Advances in Intelligent Systems and Computing 284

Ewa Pietka
Jacek Kawa
Wojciech Wieclawek *Editors*

Information Technologies in Biomedicine, Volume 4

Springer

Advances in Intelligent Systems and Computing

Volume 284

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

For further volumes:
http://www.springer.com/series/11156

About this Series

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing.

The publications within "Advances in Intelligent Systems and Computing" are primarily textbooks and proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Advisory Board

Chairman

Nikhil R. Pal, Indian Statistical Institute, Kolkata, India
e-mail: nikhil@isical.ac.in

Members

Rafael Bello, Universidad Central "Marta Abreu" de Las Villas, Santa Clara, Cuba
e-mail: rbellop@uclv.edu.cu

Emilio S. Corchado, University of Salamanca, Salamanca, Spain
e-mail: escorchado@usal.es

Hani Hagras, University of Essex, Colchester, UK
e-mail: hani@essex.ac.uk

László T. Kóczy, Széchenyi István University, Győr, Hungary
e-mail: koczy@sze.hu

Vladik Kreinovich, University of Texas at El Paso, El Paso, USA
e-mail: vladik@utep.edu

Chin-Teng Lin, National Chiao Tung University, Hsinchu, Taiwan
e-mail: ctlin@mail.nctu.edu.tw

Jie Lu, University of Technology, Sydney, Australia
e-mail: Jie.Lu@uts.edu.au

Patricia Melin, Tijuana Institute of Technology, Tijuana, Mexico
e-mail: epmelin@hafsamx.org

Nadia Nedjah, State University of Rio de Janeiro, Rio de Janeiro, Brazil
e-mail: nadia@eng.uerj.br

Ngoc Thanh Nguyen, Wroclaw University of Technology, Wroclaw, Poland
e-mail: Ngoc-Thanh.Nguyen@pwr.edu.pl

Jun Wang, The Chinese University of Hong Kong, Shatin, Hong Kong
e-mail: jwang@mae.cuhk.edu.hk

Ewa Pietka · Jacek Kawa
Wojciech Wieclawek
Editors

Information Technologies in Biomedicine, Volume 4

Springer

Editors
Ewa Pietka
Faculty of Biomedical Engineering
Silesian University of Technology
Gliwice
Poland

Wojciech Wieclawek
Faculty of Biomedical Engineering
Silesian University of Technology
Gliwice
Poland

Jacek Kawa
Faculty of Biomedical Engineering
Silesian University of Technology
Gliwice
Poland

ISSN 2194-5357 ISSN 2194-5365 (electronic)
ISBN 978-3-319-06595-3 ISBN 978-3-319-06596-0 (eBook)
DOI 10.1007/978-3-319-06596-0
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2008926730

© Springer International Publishing Switzerland 2014
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

In clinical application we deal with problems that have to be solved in a fast and objective way. However, human observation is influenced by internal (coming form the observer) as well as external (often independent from the observer) impacts. The objectivity of classification is restricted by the receptivity of human senses which are influenced by the experiences or level of training, psychological conditions (tiredness, haste, etc.), as well as external conditions (lighting, destructive noise, etc.) A failure in perception questions the entire recognition process. The recognition process itself, influenced also by the above mentioned conditions, may cause a slowdown and/or lead to a false diagnosis.

New computerized approaches to various problems have become critically important in healthcare. Computer assisted diagnosis has been extended towards a support of the clinical treatment. Mathematical information analysis, computer applications together with medical equipment and instruments have become standard tools underpinning the current rapid progress with developing Computational Intelligence. We are witnessing a radical change as technologies have been integrated into systems that address the core of medicine, including patient care in ambulatory and in-patient setting, disease prevention, health promotion, rehabilitation and home care. A computerized support in the analysis of patient information and implementation of a computer aided diagnosis and treatment systems, increases the objectivity of the analysis and speeds up the response to pathological changes.

This book aims to present a variety of state-of-the-art information technology and its applications to the networked environment to allow robust computerized approaches to be introduced throughout the healthcare enterprise. Image and signal analysis are the traditional parts that deal with the problem of data processing, recognition and classification. Bioinformatics has become a dynamically developed field of computer assisted biological data analysis. Patients' safety and shortening of the rehabilitation time requires a more rapid development of minimally invasive surgery supported by image navigation techniques. Home care, remote rehabilitation assistance, safety of the elderly require new areas to be explored in telemedicine and telegeriatrics.

This book set is a continuation of a book series. This set contains two volumes. *Information Technologies in Biomedicine, Volume 3* discusses Image analysis techniques and their applications in healthcare, as well as some Bioinformatics issues. *Information Technologies in Biomedicine, Volume 4* consists of six parts including Computer Aided Surgery, Telemedicine, Telegeriatrics,

We would like to express our gratitude to the authors who contributed their original research papers as well as the reviewers for their valuable comments.

Ewa Pietka

Contents

I Computer Aided Surgery

II Telemedicine

III Telegeriatrics

IV Experimental Bioengineering

V Signal Processing

VI Biomechanics

Part I

Computer Aided Surgery

Laparoscopic Liver Surgery

Stanislav Czudek

Lékařská Fakulta Ostravské Univerzity,
Ostrava, Česká Republika
sczudek@seznam.cz

Abstract. A survey of various laparoscopic liver surgery procedures is presented. First, depending on the tumor location, three patient positioning are discussed. Then, the resection procedures of benign and malignant lesions are presented. Pre- and intersurgical imaging techniques including 3D imaging and ultrasound procedures are often employed. Finally, the survival rate reported by various authors after nearly 15 years of experiences concludes the study.

Keywords: liver laparoscopy, liver surgery.

1 Introduction

There has been enormous growth of minimally invasive surgery since the first laparoscopic cholecystectomy. Commonly accepted laparoscopic procedures have now come to include bariatric and anti reflux procedures, distal pancreatectomy, splenectomy, hernia repair and colon resection. The earliest reports of laparoscopic liver surgery were limited to wedge resections for staging or isolated metastase [25]. Laparoscopic liver resection finally started to gain serious widespread attention after publication of Cherqui's initial thirty patient experience [10]. Since that time, the field has seen explosive growth, with over many cases now described in the world literature [29]. Despite its widespread acceptance, laparoscopic liver resection remains a daunting technical challenge suited to a relatively small number of centers that have taken the time and effort to develop concurrent expertise in both open hepatic surgery and laparoscopy. Once these hurdles are overcome; however, laparoscopic liver resection is a safe and highly effective procedure offering numerus patient benefits.

1.1 Benign Tumors

Benign liver tumors represent a diagnostic and therapeutic challenge. Traditionally, a highly conservative approach to benign hepatic tumors has been favored, owing to the historically high morbidity and mortality associated with open liver surgery. As operative and anesthetic techniques have improved, these hurdles have come down. Despite the increased safety of hepatic surgery, the indications for resection of benign hepatic tumors have changed little: symptomatic lesions,

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_1, © Springer International Publishing Switzerland 2014

asymptomatic lesions at high risk of rupture or malignant degeneration, and inability to exclude malignancy nonoperatively. Because of concerns over oncologic adequacy, benign lesions represent the ideal starting point for a laparoscopic liver surgery program. Despite the attractiveness of minimally invasive surgery; however, surgeons should be cautioned that the ability to perform a laparoscopic resection should not change the indications for operation.

1.2 Hemangioma

Hemangioma represents the most common benign liver tumor, accounting for 5-20% of liver lesions [7]. These tumors typically occur in females in the third through fifth decades. Symptoms typically do not occur until the tumors grow relatively large (>5cm), and typically consist of abdominal pain resulting from stretching of Glisson's capsule. There have been reports of spontaneous, traumatic, or iatrogenic rupture. There is no potential for malignant degeneration with hepatic hemangioma. Hemangiomas demonstrate a typical pattern of enhancement on triple phase contrast enhanced CT. The lesion appears as a well circumscribed hypodense mass with peripheral enhancement in the arterial phase that will progress toward the center of the lesion. This pattern is typically known as centripetal enhancement. Sensitivity of triple phase CT has been reported from 75-85% with specificity of 75-100% [37]. Even better results have been reported with the use of magnetic resonance imaging, with reported sensitivity and specificity of up to 95% and 100%, respectively [35]. Because of the highly vascular nature of these tumors, percutaneous biopsy of suspected hemangiomas is contraindicated. As there is no malignant potential, symptomatic disease is the only generally accepted indication for surgical resection of hemangiomas. It should again be stressed that the availability of laparoscopy should not extend the indications for operation to asymptomatic patients. If pain is the indication for surgery, a thorough diagnostic workup is imperative to rule out other sources before attributing the symptoms to the hemangioma. The indication for surgery is more clear cut for large ruptured hemangioma, with patients often presenting in shock. Because of the dire consequences of rupture of large hemangioma, some surgeons would advocate the prophylactic resection of large lesions in patients with high risk occupations in areas remote from medical care. This opinion is controversial and should not be broadly applied.

1.3 Focal Nodular Hyperplasia

Focal nodular hyperplasia (FNH) is generally thought to arise as a hyperplastic proliferation of cells arising from an arterial malformation. This malformation may be congenital in nature such as telangiectasia or arteriovenous malformation, or may result from vascular injury [30,38]. The polyclonal nature of these lesions has significant impact on the radiographic evaluation of FNH, as it is the only common benign lesion that appears hot on Technetium sulfur colloid scan. This is from increased uptake of tracer in Kuppfer cells present within the lesion. FNH is typically an incidentally discovered lesion in women of late

child bearing age, presenting most commonly from age 30 to 50. The female to male ratio has been reported at up to 8:1 [28]. Unlike hepatocellular adenoma, FNH is not influenced by oral contraceptive use. The radiographic appearance of focal nodular hyperplasia is typically diagnostic. On triple phase computed tomography, FNH will show transient enhancement on arterial phase. On delayed imaging, the characteristic central scar then becomes hyperenhancing. This central scar represents the vascular pedicle of the lesion and is pathognomonic. The most common diagnostic difficulty is distinguishing FNH from adenoma, which may best be achieved by contrast enhanced MRI. In this setting, sensitivity and specificity can reach 97% and 100%, respectively [36]. FNH is asymptomatic in upwards of 80% of cases [7]. In very rare instances, these lesions may present with hemorrhage. There are no reported cases of malignant degeneration of FNH thus far. Because of this, there is no indication for resection of asymptomatic lesions, regardless of the size and number of lesions. Surgical resection is reserved for the rare cases in which the lesion is symptomatic or when the diagnosis is not secure.

1.4 Hepatic Adenoma

Hepatic adenoma is a less common benign hepatic neoplasm, arising most commonly in women of child bearing age. There is a strong association between development of these lesions and oral contraceptive or androgenic steroid use. While the incidence is 0.1 per year per 100,000 patients who don't use oral contraceptives, there is a marked increase to up to 4 per 100,000 oral contraceptive users [30]. The introduction of modern contraceptives with lower estrogen content has led to a decrease in incidence [34]. Less common risk factors for the development of hepatocellular adenoma include glycogen storage disease type I and type III [27]. Though typically presenting as solitary lesions, adenoma may also be present as multiple lesions. Hepatic adenomas can grow quite large, with tumors of up to 30cm reported in the literature. Ultrasonography typically lacks diagnostic utility for adenomas, which can range from hypo to hyper-echoic. Reported sensitivity of ultrasound is only around 30% [14]. The CT appearance is that of a discrete, hypodense lesion showing enhancement on arterial phase followed by washout on later images. T1 weighted MRI will show a hypoto hyperintense lesion, while T2 images will show a lesion that is more isointense. Patients with hepatocellular adenoma are more likely to present with symptomatic disease than those with FNH. Epigastric or right upper quadrant pain is present in 25-50% of patients [7]. Spontaneous hemorrhage is also relatively common with these lesions, occurring in over 20% of patients. These complications are more likely to occur in men and with lesions greater than 5cm in diameter [15]. Perhaps the most feared complication of hepatocellular adenoma is malignant degeneration. The risk has been reported in the range of 8-10% [15,30]. Although 5cm is the generally accepted size at which malignant degeneration becomes a concern, cases have been reported in lesions as small as 4cm [27]. There is also a greater risk of malignant degeneration in males and in patients with the metabolic syndrome. Malignancy within adenomas is typically

discovered only after surgical resection. In the case of small adenomas in the setting of oral contraceptive use, a period of observation following the cessation of contraception is warranted. Surgical resection in this setting is then reserved for lesions which fail to regress or continue to grow after stopping the offending medication. As with other benign lesions, symptomatology that can clearly be attributed to the adenoma is also an indication for surgical resection. The presence of multiple adenomas, or adenomatosis, is an arbitrary distinction rather than a distinct pathologic subtype, thus indications for resection are the same as for solitary adenoma. Because of the well defined risk of malignant degeneration, there are also cases where resection of asymptomatic lesions is warranted. Generally accepted criteria include adenomas greater than 5cm in size, or any adenoma in a male, regardless of size [15].

2 Laparoscopic Liver Benign Lesions Surgery

There are three commonly used patient positions employed in laparoscopic liver resection: supine, lateral decubitus, and the so-called French position in which the patient is supine with the legs in stirrups and the surgeon is positioned between the patient's legs. The appropriate position is determined based on the location of the tumor, and the surgical technique to be employed. The French position has the advantage of allowing the surgeon to operate with both hands while assistants can retract from either side of the table. The supine position is best employed when approaching lesions on the left lobe or right anterior sector of the liver. The lateral decubitus position places the patient recumbent on their left side at an angle of sixty degrees. This position allows access to the posterior segments of the right liver, as the left side down positioning prevents the liver from falling dependently into the operative field. When a hand port is to be employed, it is generally placed in the right upper quadrant as dictated by the position of the tumor being resected. A number of parenchymal transection techniques have been described in the literature, with none of them showing clear superiority over the others. Which technique is ultimately chosen thus becomes dependent upon the individual surgeon's comfort level with a given technique. Here I describe two of the more common strategies: electrosurgical dissection and stapler hepatectomy.

Electrosurgical transection techniques rely upon the surgeon's ability to operate two devices simultaneously. The surgeon should use a device such as the Harmonic Scalpel or Enseal in the dominant hand. This device is used to incise Glisson's capsule and for the majority of parenchymal transection. The device should not be fully introduced into the parenchyma to prevent tearing of large vessels. When aktive bleeding is encountered, it is immediately controlled with bipolar cautery forceps which are held in the surgeon's other hand. Larger vessels require the use of laparoscopic clips. Our group has favored the use of stapler hepatectomy. This technique provides the advantage of more rapid parenchymal transection, without the need for prior control of individual hepatic vessels. The first centimeter of parenchyma is relatively devoid of major vessels, and

is incised with electrosurgical devices as described above. The dissection then proceeds using the thin blade of the stapler as a dissector. Care must be taken to avoid inadvertent manipulation of the stapler during firing, which can lead to tearing of major vessels and subsequent hemorrhage. The use of hand assistance is helpful in stabilizing the stapler to prevent such complications. We have preferred the use of a 25 mm vascular staple load for parenchymal transection. When intraoperative hemorrhage is encountered, the presence of a hand in the abdomen is highly beneficial in allowing digital control of bleeding vessels prior to attaining definitive hemostasis. Should conversion be necessary during a pure laparoscopic procedure, it should initially be to a hand assist method rather than to full laparotomy. In all cases, conversion should not be viewed as a failure or complication, but rather as a measure of prudent judgment [6].

3 Laparoscopic Resection of Liver Malignant Lesions

After becoming comfortable with resection of benign lesions, the logical progression in the development of a laparoscopic liver program is the resection of malignant lesions. These lesions require an increased degree of skill on the part of the surgeons in order to attain adequate margins and maintain oncologic adequacy. The presence of cirrhosis in the setting of HCC or steatohepatitis following neoadjuvant chemotherapy for colorectal metastasis make proper patient selection and timing of operation critical. The consideration of adjunctive techniques such as transarterial chemoemboliztion for preoperative downstaging also becomes important.

Here, we will discuss laparoscopic management of the two most common malignant hepatic tumors: colorectal metastases and hepatocellular carcinoma. Colorectal metastases are the most common malignant hepatic tumor. Results following open resection of these lesions have been excellent, with 5 year survival rates exceeding 50% in many centers [18]. Such outcomes have set a high standard by which laparoscopic resection must be measured. The adoption of laparoscopy to this field has been hindered by concerns of tumor seeding at port sates and the possibility of missing extrahepatic lesions by inadequate inspection of the peritoneal cavity [19,21]. These hurdles have slowly been brought down, and laparoscopic resection is now a standard part of the therapeutic arsenal for hepatic malignancy. Patient selection criteria for laparoscopic resection of colorectal metastases are similar to those applied for open resection. Initial evaluation requires precise definition of tumor anatomy and exclusion of extrahepatic disease. We favor triple phase CT as the initial radiographic evaluation. When combined with digital arterial reconstruction, evaluation of aberrant vascular anatomy, which can be present in nearly half of all patients, is afforded. Evaluation of baseline liver function is performed with evaluation of bilirubin, INR, and albumin. A thorough history and physical exam is necessary to assess general fitness for major abdominal surgery. Tumor resectability is defined by the SSAT as an expected negative margin resection with preservation of at least 2 contiguous hepatic segments with adequate inflow, outflow, and biliary drainage

and a future liver remnant of more than 20% for normal parenchyma [9]. The most critical factor to producing positive outcomes is the attainment of negative operative margins (R0 resection). Facility with laparoscopic intraoperative ultrasound is a must for surgeons approaching malignant liver lesions, allowing for precise definition of tumor anatomy and planning of resection planes. As long as negative microscopic margine are obtained, there does not appear to be a minimum necessary margin width [32].

The approach to synchronous disease has received considerable attention, as it will be present in up to 25% of patients with colorectal liver metastases [26]. There are three possible surgical strategies in this setting: the classic approach of colorectal resection followed by hepatectomy, a simultaneous resection of colorectal and hepatic disease, and a reverse strategy of metastasectomy followed by primary tumor resection. The drawback of the classic strategy is the delay in metastasectomy while patients receive adjuvant therapy. The combined strategy eliminates this delay, at the cost of greater surgical insult with possibly higher morbidity. Brouquet's analysis of all three strategies found similar morbidity, mortality, and survival across groups, showing that no approach is clearly superior for all patients [5].

With increasing worldwide experience of laparoscopic resection of colorectal metastases, the oncologic integrity of laparoscopy compared with open techniques has been shown to be comparable. Nguyen's review of the world literature found only one case of port site recurrence, which occurred in a case of metastatic renal cell carcinoma that ruptured prior to resection [29]. Castaing's comparison of 60 patients undergoing laparoscopic resection and 60 patients undergoing open resection provided the first evidence of long term efficacy of laparoscopic resection for colorectal metastases. Five year survival in the laparoscopic group in this series was 62%, which was comparable to the 56% five year survival in the open group. There was no difference in width of resection margins between groups, while the laparoscopic group included a greater percentage of patients undergoing combined hepatic and colorectal resection [8]. Such results confirm that laparoscopic resection is a safe and effective alternative to open surgery for hepatic colorectal metastases.

Hepatocellular carcinoma (HCC) is the sixth most common malignancy and the third most common cause of cancer death worldwide [31]. In the United States, where chronic hepatitis C infection is the main risk factor, there has been an increase in the incidence of HCC over the past several decades [16]. Most patients present with relatively advanced disease, making curative treatment such as resection and liver transplantation applicable in only 30-40% of patients. One of the major limiting factors in preventing resectability is impaired hepatic fiction. Thus, appropriate patient selection becomes paramount in achieving successful outcomes. Because of these limitations, the role of laparoscopic liver resection has remained more limited than for other disease states. Much of the patient selection process for resection of HCC centers around assessment of the underlying liver parenchyma. The Child-Pugh classification system provides a rough framework from which to base the selection process. In

generally, Child A patients are able to tolerate limited forms of resection, while Child B and C patients are typically referred for more palliative procedures such as systemic therapy or transarterial chemoembolization. In the West, assessment is directed at determining the presence of significant portal hypertension. Generally, patients with hepatic-venous pressure gradient of less than 10, esophageal varices of no greater than grade 1, and platelet counts of over 100,000 are considered acceptable risk. In addition, bilirubin levels must be normal. A common technique in Eastern centers is the assessment of indocyanine green clearance rate (ICG). This technique involves the injection of an organic dye which is then measured in the peripheral blood after a 15 minute interval. Clearance of the dye is used as a surrogate for hepatic metabolic function. ICG retention of no more than 10-20% is considered to be acceptable. Using this technique in 1056 consecutive patients with normal bilirubin and no ascites, Imamura has been able to achieve hepatic resection with zero operative mortality [20].

Advances in imaging technology have lead to the increasing use of systemic liver volumetry as a preoperative risk assessment tool. A future liver remnant to standard liver volume ratio of greater than 20% is considered safe in patients with healthy liver parenchyma, while ratios of 30-40% are considered necessary for patients with compensated cirrhosis. An insufficient future liver remnant may be addressed with the use of adjunctive techniques such as portal vein embolization, which will be discussed in greater detail in the section on resection in cirrhotics.

Tumor related factors that preclude surgical resection include extrahepatic disease and invasion of the main portal vein, vena cava, and common hepatic artery. Multinodular disease that can't be resected with an adequate future liver remnant is also a relative contraindication to resection, although there is a role for resection of the dominant lesion with radiofrequency ablation of the remaining disease in highly selected cases. Although size alone is not a criteria for resectability, there is a practical limit to the size of lesion that can be safely approached laparoscopically. The recent international position statement for laparoscopic liver surgery recommends limitation of the laparoscopic approach to tumors $<$5cm in diameter for all but the most experienced of centers [6].

Unlike the case of hepatic colorectal metastases, there does appear to be a benefit to wider surgical margins in patients with HCC. For patients with solitary HCC lacking vascular invasion, a margin of at least 2cm has proven beneficial in a randomized controlled trial setting. Furthermore, the tendency of HCC to spread via the portal venous system favors the use of planned anatomic resection in patients with adequate hepatic reserve. The inability to perform anatomic resection should not be considered a contraindication, however, as more limited resection as been shown to be beneficial in the setting of cirrhosis [33]. Despite the limitations imposed by the greater difficulties in technical resection and patient selection, laparoscopic resection has proven to be a safe and effective alternative to open surgery in appropriately selected patients. Lai has demonstrated 5 year survival of 50%, with disease free survival of 36%, while Dagher has shown 5 year overall and disease free survival of 64.9% and 32.2%, respectively

[12,23]. Others have shown laparoscopic resection to be associated with lower morbidity and postoperative ascites compared to open resection [4]. Although hepatocellular carcinoma in the setting of cirrhosis represents the most difficult of diseases to approach via laparoscopy, these results show that the technique is safe and effective when performed in centers that have acquired the appropriate experience.

4 Laparoscopic Treatment of Patients with Cirrhotic Liver

As noted above, the cirrhotic patient represents a unique challenge to the laparoscopic liver surgeon. The possibility of postoperative liver failure resulting from inadequate remnant liver function is a dreaded complication to be avoided at all costs. One technique that can potentially prevent this problem is the use of preoperative portal vein embolization (PVE). The effectiveness of PVE is based on the remarkable regenerative capacity of the liver. The technique involves occlusion of the tumor bearing segments of the liver, which induces hypertrophy in the remaining hepatic segments. Generally, reimaging 6 weeks after PVE is performed to assess the adequacy of hypertrophy to provide an adequate future liver remnant. Failure to achieve adequate hypertrophy indicates a severely diseased liver that is not amenable to resection.

A meta-analysis of PVE has been found that the procedure is safe and able to induce adequate hypertrophy to reduce post resection liver failure in a considerable proportion of patients [2]. Preoperative PVE is currently recommended in cirrhotic patients with predicted future liver remnant of less than 40%. For centers using ICG retention, values of 10-19% with a FLR of 40-60% also represents an indication for portal vein embolization [33].

For cirrhotic patients able to undergo liver resection, laparoscopy provides a number of unique benefits. The smaller incisions cause less disruption of the abdominal wall collateral circulation. As complete evacuation of ascites is not necessary for a laparoscopic procedure, intraoperative fluid shifts are lessened. This contributes to the reduction in postoperative ascites seen with laparoscopy compared to open hepatectomy [13,17]. Another unique benefit is the reduced adhesion formation following laparoscopic surgery. For patients undergoing resection of HCC, salvage transplantation remains an important option for recurrences that are within the Milan criteria. Laurent found that liver transplants following laparoscopic compared to open resection were performed in less time, with less blood loss and transfusion requirement [24]. Similarly, Belli has found repeat hepatectomy following initial laparoscopic resection to be faster and safer, with less blood loss and risk of visceral injury [3].

The recent international consensus conference on laparoscopic liver surgery has developed guidelines for the establishment and credentialing of a laparoscopic liver surgery program [6]. Prior to embarking upon beginning a program in laparoscopic liver surgery, it is necessary to acquire experience with both advanced laparoscopy and open hepatic surgery. These requirements have made

the widespread adoption of laparoscopic liver surgery appropriately slow. As advanced laparoscopy becomes an increasingly important part of general surgery training programs, these prerequisites will become less of a hurdle, with the expected more rapid acceptance of laparoscopic liver surgery. After establishing the necessary expertise in laparoscopy and open hepatic surgery, the ideal starting point is small, benign lesions in the periphery of the liver. Extensive use of hand assistance is also critical in reducing the learning curve. Koffron has described the hybrid technique, in which mobilization of the liver is performed laparoscopically, and parenchymal transection is then performed in an open fashion through the hand port incision [22]. He has termed this approach "laparoscopic liver surgery for everyone", and we agree that this approach represents an ideal starting point for a laparoscopic liver program.

Once comfortable with performing more limited resections, the next step in development is the performance of major, anatomic resections. In this setting, the left lateral segmentectomy is the ideal starting point. Although much attention is given to the parenchymal transection phase, it should be noted that the greatest risk for vascular injury and subsequent conversion to an open procedure is actually during the mobilization phase. The most commonly injured vessel in this setting is the phrenic vein, which must be carefully identified and avoided. Conversion, as we have emphasized previously, should not be viewed as a failure or complication. Instead, the decision to convert to an open or hand assisted procedure rather than continue with a potentially unsafe situation laparoscopically is a mark of good surgical judgment.

Experience with resection of lesions located in the peripheral segments of the liver provides a foundation of skills, including mobilization, transection, hemostasis, and laparoscopic ultrasound. Once this fundamental skill set has been developed thoroughly, the surgeon is then able to proceed to more difficult lesions. At this point, malignant and/or large lesions located in the right and posterior segments of the liver can then be approached in the culmination of programmatic development. We have found that facility with minor resections can be achieved in 30 to 50 cases. More difficult resections such as formal lobectomy and right posterior resection require an additional 60 to 80 cases to master. Thus, the road to development of a laparoscopic liver resection program is long and often arduous, but is highly rewarding to both the surgeon and the patient when properly travelled.

5 Robotic Surgery

Robotic surgery is a new and promising technique. Numerous authors believe it as revolutionary as laparoscopy or thoracoscopy [1]. Multiple instruments, very useful in some cases, have been introduced. They are still being developed and sometimes far from perfection. Very many centers worldwide perform robotic procedures and publish their results. The cost of the procedure versus the benefit for the patient are the most often mentioned issues. Radical prostatectomy and some cardiac surgical procedures appear to be the most frequent indications.

Robotic-assisted surgery is a new method requiring conscientious analysis of the previous results. The method is still being developed and it is difficult to state which indications are beneficial for patients. Surgeons must try to avoid marketing operations behind some of the indications of companies producing robotic instrumentation.

Advantages of robotic surgery surely include benefits of mini-invasive procedures (laparo- and thoracoscopy): small incisions, less painful sensation, better recovery and fewer complications connected with it, reduced blood loss, higher precision of the surgeon. It is also economically beneficial due to shorter hospital stay, reduced consumption of analgesics, decreased number of transfusions.

On the other hand, the high cost of a robotic procedure appears to be the biggest, current problem in the world. Although, richer countries may afford wider utilization of such techniques (the USA), scrupulous economic analysis has to be done: costs versus benefits for the patient versus health care system in the particular country [11].

6 Conclusion

Nearly 15 years after first being described, laparoscopic liver resection has been gradually gaining acceptance in a number of centers worldwide. As the necessary skills in advanced laparoscopy and hepatic surgery become more widespread, we anticipate that the further adoption of laparoscopic liver resection will increase more rapidly. The maturation of long term series have proven the oncologic adequacy of the laparoscopic approach in a variety of settings. With the development of a greater number of surgeons who are proficient in laparoscopic liver surgery, many more patients will benefit from decreased blood loss, less postoperative pain, and shorter lengths of stay. From being a novel procedure practiced in only a handful of centers worldwide, laparoscopic liver resection is now established as a safe and effective technique in the therapeutic decision tree for patients with surgical disease of the liver. We believe that this acceptance will continue to grow to the point that the laparoscopic approach will, as has been seen with colon resection, eventually be adopted as the standard of care in appropriately selected patients.

References

1. (387 articles) (January 4-May 27, 2012), `http://www.RoboticSurgeryNewsArchive.com`
2. Abulkhir, A., et al.: Preoperative portal vein embolization for major liver resection: a metaanalysis. Annals of Surgery 247(1), 49–57 (2008)
3. Belli, G., et al.: Laparoscopic redo surgery for recurrent hepatocellular carcinoma in cirrhotic patients: feasibility, safety, and results. Surgical Endoscopy 23(8), 1807–1811 (2009a)
4. Belli, G., et al.: Laparoscopic and open treatment of hepatocellular carcinoma in patients with cirrhosis. British Journal of Surgery 96(9), 1041–1048 (2009b)

5. Brouquet, A., et al.: Surgical strategies for synchronous colorectal liver metastases in 156 consecutive patients: classic, combined or reverse strategy? Journal of the American College of Surgeons 210(6), 934–941 (2010)
6. Buell, J.F., et al.: The international position on laparoscopic liver surgery: The Louisville Statement, 2008. Annals of Surgery 250(5), 825–830 (2009)
7. Buell, J.F., et al.: Management of benign hepatic tumors. Surgical Clinics of North America 90(4), 719–735 (2010)
8. Castaing, D., et al.: Oncologic results of laparoscopic versus open hepatectomy for colorectal liver metastases in two specialized centers. Annals of Surgery 250(5), 849–855 (2009)
9. Charnsangavej, C., et al.: Selection of patients for resection of hepatic colorectal metastases: expert consensus statement. Annals of Surgical Oncology 13(10), 1261–1268 (2006)
10. Cherqui, D., et al.: Laparoscopic liver resections: a feasibility study in 30 patients. Annals of Surgery 232(6), 753–762 (2000)
11. Czudek, S.: Robotic surgery - a taste of Hollywood? Videosurgery and Other Miniinvasive Techniques 8(2), 95–98 (2013)
12. Dagher, I., et al.: Laparoscopic hepatectomy for hepatocellular carcinoma: a European experience. Journal of the American College of Surgeons 211(1), 16–23 (2010)
13. Dagher, I., et al.: Laparoscopic versus open right hepatectomy: a comparative study. American Journal of Surgery 198(2), 173–177 (2009)
14. Di, S.M., et al.: Natural history of focal nodular hyperplasia of the liver: an ultrasound study. Journal of Clinical Ultrasound 24(7), 345–350 (1996)
15. Dokmak, S., et al.: A single-center surgical experience of 122 patients with single and multiple hepatocellular adenomas. Gastroenterology 137(5), 1698–1705 (2009)
16. El-Serag, H.B., Mason, A.C.: Rising incidence of hepatocellular carcinoma in the United States. The New England Journal of Medicine 340(10), 745–750 (1999)
17. Gigot, J.F., et al.: Laparoscopic liver resection for malignant liver tumors: preliminary results of a multicenter. European Study Annals of Surgery 236(1), 90–97 (2002)
18. House, M.G., et al.: Survival after hepatic resection for metastatic colorectal cancer: trends in outcomes for 1,600 patients during two decades at a single institution. Journal of the American College of Surgeons 210(5), 744–745 (2010)
19. Hsu, T.C.: Intra-abdominal lesions could be missed by inadequate laparoscopy. American Surgeon 74(9), 824–826 (2008)
20. Imamura, H., et al.: One thousand fifty-six hepatectomies without mortality in 8 years. Archives of Surgery 138(11), 1198–1206 (2003)
21. Johnstone, P.A., Rohde, D.C., et al.: Port site recurrences after laparoscopic and thoracoscopic procedures in malignancy. Journal of Clinical Oncology 14(6), 1950–1956 (1996)
22. Koffron, A.J., et al.: Laparoscopic liver surgery for everyone: the hybrid method. Surgery 142(4), 463–468 (2007)
23. Lai, E.C., et al.: Minimally invasive surgical treatment of hepatocellular carcinoma: long-term outcome. World Journal of Surgery 33(10), 2150–2154 (2009)
24. Laurent, A., et al.: Laparoscopic liver resection facilitates salvage liver transplantation for hepatocellular carcinoma. Journal of Hepato-Biliary-Pancreatic Surgery 16(3), 310–314 (2009)
25. Lefor, A.T., Flowers, J.L.: Laparoscopic wedge biopsy of the liver. Journal of the American College of Surgeons 178(3), 307–308 (1994)

26. Martin, R.C., et al.: Simultaneous versus staged resection for synchronous colorectal cancer liver metastases. Journal of the American College of Surgeons 208(5), 842–850 (2009)
27. Micchelli, S.T., et al.: Malignant transformation of hepatic adenomas. Modern Pathology 21(4), 491–497 (2008)
28. Mortele, K.J., Ros, P.R.: Benign liver neoplasms. Clinics In Liver Disease 6(1), 119–145 (2002)
29. Nguyen, K.T., et al.: World review of laparoscopic liver resection-2,804 patients. Annals of Surgery 250(5), 831–841 (2009)
30. Paradis, V.: Benign liver tumors: an update. Clinics In Liver Disease 14(4), 719–729 (2010)
31. Parkin, D.M., et al.: Global cancer statistics, 2002. CA: A Cancer Journal for Clinicians 55(2), 74–108 (2005)
32. Pawlik, T.M., et al.: Effect of surgical margin status on survival and site of recurrence after hepatic resection for colorectal metastases. Annals of Surgery 241(5), 715–722 (2005) (discussion)
33. Rahbari, N.N., et al.: Hepatocellular carcinoma: current management and perspectives for the future. Annals of Surgery 253(3), 453–469 (2011)
34. Rooks, J.B., et al.: Epidemiology of hepatocellular adenoma. The role of oral contraceptive use. Journal of the American Medical Association 242(7), 644–648 (1979)
35. Semelka, R.C., et al.: Focal liver lesions: comparison of dual-phase CT and multisequence multiplanar MR imaging including dynamic gadolinium enhancement. Journal of Magnetic Resonance Imaging 13(3), 397–401 (2001)
36. Terkivatan, T., et al.: Focal nodular hyperplasia: lesion characteristics on state-of-the-art MRI including dynamic gadolinium-enhanced and superparamagnetic iron-oxideuptake sequences in a prospective study. Journal of Magnetic Resonance Imaging 24(4), 864–872 (2006)
37. Trotter, J.F., Everson, G.T.: Benign focal lesions of the liver. Clinics In Liver Disease 5(1), 17–42 (2001)
38. Wanless, I.R., et al.: On the pathogenesis of focal nodular hyperplasia of the liver. Hepatology 5(6), 1194–1200 (1985)

Assistance in Destroying Focal Lesions in the Liver Using Image Navigation

Dominik Spinczyk

Faculty of Biomedical Engineering,
Silesian University of Technology, Zabrze, Poland
dspinczyk@polsl.pl
http://ib.polsl.pl

Abstract. Article presents methodology of computer aided destroying focal liver lesions using imaging navigation system. Methodology allows generation of personalized anatomical model of patient's liver based on CAT screening with contract media of abdominal cavity, registration of edited anatomical model relative to real patient's position on operation block, presentation of surgical tools position during treatment on the liver's anatomical model and non-direct monitoring of ablation process of lesions. Liver anatomy model comprises organ's and focal lesions' surface and central lines of vascular structures in liver. Generation of personalized patient's model has been verified based on anonymized abdominal cavity CAT with contrast media screening cases. The registration methodology has been evaluated on imaging data and patient from the Second Department of Clinical Radiology and the Chair and Department of General, Transplant and Liver Surgery at the Medical University of Warsaw after acceptance of the ethics committee of the Medical University of Warsaw.

