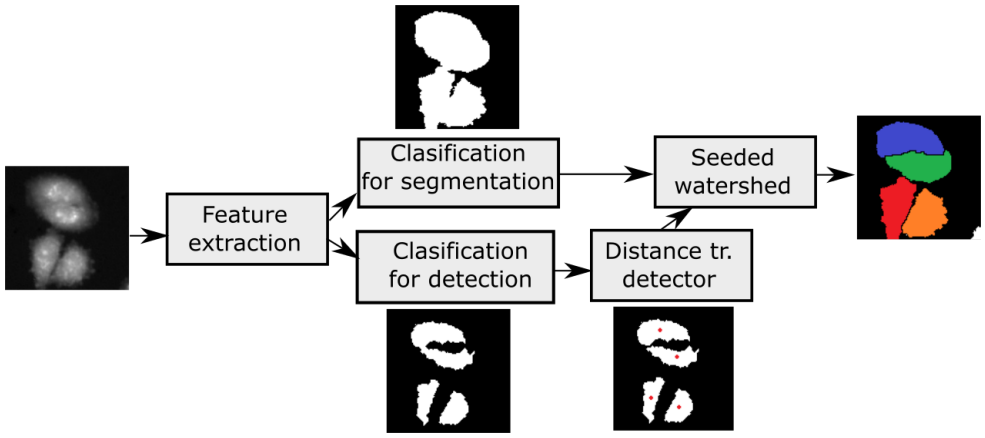


Fig. 1: Block diagram of proposed method

B. Image features

Feature selection is done so to be suitable for cell detection. There are many different features, which are commonly used for cell detection. Typically set of features with different scale is summed (or maximal response is taken) to single parametric image [8]. This parametric image is than thresholded (binary objects are detected cells) or local maxima are found and considered as cells. Simple summation of scales can lead to information loss. In this paper, different scales of different features are extracted and used as features for the classifier. Example of these features for one scale is in Figure 2.

The Laplacian of Gaussian (LoG) filter is one of the most popular methods for blob object detection in medical image analysis [3]. LoG filter is appropriate feature for our purpose and it is defined as

$$LoG(\mathbf{x},\sigma) = \nabla^2 G(\mathbf{x},\sigma) = \frac{\sigma^2 - ||\mathbf{x}||_2^2}{2\pi\sigma^6} e^{\frac{-||\mathbf{x}||_2^2}{2\sigma^2}} \quad (1)$$

where G is 2D Gaussian function, $\mathbf{x} = (x, y)$ and $||\cdot||_2$ is l2-norm. It is actually matched filter of blob shape, which produce higher response in cell location. σ is the standard deviation of the Gaussian, which should correspond to cells radii. Here, 10 scales are used for 10 feature maps extraction.

Next features are based on Hessian analysis, which is often used for cell detection [9]. Hessian is formulated as

$$H(x,y,\sigma) = \begin{pmatrix} \frac{\partial I(x,y;\sigma)}{\partial x^2} & \frac{\partial I(x,y;\sigma)}{\partial x \partial y} \\ \frac{\partial I(x,y;\sigma)}{\partial y \partial x} & \frac{\partial I(x,y;\sigma)}{\partial y^2} \end{pmatrix} \quad (2)$$

where $I(x, y; \sigma)$ is image smoothed with a Gaussian of standard deviation σ. The eigenvalues of this matrix have different properties for different image structures. For example blob-like structure is represented by two positive eigenvalues. Here, 10 different scales were used, which produced 20 feature maps.

Another cell detection method capable to produce reliable features for cell detection/segmentation is fast radial-symmetry transform (FRST) [10]. Every pixel votes in its gradient direction at distance r, so the affected pixel position is

$$P(\mathbf{x}) = \mathbf{x} + round\left(\frac{g(\mathbf{x})}{\|g(\mathbf{x})\|_2} r\right) \quad (3)$$

where $g(x)$ is gradient image. On the affected position, orientation projection image O_r is increased by 1 and magnitude projection image M_r is increased by $||g(x)||_2$. Usually are these images combined together, but in our method are both with 10 different radii used as features, which produce 20 feature maps.