Keywords: ablation of liver tumours, image navigation, personalized model of liver's anatomy, computer aided diagnostics and therapy.

1 Introduction

Primary and metastatic liver tumours make an significant challenge for contemporary medicine. Among available techniques of destroying focal liver lesions there are resection techniques,which are the most effective and ablation techniques applied in cases where resection techniques can't be used [1]. Additional difficulty in case of destroying focal liver lesions using ablation techniques is big organ's movability during breathing, which is an effect of the fact that liver is adjacent to diaph. Organ displacements can reach a few centimetres, which causes target point to displace [2,3]. Imaging navigation systems aids treatments performance through presentation of treatment run on a spatial patient's anatomy model [4,5]. Model's task is to show the most important anatomical structures from the performed treatment point of view, target point and surgical tools positions visualization on anatomical model. Nowadays in clinical practice imaging

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_2, © Springer International Publishing Switzerland 2014

navigation systems are commonly used for treatments using stiff human body skeleton's elements as referral points, for example: skull's bones, spinal cord's elements etc. Imaging navigation of parenchymal organs is more challenging due to lack of fixed referral points. The aim of this work is the presentation of imaging navigation system methodology supporting destruction of non-resectional focal lesions in the liver using ablation techniques.

2 Material and Methods

After literature studies, simulation and pilot studies methodology of aiding of destroying focal liver lesions focused on ablation techniques were proposed, especially electrical ablation with radio frequencies were considered. Methodology consists of a few stages:

- Preparation of personalized patient's anatomy model.
- Treatment planning.
- Registration of patient's position during treatment in comparison to pre-treatment patient's anatomy model.
- Operation field visualization during treatment and ablation process monitoring.

2.1 Preparation of Personalized Patient's Anatomy Model

Creation of personalized anatomical liver's model is based on patient's abdominal cavity CT examinations with contrast. Based on CAT images the following are found: liver's surface, vascular structures and volumes of focal liver lesions. For liver volume segmentation semi-automatic method proposed by Juszczyk is applied [6]. The method requires to indicate a single point within the spleen and one point inside the liver. The points indicated by the user are treated as a beginning of the rolling vector by which finding a common area in liver and spleen in two-dimensional cross-section is performed. After assigning the labels of object on the current cross-section, the shape of the image is transferred to the adjacent images. Stopping condition is to find the layer without the labeling of the object. Once you find the volume of the spleen, remove it from the initial volume and the result volume is the correct volume of the liver.

Vascular structure segmentation applies the modification of vascular Frangi filter [7], based on the analysis of the eigenvalues of the Hessian matrix. The method consists of two stages. At the first stage method determines the intensity characteristic of liver tissue, removes the structure of "hard" (e.g. high-density structures) and smoothes the image using anisotropic filtering. Then modified filter proposed by Rudzki is used [8], which is a modification of vascular Frangi filter, and introduces sigmoidal response functions of gaining a better sensitivity and higher immunity to noise.

The segmentation of lesions in the liver CT study raises a number of problems well-known and described in the literature [9] [10]. The changes, depending on the type, size, and stage of malignancy, are characterized by different levels

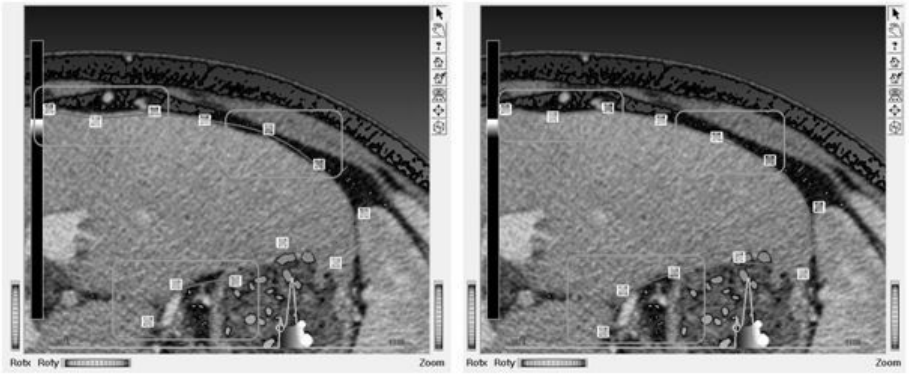


Fig. 1. Liver model evaluation: before correction - left, after correction - right

of intensity. Segmentation procedure of the growth area with adaptive criteria selection is applied, proposed by Badura [11]. Fuzzy expert system is used to monitore the shape, size and nature of the intensity of the tumor region.

Details on used segmentation methods were presented in the articles [8,6,11,12]. Building of personalized patient's model requires point inside liver parenchyma and one point for each focal lesion. Finding vascular structures does not require identifying specific points in liver's images, while applied algorithm uses the liver's mask obtained at the stage of segmentation. After indication of mentioned starting points personalized anatomical liver's model is generated automatically. Automatic results obtained in the stage of segmentation are processed to generate geometric models based on B-splines curves which makes it easier to use them at a later stage. After results receipt verified by a physician radiologist is planned, by evaluating the results of segmentation presented on the background of layered CT images the abdominal cavity. There exists possiblity to manually correct segmentation results by editing the splines' control points representing contours of segmented structures (Fig. 1). Vascular structures are presented in the form of central lines using smooth splines, they are also manually correctable (Fig. 2).

It is possible to block control points in the selected plane. This option is useful for manual correction of liver's contour segmentation results, where cuts of obtained model are presented in corresponding plane. In case of control points relocation in space, which is the case when correction of control points of vascular structures has been done (Fig. 2), movement of control points in any direction is possible.

2.2 Treatment Planning

To plan treatment personalized anatomical patient's liver model, presented in previous point, is used. Author has proposed a method of presentation of anatomical liver's model in the way suitable for treatment's planning. Model's presentation using selected layered DICOM images (2D presentation of selected cut of the chosen volume in corresponding layered CT cut - Fig. 1) and presentation

of the whole volume of abdominal cavity (volumetric imaging Fig. 3). Liver's outline and focal lesions are presented in the form of organ's surface cuts corresponding to specific DICOM images. Presentation of segmented structures on the background of layered images of abdominal cavity allows to assess resectability changes and select entry point on the patient's surface (Fig. 4), target point located inside segmented focal change (Fig. 3) and follow whether on the potential trajectory of surgical tool there are any organs which should be omitted during tools introduction. Selection of the point of entry is done by mouse on CT volume. Coordinates of indicated points are stored and used at a later stage of treatment.

2.3 Patient's Position Registration Relative to Pre-operation Anatomical Liver's Model

As registration process in the present case we understand explicit projection of position on the patient body to the corresponding position in the preoperative anatomical model built on the basis of layered CT images. Registration process takes place using markers placed on patient's abdominal cavity surface before CAT is performed. Markers placement is presented in Table 1 and Fig. 4. To perform registration it is necessary to determine position of the markers on CT images and the patient surface. Having a set of markers' coordinates in both coordinate systems Horn algorithm is used [13,14] in order to find stiff match between coordinate system of personalized anatomical liver's model and coordinate system related to real position of the patient on an operation block. Positioning of the various markers is found on CT images and stored in order to carry out calculations to find a rigid transform algorithm using Horn algorithm.

Stiff match between patient position and personalized anatomical model during treatment is not sufficient due to deformations of patient's body taking place which are effects of breathing movements and pressure of surgical tools. To take those things into consideration non-stiff registration is introduced which allows to take into the account local shape deformations. Non-stiff registration

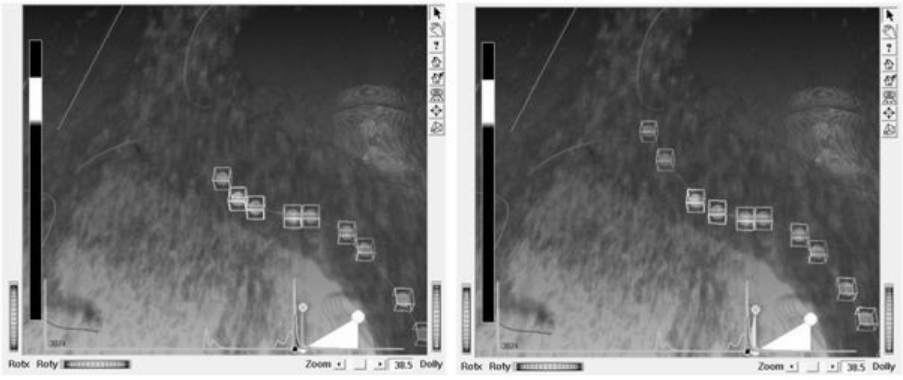


Fig. 2. Vessel skeleton evaluation: before correction - left, after correction - right

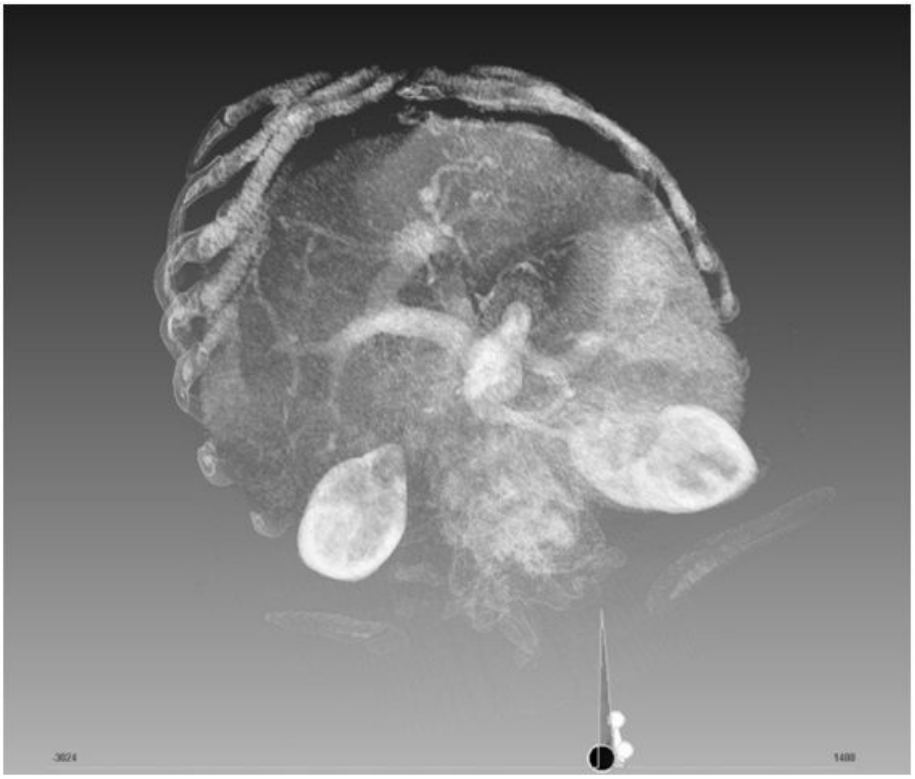


Fig. 3. Focal liver lesion - volumetric presentation

is based on so called deformation field, which is in turn based on a markers position interchanges, which are observed continuously during treatment. Based on deformation field correction of target point position is estimated. Deformation area allows to calculate target point position correction for various respiratory phases. More accurate definition of respiratory phase and the methodology for calculating deformation area and results obtained in various respiratory phases are presented in the paper [15], where obtained correction of deformation area was compared with the actual position of the target point, using as target point one of the markers, position of which was continuously tracked by position tracking system. In the paper averaged results obtained from 10 patients with any liver diseases from the Second Department of Clinical Radiology and the Chair and Department of General, Transplant and Liver Surgery at the Medical University of Warsaw, for various markers configuration in relation to target point were presented.

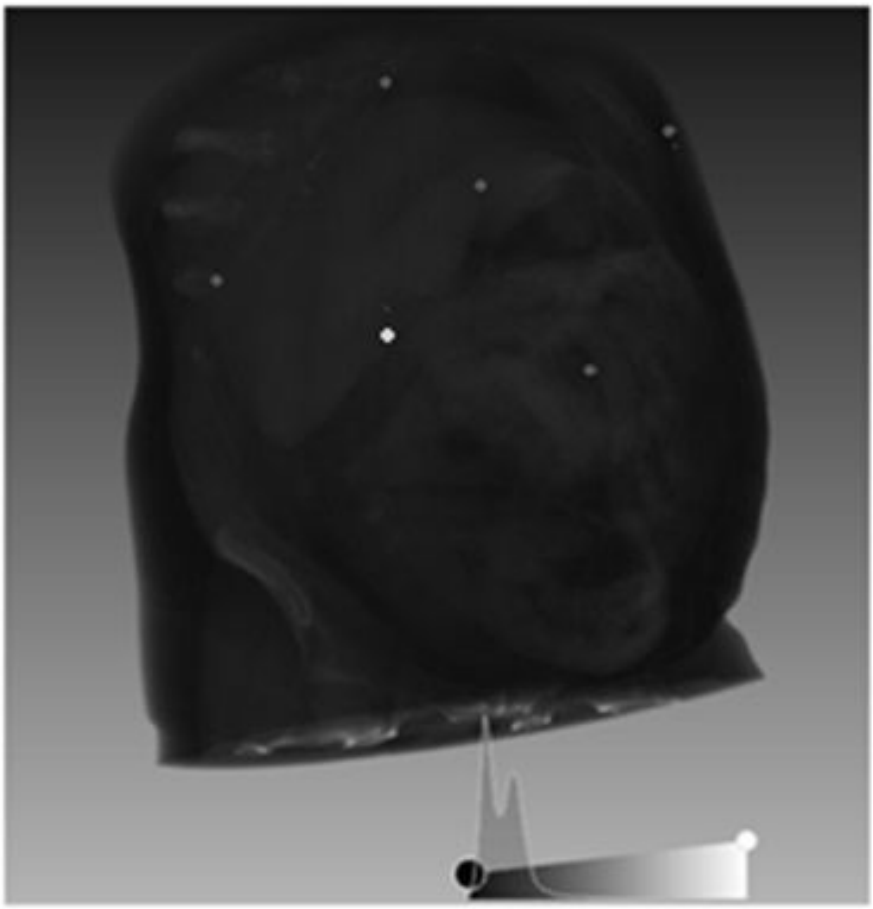


Fig. 4. Abdominal surface of the patient with markers

2.4 Operation Field Visualization During Treatment and Ablation Process Monitoring

In order to aid treatment performance graphical interface has been proposed which presents current position of surgical tools (in this ablation needle) on personalized anatomical patient's model. Current position of ablation needle is tracked by position tracking system NDI Polaris Vicra. Key element which is crucial for successful ablation treatment in starting phase is placement of ablation needle ending as close as possible to the proposed target point. For this purpose serves viewer interface (Fig. 5), which shows ending of ablation needle position relative to target point during needle introduction, and deviation of needle introduction direction relative to direction between entry and target points - found during treatment planning phase [16]. Located next to the

Table 1. Anatomical position of skin markers for manual registration

Anatomical location	Grounds
Xiphoid process	Point fixed to bony structures in the midline of the body
Intersection of midclavicular line and right costal margin	Anatomical point fixed to bony structures
Intersection of anterior axillary line and right costal margin	Anatomical point fixed to bony structures
Intersection of midclavicular line and left costal margin	Anatomical point fixed to bony structures
Intersection of anterior axillary line and left costal margin	Anatomical point fixed to bony structures
First control point on the abdominal surface	Registration accuracy validation

toolbar, indicates by its length the distance between ablation needle tip and target point. Located in the centre of the disc two points represent the ends of ablation needle onto a plane perpendicular to the direction the plane is aiming, which is visualized in a red-white plate.

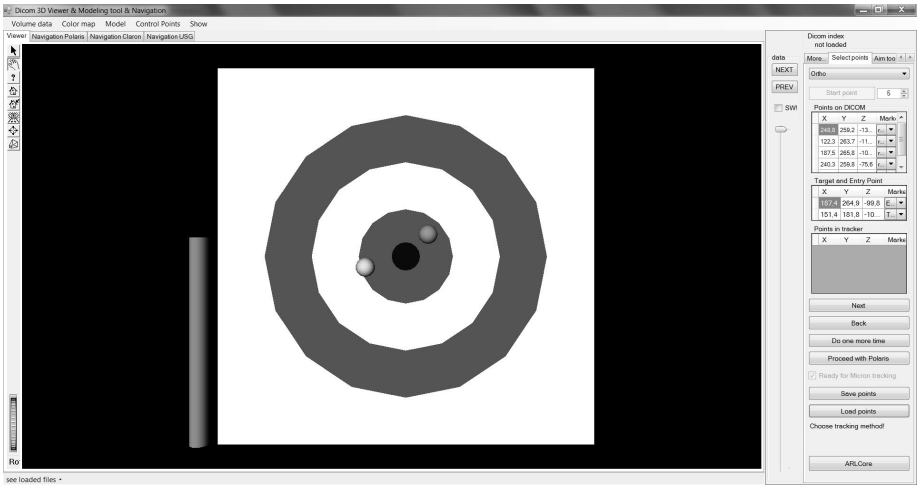


Fig. 5. Graphical user interface supported needle insertion

Second important element is to monitor tumour's ablation process. In order to do it bipolar system of electrical ablation with radio frequencies Celon produced by Olympus has been integrated. It allows for non-direct ablation process monitoring through tracking of electrical parameters: tissue resistance, momentary power and total energy delivered to the apparatus. Example of the process of animal liver tissue ablation is shown in Fig. 6. System automatically terminates the ablation process once it detects an increase in tissue resistance around ablation needle (Fig. 6 - end of lower graph).

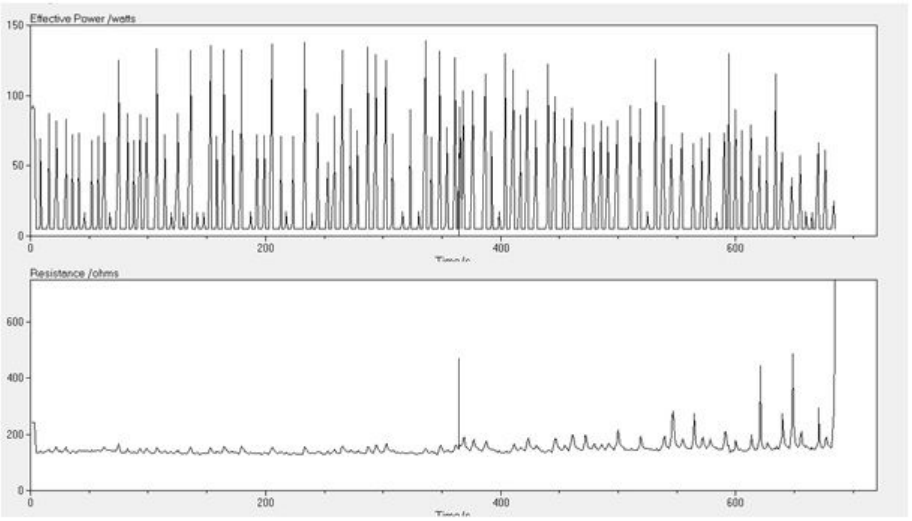


Fig. 6. Diagram of temporary power (up) and tissue resistance (bottom) during swine liver ablation from Olympus Celon device

3 Results and Conclusions

Proposed methodology for creation of personalized liver anatomical model verified on anonymized cases of abdominal cavity CT. In the proposed methodology automatically obtained segmentation results are submitted for verifications by radiologist and only then verified model serves in the next phase of planning of destroying focal lesions. Due to lack of ability to verify segmentation results directly for example in the case of vascular tree as a measure of segmentation results level of vascular structure branches has been proposed. Vascular tree segmentation has been performed successfully to third or fourth level of branches on average [8]. Results of registration were presented in previous article [15]. Two registration methods of abdominal preoperative CT and physical patient position in OR were presented and compared. This approach is being developed as a step to image guided percutaneous liver tumor ablation. The proposed equation of deformation is calculated in real time. Assuming regular breathing conducted by the respirator, recurrence of patient position and short time between CT study and procedure, the proposed solution could be used to synchronize the respiratory phase for which the static preoperative CT anatomical model was generated, which is one of the most challenging steps in RF ablation [15]. Conducted pilot studies confirmed possibility of application of proposed methodology for destroying focal liver lesions. Further studies direction is to fully integrate proposed methodology in the clinical environment and to verify correlation between external markers movement with internal organs displacement.

Acknowledgement. The author would like to acknowledge financial support from the Polish National Centre for Research and Development (NCBiR), grant no. LIDER/03/47/L-1/NCBiR/2010 'Development of planning system and computer aided minimally invasive surgery for hepatic and metastatic liver carcinoma localization, diagnosis and destruction'. The registration method has been evaluated on imaging data from the Second Department of Clinical Radiology for patients on the Chair and Department of General, Transplant and Liver Surgery at the Medical University of Warsaw after acceptance of the ethics committee of the Medical University of Warsaw.

References

1. Pereira, P.: Actual role of radiofrequency ablation of liver metastases. Eur. Radiol. 17, 2062–2070 (2007)
2. Spinczyk, D., Karwan, A., Rudnicki, J., Wróblewski, T.: Stereoscopic liver surface reconstruction. Videosurgery and Other Miniinvasive Techniques 7, 181–187 (2012)
3. Spinczyk, D., Karwan, A., et al.: In vitro evaluation of stereoscopic liver surface reconstruction. Videosurgery and Other Miniinvasive Techniques 8, 80–85 (2013)
4. Phee, S., Yang, K.: Interventional navigation systems for treatment of unresectable liver tumor. Medical & Biological Engineering and Computing 48, 103–111 (2010)
5. Peters, T., Cleary, K.: Image-Guided Interventions: Technology and Applications. Springer (2008)
6. Juszczyk, J.: Sekwencyjna segmentacja wybranych struktur anatomicznych w serii obrazów tomograficznych. PhD thesis. Silesian University of Technology (2012)
7. Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Multiscale Vessel enhancement Filtering. In: Wells, W.M., Colchester, A.C.F., Delp, S.L. (eds.) MICCAI 1998. LNCS, vol. 1496, pp. 130–137. Springer, Heidelberg (1998)
8. Rudzki, M.: Filtr wieloskalowy w detekcji struktur naczyniowych wątroby z obrazów tomografii komputerowej. PhD thesis. Silesian University of Technology (2011)
9. Xing, Z., Jie, T., Dehui, X., Xiuli, L., Kexin, D.: Interactive liver tumor segmentation from ct scans using support vector classification with watershed. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC, vol. 3, pp. 6005–6008 (2011)
10. Häme, Y.: Liver Tumor Segmentation Using Implicit Surface Evolution. The Midas Journal (2008)
11. Badura, P., Pietka, E.: 3D Fuzzy Liver Tumor Segmentation. In: Piętka, E., Kawa, J. (eds.) ITIB 2012. LNCS, vol. 7339, pp. 47–57. Springer, Heidelberg (2012)
12. Spinczyk, D.: System nawigacji obrazowej w ablacji zmian ogniskowych w wątrobie. Silesian University of Technology (2013)
13. Horn, B., Hilden, H., Negahdaripour, S.: Closed form solution of absolute orientation using orthonormal matrices. J. Opt. Soc. Am. A 5, 1127–1135 (1988)
14. Fitzpatrick, J., West, J., Mauer, C.: Predicting error in rigid-body point-based registration. IEEE Trans. Med. Imaging. 17, 694–702 (1988)

15. Spinczyk, D., Zylkowski, J., Wróblewski, T.: Continuous registration based on computed tomography for breathing motion compensation. Videosurgery and Other Miniinvasive Techniques 8, 265–272 (2013)
16. Seitel, A., Maier-Hein, L., Schawo, S., Radeleff, B., Mueller, S., Pianka, F., Schmied, B., Meinzer, H.: In-vitro evaluation of different visualization approaches for computer assisted targeting in soft tissue. Computer-Assisted Radiology and Surgery 2(suppl. 1), S188–S190 (2007)

Image Navigation in Minimally Invasive Surgery

Bartłomiej Pyciński[1], Jan Juszczyk[2], Pawel Bożek[3], Jacek Ciekalski[4], Józef Dzielicki[3], and Ewa Pietka[2]

[1] Silesian University of Technology, Faculty of Automatic Control, Electronics and Computer Science, Gliwice, Poland
bartlomiej.pycinski@polsl.pl
[2] Silesian University of Technology, Faculty of Biomedical Engineering, Zabrze, Poland
[3] Medical University of Silesia, Zabrze, Poland
[4] Sports Medicine Center CMS, Zabrze, Poland

Abstract. This study presents the development of an image navigation system dedicated to assist a miniinvasive abdominal surgery. First, the system overview is discussed. Then, a development of a mesh model of abdominal structure is presented. Next, a system calibration including the component as well as patient registration is described.The evaluation of the system has been performed separately for each component. The accuracy of calibration as well as patient registration is reported. This yields the robustnes evaluation of the overall system.

Keywords: Image navigation, visualization, calibration.

1 Introduction

The term of image navigation denotes a stereotactic surgical technique that uses preoperative CT or MRI data that match the corresponding surgical anatomy using three-dimensional coordinates to enhance intraoperative accuracy and anatomic orientation [1]. These systems are used in clinical tests since early 90's [2]. Due to the simplicity of anatomical landmarks registration onto the radiological data, neurosurgery [3,4] and orthopaedics [5] are the pioneering procedures. Moreover, location of anatomical structures corresponds to its appearance in the preoperatively acquired radiological data.

In this study an image navigation system employed in abdominal minimally invasive surgery is built. In these procedures insuflation and breath related organ motion require compensation. Some attempts have already been made to overcome these problems [6] and reconstruct the organ surface [7].

The main goal of this paper is to assess the accuracy of calibration and patient registration. The paper is organized as follows. Section 2 presents the system with the calibration procedure, patient registration, image preprocessing and display. Section 3 discusses the results. Section 4 concludes this phase of the study.

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_3, © Springer International Publishing Switzerland 2014

2 System Overview

Image navigation system described in the paper is based on the Polaris Spectra (NDI Inc., Canada [8]) optical tracker and dedicated markers attached to various elements during the procedure. Polaris Tracker is widely used by research centres for surgery that implements the navigation imaging systems [8,9]. Two items tracked during the surgery are laparoscopic instruments and the operating table.

While tracking the surgical instruments is a standard procedure, the purpose of attaching the marker to the table needs some explanation. This is an indirect way to track the position of the operating field during the surgery. This permits the camera to be moved around the table during the procedure. Without tracking the operating table, the camera cannot be displaced.

2.1 Local Coordinate Systems

Each element of the navigation system is related to a coordinate system, local for each item. The most common coordinate systems and objects related to them are presented in Tab. 1.

Table 1. Local coordinate systems commonly used

Element of the system	Referencing coordinate system
Tracker	„Global" coordinate system
Stylus	1. Fixed polaris marker
	2. Stylus tip coordinate system
Forceps	1. Fixed polaris marker
	2. Forceps tip coordinate system
Patient's body	Polaris marker fixed to the operating table
DICOM Image	DICOM coordinates
Organs' models	DICOM coordinates

All transformations between coordinate systems are performed by isometric transformations (i. e. rigid body transformation) and are stored in the 4×4 transformation matrix. Rigid body transformation is an arbitrary composition of translation and rotation of the coordinate frame.

Let P^X be coordinates of point P in the frame X, and P^Y is coordinates of point P in the frame Y. Let M_Y^X be a transformation matrix from coordinate frame Y to X, then

$$P^X = M_Y^X \cdot P^Y. \tag{1}$$

Transforming repeatedly one coordinate system into another is considered as a consecutive multiplying of the transformation matrices. For example, if coordinates of P are given in Z and matrices M_Y^X and M_Z^Y are known, new coordinates of P in X are found by:

$$P^X = M_Y^X \cdot M_Z^Y \cdot P^Z. \tag{2}$$

Note that

$$M_Y^X = \left(M_X^Y\right)^{-1}. \tag{3}$$

In order to connect all coordinate systems, transformation matrices have to be found. These matrices can be either variable or invariant. The first group denotes actual positions of Polaris markers, while invariant transformations have to be calculated at the beginning of the operation, i. e. surgical tool's pivot calibration, patient registration.

To determine the transformation between any two coordinate systems, one has to multiply all transformation matrices, according to the equation (2). Each transformation between two coordinate systems has a certain inaccuracy, which refers to the translation as well as the rotation. The scheme of transformations is shown in Fig. 1.

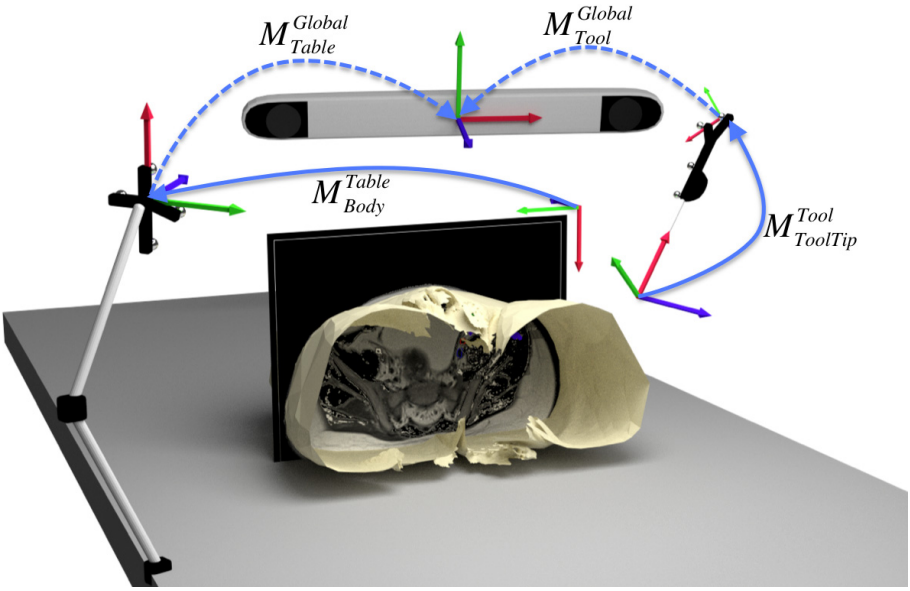


Fig. 1. Transformations between different coordinate systems; solid lines reflect invariant transformations and dashed lines reflect variable transformations

2.2 Tool Tip Calibration

Calibration of the tool tip (laparoscopic forceps and stylus) is a process of determining the invariant transformation between the coordinate system associated with the tool tip and the Polaris marker attached to this tool. The calibration involves fixation of the tip and rotation of the device around the pivot point (Fig. 2). The positions of the marker recorded in the global coordinate system are then used to determine the transformation matrix $M^{Tool}_{ToolTip}$. Once the matrix is known, the coordinates of the tip are given on-line.

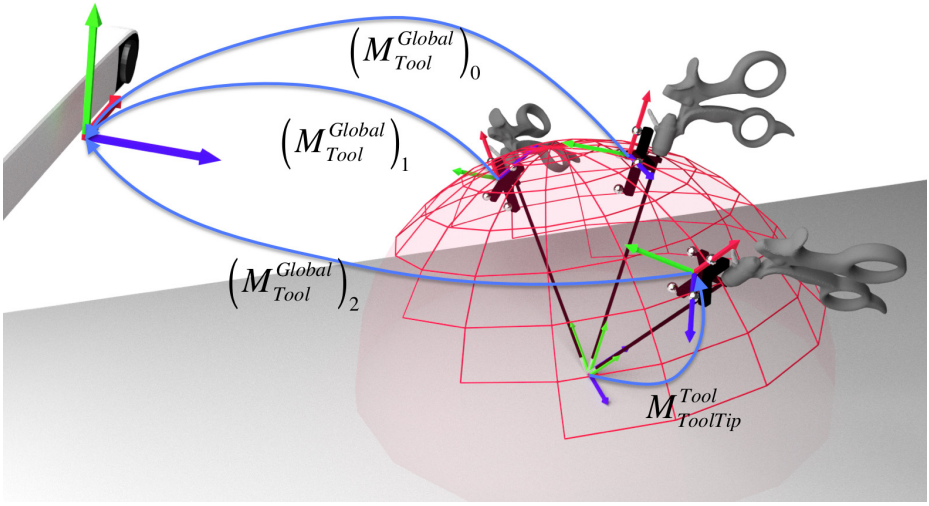


Fig. 2. Surgical tool calibration consists of recording multiple positions of the marker on tool $\left(M_{Tool}^{Global}\right)_i$ in order to compute the transformation matrix $M_{ToolTip}^{Tool}$

Calculating matrix $M_{ToolTip}^{Tool}$ on the basis of multiple registered matrices $\left(M_{Tool}^{Global}\right)_i$ is the optimization problem of overdetermined system of linear equations [10,11]. To solve it, the LSQR method [12] has been implemented.

2.3 Radiological Data Preprocessing

Most of the approaches to medical image segmentation are studies on automatic segmentation algorithms [13,14], although semi-automatic methods are also common [15,16]. However, none of them has yet succeeded in creating universal software able to facilitate the segmentation of all anatomical organs. The difficulty of the problem lies mainly in the variety of shapes, locations, and homogeneity of the segmented structures. The complexity increases when pathological structures of blurred edges are extracted.

Results of the segmentation are stored as binary DICOM masks or directly as contours for each of the slice of the image. Regardless of the format, they are in the same coordinate system as the original data. In order to convert segmented structures from raster to a mesh grid, a program based on VTK implementation of the marching cubes [17] method has been developed. In the case of manual contours delinated in OsiriX[1], built-in tools have been used. Such models require additional processing, which is performed using the program

[1] OsiriX website: `http://www.osirix-viewer.com/`

Blender[2]. The goal of processing is mainly to simplify the model (by remeshing), to remove unnecessary elements such as free vertices and edges, and to transform the position to the coordinate system compatible with 3DSlicer[3].

Both the radiological images and the models created on the basis of them are oriented in the LPS (Left-Posterior-Superior) coordinate system. This means that X, Y and Z axes increase towards left, posterior and superior direction, respectively. On the other hand, Slicer uses the RAS (Right-Anterior-Superior) coordinate system, so that the objects loaded into it have to be flipped along the X and Y axes. These models are loaded directly into the Slicer scene. The workflow is shown in Fig. 3.

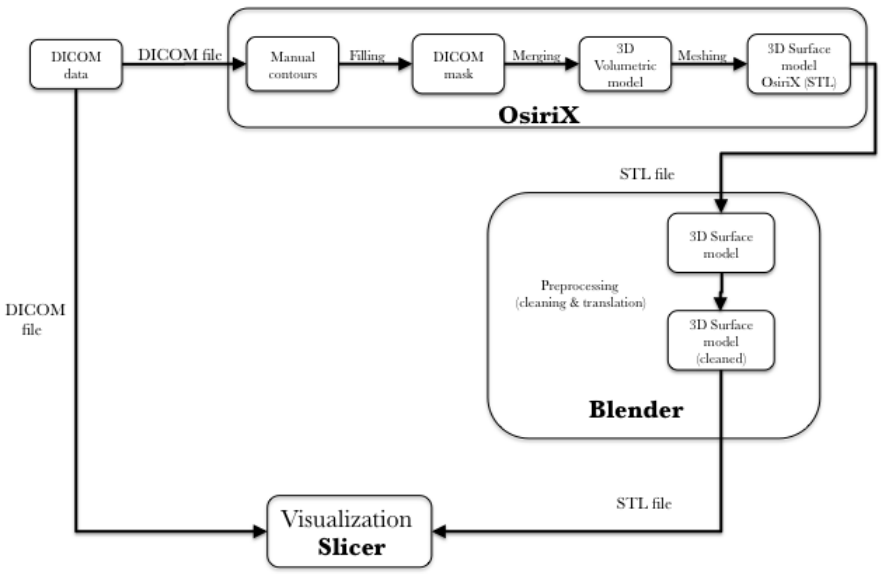


Fig. 3. Data preprocessing workflow

2.4 Patient's Registration

Patient's registration is a process of transforming points located on patent's skin and corresponding voxels into one coordinate system. At least three non-collinear points are needed to determine the transformation between the coordinate systems. Typically, the points do not overlap and determined transformation minimizes the residual sum of squares (RSS). Patient's registration is computed with Horn's algorithm of absolute orientation [18] also known as "Landmark-based registration algorithm". Measurements are performed before to determine,

[2] Blender website: `http://www.blender.org/`
[3] 3DSlicer website: `http://www.slicer.org/`

which points on the patient's body are the best to carry out the registration, i. e. their residuals are the smallest.

During the calibration only anatomical landmark points are used (no artificial landmarks have been attached to the patient's body). Although artificial landmarks could increase the registration accuracy [19], their sterility required during the surgery may not be guaranteed.

The landmark points on patient's body are located in the image and their position in DICOM coordinate system should be taken before the surgery.

2.5 Visualization

The program 3DSlicer is used to display the virtual scene during the operation. MRI or CT slices in several planes and three-dimensional virtual scene are shown simultaneously (Fig. 4). Two-dimensional views change dynamically when forceps moves showing the slice of the current tool location.

Current markers position are being sent to the Slicer in the form of transformation matrices by OpenIGTLink protocol. The communication is provided by a PLUS library [20].

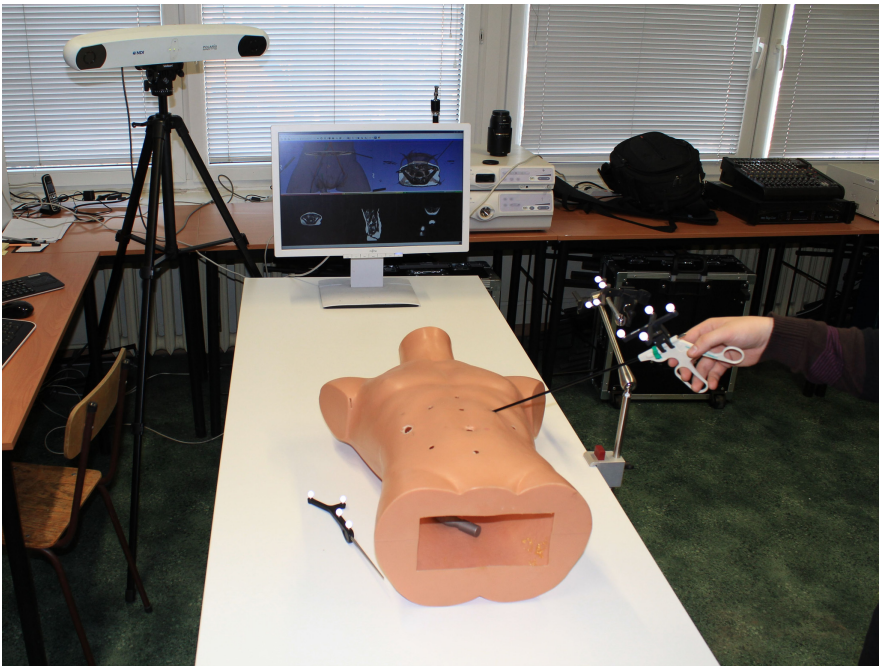


Fig. 4. Image navigation system and the visualization of models

Coronal and sagittal views of the three-dimensional patient model are displayed simultaneously (Fig. 4). If necessary, the perspective can be easily changed (e.g. to the lateral or oblique view). Two-dimensional scans are also displayed.

3 Results

Two settings have been employed for the system evaluation. The first set-up measures the calibration accuracy, the second one assesses the exactness of the patient registration. The accuracy of the Polaris tracking system is at the range of 0.2 mm.

In order to measure the performance of the calibration procedure, 30 experiments have been repeated. Each of them recorded 200 positions of the tool tip. The results (Tab. 2) indicate an influence of the tool length. The longer the distance between the marker and the pivot point, the more the inaccuracy of rotation influences the total error.

Table 2. Calibration accuracy of stylus and forceps tips

Radius length [mm]	Mean error [mm]	Std deviation [mm]
160.3	0.21	0.07
357.3	0.52	0.14

Patient registration accuracy is based on nine landmarks points on the surface of the abdomen. The following points have been chosen:

- Xiphoid process of sternum (XP)
- A point in the middle of costal arch right and left (CAR, CAL)
- Anterior superior iliac spine right and left (ASISR, ASISL)
- Greater trochanter of right and left femur (GTR, GTL)
- Umbilicus (U)
- Pubic symphysis (PS)

They have been recorded just before the CT abdomen or pelvis examination. Thus, these landmarks are acquired and marked on the CT scan. Due to the obesity of the patient, the presence of an ostomy bag on the left part of abdominal wall, bandages after recent surgeries of abdomen and pelvis, or a thick layer of clothes, the points located over trochanters and pubic symphysis are sometimes missed (Fig. 5)

After the conversion of all points, a Fiducial Registration Error (FRE) has been found [21]. It is defined as a distance between the current position of a point in the original coordinate system and its corresponding position after the conversion of the second coordinate system. The results are shown in Tab. 3.

The overall system accuracy is defined as a distance between the tool tip that points the target and its location displayed on the 3D patient model. Our first approach to the accuracy analysis yields the results of 24.17 mm (Tab. 4).

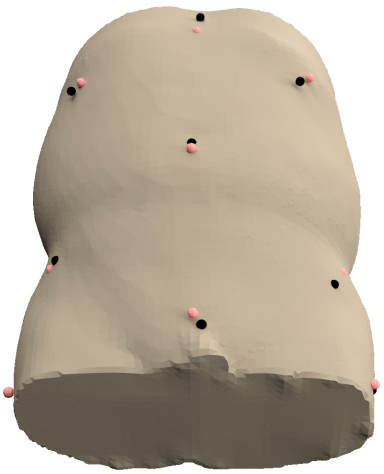


Fig. 5. Position of landmarks on the model in DICOM (light) and transformed Polaris (dark) coordinate system

Table 3. FRE values of anatomical landmark points at 9 studies [mm]

	1	2	3	4	5	6	7	8	9	mean
XP	36.52	15.77	30.11	45.78	8.12	20.46	15.38	21.26	30.22	24.85
CAR	26.81	26.52	44.81	15.96	22.50	30.85	10.35	26.37	14.53	24.30
CAL	26.70	22.43	54.65	23.76	14.84	41.03	7.03	34.09	19.46	27.11
ASISR	6.57	35.67	27.35	11.26	10.69	22.76	28.11	28.45	11.16	20.22
ASISL	13.10	23.44	23.30	26.89	24.10	36.18	29.63	36.47	29.39	26.94
GTR	–	9.08	23.74	–	28.76	23.98	16.07	–	18.11	19.96
GTL	–	35.72	34.30	–	31.01	27.38	22.58	–	12.29	27.21
U	17.34	43.54	16.41	15.27	15.06	11.01	–	36.20	11.74	20.82
PS	–	40.07	20.60	15.95	20.06	7.07	–	–	14.05	19.63

Table 4. Partial and total errors of the system [mm]

Erorr	Value [mm]
Marker track error	0.2
Tool tip calibration error	0.52
Fiducial registration error	23.45
Total	24.17

4 Summary

In this paper a methodology of navigation in minimally invasive surgery has been shown. Procedure of tools calibration and patient registration is discussed. Development of the patient abdomed model has been described. Visualization system available in the operating room is proposed. The overall accuracy of the entire procedure is described and the error is less than 24.17 [mm].

Patient registration process can be performed in the operating room, and the system is ready for testing in clinical settings. Presented system is dedicated to organs fixed to the abdominal wall and insuflation process does not cause their motion.

Acknowledgement. This research is supported by a NCN Grant No. UMO-2012/05/B/ST7/02136.

References

1. Stedman's Medical Dictionary. Lippincott Williams & Wilkins (2006)
2. Jolesz, F.A.: Image-guided procedures and the operating room of the future. Radiology 204(3), 601–612 (1997)
3. Lindseth, F., et al.: Image Fusion in US-Based Neuronavigation: Improving Overview and Interpretation by Integrating Preoperative MRI with Intraoperative 3D Ultrasound. Computer Aided Surgery 8, 49–69 (2003)
4. Salah, Z., et al.: Navigation-supported diagnosis of the substantia nigra by matching midbrain sonography and MRI. In: Proc. SPIE, vol. 8315 (2012)
5. Barratt, D.C., Penney, G.P., Chan, C.S.K., Slomczykowski, M., Carter, T.J., Edwards, P.J., Hawkes, D.J.: Self-calibrating 3D-ultrasound-based bone registration for minimally invasive orthopedic surgery. IEEE Transactions on Medical Imaging 25(3), 312–323 (2006)
6. Spinczyk, D., Żyłkowski, J., Wróblewski, T.: Continuous registration based on computed tomography for breathing motion compensation. Videosurgery and Other Miniinvasive Techniques 8(4), 265–272 (2013)
7. Spinczyk, D., Karwan, A., Żyłkowski, J., Wróblewski, T.: In vitro evaluation of stereoscopic liver surface reconstruction. Videosurgery and Other Miniinvasive Techniques 8(1), 80–85 (2013)
8. Wiles, A.D., et al.: Accuracy assessment and interpretation for optical tracking systems. In: Proc. SPIE, Medical Imaging 2004, vol. 5367 (2004)
9. Tamura, Y., et al.: Surface-based registration accuracy of CT-based image-guided spine surgery. Eur. Spine J. 14(3), 291–297 (2005) (Epub November 4, 2004)
10. Cleary, K., et al.: Electromagnetic Tracking for Image-Guided Abdominal Procedures: Overall System and Technical Issues. In: Engineering in Medicine and Biology Society, IEEE-EMBS 2005, pp. 6748–6753 (2005)
11. Detmer, P.R., et al.: 3d ultrasonic image feature localization based on magnetic scanhead tracking: In vitro calibration and validation. Ultrasound in Medicine & Biology 20(9), 923–936 (1994)
12. Paige, C.C., Saunders, M.A.: LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares. ACM Trans. Math. Software (1982)

13. Linguraru, M.G., Summers, R.M.: Multi-organ automatic segmentation in 4D contrast-enhanced abdominal CT Biomedical Imaging: From Nano to Macro. In: 5th IEEE International Symposium on ISBI 2008, pp. 45–48 (2008)
14. Hu, S., Hoffman, E.A., Reinhardt, J.M.: Automatic Lung Segmentation for Accurate Quantitation of Volumetric X-Ray CT Images. IEEE Transactions on Medical Imaging 20(6), 490–498 (2001)
15. Wieclawek, W., Pietka, E.: Fuzzy clustering in Intelligent Scissors. Computerized Medical Imaging and Graphics 36(5), 396–409 (2012)
16. Czajkowska, J., Badura, P., Pietka, E.: 4D Segmentation of Ewing's Sarcoma in MR Images. Information Technologies in Biomedicine 69(2), 91–100 (2010)
17. Lorensen, W.E., Cline, H.E.: Marching Cubes: A High Resolution 3D Surface Construction Algorithm. In: SIGGRAPH 1987 Proc., vol. 21, pp. 163–169 (1987)
18. Horn, B.K.P.: Closed-form solution of absolute orientation using unit quaternions. Journal of the Optical Society of America A 4(4), 629–642 (1987)
19. Kaspersen, J.H., et al.: Three-Dimensional Ultrasound-Based Navigation Combined with Preoperative CT During Abdominal Interventions: A Feasibility Study. Cardiovascular and Interventional Radiology 26, 347–356 (2003)
20. Lasso, A., et al.: Implementation of the PLUS open-source toolkit for translational research of ultrasound-guided intervention systems. In: MICCAI (2012)
21. Fitzpatrick, J.M., West, J.B., Maurer Jr., C.R.: Predicting error in rigid-body point-based registration. IEEE Transactions on Medical Imaging 17(5), 694–702 (1998)

The Effect of Marker Detection Efficiency on Performance of the Optical Tracking Systems in Computer Assisted Surgery

Adrian Goral

AGH University of Science and Technology
Department of Measurement and Electronics
Al. Mickiewicza 30, 30-059 Kraków, Poland
adrian.goral@agh.edu.pl

Abstract. In computer assisted surgery, the localizers can be used for tracking of an object in motion. In this paper, we propose a method for simulating the accuracy of such tracking mode for a specified efficiency of marker detection. From simulation, we obtain the relationship between marker detection efficiency and tracking precision, assuming linear interpolation of missing locations. We find, that for a typical design of the tracked tool, the marker recognition efficiency should be no less than 88% to keep the tracking precision at the level that is acceptable in computer assisted surgery. We use this criterion to determine a usable measurement volume of a physical prototype of optical localizer.

Keywords: computer assisted surgery, optical tracking system, marker detection.

1 Introduction

The task of computer assisted surgery (CAS) is to build a three-dimensional model of relevant anatomical structures and surgical tools present in operative field. For the system to provide correct indications, the model is required to be continuously updated during the surgical procedure. One possible way to update the model is to use a tracking system (localizer), that recognizes certain landmarks in the operative field and calculates their positions relative to patient's anatomy. In CAS, the localizer is used either for measurements of static points (e.g. in free-hand navigated orthopaedic interventions [1]), or for tracking of an object in motion (e.g. in robotic surgery, as a part of a control system [2]).

Currently, the most of the commercially available localizers used in CAS employ imaging devices, such as calibrated CCD cameras [3]. Therefore, they are called optical localizers or optical tracking systems. Optical localizer detects a set of characteristic locations in the images from its built-in cameras and determines their coordinates in three-dimensional space. To facilitate the detection, visually distinctive objects called markers are placed at the measured locations.

Localizers in CAS systems are expected to comply with certain requirements, most of which pertain to positioning accuracy. According to typical values given

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_4, © Springer International Publishing Switzerland 2014

in the literature, the positioning error in CAS should not exceed 1 mm [4]. However, it is rarely emphasized that apart from being accurate, the localizer should also meet requirements that concern more basic aspects of its operation, especially effective marker detection. The efficiency of marker detection can be expressed as percentage of image frames in which marker can be detected. While many research groups and standards organizations have proposed their own procedures for accuracy assessment [5,6], no guidelines concerning assessment of marker detection efficiency are present in the literature.

The efficiency of marker detection should be taken into account for two reasons. First, while tracking a moving object, it can affect the accuracy of the system. Second, it limits a usable measurement volume of the localizer, and the size of the volume determines functionality of the system in a given application. In this paper, we provide evidence for these statements, both by numerical simulations and by measurements done on a physical prototype of optical tracking system. First, we simulate tracking of an object moved along a specified trajectory. In this situation, detection efficiency of markers can increase tracking error. Based on results of the simulation, we can establish a relationship between these two parameters. Second, we run an experiment to determine size of usable measurement volume of prototype system, assuming detection efficiency as a limiting criterion.

In the usable measurement volume of a localizer, not only should a measurement be feasible, but also certain operational parameters of the system must reach a required value. In this work, we determine the usable measurement volume in terms of efficiency of marker detection. Using the relationship obtained in our simulation, we find a minimum detection efficiency acceptable for CAS as the value that corresponds to the tracking error of 1 mm. It turns out, that the markers should be detected in at least 88% of image frames, to guarantee sufficiently low value of the error while tracking an object in motion. Using this criterion, we examine spatial distribution of detection efficiency measured for a prototype of optical tracking system. We demonstrate shrinking of usable measurement volume of the system caused by decreasing detection efficiency.

The quantitative findings presented in this work are specific to the parameters of the simulation (design, trajectory and speed of the tracked tool, acquisition parameters) and to the performance of the tested prototype in the experimental part. However, using the methods we present, one can evaluate the usable measurement volume of any optical tracking system, especially of a prototype localizer for which neither the measurement volume is specified, nor the detection efficiency is guaranteed.

2 Describing Accuracy of the Localizer

In CAS, localizers are used for measuring the positions of certain anatomical landmarks. It is usually impractical to place the tracked marker directly in the measured point. Therefore, in real applications the measurement is done by means of a tool with a tip for mechanical probing. The position of the tip is

precisely known in the local coordinate system associated with the tool. The markers are in this case attached on the tool and their positions are also defined in a tool coordinate system. By locating the markers, the localizer can determine global position of whole structure, including position of the tip.

To determine position of the tip we need to know position and orientation of the local tool coordinate system in a global coordinate system of the localizer. We can describe the position and the orientation by a three-dimensional rigid transformation that yields the best linear fit of the markers' positions given in local coordinate system to the positions measured in global coordinate system [7]. The more precisely the markers are measured, the more accurate the transformation can be. Consequently, the better the transformation is, the more accurately the tip can be positioned.

Marker positionig error is insufficient measure of the tracking precision for a clinician during a computer assisted procedure. Therefore, the tip positioning error is introduced to make the surgeon aware of tip trackig precision. The tip positioning error can be estimated either analytically or experimentally. The analytic expression, derived by Fitzpatrick et al. [8], bases on the single marker positioning error and on the spatial distribution of markers on the tool. The experimental method makes no assumptions concerning the sources of the error. In this method, a set of locations on the object of a known geometry is measured using a probing tool and the tested localizer. To evaluate the tip positioning error the measurement results are compared with the reference locations. The disparity is expressed as a root mean square (RMS) distance between the corresponding points in reference and measured sets. If we denote the reference points as $\mathbf{p}_i^{(r)} = \left(x_i^{(r)}, y_i^{(r)}, z_i^{(r)}\right)$ and the measured points as $\mathbf{p}_i^{(m)} = \left(x_i^{(m)}, y_i^{(m)}, z_i^{(m)}\right)$, the RMS distance $\langle\varepsilon\rangle$ is given by the equation:

$$\langle\varepsilon\rangle = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\left(x_i^{(r)} - x_i^{(m)}\right)^2 + \left(y_i^{(r)} - y_i^{(m)}\right)^2 + \left(z_i^{(r)} - z_i^{(m)}\right)^2\right]}, \quad (1)$$

where N is the total number of points in each set. If the two sets are expressed in different coordinate systems, it is required to transform one of the sets to the cooordinate system of the other set before calculating RMS error. The transformation can be found by the best-fit algorithm which is also used for tool positioning, as it has been mentioned earlier in this section.

3 The Impact of Inefficient Detection on the Accuracy

Under certain circumstances, localizer may not detect a marker although at the moment of image acquisition the marker was present in the common field of view of its cameras. To describe this effect quantitatively, we can define efficiency of marker detection η as a ratio of the number of image frames N_d in which the marker was detected to the total number of frames acquired, N:

$$\eta = \frac{N_d}{N} \cdot 100\%. \quad (2)$$

If all N image frames were captured for a fixed marker position, the obtained detection efficiency is a site specific value. By repeating the procedure in various points of measurement volume we can determine a spatial distribution of detection efficiency.

Now, let us consider a marker moving along a specified trajectory. The localizer records the movement as a set of samples – 3-dimensional points, $\mathbf{p}(t_i)$, each representing marker position at the discrete time t_i. The lower the detection efficiency, the larger loss of samples along the trajectory. Assuming a uniform distribution of detection efficiency across the measurement volume, the number of trajectory points lost due to failed detections equals $(100\% - \eta)N$.

Lower values of marker detection efficiency can make tracking of a probing tool difficult. For any given image frame, all markers' positions are needed to calculate position of the tip. If one or more markers are not detected, then the tip position also cannot be determined. Therefore, for each marker's trajectory, the missing samples must be reconstructed from the remaining ones. The simplest method of the reconstruction is linear interpolation which bases on two closest correctly acquired samples, one preceeding and one following the missing sample within the trajectory. For linear interpolation, the missing sample $\mathbf{p}(t_{j+k})$ can be reconstructed as:

$$\mathbf{p}(t_{j+k}) = \mathbf{p}(t_j) + \frac{k}{l-j}\left[\mathbf{p}(t_l) - \mathbf{p}(t_j)\right], \tag{3}$$

where $\mathbf{p}(t_j)$ is the closest preceeding sample and $\mathbf{p}(t_l)$ is the closest following sample. The indices, j, k, l, are positive integers where $j + k < l$.

The real trajectory of the tip varies from the trajectory calculated using the reconstructed trajectories of the markers. The magnitude of the discrepancy between real and computed tip trajectory can be expressed as tip positioning error. For our simulation, the tip positioning error is the RMS distance between corresponding points of real and calculated trajectory computed using Equation 1.

4 Simulation

We simulated tracking of a moving tool with a specified detection efficiency to establish the relationship between the detection efficiency and the accuracy of the localization. The tool used in the simulation is a pointing device with a tip that can be used as a precise mechanical probe. Three markers are placed on the tool in a defined manner, so that the localizer is able to determine its absolute orientation and, consequently, the position of the tip. The distribution of the markers relative to the tip was chosen to match the geometry of tools used in CAS, to make it possible to compare the results of the simulation with accuracy requirements typical to medical applications (see Fig. 1a).

The trajectory of the tool was defined as a composition of a translation along a circle 1 m in diameter and a rotation around the main axis of the tool, defined by a unit vector $\hat{\mathbf{u}} = (0, 0, 1)^\top$ in tool coordinate system. The trajectories of the markers and the tip are closed loops to avoid the problem of interpolation of first

and last point. During the translation along the full loop the tool performs 4 revolutions around its axis (see Fig. 1b). In the simulation, the tool moves along the circle with a speed of 1.26 m/s and rotates around its axis with an angular speed of 1.60 rad/s. We simulated the acquisition at the rate of 20 frames per second which gives 100 samples per each trajectory of the marker per single loop.

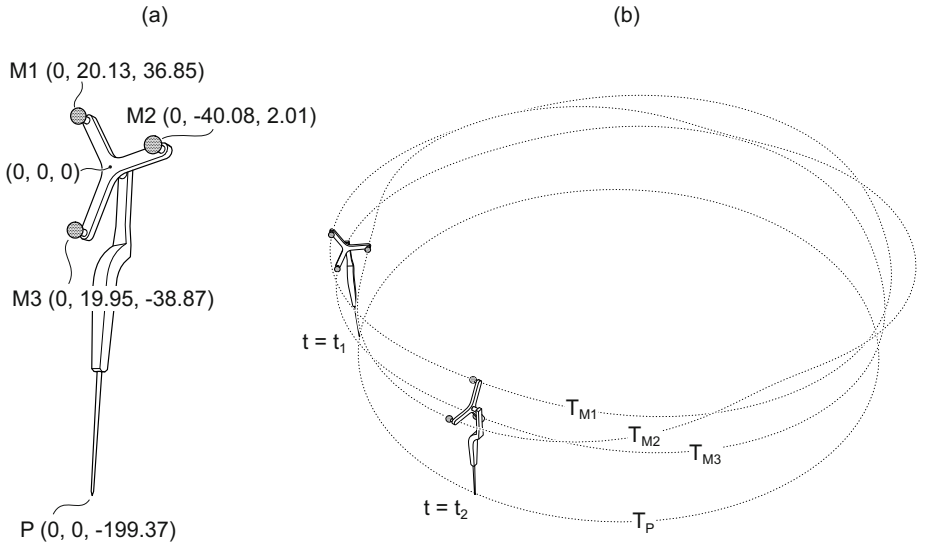


Fig. 1. Tool design (a) and trajectory (b) used to simulate the effect of marker detection efficiency on accuracy of the tracking system. Coordinates are given in millimeters in the tool coordinate system. M1...M3 denote markers, P denotes tip of the tool, T denotes the trajectory of the landmark whose symbol is given in the subscript

The whole set of the recorded positions for either single marker or for the tip builds up a discrete trajectory. (see Fig. 1b). To simulate the inefficient detection we rejected a number of randomly selected samples with uniform probability distribution along the trajectory. The total number of the rejected samples was set to match the specified detection efficiency. Next, based on decimated trajectories of the markers we reconstructed the position of the tip using linear interpolation, according to Equation 3. The resulting reconstructed trajectory of the tip differs from its original trajectory, the bigger difference meaning less accurate tracking. The discrepancy between the original and the reconstructed trajectory was calculated as RMS point-by-point distance, according to Equation 1. We examined the distance for a range of values of η, from 85% to 100%.

As the samples in the decimation stage are rejected randomly, the resulting RMS distance should be slightly different for each execution of the simulation, depending on particular distribution of rejected samples along the trajectory. Therefore, instead of a single RMS value defining the accuracy of the system for a given detection efficiency, we expected a range of values. To determine this

range, we repeated the simulation 30 times for each examined value of marker detection efficiency.

For the purpose of the simulation we adopt several simplifications. First, we assume that as long as the marker is detected, its position can be measured without error, so that we do not consider limited precision of the localizer. By this simplification we can attribute entire tool positioning error observed in the simulation to inefficient detection. Second, we assume that the efficiency of marker detection is uniform in entire measurement volume. Another simplification pertains to physical properties of the tracked tool. We assume, that spatial orientation of the tool relative to the cameras does not affect the efficiency of marker detection. In reality, this assumption is not neccessarily true, because the elements of the tool's structure may occlude the markers in the image of the camera at certain orientations. Yet, this simplification makes the results of the simulation independent from the features of tool design other than distribution of markers relative to tip position.

5 Determining Usable Measurement Volume

To demonstrate the impact of limited detection efficiency on performance of the localizer experimentally, we determined the usable measurement volume of a physical prototype. The prototype is built of two 1024×768 pixel cameras, operating in visible light, with focal length of 8 mm and maximum frame rate of 33 fps. The cameras are installed on a frame, 270 mm apart, with their optical axes at an angle of 9.1°. The markers used by the localizer are flat printed circular patterns. The diameter of the markers to be used in the system can be adjusted as required.

In the experiment, we first determined approximate dimensions of measurement volume of the prototype, without any additional requirements. We used a marker of a sufficiently big diameter, such that presence of the marker in the common field of view of cameras was the only limitting factor for successful acquisition. We tracked the marker at random positions in front of the localizer collecting a cloud of 13200 data points. Then, we fitted six planes to the surface points of the cloud. The intersections of these planes define the edges of the measurement volume. The obtained geometry was presented in Fig. 2).