Additionally gradient magnitude images with 10 different Gaussian blurs are used as features.

C. Classification for detection

This step involve to classify pixels in order to get individual cells as individual binary objects. A manually labeled mask of whole cells are morphologically eroded to obtain binary mask of cell centers only. This leads to good separation between objects in classified image.

D. Distance transform detector

This step involves separation of poorly divided cells after classification for detection. Distance transform (DT) of binary image is defined as an image, which pixel value are distance to nearest background pixel. H-maxima transform eliminate local maxima smaller than the set depth h. DT with h-maxima transform can be used for cell detection [4]. For cell detection, local maxima of the distance transform image are detected cells and h-maxima transform can be used for small maxima elimination. This detector is normally applied on thresholded binary foreground, but in our case it is applied on results of classification for detection.

E. Classification for segmentation

This step is for background segmentation and it is straightforward. Each pixel is classified as background or foreground pixel, based on its features. For simplicity (and shown to have appropriate results) same features as for cell detection are used, only difference is, that binary masks without erosion are used for training.

F. Seeded watershed

Finally seeded watershed [11] is used in order to combine detection and background segmentation results to single-cell

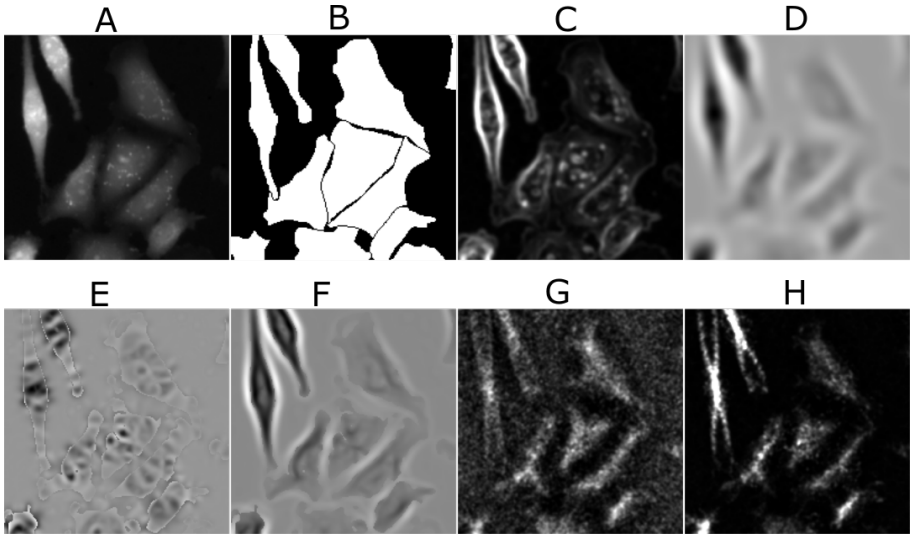


Fig. 2: Original image A, ground truth mask B, gradient magnitude C, Log filter, first and second Hessian eigenvalue E and F, FRST orientation and magnitude projection images G and H

segmentation. Seeded watershed is used on original image, where seed-points are produced by cell detection. Watershed results (boundary lines) are then used for division of foreground.

III. Results and Discussion

A. Dataset

PNT1A human cell line cultured in RPMI-1640 medium was used for method evaluation. This cell line has been selected because its popularity in research and large variability in cell shapes and intensities. Images were acquired with Tescan multimodal holographic microscope Q-PHASE with objectives Nikon Plan 10/0.30 and CCD camera Ximea MR4021 MC-VELETA. Dataset contains 18 fields of view (FOV), with 428 cells and resolution 600x600px. Three FOV are used for training of classifier and 15 FOV for evaluation of results (data example on Figure 3).

B. Parameter settings

For LoG, FRST and Hessian analysis, the scale range was set to the expected cell radii range 5 - 25 and for gradient magnitude was Gaussian standard deviation was set to range 0 - 10, all with 10 linear steps. For random forests 200 trees were used. Change of these parameters shown to have very small influence to the result. The only sensitive parameter is h for h-maxima transform in DT detector, which was experimentally determined to 2.

C. Method for evaluation

The results are reported in terms of $Precision = TP/(TP + FP)$, $Recall = TP/(TP + FN)$ and $Dice = 2TP/(2TP + FP + FN)$. For cell detection, FN are represented by ground-true masks without detected seed-point inside the mask and FP are redundant seed-points inside the mask or seed-points on the background. For the evaluation of background segmentation using of these formulas is straightforward. For evaluation of single-cell segmentation same approach as in [12] has been used. Reference cell R and segmented cell S are considered matching if $|R \cap S| > 0.5|R|$ where each S can be used only once. Without any overlap is matching function set to empty. The final values of $Precision$, $Recall$ and $Dice$ are computed as mean over all reference cells. Mask of cells at the image boundary were manually labeled and they are not evaluated.

D. Evaluation

The results were evaluated for all three steps - background segmentation, cell detection and single cell-segmentation in Table I. There was done a comparison with suitable methods for individual steps and final single-cell segmentation. For example of results see Figure 3.

For background segmentation, simple thresholding (optimal threshold was found on training images) provides only slightly worse results (0.0288 difference in dice coefficient) than proposed method, which achieve practically same results as Ilastik [6]. QPI background is very easy to segment and there is not much possibilities for improvement.

For cell detection DT detector is evaluated besides proposed method. DT is simple but still actively used method (e.g. [13]). DT is part of the proposed algorithm, but for comparison instead of use on cell detection results, it was used on background segmentation as this method usually is. [4]. Method for cell detection shown great improvement 0.0810 of dice coefficient.

For final single-cell segmentation is proposed method compared with FogBank [14]. FogBank is method for cell segmentation focusing on proper cell separation across multiple cell line and modalities. The algorithm is available with GUI, where user can interactively set parameters. Parameters were experimentally set on 3 training images to optimal value and evaluated on tested images. Proposed method is significantly better and provides 0.1646 improvement of dice coefficient

TABLE I: QPI cell image segmentation and detection results

	Background segmentation			Cell detection		Single-cell segmentation	
	Proposed	Threshold	Ilastik	Proposed	Distance tr.	Proposed	FogBank
Precision	0.9304	0.9255	0.9335	0.9504	0.8667	0.8348	0.6046
Recall	0.9466	0.8944	0.9379	0.9387	0.8603	0.8527	0.7949
Dice	0.9385	0.9097	0.9357	0.9445	0.8635	0.8351	0.6705