Using a similar method we determined the distribution of detection efficiency for several chosen diameters of markers. We tested 15 marker diameters, varying from 10.1 mm to 30.3 mm. The markers were printed on a test board. To keep the board fixed during the measurement we installed it on a stand. At a given board position we made a series of 60 image acquisitions, each followed by marker recognition. We repeated the measurement in a set of 379 board locations scattered within measurement volume. Based on the number of correct recognitions within a series, we determined a discrete spatial distribution of recognition efficiency, individually for each of the tested marker diameters.

From the results of the simulation described in the previous section, we found a minimum value of marker detection efficiency, η_{min}, for which the maximum

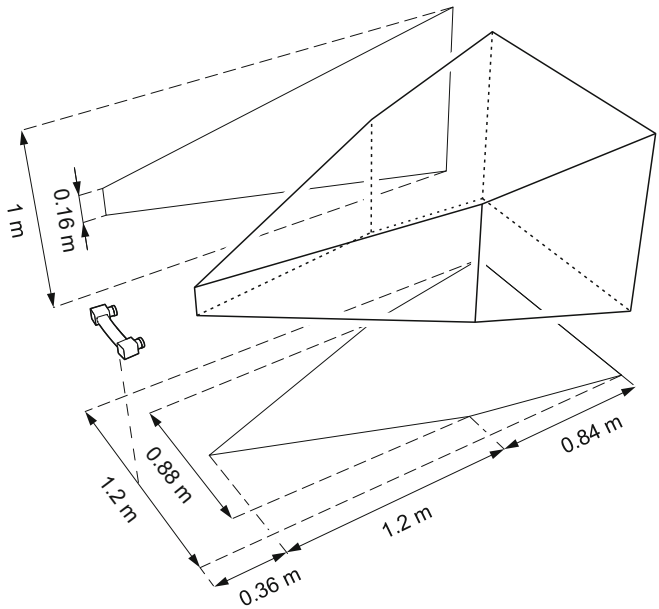


Fig. 2. Shape and dimensions of measurement volume resulting from geometry of the tested system. The volume was arbitrarily confined by a vertical plate located 2.4 m from the cameras, to specify all dimensions unambiguously

acceptable value of error is not exceeded, considering the requirements for CAS. Then, we used η_{min} to delineate usable measurement volume of the prototype for each diameter of marker. For each spatial distribution of detection efficiency we found a surface that separates the points with detection efficiency higher than η_{min} from the points with detection efficiency lower than η_{min}.

6 Results and Discussion

As a result of the simulations, we obtained a set of 30 RMS error values for each tested value of η. For a given η, the RMS values were uniformly distributed within certain interval, which is a consequence of the random decimation used to simulate failure in detection. The results of the simulation are presented in Fig. 3 as a relationship between detection efficiency and RMS positioning error calculated along the trajectory of the tool tip, T_P. For clarity, the plot contains only the minimum and maximum RMS value for each examined η, corresponding to worst- and best-case scenario of failure in detection. The maximum RMS error corresponds to the worst-case scenario in which a substantial number of consecutive marker locations is rejected in a few regions of trajectory. The interpolation algorhitm is not able to reconstruct the long section of marker's trajectory based on remaining points, which results in high error. In turn, the minimal RMS error corresponds to the best-case scenario, in which only isolated nodes are lost in

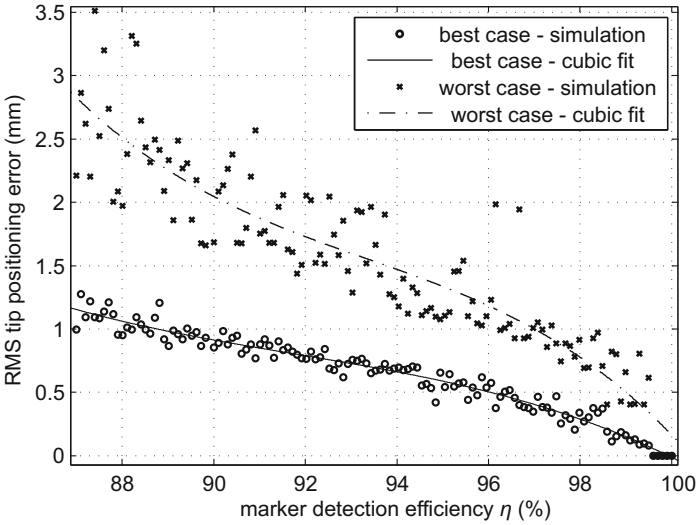


Fig. 3. The relationship between detection efficiency of the marker and the accuracy of tracking obtained in the simulation. The plot shows only the highest and the lowest RMS error value obtained for each value of η examined in the simulation, corresponding to worst- and the best-case scenario, respectively

many regions of trajectory. The linear interpolation is able to cope well with this situation, which results in relatively low RMS error.

Figure 3 shows, that the positioning error decreases as the detection efficiency improves. However, the relationship is not linear, and therefore we chose a third order polynomial, as it is a closest simple function to simulation data representing both worst- and best-case scenario. One can see, that for a given tool design it is required to keep the marker detection efficiency at least at the level between 88% and 97% (for the best and the worst case respectively) to guarantee the acceptable value of positioning error. Considering the other sources of positioning error, which were not taken into account in this simulation, the detection efficiency should be even larger.

The experimental results show that detection efficiency restricts the usable measurement volume of the tested prototype (see Fig. 4). As we increase the marker's diameter, the detection efficiency increases and the usable measurement volume expands. The maximum range of the measurement volume along z axis differs for the best and the worst case scenario. The difference falls within the range from approximately 10 cm (for small markers) to 15 cm (for large markers). The results also show, that detection efficiency on the left side of the measurement volume (negative x values) is typically better than the efficiency on the right side (positive x values), especially for larger marker diameters. This effect is caused by different focus settings of the cameras used in the system.

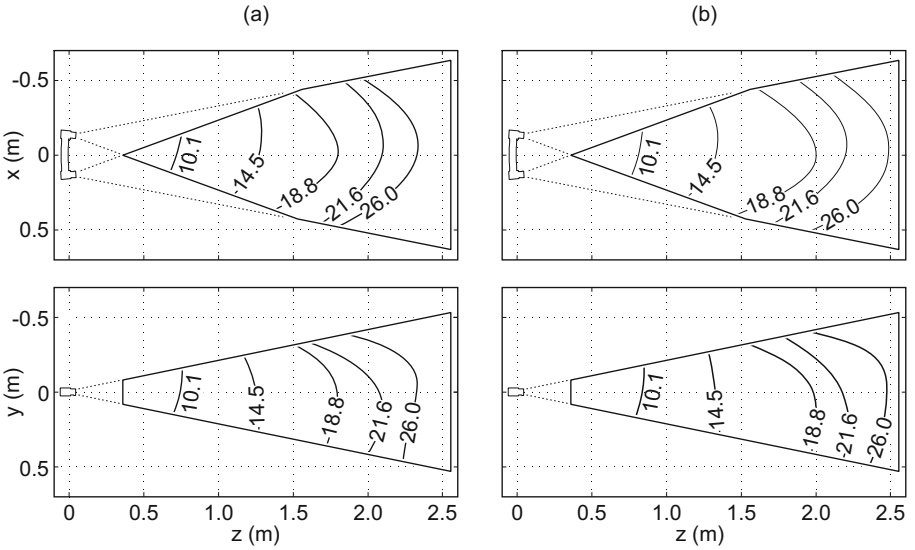


Fig. 4. The usable measurement volume of the tested prototype for markers with different diameters seen from above (top plot) and from the side (bottom plot). The isocontours correspond to: (a) worst-case scenario ($\eta_{min} = 97\%$), (b) best-case scenario ($\eta_{min} = 88\%$). The diameters of the markers are given in millimeters

Considering the results obtained from simulations, the tested prototype could be used for precise tracking of moving objects in CAS applications only using large markers, to guarantee sufficiently large usable measurement volume. The simulations show, that the minimum detection efficiency that allows for reasonable tracking precision should be higher than 88%. For the tested prototype, the detection efficiency of small markers reaches this limit only for small regions of measurement volume that are closest to the cameras.

7 Summary

In this work, we show that efficiency of marker detection is equally important operational parameter of the optical tracking system as its accuracy. To support this statement, we first set up a simulation that involves tracking of an object in motion using an optical localizer. In the simulation, we consider a situation in which the trajectory points, missing due to limited detection efficiency, are linearly interpolated using the set of remaining, correctly acquired locations. The results of the simulation prove that in such a situation the efficiency of marker detection limits the localizer's accuracy. We observe a monotonic non-linear relationship between the detection efficiency and RMS value of tool positioning error along the examined trajectory.

Next, we propose a testing procedure that can be used for experimental assessment of usable measurement volume of an optical tracking system. We use the

connection between detection efficiency and tracking precision to establish the minimum detection efficiency for localizers in CAS systems. We delineate usable measurement volume, by applying this minimum value to spatial distribution of marker detection efficiency obtained for a prototype localizer.

The detection efficiency of the tested prototype turns out to be sufficient for tracking a typical probing tool used in CAS only for large diameters of markers. For small markers, the usability of the system can be limited by small usable measurement volume. The results also allow for identifying a discrepancy in focus of the cameras working in the tested system.

Acknowledgement. I am grateful to prof. Józef Kozak for our fruitful discussions and for providing me with an opportunity to work on quality measurement equipment in laboratories of B. Braun Aesculap (Tuttlingen, Germany). I also thank prof. Janusz Gajda for his useful comments on early versions of this paper.

References

1. Delp, S., Stulberg, D., Davies, B., Picard, F., Leitner, F.: Computer Assisted Knee Replacement. Clinical Orthopaedics 354, 49–56 (1998)
2. Mozes, A., Chang, T., Arata, L., Zhao, W.: Three-dimensional a-mode ultrasound calibration and registration for robotic orthopaedic knee surgery. International Journal of Medical Robotics and Computer Assisted Surgery (9), 91–101 (2010)
3. Peters, T., Cleary, K. (eds.): Image Guided Interventions. Technology and Application. Springer Science+Business Media, New York (2008)
4. Philips, R.: The accuracy of surgical navigation for orthopaedic surgery. Current Orthopaedics 21, 180–192 (2007)
5. Elfring, R., de la Fuente, M., Radermacher, K.: Assessment of optical localizer accuracy for computer aided surgery systems. Computer Aided Surgery 15(1-3), 1–12 (2010)
6. Standard Practice for Measurement of Positional Accuracy of Computer Assisted Surgical Systems. ASTM Standard F2554-10. American Society for Testing and Materials (January 2011)
7. Horn, B.K.: Closed-form solution of absolute orientation using unit quaternions. Journal of the Optical Society of America A 4(4), 629–642 (1987)
8. Fitzpatrick, J.M., West, J.B., Maurer, C.R.: Predicting error in rigid-body point-based registration. IEEE Transactions on Medical Imaging 17(5), 694–702 (1998)

3D Bone Shape Modelling Basing on Dataset Recorded by Ultrasound Free-Hand Navigated Probe

Ewelina Świątek-Najwer[1], Katarzyna Otto, Paweł Krowicki[1], Krzysztof Krysztoforski[1], Peter Keppler[2], and Josef Kozak[3]

[1] Wroclaw University of Technology,
Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland
ewelina.swiatek-najwer@pwr.wroc.pl
http://www.biomech.pwr.wroc.pl
[2] Clinic University of Ulm, Steinhoevelstrasse 9, 89075 Ulm, Germany
[3] Aesculap BBraun, Am Aesculap Platz, 78532 Tuttlingen, Germany

Abstract. The aim of this paper was to analyze the spatial geometry of objects basing on ultrasound images recorded by free-hand probe. The data analysis consists of three phases: data recording, contour segmentation and triangulation. Several algorithms of segmentation have been tested to recognize the object's contour. To analyze the efficiency of triangulation algorithms, the similarities of obtained 3D models to the reference CT-based 3D model were evaluated. For objective and quantitative evaluation of obtained models the Hausdorff's distances were calculated for compared surfaces. The results revealed higher efficiency of greedy projection method than the 2D Delaunay's Triangulation applied for projected set of points in step by step procedure for the same cloud of points obtained as the result of user controlled segmentation procedure. The free hand ultrasonography can be applied to measure the spatial tissue shape for virtual planning of surgery without radiation.

Keywords: biomedical engineering, 3D imaging, free-hand sonography, surface similarity.

1 Introduction

Imaging techniques are nowadays one of the most important diagnostic methods to evaluate shape, size, location and even functionality of organs and pathological structures. The algorithms of image processing have a major impact on the successful diagnostics. Automatic tissue recognition preceded by reduction of artefacts, qualitative and quantitative evaluations are essential for physician to provide efficient diagnostics. The problem of image filtration is the reduction of useful information on image. The change in the image content lead to incorrect results of analysis. The recognition of tissues on images require some evidence such as statistical distribution of grey levels in particular scans, since the only technique of imaging enabling the recognition using a simple binarization with

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_5, © Springer International Publishing Switzerland 2014

constant threshold is computed tomography. In ultrasound imaging the same gray level can characterize echoes from different tissue structures. Ultrasonography is one of the most complex method of imaging to interpret. It requires experienced radiologists to provide good quality of images with high signal to noise coefficient and intended information on scan. The analysis is difficult due to so called speckles. The main advantages of ultrasonography are: noninvasiveness, improving high quality and inexpensive device.

Coupé confirmed that the main problem in 3D modelling is a non uniform dataset, and interpolations influence the reconstruction quality [1]. Zhang applied Radial Basis Functions to interpolate the pixels into a 3D set and tested it on phantom and on human hand [2], however presented results consider only one side of analysed objects and the results could not be quantitatively evaluated. Rohling applied reconstruction with registration of landmarks visible on intersecting datasets on a phantom and human gall bladder [3]. The authors avoided applying segmentation techniques, which are difficult to apply in ultrasound image interpretation. Other authors have made efforts to identify the bone contours on ultrasound scans. Daanen described the characteristics of bone-soft tissue border echo on ultrasound scans as: high echogenicity of bone contour as a result of high difference of acoustic impedances, acoustic shadow under the bone contour because of high absorption coefficient, continuous contour, contour consisting of horizontal echoes (by specular reflection of ultrasounds from bone surface) [4]. Another solution for ultrasound-based shape recognition is to apply deformable models. Nevertheless even that algorithm lead to unwanted deformations [5]. There still exists a need to develop the methods of automatic 3D reconstruction for free hand recorded ultrasound data.

The main aims of our study were: to develop method of bone spatial geometry recognition basing on segmentation of images saved by navigated ultrasound probe and to validate the results using models reconstructed basing on computed tomography scans. The article consists of three sections. Material and method section describes the developed measuring station and applied algorithms of image analysis. Results section presents segmentation procedure result, 3D models and similarity evaluation. In the discussion the results are analyzed in the aspect of the method application in computer aided surgery.

2 Material and Methods

2.1 Measuring Station

The three-dimensional ultrasound imaging can be obtained using a free-hand sonographic probe. The system consists of ultrasound system EchoBlaster 128 from Telemed with multifrequentional linear array (with selected central frequency 5MHz and width 80 mm) and optical navigation system Polaris from Northern Digital Incorporation. Optical navigation system enables tracking of ultrasound probe - its position and orientation during recording of scans [6,7].

To analyze the efficiency of developed algorithm a bone phantom (sawbone) from Synthes company was scanned. The sawbone was placed in a tank filled

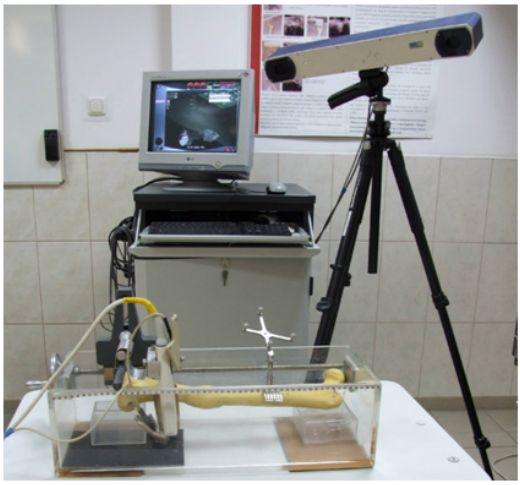


Fig. 1. Measuring station with gantry mechanism

with water. The measuring station was equipped with gantry mechanism applied to achieve possibly regular set of scans (Fig. 1). The ultrasound probe was moved along the long axis of sawbone on the lateral and medial surface, to provide complete imaging. The data was recorded using a calibrated navigated ultrasound probe together with its position and orientation in relation to the reference frame mounted on the scanned object. Using the reference frame the system takes into account movement of object and navigation system. 864 scans were recorded in the locations and orientations visualized on Fig. 2.

Fig. 2. Visualization of recorded ultrasound scans location and orientation

2.2 Tested Segmentation Algorithms

Various typical methods of segmentation were analyzed to identify the bone contour. In our self-developed software (Micorsoft Visual C++) we applied libraries from the Insight Segmentation and Registration Toolkit [8] and Visualization Toolkit [9], with implemented various methods of filtration and segmentation. We tested Otsu criterion, which estimates threshold based on image histogram, Canny Edge Detection and Watershed Segmentation.

Basing on the assumptions on bone tissue ultrasound imaging, an algorithm of bone contour identification has been developed. Because of the high echogenicity and presence of shadow under the contour, the first and second derivative (gradient and laplace transform) have been applied to identify the bone contour. The next step of processing was binarization with a user-selected threshold and defined region of interest. The contour is searched upward or downward basing on the choice of user. The software provides possibility to edit the contour to eliminate possible errors or add points of contour, if it was not recognized. The finally obtained contour is smoothed and may be decimated to reduce the number of vertexes for triangulation algorithm. This option sometimes is crucial, because the number of points can become too large to analyse it, since the ultrasound image size was 512 x 512 pixels and the number of scans was intentionally high. The decimation was equal for all scans to provide possibly regular mesh.

2.3 Tested 3D Modelling Algorithms

The final step of shape recognition was the triangulation procedure. Various algorithms were tested to analyze their efficiency. We applied Delaunay 3D Triangulation (convex hull) with alpha parameter settings and Delaunay 2D Triangulation [9]. To apply the two dimensional Delaunay's triangulation the points of cloud need to be projected on plane. The points were projected on the plane with a normal vector perpendicular to the visible contours, then the triangulation was performed. The triangulation was performed step by step for projection plane rotated and translated along the bone.

The algorithm of Hoppe's Surface Reconstruction Filter [10] enables spatial modelling basing on the unorganized set of points. The algorithm is based on calculations of zero set of estimated distance functions. The topological type of surface is calculated, including the boundary curves. The result of this algorithm is translated and scaled (in relations to the cloud of points) mesh. Therefore the mesh needs to be transformed back to the initial location.

Last applied algorithm was the fast triangulation of unordered point (greedy projection) [11,12]. The greedy projection may be applied for unorganized group of points, also in case of connected elements. It works properly for locally smooth surfaces and in case of smooth transitions between areas of different density of points. Triangulation is performed locally throughout projection of local neighbourhood of point along normal in the point and connecting of unconnected points. The algorithm has a number of parameters to set: maximal number of neighbours (size of neighbourhood), radius (maximal length of edge of each triangle), maximal and minimal angle of triangle, maximal angle of surface.

2.4 Qualitative Evaluation: Similarity of Surfaces

To create a reference 3D model the scans of the Computed Tomography were analysed. The segmentation of DICOM dataset and 3D shape reconstruction was performed using InVesalius software. The model was saved in stereolitographic format.

The result of ultrasound-based model was compared to the reference CT-based model. First step was to match the two models in one coordinate system. The matching procedure was performed using translation and rotation operator to lead the models to best fitting. Additionally the model needed to be flipped because of different orientation of local coordinate system. To compare the models of sawbone obtained using two triangulation techniques for the cloud of points obtained using the developed algorithm, the Hausdorff's distances have been calculated [13].

The Hausdorff's distance for A and B set is defined as:

$$H(A,B) = \max\{\delta(A,B), \delta(B,A)\} \tag{1}$$

where:

$$\delta(A,B) = \max\{\delta(x,B) : x \in A\} \tag{2}$$

$$\delta(B,A) = \max\{\delta(y,A) : y \in B\} \tag{3}$$

$$\delta(x,B) = \min\{d(x,y) : y \in B\} \tag{4}$$

$$\delta(y,A) = \min\{d(x,y) : x \in A\} \tag{5}$$

and d(x,y) is the distance of point x belonging to the set A from element y belonging to set B. Hausdorff's distance provide scalar characterization of surfaces similarity. If the models match the Hausdorff's distance equals zero. High value of Hausdorff's distance means that the models differ significantly.

3 Results

3.1 Results of Segmentation

First part of results section constitute description of results of segmentation of phantom and human bone on ultrasound scans:

1. Otsu criterion (Insight Segmentation and Registration Toolkit) [8]
   Fig. 3 shows, that the echo of sawbone is well recognized (A, B), whereas in case of human lower limb scanning the algorithm recognizes many additional structures (C, D).

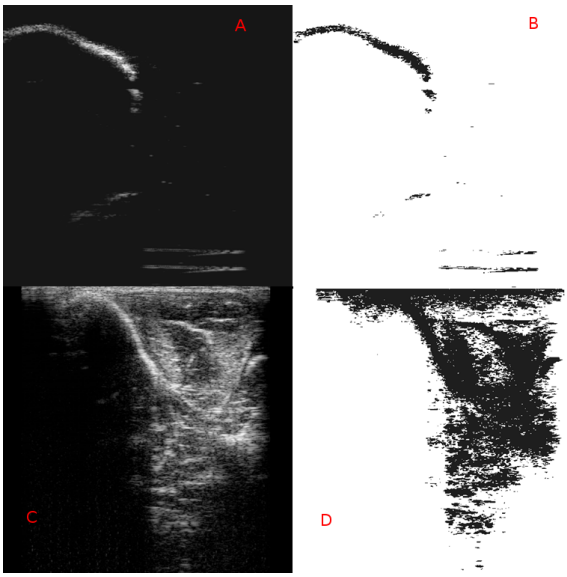


Fig. 3. Result of Otsu criterion binarization for ultrasound scan of sawbone (A,B) and human lower limb (C,D)

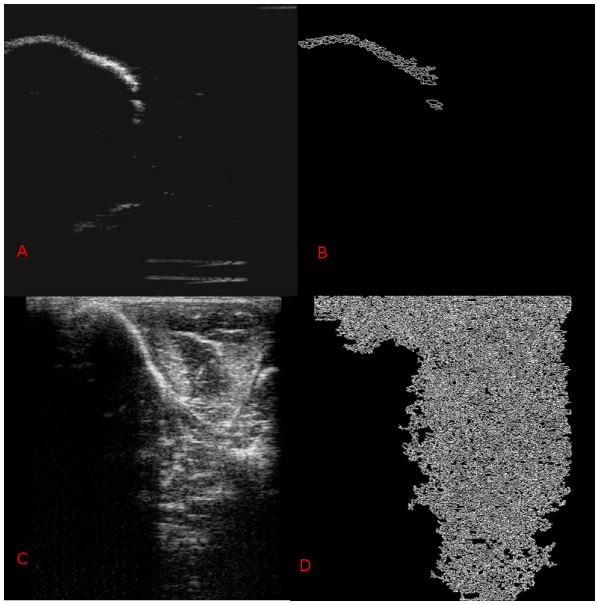


Fig. 4. Result of Canny Edge Detection algorithm for ultrasound scan of sawbone (A,B) and lower limb (C,D)

2. Canny Edge Detector (Insight Segmentation and Registration Toolkit) [8]
   Fig. 4 shows, that the Canny Edge Detector works properly for recognition of sawbone shape (A,B). For lower limb scan (Fig. 4 C, D) variance 0.1 and threshold 63 were applied. The result of segmentation contains more detected contours than it is required. Change of lower and upper threshold and variance did not improve the result of segmentation.
3. Watershed segmentation (Insight Segmentation and Registration Toolkit)
   Our tests of watershed segmentation algorithm implemented in Insight Segmentation and Registration Toolkit considered changes in conductance parameter [8]. The higher is the value of conductance the number of recognized objects is lower, but their shapes are deformed. The recognition of bone contour as a single object is impossible and the procedure is time consuming.
4. Developed algorithm
   The developed algorithm has been tested for a large dataset (864 scans recorded in various locations and orientations). In the developed software the user needs to define: the regions of interest, the thresholds, and the direction of contour searching for particular scans (upwards/downwards) (Fig. 5). All scans were segmented and the contours were manually corrected. As a result of segmentation a cloud of points has been obtained. Despite the extensive dataset, there appeared deficiencies in the modelled shape of sawbone. The lack of data is caused by no access to the part of surface because of mounted reference frame on the proximal part of shaft. The lacking contours were not manually filled.

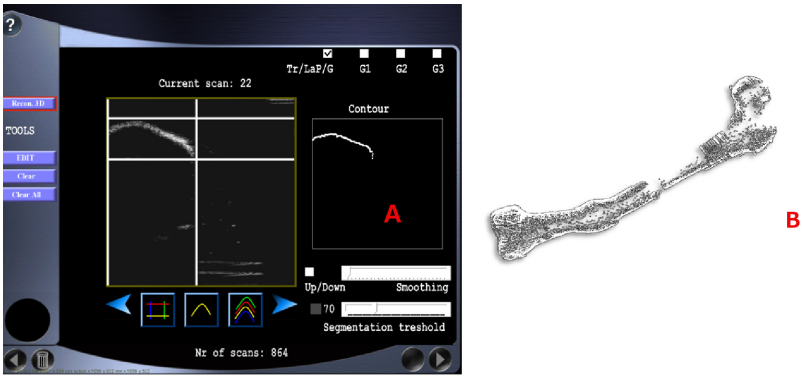


Fig. 5. Developed algorithm of segmentation: single contour (A), cloud of points (B)

3.2 Results of 3D Modelling

1. Delaunay 3D Triangulation (convex hull) (Visualization Toolkit) [9]
   The three-dimensional Delaunay triangulation without zero alpha parameter is not capable to define the triangulated mesh for the cloud of points (Fig. 6). That algorithm enables obtaining the smallest convex set containing the set

of points. The result reminds the film stretched on the set of points. Non-zero alpha coefficient enables reconstruction of proper shape, the higher is the value of alpha the mesh gets closed, but also in certain regions deformed.

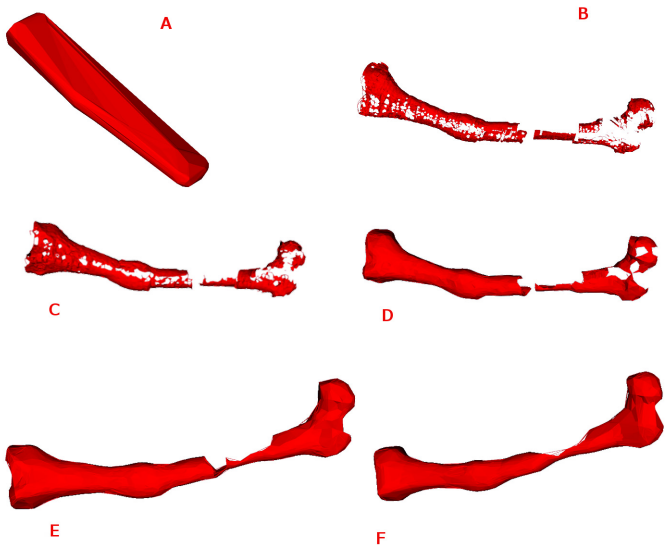


Fig. 6. Result of 3D Delaunay triangulation (A: alpha = 0, B: alpha = 1.95, C: alpha = 2.95, D: alpha = 5.95, E: alpha = 10.95, F: alpha = 15.95)

2. Delaunay 2D Triangulation (Visualization Toolkit) [9]
   The procedure of Delaunay 2D triangulation was partially manual and time consuming, however the results of triangulation were visually satisfactory (Fig. 7).
3. Hoppe's Surface Reconstruction Filter (Visualization Toolkit) [10,9]
   The tests of Hoppe's algorithm revealed that in case of irregular distribution of points the estimated surface is deformed. In the middle of surface a segment of sawbone shaft is visible. However both shaft and epiphyses are highly deformed as a side effect of reconstruction process (Fig. 8A).
4. Fast triangulation of unordered point (Point Clouds Library) [11,12]
   The result of greedy projection presented on Fig. 8 B has been obtained for following parameters: radius 250, scaling factor 2.5, maximal number of neighbours 200, maximal angle of surface 45 degrees, minimal and maximal angle of triangle 10 and 120 degrees respectively. Greedy projection provides good result of 3D reconstruction. It does not require assistance of user, however it is necessary to provide proper dataset to visualize the shape of object.

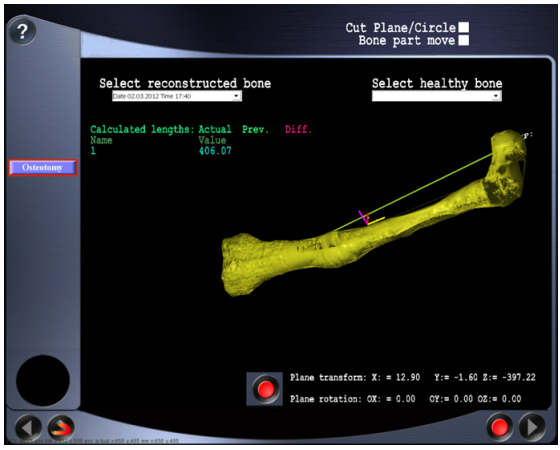


Fig. 7. Result of 2D Delaunay's triangulation

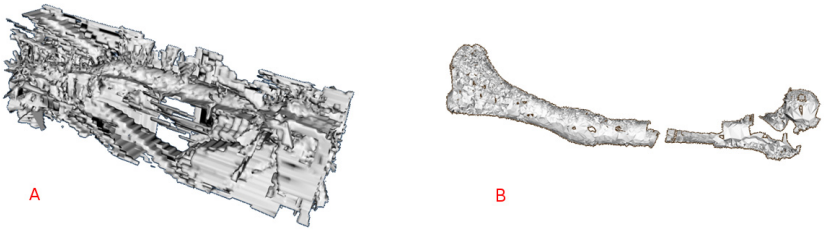


Fig. 8. Result of Hoppe's Surface Reconstruction Filter (A) and Fast triangulation of unordered point (B)

3.3 Quantitative Evaluation: Similarity of Surfaces

To evaluate the results of shape recognition quantitatively, the CT-based model was matched with two ultrasound-based models obtained using greedy projection and 2D Delaunay triangulation (Fig. 9). The similarity analysis of models was performed using Hausdorff distance parameter. The results of calculations are presented in Tab. 1.

Smaller value of Hausdorff's distance has been obtained for greedy projection triangulation technique, and higher for Delaunay's 2D triangulation.

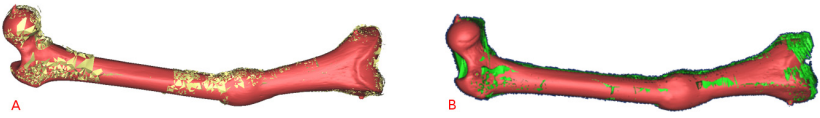


Fig. 9. View of CT and ultrasound-based models after matching (A: greedy projection, B: 2D Delaunay triangulation models matched with CT model)

Table 1. Results of similarity analysis for CT-based and ultrasound-based models depending on applied triangulation method

Triangulation method	Hausdorff's distance
Delaunay's triangulation 2D	27.56 mm
"Greedy" projection triangulation	21.98 mm

4 Discussion

The ultrasound imaging using free hand probe provide possibility to measure a three-dimensional shape of small objects, providing large dataset with possibly uniform distribution. The ultrasound imaging could become an alternative to expensive (as magnetic resonance imaging) or invasive (as computed tomography) to get the shape of bone surface useful for virtual planning of orthopaedic surgeries, such as bone deformities correction.

The aim of this study was to analyse and test the efficiency of surface identification methods for ultrasound scans. An important task was to test the algorithm for a large dataset recorded for femur sawbone. The identification and triangulation algorithms were tested. Selected and developed segmentation algorithm involving contour segmentation using gradient, laplace and thresholding procedure for a defined region of interest as well as 2D/3D Delaunay and greedy projection triangulation were analysed.

Basing on evaluated Hausdorff's distance, the efficiency of two methods of triangulation applied for the best cloud of points dataset were compared. In both cases the obtained model is deformed comparing to the CT-based model. Some influence on the obtained Hausdorff's distance had the matching procedure, leading to fitting of both models in the same coordinate system. A great influence on the models similarity had the lack of scanning in the region of mounted reference frame. The real problem is the possibility to record appropriate ultrasound dataset. A great influence on the obtained results of shape reconstruction had the setting of ultrasound probe regarding the scanned sawbone. The probe should be perpendicular to the scanned surface. Non-smooth surface in some regions of model was caused by improper position of probe regarding the sawbone surface. The probe pressing was not necessary, because the measurement was performed in the tank filled with water.