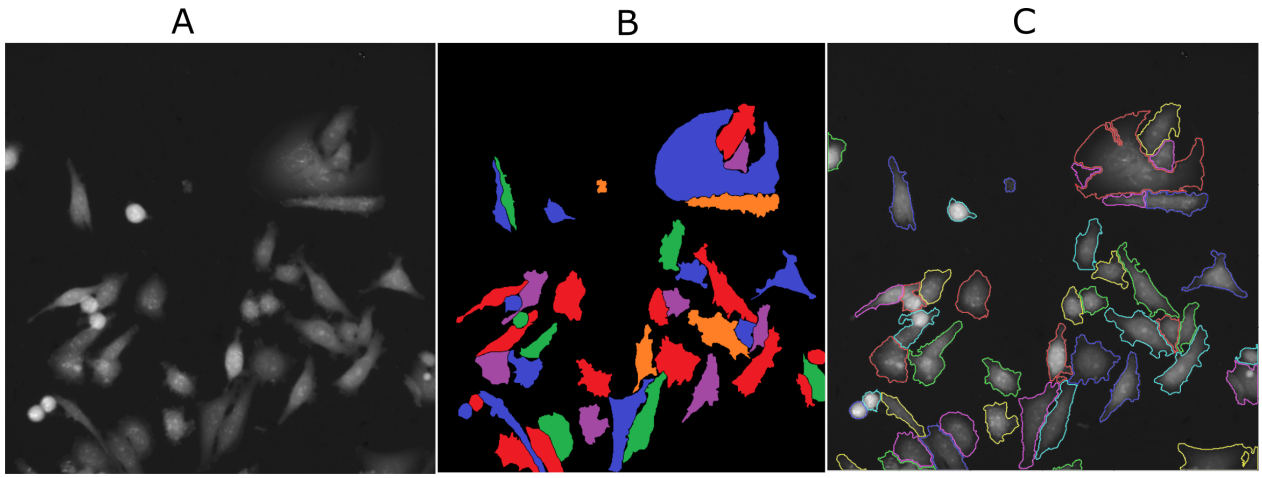


Fig. 3: Original image A, ground truth mask B, final segmentation C

against FogBank. Difference is mainly caused by better cell detection, leading to better cell division.

IV. Conclusion

A novel method for robust single-cell segmentation of QPI images were proposed. The method is based on pixel-wise classification for both background segmentation and cell detection using suitable set of features. Cell detection was refined with DT detector. Background segmentation and cell detection are then combined together with seeded watershed to the final segmentation. The method includes a completely new approach for cell detection which provides significantly better results against common DT detector alone. It also showed strong robustness against cell shape variation and of clustered cells.

References

[1] M. Lostak, R. Chmelik, M. Slaba, and T. Slaby, "Coherence-controlled holographic microscopy in diffuse media," *Opt. Express*, vol. 22, pp. 4180–4195, Feb 2014.

[2] J. Balvan, A. Krizova, J. Gumulec, M. Raudenska, Z. Sladek, M. Sedlackova, P. Babula, M. Sztalmachova, R. Kizek, R. Chmelik, *et al.*, "Multimodal holographic microscopy: distinction between apoptosis and oncosis," *PloS one*, vol. 10, no. 3, p. e0121674, 2015.

[3] Y. Al-Kofahi, W. Lassoued, W. Lee, and B. Roysam, "Improved automatic detection and segmentation of cell nuclei in histopathology images," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 4, pp. 841–852, 2010.

[4] K. Thirusittampalam, M. J. Hossain, O. Ghita, and P. F. Whelan, "A novel framework for cellular tracking and mitosis detection in dense phase contrast microscopy images," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 3, pp. 642–653, 2013.

[5] I. Arganda-Carreras, V. Kaynig, C. Rueden, K. W. Eliceiri, J. Schindelin, A. Cardona, and H. S. Seung, "Trainable weka segmentation: a machine learning tool for microscopy pixel classification," *Bioinformatics*, 2017.

[6] C. Sommer, C. Straehle, U. Kthe, and F. A. Hamprecht, "Ilastik: Interactive learning and segmentation toolkit," in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 230–233, March 2011.

[7] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[8] H. Peng, X. Zhou, F. Li, X. Xia, and S. T. C. Wong, "Integrating multi-scale blob/curvilinear detector techniques and multi-level sets for automated segmentation of stem cell images," *Proceedings. IEEE International Symposium on Biomedical Imaging*, vol. 2009, pp. 1362–1365, 2009.

[9] J. Liu, J. M. White, and R. M. Summers, "Automated detection of blob structures by hessian analysis and object scale," in *2010 IEEE International Conference on Image Processing*, pp. 841–844, Sept 2010.

[10] G. Loy and A. Zelinsky, *A Fast Radial Symmetry Transform for Detecting Points of Interest*, pp. 358–368. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002.

[11] A. Pinidiyaarachchi and C. Wählby, *Seeded Watersheds for Combined Segmentation and Tracking of Cells*, pp. 336–343. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005.

[12] M. Maska, V. Ulman, D. Svoboda, P. Matula, P. Matula, C. Ederra, A. Urbiola, T. España, S. Venkatesan, D. M. Balak, *et al.*, "A benchmark for comparison of cell tracking algorithms," *Bioinformatics*, vol. 30, no. 11, pp. 1609–1617, 2014.

[13] X. Ji, Y. Li, J. Cheng, Y. Yu, and M. Wang, "Cell image segmentation based on an improved watershed algorithm," in *2015 8th International Congress on Image and Signal Processing (CISP)*, pp. 433–437, Oct 2015.

[14] J. Chalfoun, M. Majurski, A. Dima, C. Stuelten, A. Peskin, and M. Brady, "Fogbank: a single cell segmentation across multiple cell lines and image modalities," *Bmc Bioinformatics*, vol. 15, no. 1, p. 431, 2014.

Listen to Your Crowd - Analysis of User Comments of a Crowdsourced QoE Study

Ondrej Zach
Dept. of Radio Electronics
Brno University of Technology
Technicka 12, 616 00 Brno, Czech Republic
Email: ondrej.zach@phd.feec.vutbr.cz

***Abstract*—Crowdsourcing-based studies in the field of Quality of Experience usually focus only on one result which is the impact of the conditions tested on the final recipient of the multimedia content (the viewer). However, much more information is often gathered, not just the quality scores. In this paper, we analyze the crowdworkers' feedback based on comments they left after finishing the given task. Based on the analyses, recommendations to improve the tasks are made.**

I. Introduction

In the field of Quality of Experience (QoE) crowdsourcing is an often used tool to gather opinions of the users on the quality of the presented content. Where traditional studies in controlled environment may be both cost and time demanding, crowdsourcing can bring significant savings on both sides. However, the whole work flow of preparation of a crowdsourcing-based study on QoE is considerably different from lab-based studies and hence must be carried out with care in order to gather correct results. In a crowdsourcing environment, the research has no control of the users' equipment and therefore such a platform must be designed in order to run as flawlessly as possible at wide variety of devices.

The main difference between the traditional approach in QoE evaluation and using of crowdsourcing is in the final processing of the data gathered. In a crowdsourcing study, the gathered answers have to be filtered out not only based on outliers, but also on the consistency of the answers of the user and possible cheating. Therefore, much more information is gathered from the users to be used afterwards to identify possible cheating behavior.

The focus of the paper is to analyze the feedback of the users in crowdsourcing QoE study to address the possible issues or recommendations which might be beneficial for future research.

The rest of the paper is divided as follows. Section 2 describes the studies during which we gathered the data we use. In Section 3 we analyze the comments the participants left after completion of the task and the last Section concludes.

II. The Study

In [3]–[6] results of our previously published crowdsourcing-based QoE experiments are described. In [5] we focused on the issue of using crowdsourcing for evaluation quality of video sequences encoded using High Efficiency Video Coding (HEVC). We faced with the issues of smooth video playback on the side of the crowd workers. The research published in [4] is focused on a comparison of Advanced Video Coding (AVC) and High Efficiency Video Coding (HEVC) coded video content. An innovative feature of the study was an adaptive approach to allocation of the sequences to the users. The sequences to be played to the next participant were chosen according to the length of the confidence interval (CI) of the previously gathered quality scores of that respective sequence. Such an approach may lead to faster convergence of the scores to values with lower CIs. In [3] we evaluate the performance of the adaptive approach compared to conventional random or pseudo-random approach. For this study, we used AVC coded content only. We also suggested the best approach to content and quality level selection based on criteria of better statistical properties of the results (e.g. CI length, distribution etc.).

The video database created for the study in [3] consisted of 5 different contents, each encoded to 5 different bitrate levels. This resulted in 25 different AVC encoded processed video sequences (PVS) and 25 HEVC encoded PVSs. These PVSs were then used for the crowdsourced studies. The length of each sequence was adjusted to 10 s. The study was ran using a web-based application similar to [7]. The framework included wide variety of features to monitor user's activity and to detect possible cheating. For example, a set of simple user question was presented to the participant twice, at the beginning and at the end of the test, with different order of the answers. If the user answered some of these question differently, his ratings were marked as inconsistent. Also, the framework could monitor if the user switched to another tab of the web browser etc. The participants watched five PVSs with different content each. The study was available as a micro job on the Microworkers[1] crowdsourcing platform. Every user could participate and was rewarded with 0.30$ upon completion of the test.