Tissue structures identification is a time consuming procedure, because of complex mathematical calculations and manual corrections. In clinical practice an automated tool is necessary to recognize shape of tissue structures in reasonable time. The most difficult issue in ultrasound data analysis is that certain scan has different greyness distribution and any settings of segmentation procedure need to be adjusted for each scan separately.

Obtained results of investigation helped to evaluate the efficiency of possible algorithms, however there is no ideal solution so far. The result of reconstruction depend on the analysed dataset. The designed algorithm of data analysis can be applied for small objects, for large objects it is impossible to record and

analyse that amount of data in reasonable time. For small objects and uniform distribution of points the algorithm is efficient.Despite of precise planning of data recording using gantry mechanism and registration of large dataset (864 scans), still it was not possible to record data in some area, so the dataset was not optimal. Currently the computed tomography is the best method of scanning with high resolution and uniformly. To obtain that type of imaging it would be necessary to build a kind of ultrasound tomograph with acoustic medium providing the ultrasound imaging in reflection or transmission mode.

Acknowledgement. We express special thanks to Prof. Josef Kozak and Aesculap BBraun (Tuttlingen, Germany) for providing all the equipment applied during the tests. This work benefited from the use of the Insight Segmentation and Registration Toolkit and Visualization Toolkit, developed as an initiative of the U.S. National Library of Medicine, as well as freeware: Point Cloud Library, InVesalius and MeshLab.

The work was partially supported by fellowship co-financed by European Union within European Social Fund.

References

1. Coupé, P., Hellier, P., Azzabou, N., Barillot, C.: 3d freehand ultrasound reconstruction based on probe trajectory. In: Duncan, J.S., Gerig, G. (eds.) MICCAI 2005. LNCS, vol. 3749, pp. 597–604. Springer, Heidelberg (2005)
2. Zhang, Y., Rohling, R., Pai, D.K.: Direct surface extraction from 3d freehand ultrasound images. In: Proceedings of the Conference on Visualization 2002, VIS 2002, pp. 45–52. IEEE Computer Society, Washington, DC (2002)
3. Rohling, R.N.: 3D Freehand Ultrasound: Reconstruction and Spatial Compounding. University of Cambridge (1998)
4. Daanen, V., Tonetti, J., Troccaz, J.: An information fusion method for the automatic delineation of the bone-soft tissues interface in ultrasound images. In: Sonka, M., Kakadiaris, I.A., Kybic, J. (eds.) CVAMIA/MMBIA 2004. LNCS, vol. 3117, pp. 218–229. Springer, Heidelberg (2004)
5. Krowicki, P., Krysztoforski, K., Świątek-Najwer, E., Będziński, R.: New approach to bone surface reconstruction from 2.5d sonographic dataset. In: Jorge, R.M.N., Tavares, J.M.R.S., Barbosa, M.P., Slade, A.P. (eds.) Technology and Medical Sciences, pp. 321–324. A Balkema book, Taylor & Francis (2011)
6. Świątek-Najwer, E., Krysztoforski, K., Dragan, S.L., Będziński, R.: The investigation of the lower limb geometry using 3d sonography and magnetic resonance. Measurement 45(4), 702–710 (2012)

7. Krysztoforski, K., Krowicki, P., Świątek-Najwer, E., Będziński, R., Keppler, P.: Noninvasive ultrasonic measuring system for bone geometry examination. The International Journal of Medical Robotics and Computer Assisted Surgery 7(1), 85–95 (2011)
8. Ibanez, L., Schroeder, W., Ng, L., Cates, J.: The ITK Software Guide, 1st edn. Kitware, Inc. (2003), http://www.itk.org/ItkSoftwareGuide.pdf ISBN 1-930934-10-6
9. Schroeder, W., Martin, K., Lorensen, B.: The Visualization Toolkit–An Object-Oriented Approach To 3D Graphics, 4th edn. Kitware, Inc. (2006)
10. Hoppe, H., DeRose, T., Duchamp, T., McDonald, J., Stuetzle, W.: Surface reconstruction from unorganized points. SIGGRAPH Comput. Graph. 26(2), 71–78 (1992)
11. Marton, Z.C., Rusu, R.B., Beetz, M.: On Fast Surface Reconstruction Methods for Large and Noisy Datasets. In: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Kobe, Japan, May 12-17 (2009)
12. Rusu, R.B., Cousins, S.: 3D is here: Point Cloud Library (PCL). In: IEEE International Conference on Robotics and Automation (ICRA), Shanghai, China, May 9-13 (2011)
13. Alt, H., Buchin, M.: Can we compute the similarity between surfaces? Discrete and Computational Geometry 43(1), 78–99 (2010)

Hip Joint Centre Localization: Evaluation of Formal Methods and Effects on Joint Kinematics

Magdalena Żuk, Ewelina Świątek-Najwer, and Celina Pezowicz

Wroclaw University of Technology, Wyb. Wyspiańskiego 27,50-370 Wrocław, Poland
magdalena.zuk@pwr.wroc.pl
http://www.biomech.pwr.wroc.pl

Abstract. Accurate methods of hip joint centre (HJC) localization are necessary in gait analysis. It was shown that current methods could involve large mislocation errors, what affect both kinematics and kinetics. The purpose of present study was to compare three different HJC localization methods: predictive methods, functional calibration methods, and medical imaging analysis technique, as well as to assess of its effect on joint kinematic variables during gait on population of three able-bodied subjects. Significant deviations were observed for HJC determined with predictive method compared to ultrasound technique (44 ± 7 mm), resulted in errors propagating into calculated joint angles (mean 2.7°). While lower deviations observed for functional method comparing to ultrasound technique (mean 23 ± 6 mm) results in negligible joint angle differences (mean 0.6°). Therefore, functional methods are highly recommended in the absence of imaging technique.

Keywords: hip joint centre, gait analysis, sphere fitting, freehand ultrasound.

1 Introduction

An accurate localization of hip joint centre (HJC) is of interest across wide range of applications, especially in lower limb movement analysis and computer aided surgical intervention to determine alignment of lower limb anatomical axes [1,2,3]. In gait analysis, location of HJC affects both kinematic and kinetic variables [4,5]. HJC is assumed to be ball-and-socket joint with rotation centre is coincident with centre of femur head. Therefore there are three different approach to HJC localization: predictive methods, functional calibration methods, and medical image techniques in recent years [4,6].

The most widely used methods in gait analysis are predictive methods based on anthropometric estimation. There are available several regression equation [7,8,9,10] based on empirical relation between position of palpable pelvis landmarks and HJC. Currently, the most widely applied is Davis equation [7], which estimates position of HJC from positions of pelvis markers and lower limb length. This regression equitation was developed on the basis of planar X-rays of very limited and specific population of living subjects (25 subjects), which obviously

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_6, © Springer International Publishing Switzerland 2014

lead to inaccurate results [11,6]. Although, later equations [8,9], stemmed from medical images from wider population, with proven greater accuracy [6,11] were proposed, mainly Davis equation is implemented in commercial software.

The next type are functional methods, which determine centre of rotation from pelvis and thigh marker positions during calibration movements. For this purpose, there are two different approaches, sphere fitting techniques and transformation techniques. In first approach, centre and radius of sphere as optimized to segment marker trajectories during movement of femur relative to pelvis [1,8,12,13,14]. In transformation techniques, pelvis and thigh local coordinate systems are defined and joint centre is approximated as fixed point in both local coordinate systems [1,15,16]. Both approaches were compared and evaluated, simulation studies gave better results for transformation techniques [1] and an in-vitro analysis [17], while sphere fitting techniques performed better in in-vivo study (cadaver study [18] and living subject analysis using medical imaging [6]). Although, choice of appropriate approach among functional methods, seems to remain unresolved, in general functional methods perform better than predictive methods [6,11].

Medical imaging techniques, such as MRI [9], bi-plane X-rays [11,19,20] and three dimensional ultrasound [6,21], are used rather as a gold standard measurement tools in the evaluation of the other methods, than in clinical gait analysis.

Many different predictive and functional methods were compared and evaluated using medical imaging data in previous papers [1,6,11,17,18,20,22]. It has been shown, that hip joint centre position discrepancy alter both kinematics and kinetics [4,5]. In these studies, different HJC position errors from specific range were introduced to kinematic model to calculate corresponding kinematic and kinetic variables. Analysed errors were not actual HJC mislocation obtained for the same population as kinematic description. Direct effect of HJC location estimated using different methods on kinematic and kinetic variables have not been investigated.

The purpose of the present study was to compare three different HJC localization methods: regression equation, functional method and ultrasound technique, as well as assessment of direct effect of HJC mislocation on joint kinematic variables during gait.

2 Materials and Methods

Three able-bodied subjects (two female and one male) without a walking disability were analyzed (A: female, height: 167 cm, BMI: 26.89 kg/m^2, age: 33 years, B: male, height: 188 cm, BMI: 24.62 kg/m^2, age: 25 years, C: female, height: 158 cm, BMI: 18.43 kg/m^2, age: 26 years). Subjects were on purpose diverse in terms of body mass index. Participants underwent three different anatomical calibration procedure: 1) palpation of external bone landmarks of lower limb 2) functional calibration movements 3) ultrasound examination and gait analysis. Hip joint centers localization were estimated using three different methods: freehand ultrasound measurement (medical image technique), least squared sphere fitting (functional method) and Davis equation [6] (predictive method).

2.1 Free-Hand Ultrasound Measurement

The ultrasound measuring system, described and validated in previous paper [2,23], consisting of ultrasound probe EchoBlaster 128 (Telemed, Lithuania) and infrared optical tracking system – Polaris (NDI, Canada) was applied (Fig. 1). 2D ultrasound images are transformed to 3D coordinates of each pixels through a calibrated ultrasound probe equipped with active markers. Thereby the system measures spatial geometry of bones and soft tissues with a high accuracy.

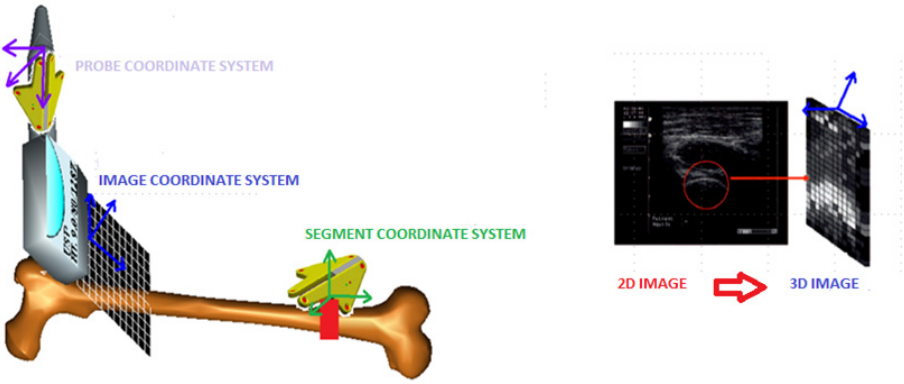


Fig. 1. Principle of free-hand ultrasound system

Six scans of femur head of right lower limb for different probe orientations were recorded for each subject. Hip joint centre position was estimated using a simplified method – as the centre of circle matched to femur head contour averaged over four selected scans. Position of HJC was determined relative to the marker cluster placed on distal part of thigh.

2.2 Functional Method

The functional calibration movement was arc movement consisting of a hip flexion, a half circumduction and a extension to the neutral position. Each subject was asked to perform series of calibration movement at a self-selected speed, while motion capture system (Optotrak Certus, NDI, Canada) tracked position and orientation of clusters of non-collinear active markers (Optotrak Smart Marker Rigid Body, NDI, Canada) placed on the pelvis and distal part of thigh.

Algebraic sphere fitting method together with some of transformation techniques were applied to estimate centre of hip rotation from collected data. First, the origin of thigh cluster coordinate system was transformed into a coordinate system of pelvis cluster.

It was assumed, that origin of femoral segment cluster move on the surface of sphere with specific radius around common centre (pelvic centre of rotation).

Centre of rotation was estimated using algebraic sphere approach, called Kasa-Delonge method [1], which in this case minimizes the following sum:

$$f_{algebraic}(c_p, r_1, ..., r_n) = \sum_{i=1}^{n} (\parallel p_{ci} - c_p \parallel^2 - r^2)^2 \tag{1}$$

where c_p is the centre and r is the radius of sphere, p_{c_i} is the origin of femoral cluster in pelvis coordinate system in the time frames $i = 1, ..., n$. This minimization task has simply close solution (opposed to geometric sphere fitting method, where minimization is a non-linear problem solved iteratively).

The centre of rotation and radius of sphere were calculated from coordinates of the thigh cluster origin during functional movement using a Matlab script [24]. Root-mean-square error (RMS error) and normalized RMS (radius RMS error divided by mean radius) for radius were calculated. HJC position in pelvic cluster coordinate system (c_t) was transformed into thigh cluster coordinate system using:

$$c_t = \frac{1}{n} \sum_{i=1}^{n} c_{ti} = \frac{1}{n} \sum_{i=1}^{n} {T_{t_i}}^{-1} T_{p_i} {c_p}', \tag{2}$$

where c_{ti} is instantaneous HJC position in the thigh cluster coordinate system, T_{ti} is global transformation of thigh cluster coordinate system , T_{p_i} is is global transformation of pelvis cluster coordinate system in $i = 1, ...n$ time frames and c_p is sphere centre in pelvis coordinate system.

2.3 Predictive Method

Position of hip joint centre was estimated using Davis method [7] as widely used predictive method in gait analysis. HJC position was calculated on the basis of positions of sacrum, anterior superior iliac spine right and left (ASIS right and left) and lower limb length. Anatomical landmark locations were measured as in 2.4. HJC was defined in pelvic anatomical coordinate system and recalculated to femur cluster coordinate system for neutral, standing position.

2.4 Gait Analysis

Right lower limb motion during gait was tracked using motion capture system (Optotrak Certus, NDI, Canada) consisted of single position sensor with three cameras (Fig. 2). Data acquisition, joint kinematic calculations, the gait visualization and data recording were performed with a software developed by the author [25].

Four clusters of three non-collinear active markers (Optotrak Smart Marker Rigid Body, NDI, Canada) on rigid plates were placed on each segment (pelvis, thigh, shank and foot) enables tracking of each lower limb segment independently. Fully anatomical gait analysis protocol based on anatomical joint coordinate system specified by twelve palpable anatomical bony landmarks of right lower limb and HJC was used (according to ISB recommendation [26]).

Anatomical landmarks were indicated using a navigated pointer (equipped with markers)[25], while its positions were registered in the reference of corresponding cluster. Positions of thus defined virtual markers were calculated on the basis of an actual clusters position and anatomical calibration data.

Relative orientation of adjacent lower limb segments was defined as the relative orientation of anatomical coordinate systems determined using Cardan's angular convention [27]. The three Cardan angles were used to describe the hip and knee joint action of a flexion/extension, an adduction/abduction, and an internal/external rotation.

Subjects walked barefoot with low speed. Three gait cycles for each subjects were selected from the series of recorded cycles. Three different HJC relative position, obtained using predictive, functional and ultrasound methods, were introduce to kinematic model together with calculated trajectories of other anatomical landmarks. The results are the three kinematic descriptions of gait (corresponding to three hip joint centre localization methods) for the same set of gait cycles. Data was processed, including filtering of marker trajectories with 4th order low-pass Butterworth filter (cut-off 6Hz) and joint angle normalisation to 100 point per cycle using Matlab.

Fig. 2. Gait analysis system

For visual comparison of three kinematic description of gait, joint angle curves (mean over 3 cycles) were plotted for each subject. Differences between two pairs of angles (angles for Davis HJC estimation and functional method compared to the kinematic variables for ultrasound measurement) were averaged over gait cycle for each subject and presented in box plots.

3 Results

In free-hand ultrasound method, position of HJC was estimated as a mean circle centre matched to four femur head contours from different scans. Spread of circle centres around average differs among subjects, RMS errors were from 4,9 mm (subject C), 6.1 mm (subject B), to even 13.7 mm (subject A).

Thigh cluster trajectory (consisted of at least 645 data points for subject A) were fitted to sphere with relatively high accuracy (Fig. 3, Fig. 4, Fig. 5). Calculated RMS errors for radius were from 3 mm (subject C) to 6 mm (subject A). Range of motion (ROM) during functional calibration was substantial.

There were differences between HJC positions estimated with three procedures (Tab. 1) observed especially in a transverse plane. Generally HJC position estimated by functional method is more comparable to free-hand measurement (mean linear distance 23 ± 6mm) than predictive method (mean linear distance 44 ± 7 mm). The greatest discrepancies were observed in subject A.

Observed HJC location differences altered values of kinematic variables (Figures 6, 7). Kinematics obtained for functional HJC are more comparable to ultrasound HJC than those for predictive method. The impact of the HJC position on kinematic parameters vary across the subjects.The greatest discrepancies were observed for flexion/extension angles in subject A and for ab/adduction angles in subject C. HJC mislocations did not almost neither affect knee nor hip rotation.

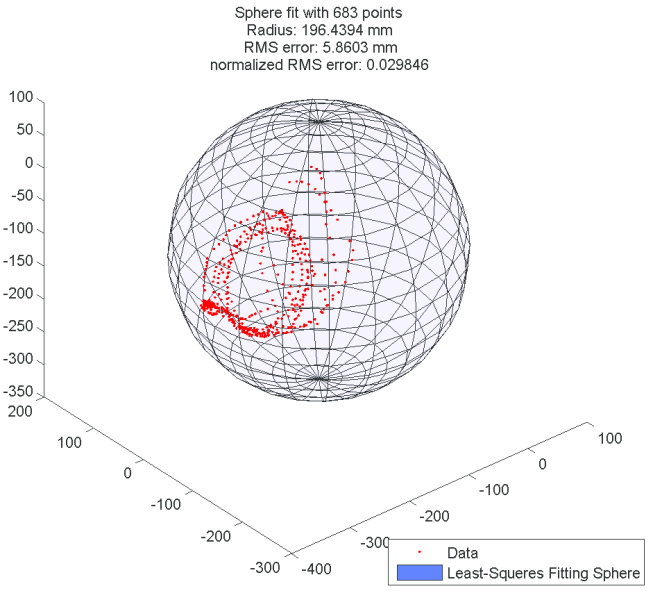


Fig. 3. Sphere fitted to thigh cluster trajectory of subject A

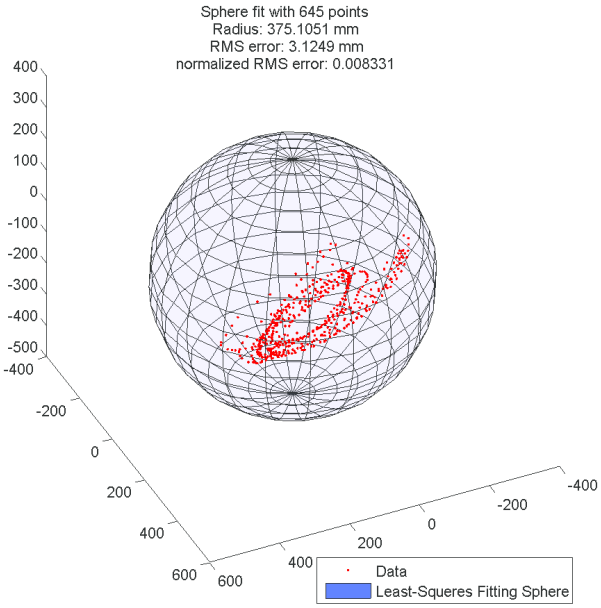


Fig. 4. Sphere fitted to thigh cluster trajectory of subject B

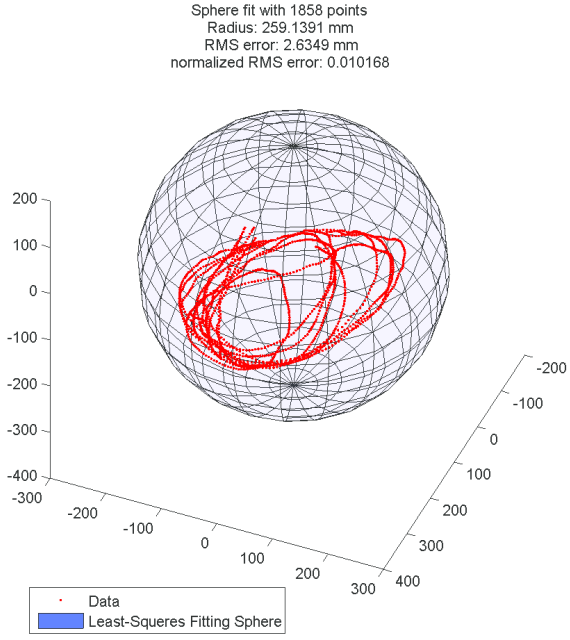


Fig. 5. Sphere fitted to thigh cluster trajectory of subject C

Table 1. Linear distance and distance in each plane between hip joint centre positions indicated by three different methods (regression equation, free-hand ultrasound measurement, functional calibration); coordinates differences should be interpreted as follows: in x anterior, y superior, z lateral in the femur coordinate system (defined by lateral and medial epicondyles indicated with pointer and femur head indicated by ultrasound measurement)

	Subject	ΔX	ΔY	ΔZ	Δ [mm]
Davis equation vs. ultrasound measurement	A	-41	28	-4	50
	B	33	-1	14	36
	C	17	-3	41	44
				mean	43 ± 7
Sphere fitting vs. ultrasound measurement	A	2	24	16	29
	B	19	4	9	21
	C	9	-4	-15	18
				mean	23 ± 6
Sphere fitting vs. Davis equation	A	-43	-4	-21	48
	B	14	-4	6	16
	C	8	0	55	56
				mean	40 ± 21

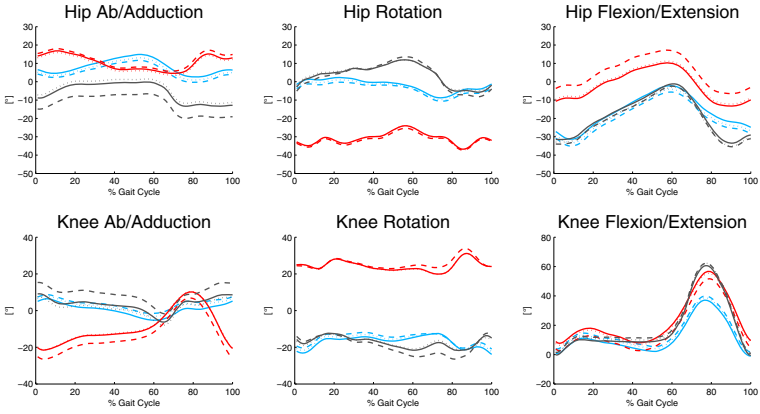


Fig. 6. Kinematic variables as calculated for three HJC localization procedures: Davis equation(dashed line), functional method (doted line) and ultrasound procedure (solid line), averaged across three cycles for three subject: A(red), B(blue), C (grey)

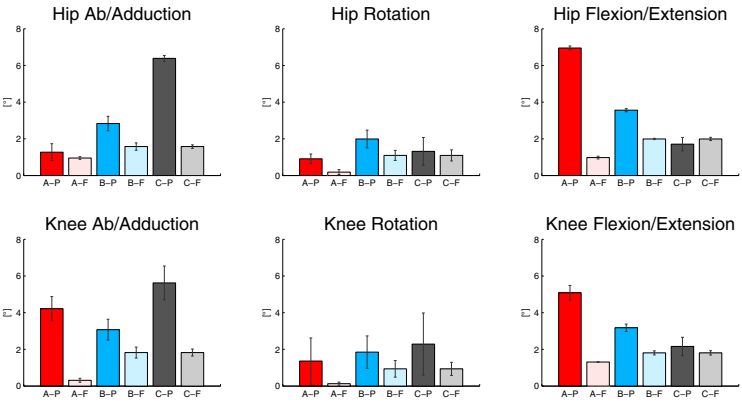


Fig. 7. Differences between joint angles calculated for three HJC localization procedures averaged over gait cycle for each subject A(red), B(blue), C (grey). Kinematic variables corresponding to Davis predictive (symbol P, darker bar) and functional (symbol F, lighter bar) methods were compared to those for ultrasound procedure

4 Discussion

Three procedures were applied: ultrasound-based method, functional (sphere fitting algorithm) and predictive (Davis equation) to estimate HJC localizations of three subject very diverse in terms of physique. Although applied ultrasound measurement system enables measurement of bone and tissue geometry with the high accuracy, applied HJC estimation method (by circle matching to the contour) is simplified and hence its accuracy is limited (as shown in 3). This ultrasound method should be improved in the future, especially that high accuracy of such method is achievable as it was shown in previous paper [6]. However it is reasonable to assume, that estimated HJC is inside femur head. The disadvantage of using this method in gait analysis is a need for additional examination using further equipment (properly calibrated ultrasound probe). Furthermore, interpretation of ultrasound images is difficult, and the bone contour is not always distinguishable from the surrounding tissues. Despite these limitations, the high reproducibility of ultrasound musculoskeletal geometry measurement is possible.

Davis equation is based on data from limited and specific population, it depends on accuracy of pelvis anatomical landmarks palpation. It is reported to performed badly [6,11]. Also in this study, results significantly differ from the other, which also reveals in kinematic data. The main advantage of using this method in the gait analysis is no need for any additional anthropometric measurements, equipment or performance of calibration movement.

Algebraic sphere fitting method together with some of transformation techniques were implemented to estimate HJC on the basis of current marker set (marker clusters). Functional methods are independent from anatomical calibration accuracy, however it is prone to soft tissue artifacts, especially during

extensive calibration movement. In previous paper was shown, that functional methods outperformed predictive methods [6,4]. Implemented sphere fitting algorithm estimated centers of rotation with relatively low radius RMS errors. Functional procedure and ultrasound method gave similar results. Functional methods do not require any additional measurements or equipment. Calibration movements are not time consuming, however can not be performed by patients with reduced hip mobility.

HJC mislocations, obviously propagate to hip and knee kinematic variables. HJC mislocations almost din not affect observed knee and hip rotation, what is consistent with anatomical coordinate frames definitions. HJC mislocation in mediolateral direction affects add/abduction angles, while discrepancy in anterior-posterior direction influences flexion/extension angles.

Deviation in HJC location using functional and ultrasound methods (mean 23±6 mm) can be overestimated due to limited ultrasound method (reported deviations are smaler [21,6,17,11]). Despite the deviation in HJC location, the corresponding kinematic data sets are similar, while kinematic data for Davis method differ considerably.

Acknowledgement. Measurements were performed using equipment provided by Aesculap BBraun (Tuttlingen, Germany). We express special thanks to Prof. Josef Kozak for providing the devices for tests.

References

1. Ehrig, R., Taylor, W., Duda, G., Heller, M.: A survey of formal methods for determining the centre of rotation of ball joints. Journal of Biomechanics 39(15), 2798–2809 (2006)
2. Świątek-Najwer, E., Będzinski, R., Dragan, S.Ł., Krysztoforski, K.: Investigation of lower limb mechanical axis using 3d sonography and magnetic resonance. Measurement 45, 702–710 (2012)
3. Świątek-Najwer, E., Będzinski, R., Krowicki, P., Krysztoforski, K., Keppler, P., Kozak, J.: Improving surgical precision - application of navigation system in orthopedic surgery. Acta of Bioengineering and Biomechanics 10, 55–62 (2008)
4. Stagni, R., Leardini, A., Cappozzo, A., Benedetti, M., Cappello, A.: Effects of hip joint centre mislocation on gait analysis results. Journal of Biomechanics 33, 702–710 (2000)
5. Croce, U., Leardini, A., Chiari, L., Cappozzo, A.: Human movement analysis using stereophotogrammetry: Part 4: assessment of anatomical landmark misplacement and its effects on joint kinematics. Gait & Posture 21, 226–278 (2005)
6. Sangeux, M., Peters, A., Baker, R.: Hip joint centre localization: Evaluation on normal subjects in the context of gait analysiss. Gait & Posture 34, 324–328 (2011)
7. Davis, R., Õunpuu, S., Tyburski, D., Gage, R.: A gait analysis data collection and reduction technique. I Human Movement Science 10, 575–587 (1991)
8. Bell, A., Pedersen, D., Brand, R.: A comparison of the accuracy of several hip center location prediction methods. Journal of Biomechanics 23, 617–621 (1990)
9. Harrington, M., Zavatsky, A., Lawson, S., Yuan, Z., Theologis, T.: Prediction of the hip joint centre in adults, children, and patients with cerebral palsy based on magnetic resonance imaging. Journal of Biomechanics 40, 595–602 (2007)

10. Seidel, G., Marchinda, D., Dijkers, M., Soutas-Little, R.: Hip joint center location from palpable bony landmarks–a cadaver study. Journal of Biomechanics 28, 995–998 (1995)
11. Leardini, A., Cappozzo, A., Catani, F., Toksvig-Larsen, S., Petitto, A., Sforza, V., Cassanelli, G., Giannini, S.: Validation of a functional method for the estimation of hip joint centre location. Journal of Biomechanics 32, 99–103 (1999)
12. Cappozzo, A.: Gait analysis methodology. Human Movement Science 32, 27–50 (1984)
13. Piazza, S., Okita, N., Cavanagh, P.: Accuracy of the functional method of hip joint center location: effects of limited motion and varied implementation. Journal of Biomechanics 34, 967–973 (2001)
14. Halvorsen, B.: Bias compensated least squares estimate of the center of rotation. Journal of Biomechanics 36, 999–1008 (2003)
15. Schwartz, M., Rozumalski, A.: A new method for estimating joint parameters from motion dat. Journal of Biomechanics 38, 107–116 (2005)
16. Piazza, S., Erdemir, A., Okita, N., Cavanagh, P.: Assessment of the functional method of hip joint center location subject to reduced range of hip motion. Journal of Biomechanics 38, 349–356 (2004)
17. Lopomo, N., Sun, L., Zaffagnini, S., Giordano, G., Safran, M.: Evaluation of formal methods in hip joint center assessment: An in vitro analysis. Clinical Biomechanics 25, 206–212 (2010)
18. Cereatti, A., Donati, M., Camomilla, V., Margheritini, F., Cappozzo, A.: Hip joint centre location: An ex vivo study. Journal of Biomechanics 42, 818–823 (2009)
19. Pillet, H., Sangeux, M., Hausselle, J., Rachkidi, R., Skalli, W.: A reference method for the evaluation of femoral head joint center location technique based on external markers. Gait & Posture 39, 655–658 (2014)
20. Kirkwood, R.N., Culham, E., Costigan, P.: Radiographic and non-invasive determination of the hip joint center location: effect on hip joint moments. Clinical Biomechanics 14, 227–235 (1999)
21. Hicks, J., Richards, J.G.: Clinical applicability of using spherical fitting to find hip joint center. Gait & Posture 22, 138–145 (2005)
22. Speirs, A., Benoit, D., Beaulieu, M., Lamontagne, M., Beaulé, P.: The accuracy of the use of functional hip motions on localization of the center of the hip. HSS Journal 8, 192–197 (2012)
23. Krysztoforski, K., Krowicki, P., Świątek-Najwer, E., Będzinski, R., Keppler, K.: Noninvasive ultrasonic measuring system for bone geometry examination. International Journal of Medical Robotics and Computer Assisted Surgery 7, 85–95 (2011)
24. Jennings, A.: Sphere fit (least squared) (2013), `http://www.mathworks.com/matlabcentral/fileexchange/34129-sphere-fit-least-squared`.
25. Żuk, M., Krysztoforski, K.: Marker set for gait analysis and its implementation in custom software. In: Proceedings of Conference Biomechanics, Bialystok, pp. 333–334 (2011)
26. Wu, G.: Isb reccommendation on definitions of joint coordinate system of various joints for reporting of human joint motion-part i:ankle,hip, and spine. Journal of Biomechanics 35, 543–548 (2002)
27. Tupling, T.: Use of cardan angles to locate rigid bodies in three-dimensional space. Med. & Biol. Eng. & Comput. 25, 527–532 (1987)

Radiological Atlas for Patient Specific Model Generation

Jacek Kawa, Jan Juszczyk, Bartłomiej Pyciński,
Paweł Badura, and Ewa Pietka

Silesian University of Technology, Faculty of Biomedical Engineering,
Zabrze, Poland
jacek.kawa@polsl.pl

Abstract. The paper presents the development of a radiological atlas employed in an abdomen patient specific model verification.

After a patient specific model introduction, the development of a radiological atlas is discussed.

Unprocessed database, containing DICOM images and radiological diagnosis presented. This database is processed manually to retrieve the required information. Organs and pathologies are determined and each study is tagged with specific labels, e.g. 'liver normal', 'liver tumor', 'liver cancer', 'spleen normal', 'spleen absence', etc. Selected structures are additionally segmented. Masks are stored as gold standard.

Web service based network system is provided to permit PACS-driven retrieval of image data matching desired criteria. Image series as well as ground truth images may be retrieved for benchmark or model-development purposes. The database is evaluated.

Keywords: patient specific model, radiological database, abdomen.

1 Introduction

Patient specific model is a popular tool of modern health care. It is used in different areas of medicine and biology to permit customized treatment of a diseases, laboratory testings etc. In a surgery, especially its minimally-invasive branch, patient specific model (PSM) includes current locations and forms of organs and lesions in the operating zone. It permits detailed planning of intervention (including points of entry and marking critical zones) and navigation during surgical procedure. PSM in this area is therefore mainly customized to patients anatomy and may neglect physiology. In process of PSM development, expert knowledge (e.g. 'most human have two kidneys located symmetrically in abdomen, liver located in right part of abdomen, spleen located in opposite area etc.'), sometimes presented as radiological template or phantom, is transformed into detailed 3D model with each organ described in terms of position, shape and possibly additional parameters as density, resilience etc.

As current location and shape of organs and lesions is visible on medical images, creation of PSM usually involves processing of multiple slices of either high

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_7, © Springer International Publishing Switzerland 2014

resolution Computed Tomography (CT) images or different sequences of Magnetic Resonance (MR) images to: (1) segment and recognize organs and lesions, (2) find anatomical landmarks, and (3) estimate possible tolerances of parameters based on image meta-data. Quality of each step (and thus PSM generation procedure) and level of automation is usually higher in atlas-based method, where location or (and) shapes are initially set and only modified thorough procedure. Such approach reduces many false positive and false negative-kind of errors that would otherwise have to be corrected in error-prone or time-demanding procedures. It also provides some initial information, that may be used in image processing algorithms.

Development of a robust generation method of PSM of human abdomen to be employed in a non-invasive surgery is also one of a long-term goals of a current study [1]. PSM in combination with calibration phantom and tracking devices should permit navigation and fusion of live ultrasonography (USG) images with pre-intervention CR and MR volumes.

Atlas based approach had been chosen as a basis for both segmentation (some cases will be used to build segmentation templates and models) and verification phases of PSM development (other cases will be used for verification). The construction of an atlas (gathering exemplary cases, analysis, quantitative description and gold standard images) is a first phase of the study.

1.1 Medical Databases and Atlases

Past two decades witnessed a rapid growth in development of computer-aided diagnosis (CAD) systems. Progress in medical imaging techniques as well as computational power enables advanced multidimensional data processing. Cooperation of engineering and medical research institutes is more common nowadays. That implies the need to establish reliable benchmark datasets. Acquisition and sharing of image data is not uncommon, yet annotating such data requires a lot of effort. Some of existing databases were considered as a basis for the PSM model generation task.

First attempts of creation an annotated database have been done in 1990's in mammography [2,3,4]. For example, the Digital Database for Screening Mammography (DDSM) [4] contains 2620 screening mammograms with lesions manually delineated and annotated by expert radiologists. Moreover, some software has been added to provide basic image and metadata processing and presentation. The annotated chest radiograms have also been collected into large databases to provide resource for lung cancer diagnosis. Shiraishi et al. [5] report 247 images inspected by 20 radiologists.

Benchmark databases of 3D images employed in PSM require much more attention. Several thoracic CT databases have been collected, mostly designed for lung cancer diagnosis [6,7]. A database of annotated CT studies has been developed and published by the Lung Image Database Consortium (LIDC). The first attempt contained 23 cases of lung nodules delineated by 6 radiologists in 3 blinded and unblinded reviews [8]. Such an approach enabled the authors to prepare a voxel probability map for each nodule. Furthermore, each expert

has assessed each nodule assigning numerical rates (concerning e.g. likelihood of malignancy, calcification, sphericity, etc.) A few years later a larger set has been completed under a LIDC/IDRI (Image Database Resource Initiative) affiliation [9]. CT of 1018 cases have been processed by 4 experts to produce compact annotations on over 7000 nodules. The XML standard has been employed to store the information. Authors provided a precise specification of an XML file. No software is attached to interpret and present the content. To confirm the former opinion on annotating 3D studies note, that the National Cancer Institute offers an access to a variety of imaging databases [10], yet only LIDC/IDRI studies are tagged as annotated. Due to limited scope, however, this set cannot be used in the PSM development.

Various databases have also been designed for abdominal region and contain mostly CT images. They often have a form of a test dataset in a detection/segmentation competition. The SLIVER07 database [11,12] may server as an example. It is used to evaluate and compare liver segmentation methods. A database prepared for comparative study on multiple abdominal and thoracic organs – 3D Image Reconstruction for Comparison of Algorithm Database (3D-IRCADb) – contains 20 CT studies with delineations of selected organs, vasculature and tumours [13]. The authors provided also the masks as surface meshes in VTK (Visual Toolkit) format. Included 3D contours provide robust starting point for modelling various organs or lesions. Diversity of cases has yet still been considered insufficient for the PSM generation task. In this study supplementary Atlas has therefore been created to provide additional information about large variety of cases not covered by the existing datasets. Methodology has been proposed to select studies, label them with regard to visible organs and pathologies. IT infrastructure has been developed to accommodate the data and permit gathering gold standard outlines.

The paper is organized as follows. Section 3 describes image data. The methodology of the image data and radiological reports analysis as well as the database anonymization is introduced in Sections 3. Section 4 presents the results. Section 5 concludes the paper.

2 Material

The image Atlas has been prepared using clinical data (radiological descriptions and corresponding DICOM image study) selected from Radiological Information System (RIS) and PACS (Picture Archiving and Communication System) of a medical institution with laboratories located in three different medical facilities. The examinations were registered between 2008 and 2013 year. Currently the dataset contains 490 cases: 395 CT exams and 95 MRI exams. Number of examinations in each year is presented in Table 1. CT exams were registered with a GE LightSpeed16 and MRI with a GE Sigma HDxt 1.5T.

Following selection criteria have been used:

- different cases should be included varying in age of patient, pathologies etc.
- cases with more than one examination available (e.g. follow-up) are preferred,

Table 1. Amount of examination in each year

	CT	MRI
2008	25	10
2009	64	1
2010	67	26
2011	53	18
2012	74	26
2013	112	14

- normal cases are included as a reference,
- CT as well as MR scans are included with different acquisition protocols (for CT scans, three-phases contrast enhanced images used in liver examination are preferred.)

Anatomically, all of the image series in database show at least the abdomen. Radiological description refers pelvis as well as chest images.

The source datasets (DICOM images with radiological descriptions) have been anonymized during the export procedure. Personal information has been replaced, but link between the base and follow-up examinations has been preserved:

- each patient has a new, semi-random, unique ID number assigned and used consistently in all source images and radiological descriptions data,
- each patient ID number has a new name and social ID assigned,
- year of birth, date of examination and sex of the patient have not been changed in order to permit the statistical analysis.

At the next step, the image data has been exported into local PACS database and processing has been started in order to form Atlas Database. Statistical information about cases that have already been processed and included in current version of Atlas as well as corresponding patients information and pathological cases may be found in Section 4.

3 Methodology

Clinical cases containing image data and radiological report have been analyzed. Then, a search service has been designed to permit a Web access.

3.1 Case Analysis

Clinical reports are included in the dataset in a free text form. They have been analyzed to provide standardized presentation.

In each case, the diagnosis has been considered as a primary data source. Based on the report a basic set of features (including anatomical structures

and pathologies) has been assigned to the study (including visible organs and pathologies). Additional features, if necessary, might be assigned after image examination.

Two physicians have read the radiological report and performed a two-step analysis. Each study has first been processed by a physicians with access to a DICOM viewer. Extracted information and delineated anatomical structures have been verified by a second physician using the same input data. Discrepancies have been subsequently discussed and eliminated. Any outlines produced during this process have also been evaluated. Each volume with quality of reference data has been stored as a DICOM series in PACS. In the second step, employing a spreadsheet, features have been assigned to each study.

Once the set of features have been selected, a standardized XML (eXtensible Markup Language), final form of database has been defined. It matches two main requirements. First provides a portable and less ambiguous way of propagation of the database, then provides a mapping between features and the case record allowing an automated export.

The database format is described by XSD (XML Schema Definition):

- XSD provides strict formal definition of each record and value; if designed properly, permits each future and each case to be described unambiguously,
- new features can be added into a formal description without breaking the backward compatibility,
- number of data records (cases) is not constrained by the format limitations,
- each version of database can be easily archived and made available for external research,
- XML is system independent and can be processed on all software platforms,
- a large number of tools and libraries is available for data processing.

The main disadvantage, noticeable when the number of records grows, is a bad scaling for operations performed on a whole database (including for example parametrized case search).

3.2 Database Format

XML database contains current information about all processed studies and series. Each record is stored in a separate file. Its format is defined by a single XSD (XML Schema Definition) template file. In the current workflow, the XML files are generated automatically based on spreadsheet file, but alternatively XSD-driven XML generator can also be used to describe new cases directly.

Each file contains combined information from two sources of data: (1) radiologist's description processed and stored temporarily in spreadsheet file, and (2) DICOM tags of image files:

- Set of unique keys identifying original DICOM study and description as well as selected information extracted from the source database: patient's personal data, anatomical regions, modality, study/series description etc.
- List of organs and their pathologies as defined by referring physicians.

```
<study xmlns="http://ib.polsl.pl"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://ib.polsl.pl Study.xsd"
date_added="2014-01-01T12:00:00"
date_last_modified="2014-01-01T12:00:00"
AccessionNumber="118/10/CT"
StudyInstanceUID="1.2.840.11..."
StudyID="21899"
PatientID="390312ABCDE" id="129"
patient_id="45722" exam_id="41876"
patient_first_name="Bfq"
patient_last_name="Abcevx"
patient_birth_date="1939-09-08"
PatientsSex="M" OperatorsName="3"
exam_perf_date="2010-01-11T00:00:00"
exam_name="TK - jamy brzusznej wielofazowe"
PerformingPhysiciansName="35604"
study_complete="1" series_count="4">
```

(a) `study` element

```
<series index="1"
isMask="0"
SeriesInstanceUID="1.2.840.11..."
SeriesDescription=""
InstanceCreationDate="2010-01-11"
InstanceCreationTime="11:50:34"
pixelSpacingX="0.765625"
pixelSpacingY="0.765625"
sliceThickness="2.5"
rows="512" columns="512"
slices="164"/>
```

(b) `series` element

Fig. 1. XML elements with attributes

DICOM information on study level is stored as a set of attributes of the root tag `study` (Fig. 1(a)).

The attribute `study_complete` distinguishes XML files describing a whole study (description is valid for all series with the same DICOM StudyInstanceUID ID and all these series are defined in subsequent `series` elements) or selected series only (e.g. different description is present for each series and more XML files are present with same StudyInstanceUID; all affected series are defined in subsequent `series` elements). The latter is more frequently used when gold standard outlines are available in image database. In both cases, `series` element describe single DICOM series (Fig. 1(b)).

The main part of the file contains a description of the anatomical structures of the abdominal cavity and their pathologies (Fig. 2).

```
<gallbladder>
    <normal>0</normal>
    <absence description="resection in 2008">1</absence>
    <lithiasis>0</lithiasis>
    <cancer>0</cancer>
    <other>0</other>
</gallbladder>
```

Fig. 2. `gallbladder` element with attributes

The list of pathologies valid for each organ is defined in the XSD template file. It includes most common diseases (particularly those that are essential for creating patient specific model), and some less common, but already encountered during processing the radiological reports. Pathology indicator may is set to true or false (boolean type) and optional description can be included (Fig. 3). The list of pathologies is open and new elements may be added without breaking backward compatibility (the XML file compatible with previous version of schema is also compatible with a new one).

```
<xs:complexType name="booleanWithDesription" >
    <xs:simpleContent>
        <xs:extension base="xs:boolean">
          <xs:attribute name="description" />
        </xs:extension>
    </xs:simpleContent>
</xs:complexType>
```

Fig. 3. XSD definition of a new type `booleanWithDesription` used for describing the pathologies

Current schema contains a total of 29 anatomical structures. Each one is described by two to ten elements. All can be used to define the searching criteria.

3.3 Search Engine

To simplify the XML database access, a search engine has been designed. Both available interfaces: web service (machine interface) as well as a web page (human interface) have similar capabilities and permit for a search with respect to specified criteria and return unique DICOM keys for image retrieval.

Basic requirements has been defined as:

- XML atlas files should be used as a primary data source.
- Searching attributes have to match the XML schema definition (XSD).
- Returned entries should point unambiguously a single DICOM study or series and permit the image retrieval.

The most straightforward solution with direct XML parsing has been rejected as inefficient for searching purposes. Solution with auxiliary SQL database has been chosen. Processing unit has been developed in PHP (PHP scripting language) to permit an easy implementation in web service and a web page.

SQL database is automatically recreated each time the XSD is updated. All available XML files are then imported[1]. Three main tables are defined (Fig. 4):

1. `Study table` – containing DICOM identification data (StudyInstanceUID) and defining inner StudyID identifier.
2. `Series table` – containing series available in each study.
3. `Attributes table` – containing matching attributes assigned to each SeriesID (if no distinction is made between series on XML level, all series are still imported separately with identical attributes).

Please note, that this structure is suboptimal in terms of data management, but permits a fast search:

- search criteria can be easily evaluated on series level; all series belonging to the same study can be identified,

[1] New cases can also be imported later without recreating the database.

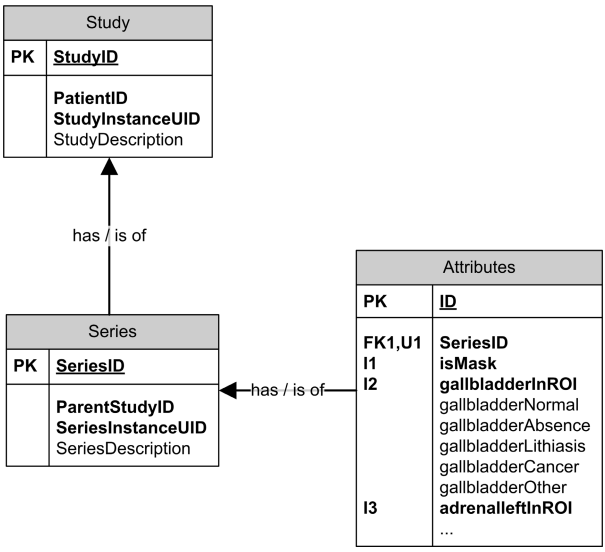


Fig. 4. Database schema

– available series are always separately tagged; gold standard masks are appropriately marked.
– both a series for given study and a study for a given series can easily be found.

Only limited subset of DICOM-related attributes is imported (and available as a result of search). Study details (e.g. PixelSpacing, InstanceCreationDate etc.) unsuitable for meaningful search criteria have been deliberately omitted. They can still be easily accessed directly, though, using separate PACS query and available StudyInstanceUID.

Columns in `Attributes` are defined using `boolean` type (no description is imported from XML). Valid elements match definitions given in reference XML schema definition.

SQL database is used by PHP search engine designed as a PHP class processing series queries. Each query is handled as an XML record matching XSD definition of a single case (Fig. 5). Valid XML record must be provided by an engine user or a query syntax error is raised. Helper static method permits JSON (JavaScript Object Notation) queries to be converted into XML counterparts.

Query is analyzed by the search engine in following manner:

1. Optional XML elements may be skipped in a query and do not influence the results.
2. Value of each provided, feature-related element is treated as a complex boolean AND-type filter with a meaning defined by one of two keywords

```
//initialize database connection and
//read XSD
$s = Search::getInstance();

// convert QUERY into mandatory form
if ($isJSON) {
  $query = Search::JSON2XML($query);
}

try {
  $results = $s.searchStudy($query);
  $returnCode = Search::SEARCH_OK;
  $returnMessage = Search::SEARCH_OK_TXT;
} catch (Exception $e) {
  $results = array();
  $returnCode = Search::SEARCH_NOK;
  $returnMessage = Search::SEARCH_NOK_TXT . " (" . $e->getMessage() .")";
}
```

Fig. 5. Search engine (simplified error handling)

included within `description` attribute (custom `booleanWithDesription` type):

- `description="enable"` - case can be returned only if it has this attribute defined and if the value of element matches specified in query (this is default behaviour if not description attribute present),
- `description="disable"` - is not be evaluated (can either be true or false)

In other words query containing `gallbladder` element (Fig. 6) while evaluated by the engine matches only studies with visible, anatomically *normal* gallbladder, *without* lithiasis and *without* cancer. Other description may or may not be present; "absence" element (here disabled) will be ignored.

```
<gallbladder>
    <normal description="enable">true</normal>
    <absence description="disable">0</absence>
    <lithiasis>0</lithiasis>
    <cancer description="enable">false</cancer>
    <other description="disable">1</other>
</gallbladder>
```

Fig. 6. Search-query element

In response, XML record is returned containing DICOM `StudyInstanceUIDs` and `SeriesInstanceUID` (Fig. 7). `Series_count` attribute indicates a number of matching series found in database for each returned case. Helper method is also provided to convert XML response to JSON counterpart.

```
...
<study series_count="2" StudyInstanceUID="111.111.111.1">
<series SeriesInstanceUID="11.11.11.11.1" isMask="false" />
<series SeriesInstanceUID="22.22.22.22.2" isMask="true" />
</study>
...
```

Fig. 7. Search response (XML format)

Web Service Interface. Search engine machine interface has been defined as a simple REST XML service. Each HTTP POST request sent to `http://server_IP/query` and encoded an XML or JSON (JavaScript Object Notation) is validated and fed to the search engine. Selection bases on HTTP Content-Type header. Errors in format or validation detected at this step are communicated to the client using the empty HTTP 400 response. Similarly, invalid method is indicated by the 405 response code.

Properly validated query in the XML form is processed by the search engine and – if the format indicated by a client in Accept-header permits the XML – produced response is directly returned to client with HTTP 200 response code. If JSON response is expected instead translation is performed. The returned identifiers of studies allow unambiguous retrieval of selected studies/series from PACS server using *Study Root Query Retrieve Information Model* and contain information whether series is or isn't the gold standard outline.

WWW Page Interface. WWW search provide a user-friendly access to the search facility. Upon request a form is generated including all possible search options converted to input fields and grouped. Each possible organ can be marked and included in the query. Each criteria available for the included organ can be selected or deselected.

Responses are directly converted into HTML tables and links are included to the proxy service to permit image retrieval in form of a ZIP archive.

Web page provides also option for query export and import. Current state of the query can be saved as an XML file and loaded later serving as a query template.

4 Results

At all levels, including spreadsheet, XML and search engine, the database contains coherent information about the number of medical cases. These cases can be used in the development of a Patient Specific Model (gold standard outlines may provide starting points or contours, parameters of tissue may be estimated, etc.) or in benchmarking. New data is continuously processed. In this section, a current contents of a database is shown.

Current version of database contains 48 CT exams from 39 patients: 17 women and 22 men. Years of birth are between 1927 and 2006 (women: 1933 to 2006, men: 1927 to 2001). Also 6 MRI exams are available. All have matching image data in the local PACS system. Classification of cases with respect to selected

organs and pathologies unambiguously[2] described is shown in Table 2. If one or more pathologies are detected in a single organ, it is always included as a single pathological case.

Table 2. Number of described cases with respect to organs' pathology

organ	normal	pathological	all
lungs	19	5	24
liver	20	19	39
gallbladder	27	7	34
spleen	33	6	39
pancreas	36	1	37
stomach and bowels	14	1	15
great vessels	18	3	21
left kidney	20	16	36
right kidney	11	25	36
urinary bladder	6	4	10
spine	5	19	24
other	178	137	315

Each of 29 organs is described with respect to common pathologies (or tagged as normal or nonexistent). Full summary with respect to right kidney is shown in Table 3.

Table 3. Right kidney tags and number of cases

Tag	Case number
normal	11
displaced	1
absence	2
enlarged	2
hypotrophic	2
tumor	5
cystis	9
cancer	1
inflammation	6
lithiasis	1
retention	3
hydronephrosis	3
other	2

The most common are spine pathologies (19 cases) and liver cystis (13 cases). The majority of the pathologies occur in patients of age between 65 and 75 years (Fig. 8).

[2] Only cases explicitly described in source radiological report.

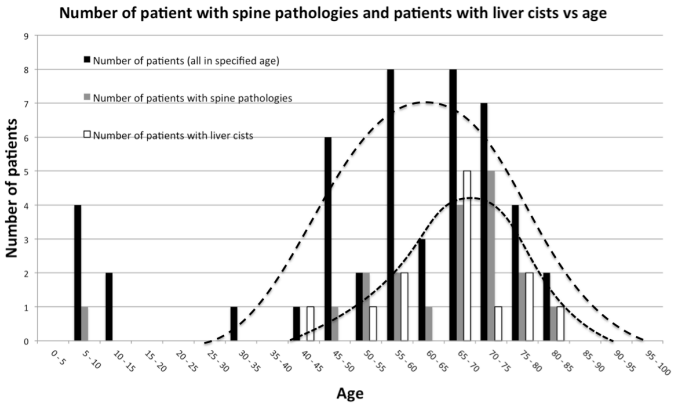


Fig. 8. Number of patients with spine pathologies and patients with liver cystis vs age

In most categories men have more pathological organs described than women. The sole exception is stomach. Table 4 shows percent of pathological cases for selected organs.

Table 4. Percent of organs with pathologies

sex	liver	spleen	kidneys	pancreas	stomach
female	53%	18%	71%	18%	71%
male	68%	43%	79%	32%	68%

Average patients age is 56 years. Most patients, however, are between 40 and 85 year of life. Only 6 patients were below 15 at the moment of examination and no cases came from patients between 15 and 30 year. Majority of pathologies had been diagnosed in patients between 45 and 75 years: 72%. 9% affects patients younger than 10 years old. That unusual distribution is a result of only two patients with a very young age (4 studies) with total 21 pathologies described. Full summary is shown in Table 5.

Table 5. Distribution of pathologies

Interval (years)	Percentage of pathologies	Interval (years)	Percentage of pathologies
0 – 5	0%	45 – 50	12%
5 – 10	9%	50 – 55	6%
10 – 15	1%	55 – 60	11%
15 – 20	0%	60 – 65	7%
20 – 25	0%	65 – 70	18%
25 – 30	0%	70 – 75	18%
30 – 35	1%	75 – 80	9%
35 – 40	0%	80 – 85	5%
40 – 45	3%	85 – 90	0%

5 Conclusion

In the paper a work in progress has been presented on Radiological Atlas for Patient Specific Model (PSM) generation. First, a Patient Specific model has been introduced. Connections between the model development and atlas are explained. Next, a workflow has been presented. Final XML format and case search engine are outlined. Database has been evaluated and number of cases of each pathology is given.

The processing work is being continued. New cases are added to the database and gold standard outlines are created. Once the database is completed, the developed infrastructure will serve as a source of validated data for generation of PSM (selected outlines will be retrieved and used to construct customized 3D model) and benchmarking (only image studies or series matching desired criteria will be used for evaluation of the model).

Acknowledgement. This research is supported by the NCN Grant no. UMO-2012/05/B/ST7/02136.

References

1. Juszczyk, J., Piętka, E.: Automatic generation of initial points for CT abdominal organ segmentation. In: Proc. of 27th Int. Congress and Exhibition of Computer Assisted Radiology and Surgery (CARS), Heidelberg, Germany, vol. 8(suppl. 1), pp. 303–304 (June 2013)
2. Nishikawa, R.M.: Design of a common database for research in mammogram image analysis. In: Proc. SPIE, vol. 1905, pp. 548–549 (1993)
3. Kallergi, M., Clark, R.A., Clarke, L.P.: Medical image databases for CAD applications in digital mammography: Design issues. Stud. Health Technol. Inform. 43, 601–605 (1997)
4. Heath, M., Bowyer, K., Kopans, D., Moore, R., Kegelmeyer, W.P.: The digital database for screening mammography. In: Yaffe, M.J. (ed.) Digital Mammography 2000: Proceedings of the Fifth International Workshop on Digital Mammography, pp. 431–434. Medical Physics, Madison (2000)
5. Shiraishi, J., et al.: Development of a digital image database for chest radiographs with and without a lung nodule: Receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. AJR, Am. J. Roentgenol. 174, 71–74 (2000)
6. Clark, K.W., et al.: Creation of a CT image library for the Lung Screening Study of the National Lung Screening Trial. J. Digit. Imaging 20, 23–31 (2007)
7. Reeves, A.P., et al.: A public image database to support research in computer aided diagnosis. In: 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 3715–3718 (2009)
8. Armato III, S.G., et al.: Lung Image Database Consortium: Developing a resource for the medical imaging research community. Radiology 232, 739–748 (2004)

9. Armato III, S.G., et al.: The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans. Med. Phys. 38(2), 915–931 (2011)
10. The Cancer Imaging Archive, http://www.cancerimagingarchive.net/ (access: January 2014)
11. Heimann, T., et al.: Comparison and Evaluation of Methods for Liver Segmentation From CT Datasets. IEEE Transactions on Medical Imaging 28(8), 1251–1265 (2009)
12. http://www.sliver07.org/ (access: January 2014)
13. 3D-IRCADb database, http://www.ircad.fr/softwares/3Dircadb/3Dircadb.php (access: January 2014)

Part II

Telemedicine

Teledermatology & Teledermoscopy: Optimizing Technology & Information by Optimizing Users

Elizabeth A. Krupinski

Department of Medical Imaging, University of Arizona, Tucson, AZ USA
Krupinski@radiology.arizona.edu

Abstract. The main method of skin cancer detection is direct examination of the skin by a trained professional. There is a shortage of dermatologists, a situation that could influence timely diagnosis and management of skin cancer. Teledermatology and teledermoscopy could overcome some of these barriers. To develop training programs we need to understand how dermatologists view digital skin images. Two dermatologists and 4 residents viewed 20 photographs and dermatoscopic images of benign and malignant lesions while eye position was recorded. They then completed an online training course and returned to repeat the cases. There were some differences in search and diagnoses although more as a function of experience than the online course. Further work is needed but we can characterize visual search in teledermatology settings.

Keywords: teledermatology, dermoscopy, training, diagnostic accuracy, eye-tracking.

1 Introduction

Skin cancer varies, but an atypical pigmented skin lesion can be melanoma, the most serious and potentially deadly form. Malignant melanoma (MM) ranks as the sixth most common cancer in the United States and the most common fatal malignancy among young adults. [1] The American Cancer Society estimated 76,250 new cases of MM and more than 9,180 deaths in 2012. [1] Between 1950 and 2000, the National Cancer Institute (NCI) Surveillance, Epidemiology and End Results (SEER) database documented a 619% increase in annual incidence of MM and a 165% increase in annual mortality. [2] Early detection is essential to patient survival and better prognosis [3,4]. The 5-year survival associated with MMs diagnosed and excised when less than 1mm thick is 95%, whereas for ulcerated MM greater than 4mm and distant metastatic disease the survival rate is 45% and 12%. More than 60,000 new MM cases per year in 2005 resulted in $60B projected-year-of-life lost in the US, a significant burden to society. [5] Skin cancer detection is mostly by direct examination of the skin by a trained professional. [6] Dermatologists are the most skilled, [7] but there is a shortage of dermatologists around the world [8]. Patients with urgent problems, such as pigmented changing lesions, report wait times as long as patients with routine problems. [9] Results of a recent survey revealed that the mean wait time

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_8, © Springer International Publishing Switzerland 2014

was 33 days. Also evident was an increasing trend for dermatology time spent in cosmetic and surgical dermatology over medical dermatology. [10] This has implications for timely diagnosis, particularly in melanoma high-risk patients with multiple moles, atypical moles or a personal or family history of melanoma. [11] Teledermatology and teledermoscopy could overcome some of barriers such as geographic distance, availability of dermatologists, poor communication with specialists, and coordination of care.

Teledermatology uses store and forward, mobile, interactive, digital photography, and video conferencing. It ranges from triage of neoplastic and non-neoplastic skin lesions to skin cancer diagnosis. A recent summary indicated that in 2012 there were 37 teledermatology active programs in the US. [12] Store-and-forward was the most common delivery modality (81%). Most were based at academic institutions (49%), Veterans Administration hospitals (27%), private practice (16%), and health maintenance organizations (8%). In 2011, the median number of consultations per program was 309. For reimbursement, most were private payers, followed by self-pay, Medicaid, Medicare, and HMOs.

A recent literature review on the accuracy of teledermatology indicated that diagnostic concordance of store and forward with clinic dermatology was good; and concordance rates for live interactive and clinic dermatology were higher, but based on fewer patients. Overall rates of management accuracy were equivalent, but teledermatology and teledermatoscopy were inferior to clinic dermatology for malignant lesions. Patient satisfaction and preferences were comparable, and teledermatology significantly reduced time to treatment and clinic visits and was generally cost-effective. [13,14,15,16] The American Telemedicine Association's developed practice guidelines for teledermatology. [17]

Detection is enhanced by dermoscopy. [18,19] A dermatoscope provides a 10x magnification and illuminates a skin lesion without reflected light, allowing morphologic classification based on accepted dermoscopic features. It enables more efficient monitoring, particularly of pigmented lesions, over time. [20] Use of dermoscopy by trained PCPs has been shown to improve detection of skin cancers and reduce the excision or referral of benign pigmented lesions by one-half. [21,22]

Eye-position recording is used to explore perceptual processes involved in medical image perception. It assumes that observers direct high-resolution foveal vision to areas in images so that attention and information-processing resources extract and process data to render decisions. The initial glance provides a global impression that includes the processing and recognition of content such as symmetry, anatomy and color that are processed with information in long-term memory. Key indications of experts are consistent, accurate and efficient diagnostic performance, which require not only dedicated training and experience but some degree of talent, aptitude and motivation. [23] A question is what are the best training methods and what types of experiences do trainees (e.g., residents) require to develop their diagnostic skills?

Experts search, process, and interpret larger perceptual units than those with less skill because they are better able to recognize these units more efficiently

and effectively as configurations or chunks of information rather than individual pieces. The mechanism behind this phenomenon is thought to be due to perceptual and/or cognitive tuning to the visual task as the observer encounters more exemplars over time with suitable training and feedback. [23] In 1963, Llewellyn-Thomas and Lansdown [24] conducted the first reported eye-position study in medical imaging. It demonstrated that search patterns are somewhat unique to the individual and tend not to be uniform in image coverage. Since then, a number studies have examined such issues as why errors occur (false negatives and false positives) [25,26,27,28,29], how experts differ from novices [29,30,31], and how different display parameters affect diagnostic accuracy and visual search efficiency. [32,33] We have been studying this issues with eye-tracking for over 25 years, mostly in radiology [26,28,30,31,32,33] but more recently in pathology with the advent of whole slide imaging. [34,35,36]

2 Materials and Methods

20 cases were compiled in collaboration. Each case had a photo of a single pigmented skin lesion (PSL) and its dermoscopy image. Each case was rated on 1-10 scale by 2 expert dermatologists, where 10 = highly suspicious & 1 = benign (Figure 1). For the analyses, 1-5 ratings were considered to be benign (n = 10) and the 6-10 ratings were malignant (n = 10). The cases spanned the range of ratings.

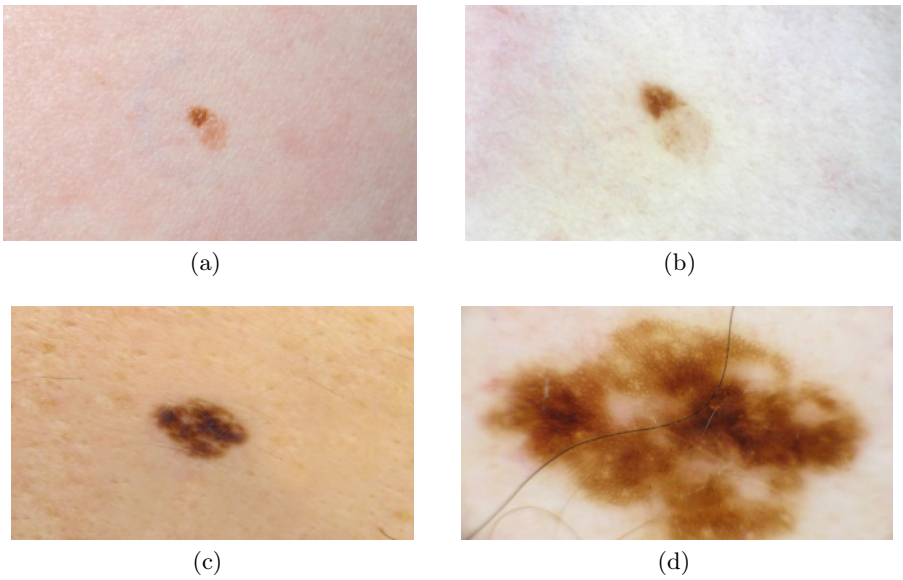


Fig. 1. Examples of photo (left) and dermatoscopic (right) images of benign (top) and malignant (bottom) lesions

Images were cropped so lesions were in the center and there was some surrounding skin visible. They were saved as jpeg and a PowerPoint presentation was created (all the same width & height, 672 x 448 pixels) with images centered in the slide with a black background. The photograph was displayed first for each case, followed by the dermoscopy image.