In [6] we focused on other factors with possible impact on the perceived QoE. For this study, we ran four crowdsourcing experiments dealing with the impact of displayed advertisements, initial delay and presence of poster image prior the video playback. In this case, the users were shown 3 videos with different conditions and were rewarded with 0.20$ upon completion.

In this paper, we focus on the feedback of the participants only in order to obtain possible recommendations for future

[1] https://microworkers.com

experiments. Overall, we had data from 3,545 user sessions. The users left feedback in 633 cases. For further analysis, only the sessions with user's comments are considered.

III. Analysis of Users' Feedback

After the users completed the micro task in our crowdsourcing platform, they were given the option to leave a comment or idea, as depicted in Fig. 1. As leaving a comment was optional, only 633 comments in total were obtained, which represents approximately 18% of all user sessions. Each of the comments was then manually processed and the comments were divided in several groups based on their content.

A. The comments

Based on the content, the comments can be divided in 6 groups:

A) **Audio issues** - comments on missing audio track,
B) **Funny comments** - comments that could be interpreted as a joke,
C) **Quality issues** - comments on quality of specific video sequence or all sequences,
D) **No comment** - users did not leave the comment field blank but wrote *NA* or similar instead,
E) **Real comments** - real and useful feedback.
F) **Appreciation** - comments like *Thank you*, *Good job*, *I liked the job* etc.,

The count of comments belonging to different groups is depicted in Fig. 2. One clear conclusion can be made - the majority of the commenting users left feedback representing the *Appreciation* group (F). These comments unfortunately do not bring any useful feedback and much more represented the effort of the participant to be included in future micro jobs.

A special case are comments belonging to group A. In this group all the comments on the issues of missing audio track in the video sequences were put. As it is a common convention in the field of video quality evaluation not to add the audio track to the final processed video sequence, also the sequences used in all of our previous studies did not contain audio. However, even as we clearly stated this at the beginning of the micro job, the users did not understand why are the tested sequences mute and demanded sound. This may lead to suggestion, that in crowdsourced environment, the users are used to traditional video streaming and therefore the presence of the audio track is an important issue to them.

In 11 cases, the participants of the studies left a comment that can be interpreted as a joke and brings much more a smile than some useful feedback. Examples of such comments can be as follows (comments are displayed as recorded by our framework without any further editing):

"All praise to the test",

"Give me bonus dollars, sir, thanks",

"I am good student& like it",

"You university guys, it's a cheetah, not a leopard".

Where majority of such comments do not bring any help for developing of future tasks, the last one, on the other hand, represents that the user really focused on the task and noticed an incosistency.

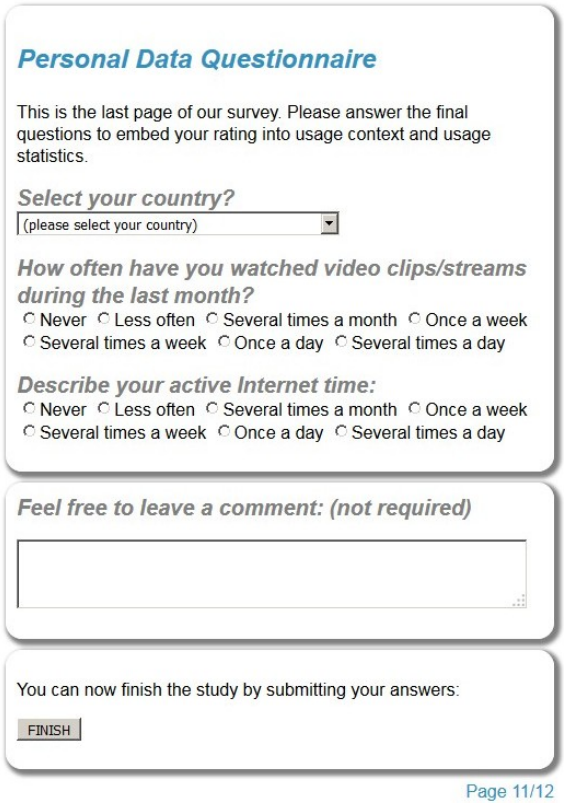


Fig. 1: The crowdsourcing framework with the comment section.