The images were displayed on a ViewSonic VA2703 LCD color monitor (1920 x 1080; 27" screen; contrast ratio 1200:1; max luminance 300 cd/m^2; sRGB gamma 2.2). Room lights were 40 lux.

The study was IRB approved. Two dermatologists with dermoscopy experience and two 1st year dermatology residents were recruited and paid. The mean age of dermatologists (both male) was 57 and the mean age of the residents (1 male, 1 female) was 28. All had knowledge of both photographic and dermatoscopic image assessment. All were given Nishihara's test for color blindness prior to the start of the study and all passed. All subjects wore corrective lenses.

The study had 5 parts. Part I was a baseline test. Subjects viewed the 20 cases and their performance and visual search parameters were measured. He/she reported whether the case was benign or malignant, and gave their confidence as definite, probable or possible. Then the dermoscopy image was presented and they could keep the previous decision and confidence or change them. Viewing time was unlimited.

Within 1 week, they completed an on-line dermoscopy program (Part II) where they reviewed 35 PSL dermoscopy cases for asymmetry, atypical network, blue & white veil (3 point check list). No feedback was provided. The scores were stored automatically. In Part III, they completed an on-line tutorial "Introduction to Dermoscopy" that reviews the 3 assessment criteria. There is an interaction session where they rated images on the 3 criteria and got feedback. Scores were automatically recorded. After 1 week they repeated the cases (Part IV) without feedback and scores were recorded. In Part V after 3-6 weeks, they repeated Part I.

The ASL SU4000 Eye-Tracker system (Applied Science Labs, Bedford, MA) recorded visual search. Prior to viewing, each subject is calibrated. The system computes line of gaze and dwell time based on pupil and corneal reflection parameters. An infrared (IR) light-emitting diode and phototransistor detector are mounted on the headband. IR light is emitted and reflects off a reflective visor into the left eye, reflecting back off the pupil and cornea to the visor, which then reflects it back to a charge-coupled device camera. The eye-position data were analyzed using standard methods. Briefly, the accuracy of the system (spatial error between true eye position and computed measurements) is less than 1 degree. The SU4000 samples eye positions every 1/60 of a second to generate raw x-, y- coordinate eye-position data. Fixations are formed by grouping x- and y-coordinates of the raw data using a running mean distance calculation having a 0.58 radius threshold. Dwell time can be calculated for each fixation, summed across fixations, then associated with a given region of interest or location in the stimulus image.

3 Results

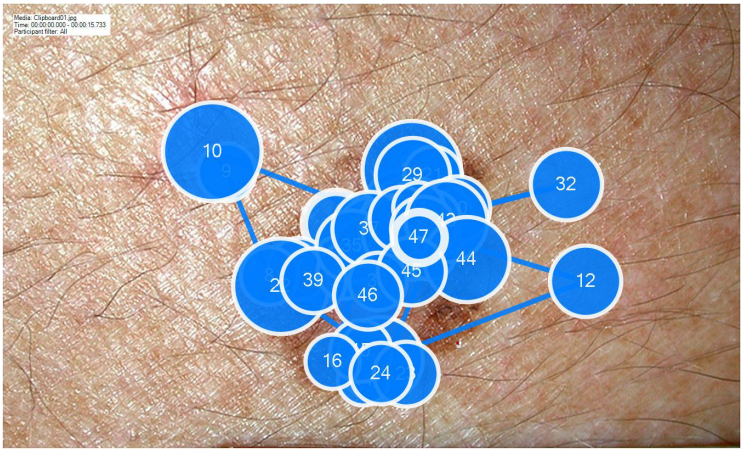
There were no differences between pre-training, interactive training, and post-training test scores for asymmetry (X^2=1.95) or atypical network (X^2=0.30), but there was for blue & white veil (X^2=4.35, p<0.05). There were no differences between pre-training, interactive training, and post-training test scores for asymmetry (X^2=1.95) or atypical network (X^2=0.30), but there was for blue & white veil (X^2=4.35, p<0.05).

Total photo time did not differ by session (F = 0.935, p = 0.335), or decision/confidence change (F = 1.452, p = 0.230), but did for training (F = 6.642, p = 0.0188) as dermatologists took less time (mean = 12.56 sec, sd = 5.68 vs 15.61, sd = 7.55). There were no differences whether decisions were correct or not. Total dermoscopy time did not differ by session (F = 0.141, p = 0.708), but did for decision/confidence (F = 3.338, p = 0.0212) with benign to malignant longer (mean = 15.39 sec, sd = 6.77) than other decisions (mean = 11.27, sd = 5.78 no change, mean = 11.92, sd = 6.79 confidence up, mean = 11.75, sd = 6.32 confidence down), and training (F = 4.819, p = 0.0298) with dermatologists taking less time (mean = 10.23, sd = 5.93 sec vs 14.21, sd = 6.25). There were no differences whether decisions correct or not.

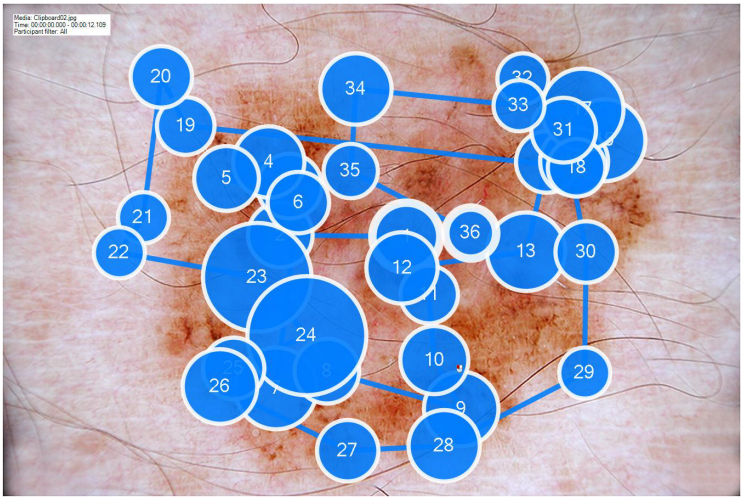
With fixations on photos, there were no differences for decision/confidence change or decision correctness. For mean dwells training was significant (F = 3.953, p = 0.0487) with dermatologists having shorter (mean=168.80 msec, sd=66.37 vs mean = 198.09, sd = 75.87) dwells. For dermoscopy number of fixations as function of change decision/confidence, there were significant differences with training (F = 7.771, p = 0.006) with dermatologists generating fewer fixations (mean = 17.73, sd = 10.68 vs mean = 24.94, sd = 10.67) than residents; and for decision change (F = 4.490, p = 0.0048) with benign to malignant/vice-versa being longer (mean = 27.24, sd = 9.85) than no change (mean = 20.01, sd = 10.48), confidence up (mean = 20.21, sd = 11.97) and down (mean = 19.63, sd = 11.05). For decision correctness training (F = 6.735, p = 0.0104) was significant with dermatologists having fewer fixations than residents.

With mean dermoscopy dwells for change in decision, there was significant difference for training (F = 5.215, p =0.0239) with dermatologists having shorter dwells (mean = 129.38, sd = 67.96 msec vs mean = 178.18, sd = 69.30); and change in decision (F = 3.549, p = 0.0161) with benign to malignant (and vice-versa) being longer (mean = 188.69, sd = 61.91) than no change (mean=144.78, sd = 70.91), confidence up (mean = 146.32, sd = 79.07) & down (mean=151.50, sd = 70.69). For decision correctness training (F = 6.635 p = 0.0110) was significant with dermatologists having shorter mean dwells.

Figure 2 shows typical patterns generated by one of the subjects in the study on one of the photo and dermatoscopic cases. Figure 3 shows how certain areas or *hot spots* receive more attention than others.



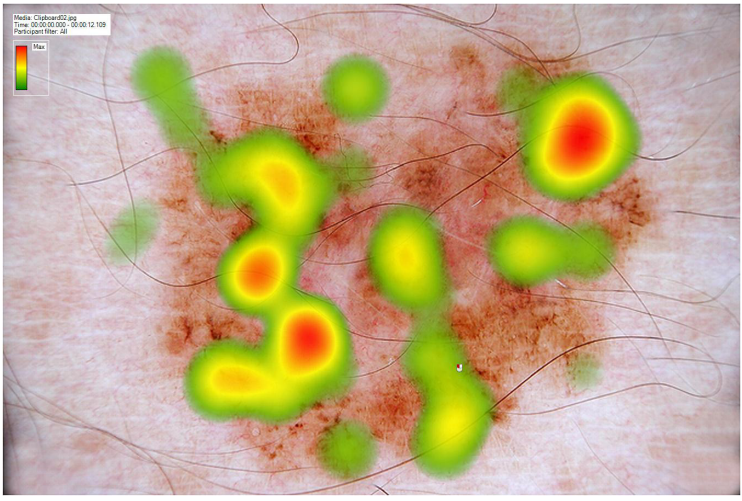
(a)



(b)

Fig. 2. Typical pattern of a subject on one of the photo (a) and dermatoscopic (b) cases

(a)



(b)

Fig. 3. *Hot spots* receive more attention than others

4 Discussion

The data suggest that the subjects benefited somewhat from training because their interactive scores increased from pre-test scores, especially for blue & white veil. However, even though some training effects were maintained, overall post-test performance decreased although not back to pre-test levels. A single training session may be useful, but it may take much more time to educate residents and maybe even dermatologists on the correct interpretation of dermoscopy. Future studies can determine how many cases it might take for someone to retain the lessons from training and then generalize and maintain them during daily interpretation of image-based cases. It may not however, be simply a matter of more cases since expertise in interpreting medical images is more complex. [36,37] A single session also does not seem to impact accuracy or visual search parameters. A consistent finding was that the dermatologists overall had more efficient search than residents generating fewer fixations with lower dwells. This has been observed in radiology and pathology and is considered one of the hallmarks of expertise. [23,24,25,26,27,28,29,30,31].

The fixations and dwells associated with decisions changing from benign to malignant or vice-versa from photo to dermatoscopic viewing were longer than any other decision, indicating increased visual processing. It seems that when someone is certain of a diagnosis with the photo and the dermoscopic image confirms that diagnosis it does not take much time to search the dermatoscopic image. If there is a shift in confidence but not benign vs malignant, the dermoscopic image serves as a confirmation and does not require extensive search. If it does give new information that contradicts the decision made with the photo, it does seem to generate in-depth scanning and processing to decide that there truly is a different diagnosis. This needs more investigation as this study only had 20 images and the majority resulted in no decision changes. Larger data sets with more observers at different levels of expertise may yield some very useful data and improve our understanding of what factors impact decisions. We also did not ask the subjects what features in the dermoscopic images made them reconsider their decisions (i.e., asymmetry, atypical network, blue & white veil something else), and that would also be useful to know.

Differences in search may have implications for developing tools to teach dermatologists and residents about how to better interpret dermoscopy features in clinical practice. Future studies can correlate dwell with visual parameters if possible then tie these to decisions. This study revealed some interesting things about the impact of dermoscopic training on performance accuracy. It also revealed, perhaps for the first time how photographic and dermoscopic images are searched during the interpretation process. Understanding these visual search strategies will help us better understand the diagnostic interpretation process and could in the future help us design better training tools that capitalize on the ways that residents and dermatologists process the information contained in photographic and dermoscopic images.

References

1. Cancer Facts & Figures 2012. American Cancer Society, http://www.cancer.org/acs/groups/content/epidemiologysurveilance/documents/document/acspc-031941.pdf (last accessed January 17, 2014)
2. Surveillance Epidemiology and End Results (SEER), http://seer.cancer.gov/statfacts/html/melan.html (last accessed January 17, 2014)
3. Denoix, P.F.: Enquete permanent dans les centres anticancereaux. Bull. Inst. Nat. Hyg. 1, 70–75 (1946)
4. National Cancer Institute Fact Sheet, http://www.cancer.gov/cancertopics/factsheet/detection/staging (last accessed January 17, 2014)
5. Tsao, H., Atkins, M.B., Sober, A.J.: Management of cutaneous melanoma. N. Eng. J. Med. 351, 998–1012 (2004)
6. Harris, J.M., Salasche, S.J., Salasche, S.J., Harris, R.B.: Can internet-based continuing medical education improve physicians' skin cancer knowledge and skills? J. Gen. Intern. Med. 16, 50–56 (2001)
7. Chen, S.C., Pennie, M.L., Kolm, P., et al.: Diagnosing and managing cutaneous pigmented lesions: primary care physicians versus dermatologists. J. Gen. Intern. Med. 2, 678–682 (2006)
8. Kimball, A.B., Resneck, J.S.: The US dermatology workforce: a specialty remains in shortage. J. Am. Acad. Dermatol. 59, 741–745 (2008)
9. Tsang, M.W., Resneck, J.S.: Even patients with changing moles face long dermatology appointment wait- times: a study of simulated patient calls to dermatologists. J. Am. Acad. Dermatol. 55, 54–58 (2006)
10. Kimball, A.B., Resneck, J.S.: The US dermatology workforce: a specialty remains in shortage. J. Am. Acad. Dermatol. 59, 741–745 (2008)
11. Banky, J.P., Kelly, J.W., et al.: Incidence of new and changed nevi and melanomas detected using baseline images and dermoscopy in patients at high risk for melanoma. Arch. Dermatol. 141, 998–1006 (2005)
12. Armstrong, A.W., Wu, J., Kovarik, C.L., Goldyne, M.E., Oh, D.H., McKoy, K.C., Shippy, A.M., Pak, H.S.: State of teledermatology programs in the United States. J. Am. Acad. Derm. 67, 939–944 (2012)
13. Warshaw, E.M., Hillman, Y.J., Greer, N.L., Hagel, E.M., MacDonald, R., Rutks, I.R., Wilt, T.J.: Teldermatology for diagnosis and management of skin conditions: a systematic review. J. Am. Acad. Derm. 64, 759–772 (2011)
14. Krupinski, E.A., LeSueur, B., Ellsworth, L., Levine, N., et al.: Diagnostic accuracy and image quality using a digital camera for teledermatology. Telemed. J. 5, 257–263 (1999)
15. Krupinski, E.A., Webster, P., Dolliver, M., Weinstein, R.S., Lopez, A.M.: Efficiency analysis of a multi-specialty telemedicine service. Telemed. J. 5, 265–271 (1999)
16. Krupinski, E., Barker, G., et al.: Telemedicine versus in-person dermatology referrals: an analysis of case complexity. Telemed. J. e-Health 8, 143–147 (2002)
17. Krupinski, E., Burdick, A., et al.: American Telemedicine Association's Practice Guidelines for Teledermatology. Telemed. e-Health J. 14, 289–302 (2008)
18. Argenziano, G., Soyer, H.P.: Dermoscopy of pigmented skin lesions–a valuable tool for early diagnosis of melanoma. Lancet. Oncol. 2, 443–449 (2001)
19. Massone, C., Di Stefani, A., Soyer, H.P.: Dermoscopy for skin cancer detection. Curr. Opin. Oncol. 17, 147–153 (2005)
20. Soyer, H.P.: Dermoscopy and recently developed imaging techniques. Introduction. Semin. Cutan. Med. Surg. 28, 141 (2009)

21. Argenziano, G., Puig, S., Zalaudek, I., et al.: Dermoscopy improves accuracy of primary care physicians to triage lesions suggestive of skin cancer. J. Clin. Oncol. 24, 1877–1882 (2006)
22. Vestergaard, M.E., Macaskill, P., et al.: Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: a meta-analysis of studies performed in a clinical setting. Br. J. Derm. 159, 669–676 (2008)
23. Nodine, C.F., Mello-Thoms, C.: The role of expertise in radiologic image interpretation. In: Samei, E., Krupinski, E. (eds.) The Handbook of Medical Image Perception and Techniques, pp. 139–156. Cambridge University Press, New York (2010)
24. Llewellyn-Thomas, E., Lansdown, E.L.: Visual search patterns of radiologists in training. Radiol. 81, 288–291 (1963)
25. Kundel, H.L., Nodine, C.F., Carmody, D.P.: Visual scanning, pattern recognition and decision-making in pulmonary tumor detection. Invest. Radiol. 13, 175–181 (1978)
26. Kundel, H.L., Nodine, C.F., Krupinski, E.A.: Searching for lung nodules: visual dwell indicates locations of false-positive and false-negative decisions. Invest. Radiol. 24, 472–478 (1989)
27. Nodine, C.F., Mello-Thoms, C., Kundel, H.L., et al.: Time course of perception and decision making during mammographic interpretation. Am. J. Roentgen. 179, 917–923 (2002)
28. Krupinski, E.A.: Visual scanning patterns of radiologists searching mammograms. Acad. Radiol. 3, 137–144 (1996)
29. Nodine, C.F., Kundel, H.L., Lauver, S.C., Toto, L.C.: Nature of expertise in searching mammograms for breast masses. Acad. Radiol. 3, 1000–1006 (1996)
30. Krupinski, E.A.: Visual search of mammographic images: influence of lesion subtlety. Acad. Radiol. 12, 965–969 (2005)
31. Lesgold, A.M., Rubinson, H., et al.: Expertise in a complex skill: diagnosing x-ray pictures. In: Chi, M.T.H., Glaser, R., Farr, M.J. (eds.) The Nature of Expertise, pp. 311–342. Erlbaum Publishers, Hillsdale (1988)
32. Krupinski, E.A., Roehrig, H.: The influence of a perceptually linearized display on observer performance and visual search. Acad. Radiol. 7, 8–13 (2000)
33. Krupinski, E.A., Roehrig, H.: Pulmonary nodule detection and visual search: P45 and P104 monochrome versus color monitor displays. Acad. Radiol. 9, 638–645 (2002)
34. Krupinski, E.A., Tillack, A.A., et al.: Eye-movement study and human performance using telepathology virtual slides. Implications for medical education and differences with experience. Hum. Path. 37, 1543–1556 (2006)
35. Krupinski, E.A.: Virtual slide telepathology workstation of the future: lessons learned from teleradiology. Hum. Path. 40, 1100–1111 (2009)
36. Krupinski, E.A., Graham, A.R., Weinstein, R.S.: Characterizing the Development of Visual Search Expertise in Pathology Residents Viewing Whole Slide Images. Hum. Path. 44, 357–364 (2012)
37. James, N., Dusza, S.W., Quigley, E.A., et al.: Influence of time on dermoscopic diagnosois and management. J. Derm. 54, 96–104 (2013)

e-Melanoma Diagnosing and Learning System. Current Status

Zdzislaw S. Hippe

University of Information Technology and Management, 35-225 Rzeszow, Poland
zhippe@wsiz.rzeszow.pl

Abstract. A new web tool for distant diagnosing and learning on melanocytic skin lesions (**e-IMDLS**) was developed and tested. Currently, two separated versions of the system (Polish and English) are implemented and accessible by Internet. Here, the main features of the English version of **e-IMDLS** system are briefly discussed. Synthesis of digital images of selected skin lesions by the system is discussed; some possibilities and limitations of this algorithmic process are commented. Finally, the way of using the system by experienced dermatologists and/or family physicians is also described.

Keywords: melanoma, skin lesions, computer synthesis of digital images.

1 Introduction

Experience of many research teams, over many years, indicates that an access to the data bases containing color pictures of melanocytic skin lesions is extremely important for education of future dermatologists and family physicians. An access to such large and reliable data bases may support students of medical schools and physicians with developing an *algorithmic approach* to recognition and classification of skin lesions and may support abandoning a *scientific guess* approach based on quick, routine and ad hoc diagnosis. However, for time being, practically there is no access to reliable data bases with color images of melanocytic skin lesions. Moreover, some centers of medical informatics that were publically accessible ceased their service while other centers offer the same standard collection of images.

In our research along these lines we develop a computer tool called **Instant-Nevi-Painter.PL** (**INP** for short), which can be treated as a core of the e-**I**nternet **M**elanoma **D**iagnosing and **L**earning **S**ystem (**e-IMDLS**). Currently, two separated versions of **e-IMDLS** (Polish and English) are implemented and accessible via Internet. The **INP** tool - as it was stated previously - being a kernel of **e-IMDLS**, is based on our own algorithms for computer synthesis of the digital images of some selected skin lesions [3]. Now we like to describe the last step of our investigations, namely the network implementation of the **INP**. It is publicly available at:

http://synthesis.melanoma.pl/

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_9, © Springer International Publishing Switzerland 2014

A complementary objective of this work was to develop a multi-category data base, containing synthesized images of selected types of melanocytic skin lesions. Additionally, going beyond the scope of our preliminary objectives that were settled with the National Science Center in Cracow, Poland (`www.ncn.gov.pl`), our data base contains also some real digital images of lesions. The entire computer program system and some of its internal links will be briefly explained.

2 Background Theory

Classification of melanocytic skin lesions and an associated evaluation of the degree of melanoma risk should be conducted using some systematic procedures known as melanoma algorithms (or strategies). Using this kind of strategy enables an elimination of ad hoc physician diagnosis and introduces an algorithmic classification of analyzed cases. In our previous research, the original **ABCD** rule [4] has been accepted, where the lesion type is recognized on the basis of values of four symptoms: asymmetry **A**, border type **B**. the number of colors **C**, and the number and type of diversified structures of the lesion surface **D**. An application of the improved **ABCD** rule [1,2] was implemented in our previous diagnosing system, a result of the joint research of the **U**niversity of Information **T**echnology and **M**anagement (**UITM**) in Rzeszow, Poland, and the University of Kansas, Lawrence, KS, USA. However, current research conducted by both groups resulted in an observation that a better recognition of the lesion type (and, as a result, better image synthesis) may be accomplished by a modification, known as **ABK**, of the **ABCD** rule, where **K** denotes a linear combination of colors and diversified structures. The **K** component was defined as a result of extensive statistical research on the analysis of symptom coexistence in the actual digital images of melanocytic skin lesions. Up to the date **53** possible values of **K** are recognized (see Fig. 1).

2.1 Investigated Types of Melanocytic Lesions

Algorithms of image synthesis, currently implemented in **INP**, cover two groups of the most dangerous melanocytic skin lesions: ***Nevus*** and ***Melanoma***. All kinds of lesions, from both groups, were included for such synthesis. Thus, in the group ***Nevus***, synthesized are Junctional nevus, Junctional and dermal nevus, Atypical/dysplactic nevus, Dermal nevus and Palmo-planar nevi, while in the group ***Melanoma*** synthesized are images of the type Superficial melanoma and Nodular melanoma.

A data base, called **Lesions**, currently contains images of synthetic and actual melanocytic skin lesions. The synthetic lesion images were created by the **INP** tool. Such images were saved to this data base after careful quality control by authorized diagnosticians. On the other hand, actual images were acquired using standard procedure of diagnosis of melanoma. Both types of images were ordered according to risk of melanoma:

benign lesion -> blue lesion -> suspicious lesion -> melanoma.

3 Selected Features of the System

Using the **INP** tool, a working set of synthetic images may be created, however, with no saving possibility. Such a set may be useful in (**1**) the process of diagnosis, as a result of a comparison of the patient lesion with the synthetic lesion and in (**2**) the process of computer-supported learning of some courses in medical education (dermatology) and paramedical education (cosmetology, health care and public health). Selected features of the developed system are further explained by referring to the most important internal links.

Link ***Tutorial***. Gives the owner's manual for the **INP** tool.

Link ***Image Synthesis***. The standard purpose of using the **INP** tool is creation of a working set of images of the melanocytic skin lesions. This set is created by a template, where values of three parameters (**A**symmetry, **B**order, and **K** - combination of colors and diversified structures) are requested. The resulted image is defined by these parameter values. A single use of the **INP** tool creates **eight** lesion images (Fig. 2). Some of these images are identical with each other except that they are tilted by 90 degrees. Tilted images represent a possibility to see by a student/physician an actual lesion from different angles. Then, by using the ***Show Symptoms*** button, it is possible to get information about selected symptoms in the lesion image (Fig. 3). There exists a possibility of more advanced computer-supported learning of the symptom diagnosis by generating a lesion with symptoms known only to the instructor, who may want a trained person to describe these symptoms and their localization.

Link ***Lesions***. There is another possibility of using a special data base with both synthetic and actual melanoma skin lesion images. These images were evaluated by a reliable team of diag-nosticians participating in the project. Both synthetic and actual images of the melanocytic skin lesions, stored in this data base, were carefully screened under control of very experienced diagnosticians. Only these diagnosticians were authorized to save image in this data base.

Link ***Research Results***. A publication list (with high **IF**) of our papers on computer-aided identification/classification of melanocytic skin lesions, containing 38 papers, is accessible.

Link ***Image Saving***. In order to run the module of saving digital images of lesions it is necessary to log in. Use the internal link loginpage. This internal link is accessible to our diagnosticians only.

4 Future Works

Further experiments, for extending of the number of values of K and for improving the **INP** tool, are currently conducted in the **UITM**, Rzeszow, Poland, in

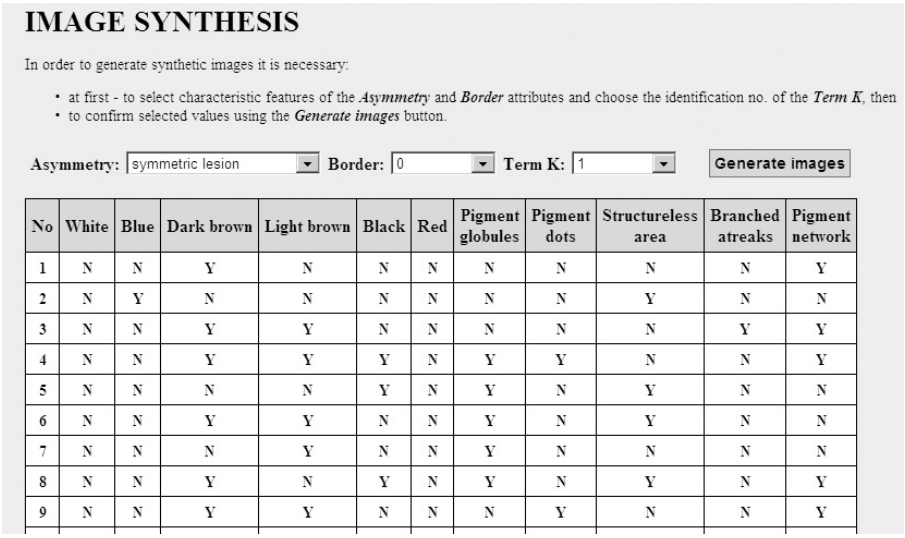


IMAGE SYNTHESIS

In order to generate synthetic images it is necessary:

- at first - to select characteristic features of the *Asymmetry* and *Border* attributes and choose the identification no. of the *Term K*, then
- to confirm selected values using the *Generate images* button.

Asymmetry: symmetric lesion Border: 0 Term K: 1 Generate images

No	White	Blue	Dark brown	Light brown	Black	Red	Pigment globules	Pigment dots	Structureless area	Branched streaks	Pigment network
1	N	N	Y	N	N	N	N	N	N	N	Y
2	N	Y	N	N	N	N	N	N	Y	N	N
3	N	N	Y	Y	N	N	N	N	N	Y	Y
4	N	N	Y	Y	Y	N	Y	Y	N	N	Y
5	N	N	N	N	Y	N	Y	N	Y	N	N
6	N	N	Y	Y	N	N	Y	N	Y	N	N
7	N	N	N	Y	N	N	Y	N	N	N	N
8	N	N	Y	N	Y	N	Y	N	Y	N	Y
9	N	N	Y	Y	N	N	N	Y	N	N	Y

Fig. 1. Specification (fragment) of the term **K**

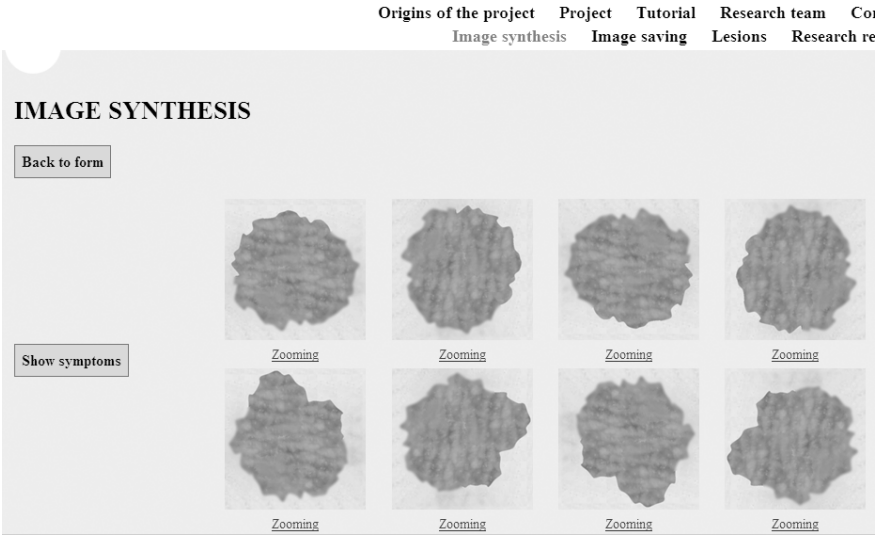


Fig. 2. Synthesized digital images of melanocytic skin lesions

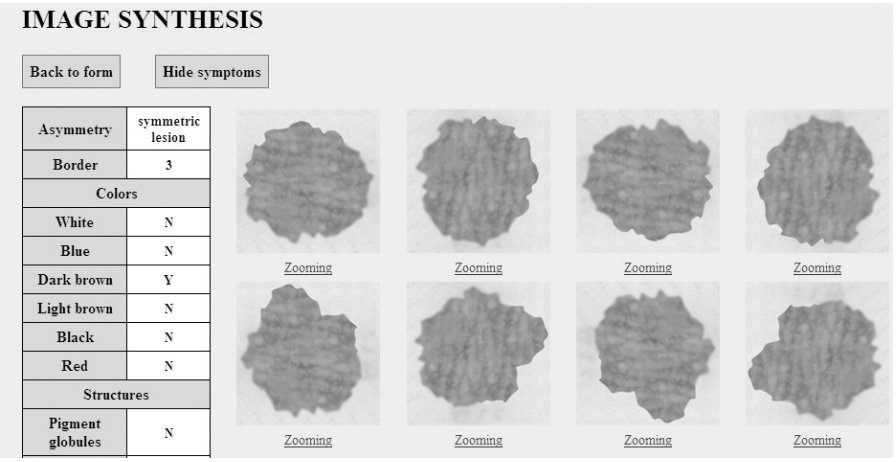


Fig. 3. Synthesized digital images of melanocytic skin lesions. Symptoms and diagnose are shown

strict cooperation with the Department and Clinic of Dermatology, Venereology and Allergology, Wroclaw Medical University, Wroclaw, Poland.

Acknowledgement. This research has been supported by the grant No. N N516 482640 from the National Research Center in Cracow, Poland. Furthermore, a special thank goes to Andrzej Bieniek, D. Sc., Ph. D. med. and Piotr Nockowski, Ph. D. med. (from Department and Clinic of Dermatology, Venereology and Allergology, Wroclaw Medical University), for contributing and valuable thoughts. And finally, I would like to express my thanks to Colleagues from the Department of Artificial Intelligence and Expert Systems for their professional and devoted work done during our research.

References

1. Bajcar, S., Grzymała-Busse, J.W., Hippe, Z.S.: A Comparison of Six Discretization Algorithms Used for Prediction of Melanoma. In: Kłopotek, M.A., Wierzchoń, S., Michalewicz, M. (eds.) Advances in Soft Computing, pp. 3–12. Springer, Berlin (2002)
2. Cudek, P., Grzymała-Busse, J.W., Hippe, Z.S.: Further Research on Automatic Estimation of Asymmetry of Melanocytic Skin Lesions. In: Hippe, Z.S., Kulikowski, J.L., Mroczek, T. (eds.) Human - Computer Systems Interaction. AISC, vol. 99, pp. 125–129. Springer, Heidelberg (2012)
3. Piatek, L.: Research on algorithms to synthesise the static medical images - an example of melanocytic skin lesions. Ph. D. Thesis, Gdańsk University of Technology, Gdańsk (2009) (in Polish)
4. Stolz, W., Braun-Falco, O., Bilek, P., Landthaler, M., Burgdorf, W.H., Cogneta, A.B.: Color Atlas of Dermatoscopy. Blackwell Wissenschafts-Verlag, Berlin (2002)

Telemetry Recording of the Electromyographic Activity of Female Reproduction Tract

Ewelina Brzozowska[1], Edward Oczeretko[1],
Bartosz Pawliński[2], and Zdzisław Gajewski[2]

[1] Bialystok University of Technology, Faculty of Mechanical Engineering,
Department of Materials and Biomedical Engineering,
Wiejska 45C, 15-351 Białystok
e.brzozowska@doktoranci.pb.edu.pl, e.oczeretko@pb.edu.pl
[2] Warsaw University of Life Sciences,
Department of Large Animal Diseases with Clinic,
Nowoursynowska 100, 02-797, Warsaw

Abstract. Bioelectrical activity of the female reproductive tract in different physiological and pathological states is vitally important in clinical practice. Telemetry allows for sending signals from the myometrium and other parts of the reproductive tract to computer system for further analysis. Telemetry system with incorporated decision support software presents telemedicine. Monica AN24 system lets us perform analysis of the uterine muscle contractions, fetal and maternal heart rate. Pregnant woman might constantly be monitored, even in the course of performing various tasks at home. Radiotelemetry enables wireless measurements of biological signals from freely moving conscious animals. The greatest advantage is that the experimental animals are neither stressed, nor are they bound in any way. No human presence is also necessary during recording of the data.