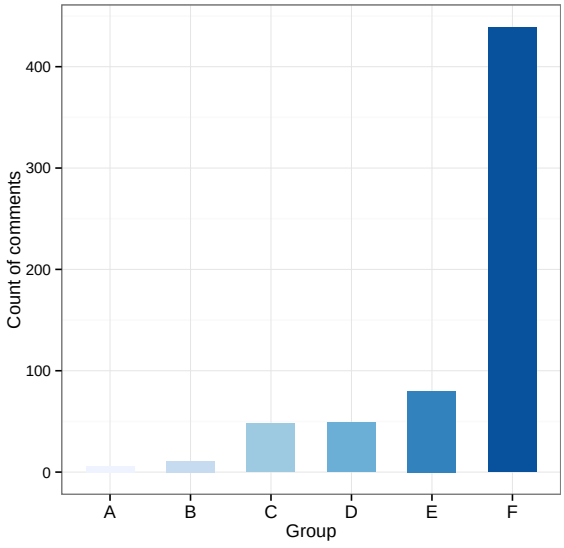


Fig. 2: Count of comments belonging to different comment groups.

B. Useful feedback

Comments containing useful feedback are the most valuable ones for a researcher. During all of our previous studies, useful comments take approximately 13% of all comments written by the participants.

The most common useful comments addressed some kind of a technical issue the participants experienced, e.g. problems with video loading, framework crashing or problems in different browsers. Although the framework was designed in order to run smoothly at as many platforms as possible and was tested thoroughly before each study, such a situation can always occur. An example of similar comments follows:

"Movie 3 and 4 took many reloads before it finally played",

"In the basketball clip there was a big stop in the middle. Other videos did not have it",

"I had some problems playing videos with Chrome Had to use Mozilla instead",

..."I tried in UC browser video 5/5 hanged and again i tried the same, in Mozilla firefox it was played good".

These comments can help the developer to focus on additional testing under specific conditions, e.g. in other browsers. On the other hand, as presented in [8], the share of the 2 most frequently used web browser is approximately 90% and therefore it might be more efficient to focus on the support of majority of the browser at the cost of rare cases of technical issues.

Several users also left comments regarding the selection of the video content and the lengths of the sequences. In our studies, we used the typical length of the sequences 10 s. However, as users are now to used videos with durations of several minutes at least, some reported they could not focus on such a short sequences and suggested longer duration:

"For a better video test result video length should be minimum 30 seconds. Thanks".

As many of the users stated in the questionnaire at the beginning of the test, they watch videos online on a daily basis and are therefore used to watching the videos in full screen. Our framework did not allow this feature as the additional upscaling of the video could have influence on the perceived QoE. This was also one of the issues users commented. In this case, information about this prior the test could help the user understand the reason behind. Another option would be to offer the participant a dummy task to get familiar with what he is asked to do.

C. Comments and consistency

One of our questions was also, if the users who gave us consistent and reliable answers would also leave a feedback. The results of the analysis showed, that in total we had the overall consistency rate of approx. 46%, in the case of comments, the consistency rate was only 35%. Hence, the reliable users were not more likely to leave an optional comment.

In Fig.3 we can see the consistency rate in percents among the different groups of comments. We can observe, that the lowest consistency rate was monitored in the group of the users, who left simple comment as *Thanks*, *Good job* etc. where the highest is in the users who left comments regarding the actual experienced they had during the fulfilling the task.

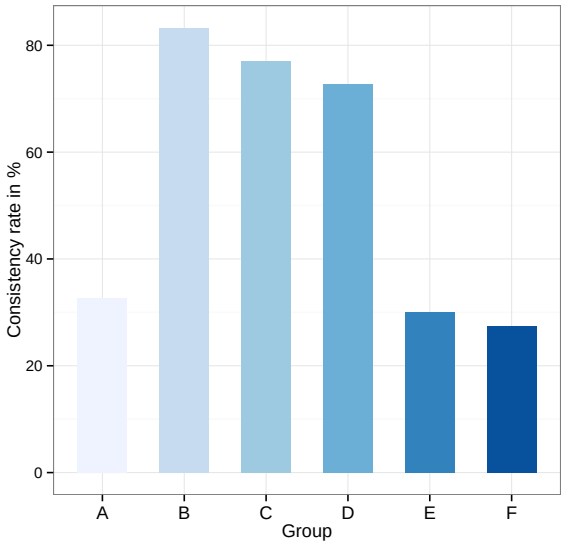


Fig. 3: Consistency rate based on groups of comments.

IV. Conclusion

In this paper, we analyzed the comments the participants left after fulfilling a crowdsourced task on QoE. We found out, that for future studies, it might be beneficial to better inform the user about what he will do and why. For special cases, even a demo of how the task looks like might be recommended.

Acknowledgment

This paper was supported by the BUT project no. FEKT-S-14-4426. The described research was performed in laboratories supported by the SIX project; no. CZ.1.05/2.1.00/03.0072, the operational program Research and Development for Innovation. Research described in this paper was financed by Czech Ministry of Education in frame of National Sustainability Program under grant LO1401. For research, infrastructure of the SIX Center was used.

References

[1] HOSSFELD, Tobias, Christian KEIMEL, Matthias HIRTH, Bruno GARDLO, Julian HABIGT, Klaus DIEPOLD and Phuoc TRAN-GIA. Best Practices for QoE Crowdtesting: QoE Assessment With Crowdsourcing. *IEEE Transactions on Multimedia*. 2014, **16**(2), 541-558. DOI: 10.1109/TMM.2013.2291663.

[2] ANEGEKUH, Louis, Lingfen SUN a Emmanuel IFEACHOR. A screening methodology for crowdsourcing video QoE evaluation. *2014 IEEE Global Communications Conference*. IEEE, 2014, , 1152-1157. DOI: 10.1109/GLOCOM.2014.7036964.

[3] SEUFERT, Michael, Ondrej ZACH, Tobias HOSSFELD, Martin SLANINA and Phuoc TRAN-GIA. Impact of test condition selection in adaptive crowdsourcing studies on subjective quality. *2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2016, , 1-6. DOI: 10.1109/QoMEX.2016.7498939.

[4] ZACH, Ondrej. Let the Crowd Be Heard - A Crowdsourcing- based Study on Video Coding. *Proceedings of the 22nd Conference STUDENT EEICT 2016.* Brno, 2016. pp. 390-394.

[5] ZACH, Ondrej, SEUFERT, Michael, HIRTH, Mathias, SLANINA, Martin and Phuoc TRAN-GIA. On use of crowdsourcing for H.264/AVC and H.265/HEVC video quality evaluation,*27th International Conference Radioelektronika (RADIOELEKTRONIKA)* Brno, 2017, pp. 1-6. DOI: 10.1109/RADIOELEK.2017.7937581

[6] SEUFERT, Michael. ZACH, Ondrej.; SLANINA, Martin, and Phuoc TRAN-GIA, Unpertrubed Video Streaming QoE Under Web Page Related Context Factors. *Proceedings of 9th International conference QoMEX 2017*. 2017. p. 1-6. 10.1109/QoMEX.2017.7965675.

[7] HOSSFELD, Tobias, Michael SEUFERT, Matthias HIRTH, Thomas ZINNER, Phuoc TRAN-GIA and Raimund SCHATZ. Quantification of YouTube QoE via Crowdsourcing. *2011 IEEE International Symposium on Multimedia*. IEEE, 2011, , 494-499. DOI: 10.1109/ISM.2011.87.

[8] ZACH, Ondrej and Eva KLEJMOV, E. When the Numbers Speak - Hidden Statistics of a Crowdsourced QoE Study. *Proc. of student conference Blansko 2016*. Brno, 2016. pp 102-105. ISBN: 978-80-214-5389- 0.