Keywords: telemedicine, telemetry, uterine activity, electromyography.

1 Introduction to Telemedicine and Telemetry

The word "telemedicine", in other words "the medicine at a distance", originates from two words "tele" and "medicine". The first formal definition of telemedicine can be found in Bird [1]. He defines telemedicine as "the practice of medicine without the usual physician-patient confrontation, via interactive audio-video communications systems". Treatment at a distance means the use of information and communication technologies (ICT) to exchange medical information in order to improve treatment outcomes [2]. Many researchers applied their own definitions [3]. Because of that, in 2007 the World Health Organization (WHO) adopted the following broad description as the ultimate definition of telemedicine: "The delivery of health care services, where distance is a critical factor, by all health care professionals using information and communication technologies for the exchange of valid information for diagnosis, treatment and

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_10, © Springer International Publishing Switzerland 2014

prevention of disease and injuries, research and evaluation, and for the continuing education of health care providers, all in the interests of advancing the health of individuals and their communities" [4].

Telemedicine is a relatively new and fast developing area constantly adapted to the changing health needs and technological advances [5]. Telemedicine systems were first made and used by NASA during American astronaut John Glenn's space flight, when the physiological parameters were remotely monitored and the data were transferred to the Earth [6]. The consequence of these pioneering experiments was the introduction of medical technology helping remote doctors to have indirect contact with the absent patient. Telemedicine is widely used by surgeons to perform operations "at a distance". Modern technology, using fast processors and algorithms for digital signal processing and compression allows transmission of high-resolution images and interactive audiovisual transmission with extreme accuracy and in real time. This very fact enables scientific collaboration at a distance, too [7]. The American Telemedicine Association (ATA), which is a nonprofit organization, develops practice standards and guidelines in the fields of telemedicine. In 2012 there were over 500 documents with various standards downloaded from the ATA [8].

Telemetry, a part of telemedicine, has revolutionized the way medicine is practiced. In the most general sense, it is a branch of technology, which deals mainly with performance of automatic measurements and wireless transmission of results at a distance. The telemetric system consists of two main subsystems: the measuring subsystem, which is used to perform the direct measurements and the data transmission subsystem which is responsible for the transmission at a distance. Telemetry system with incorporated decision support software present telemedicine [9]. Telematics is the branch of science concerned with the use of technological devices to transmit information over long distances. Using safe and secure telematics channels, data can be transmitted from measurement devices to the place where measurements will be collected and processed. It is also possible to create a return channel, which enables the operator to affect the measuring subsystem [10,11]. The Internet, modems, radio systems and mobile phone networks are primarily used to transmit all measurements [12].

The aim of this paper is to present telemetric systems for monitoring bioelectric activity of the female reproductive tract.

2 Uterine Electromyographic Activity Measurements

By means of various kinds of telemetric systems we can monitor the electrophysiological activity of the uterus which corresponds to the mechanical activity of the myometrium. Uterine electromyography (EMG) or electrohysterography (EHG in human), was developed 70 years ago [13]. It is a non-invasive technique which uses maternal abdominal surface electrodes, allowing for measurement of electrical activity of uterine muscles at rest and during contraction [14]. It can

be the greatest tool in characterizing parturition [15]. Of course there are controversies if uterine electromyography may identify high risk of preterm labor [16,17].

The other studies have proved the usefulness of this non-invasive technology in the prediction of preterm delivery and monitoring uterine contractility during pregnancy [13,18], [19]. For human, telemetric system Monica AN24 (Monica Healthcare Ltd, Nottingham, UK) performs measurements of the uterine muscle contractions and fetal and maternal heart rate [20,21,22,23]. Trimed Ltd. is a representative of Monica Healthcare Ltd. in Poland [24].

Telemetry is also used in investigation of the electrical activity of the reproductive tracts of animals: mice [25], pregnant rats [26], monkeys [27], sows [28,29]. Values of the measured EMG signals are recorded by telemetric bio-potential sensors and a radio-telemetry transmitter.

2.1 *Monica AN24* – System for Women Monitoring

Monica AN24 is a lightweight, wearable and easy-to-use battery-powered telemetric system, suitable for extended monitoring of fetal well-being and uterine contractions. The measuring electrodes are placed on the abdominal skin (Fig. 1). After the measurements, it allows to transmit the results immediately, for particular patient ID, to an online system database ktg24 [30] and then to Monica Design Kit (DK) software. Monica decision support software was developed for remote doctors to evaluate decision making during (or after) real-time recording. DK software shows:

- fetal electrocardiography (ECG) morphology,
- beat to beat fetal heart rate variability (FHR),

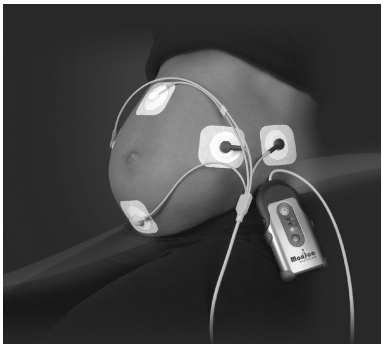


Fig. 1. Portable transdominal *Monica AN24* device (due to support given by Trimed Ltd); five electrodes are attached to the mother's abdomen

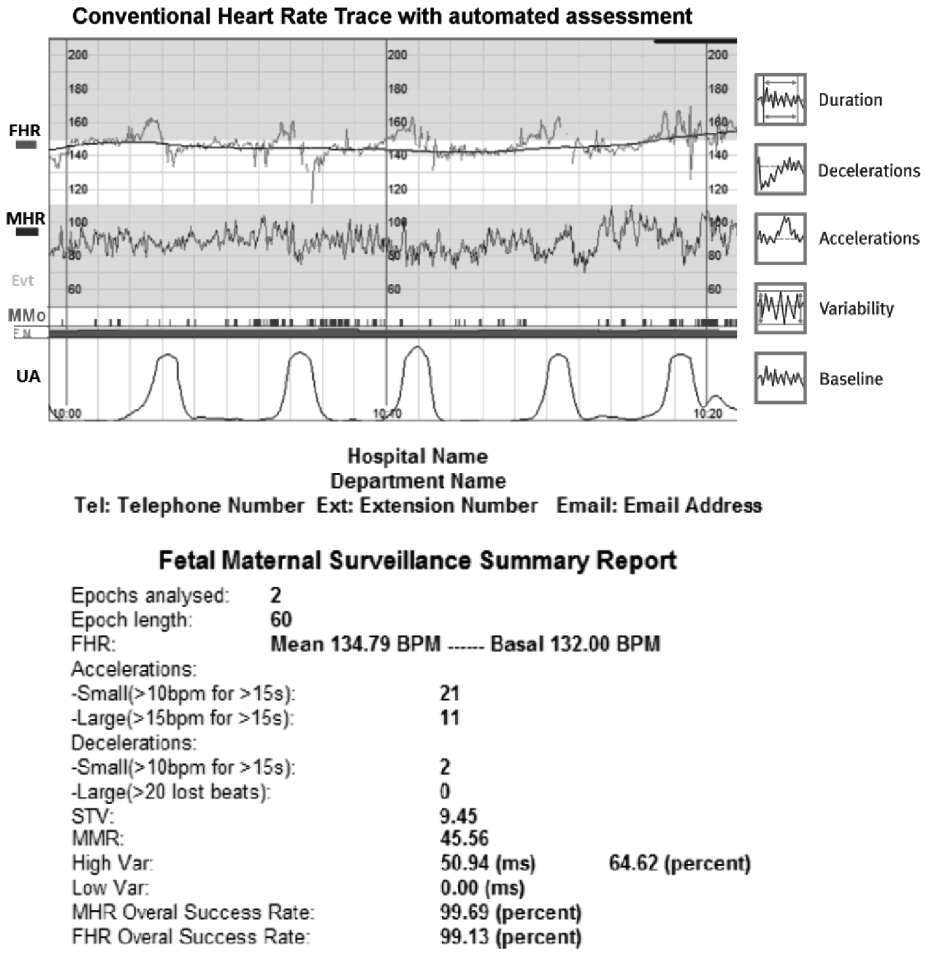


Fig. 2. Results obtained by means of the Monica AN24 device: fetal (FHR), maternal (MHR) heart rate, maternal movements (MMo) and uterine contractions (UA)

- beat to beat mother's heart rate variability (MHR),
- electrohysterogram (EHG) of the uterine muscles,
- arrhythmia identification,
- advanced FHR analysis features,
- access to the raw abdominal electrophysiological signals from a continuous 20-hour-long recording,
- data export for further statistical analysis.

Monica An24 operates in two modes: retrospective mode and real time mode. The device records data for retrospective upload (via USB) and on-screen viewing. In real time mode data are transferred using Bluetooth wireless transmission. Monica system is very good for scientific research, because it gives access to raw test results [31,32] and could increase maternal satisfaction, because it is very easy to use [33]. Fig. 2 shows the results obtained by means of the Monica AN24 device: fetal (FHR) and maternal (MHR) heart rate, maternal movements (MMo) and uterine contractions (UA). The data were taken from ktg24 website. In February 2011 the Monica system received acceptance of the Agency for Food and Drug Administration (FDA), they also underlined excellent collaboration with obese mothers confirmed by numerous clinical studies [34].

2.2 Telemetry Recording of the Reproductive Tract in the Sows

Our knowledge of uterus and oviduct electrical and contraction activity is limited though it is crucial for understanding the physiology and pathophysiology of the reproductive system. A radio-telemetric system was processed and customized at Warsaw University of Life Sciences, which allows for transforming biological signals from animal body into radio waves [29], [35]. The EMG activity signals were recorded from bulb and isthmus right oviduct and right uterus corn during estrus and early pregnancy in pigs. During the experiment the sows were kept in metabolic cages. The 3-channel telemetry implants (TL10M3-D70-EEE, DSI, USA) were placed surgically between the abdominal muscle layers and connected with three bipolar silver electrodes sutured on different parts of the uterus and oviducts (Fig. 3). The animals were recovered from surgery quickly and without complications. Research of reproduction tract carried out post mortem indicated that telemetry recording system and material (bipolar electrodes) didn't have negative influence on physiological processes in the studied animals. Study protocol was approved by the Local Ethics Committee.

The bio-signals received by electrodes from the animals body were transferred to the transmitter, from which they were sent to the receiver (RMC-1) located around $50cm$ away from the cage using radio waves. The received signal was amplified (BioAmps, ADInstruments), and was recorded afterwards using (PowerLab 4e, ADInstruments, Australia). *Chart v. 4.0* software was used for saving, storing and analysing the data. Fig. 4 shows the scheme of telemetric system for electromyographic recording in pigs. Telemetric implant is shown in Figure 5.

The EMG signal was checked every time, directly after the surgery to evaluate the correctness of the surgery. Generally, a disturbance-free signal was registered on the 3rd-5th day after the surgery. Regular registrations began on day 3. The electrodes received clear, artifact-free signal usually 5-6 weeks after the surgery. Due to the limited source of energy, after each registration session, the transmitter was turned off with the application of a magnet.

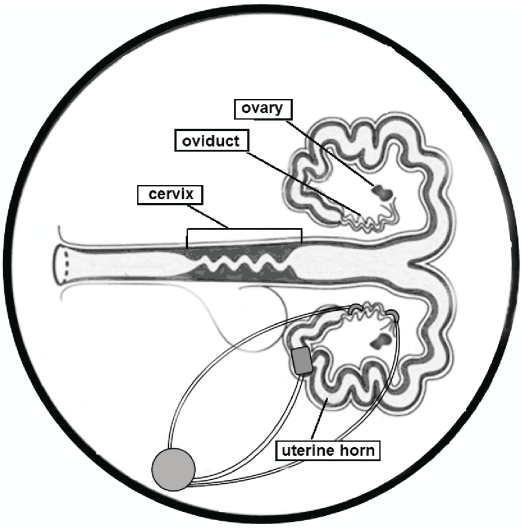


Fig. 3. Reproductive tract of the pig; electrodes are placed in the uterus corn, isthmus oviduct and bulbs oviduct

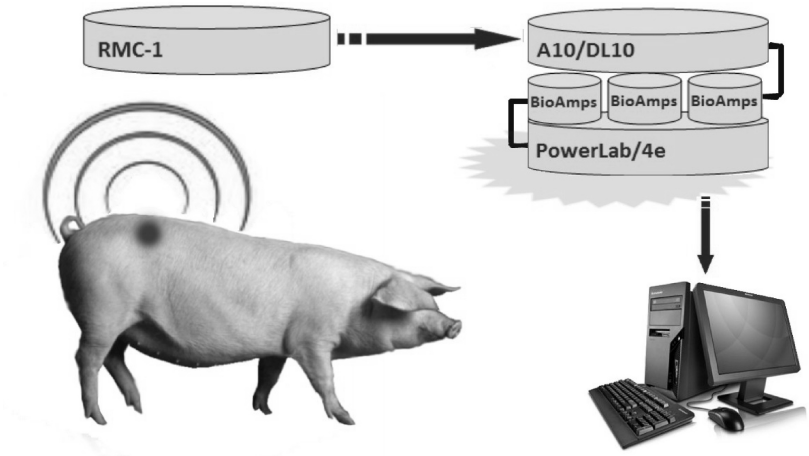


Fig. 4. The scheme of telemetry system for electromyographic recording in the sows

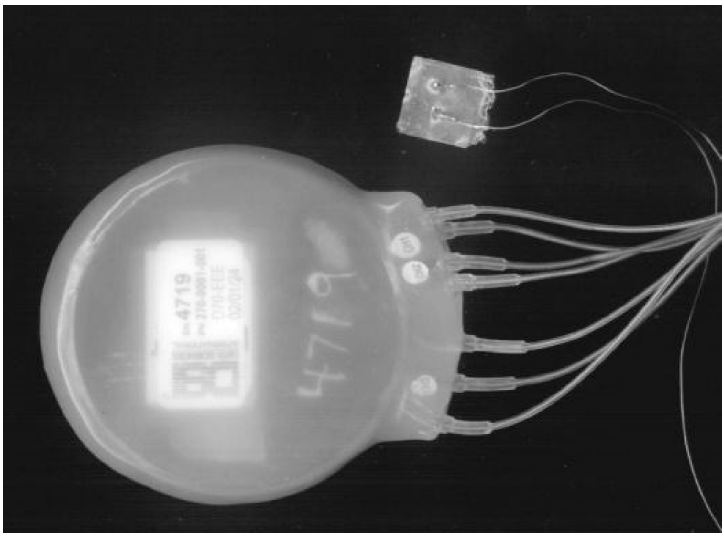


Fig. 5. Telemetric implant TL10M3-D70-EEE, DSI, USA

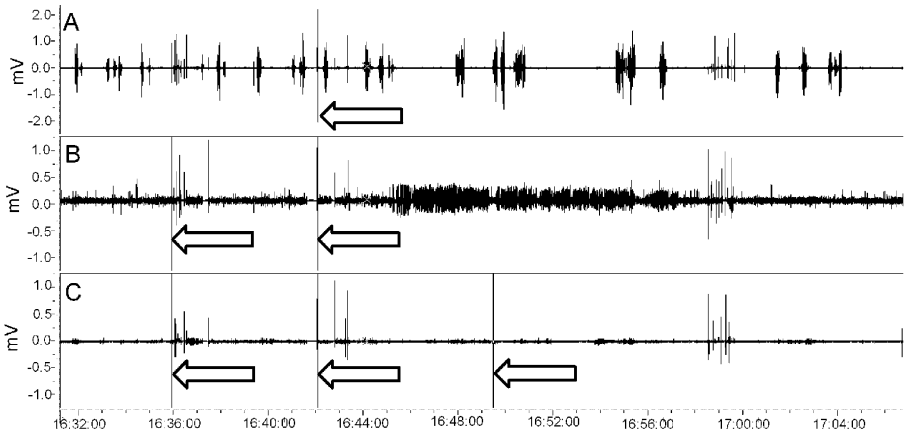


Fig. 6. Representative signals (30 minutes) of the uterus corn (A), isthmus oviduct (B) and bulbs oviduct (8 hours after first artificial insemination) of myoelectric activity in pigs recorded using telemetry; the x-axis represents real time scale ($hh : mm : ss$); artifacts are marked with arrows

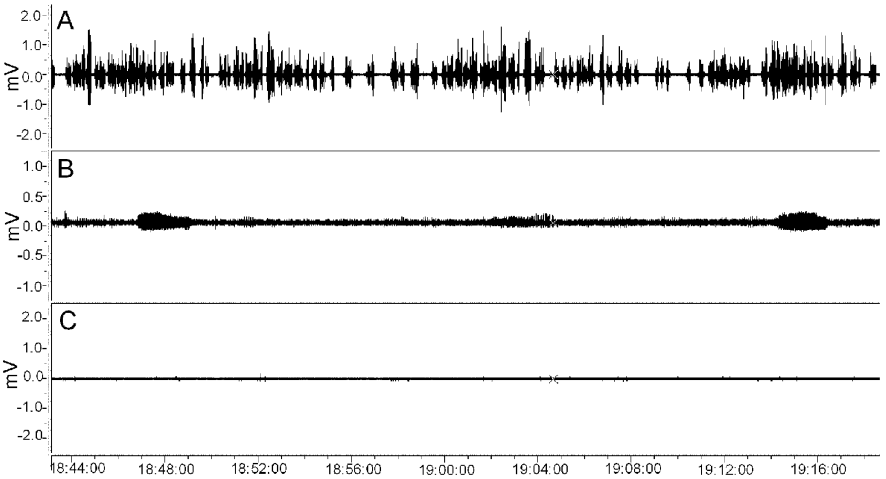


Fig. 7. Representative signals (30 minutes) of the uterus corn (A), isthmus oviduct (B) and bulbs oviduct (during 6th day in early pregnancy) myoelectric activity in pigs recorded using telemetry; the x-axis represents real time scale ($hh : mm : ss$)

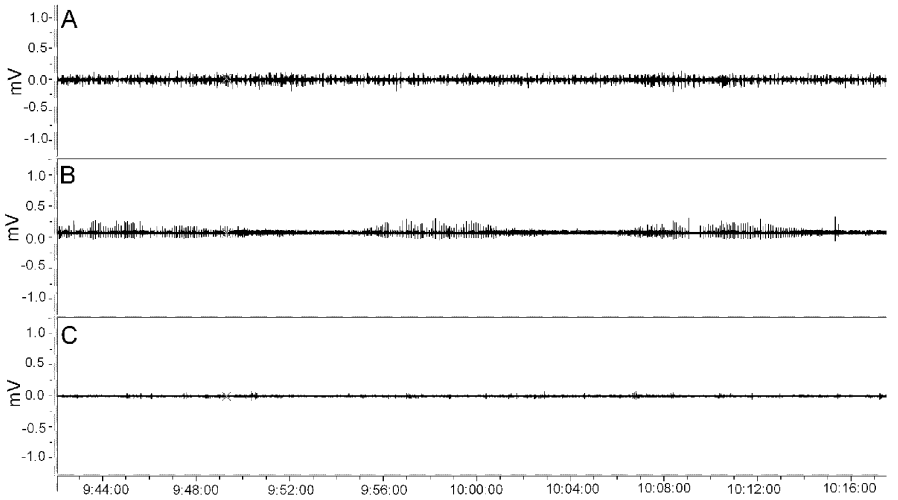


Fig. 8. Representative signals (30 minutes) of the uterus corn (A), isthmus oviduct (B) and bulbs oviduct (during 18^{th} day in early pregnancy) of myoelectrical activity in pigs recorded by telemetry; the x-axis represents real time scale ($hh : mm : ss$)

Representative EMG signals are shown on figures 6, 7, 8. Electromyographic activity shows different patterns in the uterus, bulbs oviduct in relation to time after successful insemination. In recording from uterus corn (Fig. 6A) numerous separate spikes of biopotential activity with lowered amplitude and shortened duration are visible. In isthmus oviduct EMG stimulation syndrome characterized by extended duration and constant amplitude can be observed. What is more, single biopotential activities are also visible (Fig. 6B). In bulbs oviduct both single biopotential activities and extended periods with no EMG activity were registered (Fig. 6C). Fig. 7 shows an increase in number of single biopotential activities and increase in amplitude of separated spikes of discharges which were recorded in uterus corn. In isthmus oviduct single biopotential activities with extended duration and constant amplitude were visible. No EMG activity was registered in bulbs oviduct. Lack of separate spikes of biopotential activities in uterus and bulbs oviduct in recorded signals (Fig. 8A and B). Separate bunches characterized by low amplitude can be observed in isthmus oviduct (Fig. 8C). The results are consistent with the current state of knowledge regarding reproductive sows.

3 Discussion

In studies of uterine contractile activity in humans we see an increasingly important role of telemetric systems [22], [36,37]. The *Monica AN24* system is the future of preterm birth detection, because it allows for non-invasive medical diagnostics which can be done outside hospital. It provides a great number of data for research, allowing the mothers' freedom of movement at the same time. Test results may be electronically sent for interpretation to a physician which can account for reduced waiting time to get medical advise. That is a great advantage over classical methods which require diagnosing done while staying in hospital bed. *Monica AN24* detects the best channel for monitoring fetal (FHR) and mothers heart rate (MHR), reduces the possibility of confusing signals, which has the advantage over Doppler CTG. Eight out of ten women would prefer the *Monica AN24* monitor system over the traditional Doppler investigation [38]. Incorporated decision support makes *Monica AN24* real a telemedical system.

In animals, telemetry offers the stress-free monitoring of their reproductive tracts for a long period of time. The risk of interference caused by stress with physiological processes such a spontaneous electrical activity is minimal [39]. Without telemetry, obtaining the results presented in paragraph 2.2 would have been impossible.

References

1. Bird, K.T.: Teleconsultation: a New Health Information Exchange System. Third Annual report to the Veterans Admnistration, Washington, DC (1971)
2. Murphy, R.L.H., Bird, K.T.: Telediagnosis: a New Community Health Resource. Observations on the Feasibility of Telediagnosis Based on 1,000 Patient Transactions. Am. J. Public Health 64, 113–119 (1974)

3. Strehle, E.M., Shabde, N.: One Hundred Years of Telemedicine: Does this New Technology Have a Place in Paediatrics? Arch. Dis. Child. 91, 956–959 (2006)
4. World Health Organization.: Telemedicine: Opportunities and Developments in Member States: Report on the Second Global Survey on eHealth. World Health Organization, http://www.who.int/goe/publications/ehealth_series_vol2/en/
5. Huston, T.L., Huston, J.L.: Is Telemedicine a Practical Reality? Commun. ACM 43, 91–95 (2000)
6. Gilbert, B.K., Mitchell, M.P., Bengali, A.R., Khandheria, B.K.: NASA/DARPA Advanced Communications Technology Satellite Project for Evaluation of Telemedicine Utreach Using Next-Generation Communications Satellite Technology: Mayo Foundation participation. Mayo Clin. Proc. 74, 753–757 (1999)
7. Teasley, S., Wolinsky, S.: Communication Scientific Collaborations at a Distance. Science 292, 2254–2255 (2001)
8. Krupinski, E.A., Antoniotti, N., Bernard, J.: Utilization of the American Telemedicine Association's Clinical Practice Guidelines. Telemed J. e-Health 19, 846–851 (2013)
9. Klonoff, D.C., True, M.W.: The Missing Element of Telemedicine for Diabetes: Decision Support Software. J. Diabetes Sci. Technol. 3(5), 996–1001 (2009)
10. Duplaga, M.: Systemy Telemedyczne. In: Zajdel, R., Kącki, E., Szczepaniak, P., Kurzyński, M. (eds.) Kompendium Informatyki Medycznej, pp. 519–527. Alfa Medica Press, Bielsko-Biała (2003)
11. Dumas, P., Chiche, D., Tremblay, J., Seda, O., Peng, J., Pavel, H.: Radio-Telemetry in Biomedical Research - Radio-Telemetry Blood Pressure Measurements in Animal Models of Hypertension, How It Revolutionized Hypertension Research. In: Krejcar, O. (ed.) Modern Telemetry, pp. 117–144. InTech, Croatia (2011)
12. Bashshur, R.L., Reardon, T.G., Shannon, G.W.: Telemedicine: a New Health Care Delivery System. Annu. Rev. Public Health 21, 613–637 (2000)
13. Vasak, B., Graatsma, E.M., Hekman-Drost, E., Eijkemans, M.J., van Leeuwen, J.H.S., Visser, G.H., Jacod, B.C.: Uterine Electromyography for Identification of First-Stage Labor Arrest in Term Nulliparous Women with Spontaneous Onset of Labor. Am. J. Obstet. Gynecol. 209, 232.e1-8 (2013)
14. Devedeux, D., Marque, C., Mansour, S., Germain, G., Duchene, J.: Uterine Electromyography: a Citical Rview. Am. J. Obstet. Gynecol. 169, 1636–1653 (1993)
15. Schlembach, D., Maner, W.L., Garfield, R.E., Maul, H.: Eur. J. Obstet. Gynecol. Reprod. Biol. 144(suppl. 1), S33–S39 (2009)
16. Gupta, D.: Electrohysterography May Be Integral to Remifentanil Labor Analgesia. J. Matern. Fetal Neonatal Med. 25(2), 151–152 (2012)
17. Reinhard, J., Hayes-Gill, B.R., Schiermeier, S., Loser, H., Niedballa, L.M., Haarmann, E., Sonnwald, A., Hatzmann, W., Heinrich, T.M., Louwen, F.: Uterine Activity Monitoring During Labour - a Multi-Centre, Blinded Two-Way Trial of External Tocodynamometry against Electrohysterography. Z. Geburtshilfe Neonatol. 215, 199–204 (2011)
18. Buhimschi, C., Boyle, M.B., Garfield, R.E.: Electrical activity of the human uterus during pregnancy as recorded from the abdominal surface. Obstet. and Gynecol. 90, 102–111 (1997)
19. Lucovnik, M., Kuon, R.J., Chambliss, L.R., Maner, W.L., Shi, S.Q., Shi, L., Balducci, J., Garfield, R.E.: Use of uterine electromyography to diagnose term and preterm labor. Acta Obstetricia Et Gynecologica Scandinavica 90(2), 150–157 (2011)
20. Monica Healthcare Ltd., http://www.monicahealthcare.com

21. Kerner, R., Yogev, Y., Belkin, A., Ben-Haroush, A., Zeevi, B., Hod, M.: Maternal Self-Administered Fetal heart Rate Monitoring and Transmission From Home in High-Risk Pregnancies. Int. J. Gynecol. Obstet. 84, 33–39 (2004)
22. Ungureanu, G.M., Gussi, I., Wolf, W., Taralunga, D., Pasca, S., Strungaru, R.: Prenatal Telemedicine - Advances in Fetal Monitoring. In: Graschew, G. (ed.) Advances in Telemedicine: Applications in Various Medical Disciplines and Geographical Regions, pp. 97–120. InTech (2011)
23. Rauf, Z., O'Brien, E., Stampalija, T., Ilioniu, F.P., Lavender, T., Alfirevic, Z.: Home Labour Induction with Retrievable Prostaglandin Pessary and Continuous Telemetric Trans-Abdominal Fetal ECG Monitoring. PLoS One 6(11), e28129 (2011)
24. Trimed Ltd., http://www.trimed.pl
25. Pierce, S.L., Kutschke, W., Cabeza, R., England, S.K.: In Vivo Measurement of Intrauterine Pressure by Telemetry: a New Approach for Studying Parturition in Mouse Models. Physiol. Genomics. 42, 310–316 (2010)
26. Mackay, L.B., Shi, S.Q., Garfield, R.E., Maner, W.L.: The Effect of Bilateral Pelvic Neurectomy on Uterine and Abdominal Electrical and Pressure Activity, as Measured by Telemetry in Conscious, Unrestrained Pregnant Rats. J. Perinat. Med. 37, 313–319 (2009)
27. Chellman, G.J., Bee, W.H., Resendez, J.C., Meyer, J.K., Bernal, J., Moddrelle, D.S., Haluska, G.J., Reinheimer, T.M.: Telemetric Uterine Contraction Model in Preterm Cynomolgus Monkeys. Reprod. Toxicol. 18, 285–293 (2004)
28. Pawlinski, B.: Sow oviduct and uterus myoelectrical activity in reproduction cycle recorded by telemetry. PhD Thesis, pp. 1–109 (2009)
29. Gajewski, Z., Pawlinski, B., Zabielski, R.: Sow Oviduct and Uterus Myoelectrical Activity During Early Pregnancy in Pigs Using Telemetry Implants. Reprod. Domestic Anim. 47(suppl. 4), 481–485 (2012)
30. KTG24., http://www.ktg24.pl
31. Mikkelsen, E.: Electromyographic Characterization of Uterine Contractions during Labor Using the Commercially Available Equipment Monica. Doctoral dissertation, Faculty of Health Sciences, University of Aarhus (2012)
32. Oczeretko, E.: Przetwarzanie sygnałów cyfrowych z wybranymi zastosowaniami do badań czynności skurczowej macicy, pp. 1–191. Oficyna Wydawnicza Politechniki Białostockiej, Białystok (2013)
33. Boers, K.E., Vijgen, S.M., Bijlenga, D., van der Post, J.A., Bekedam, D.J., Kwee, A.: Induction Versus Expectant Monitoring for Intrauterine Growth Restriction at Term: Randomised Equivalence Trial (DIGITAT). BMJ 341, c7087 (2010)
34. Dolan, B.: FDA Clears Monica's Wireless Fetal Heart Rate Monitor (2011), http://www.mobihealthnews.com/10264/fda-clears-monicas-wireless-fetal-heart-rate-monitor
35. Wolinski, J., Ziecik, A., Gajewski, Z., Korczynski, W., Zabielski, R.: A Method of Recording the Oviduct and Uterus Myoelectrical Activity in Gilts Using Implantable Telemetry. J. Anim. Feed Sci. 12, 359–369 (2003)
36. Kerner, R., Yogev, Y., Belkin, A., Ben-Haroush, A., Zeevi, B., Hod, M.: Maternal Self-Administered Fetal Heart Rate Monitoring and Transmission from Home in High-Risk Pregnancies. Int. J. Gynecol. Obstet. 84, 33–39 (2004)
37. Aslanidi, O., Atia, J., Benson, A.P., van den Berg, H.A., Blanks, A.M., Choi, C., Gilbert, S.H., Goryanin, I., Hayes-Gill, B.R., Holden, A.V., Li, P., Norman, J.E., Shmygol, A., Simpson, N.A.B., Taggart, M.J., Tong, W.C., Zhang, H.: Towards a Computational Reconstruction of the Electrodynamics of Premature and Full Term Human Labour. Prog. Biophys. Mol. Bio. 107, 183–192 (2011)

38. Reinhard, J., Hatzmann, H., Schiermeier, S.: Fetales Elektrokardiogramm (EKG) als Alternative der Doppler-Kardiotokografie (CTG) zur antepartualen Uberwachung des Feten - erste Ergebnisse. Z. Geburtsh. Neonatol. 212, 226–229 (2008)
39. Carbonne, B., Cabrol, D., Clerget, M.S., Germain, G.: Nomegestrol Acetate Effects on Spontaneous and Sulprostone-Induced Uterine Contraction in Pregnant Cynomolgus Monkey Monitored by Telemetry. Am. J. Obstet. Gynecol. 178, 150–155 (1998)

Part III

Telegeriatrics

The Application of Adapted TICS (Telephone Interview for Cognitive Status) for Diagnostics of Cognitive Function Disturbances in Elderly Patients. A Pilot Study

Jarosław Derejczyk[1], Olga Stępień-Wyrobiec[1], Aneta Hanusiak[1], Adam Bednorz[1], Edyta Karasek[2], Katarzyna Hornowska-Suchoń[1], Jan Szymszal[3], and Agnieszka Pogonowska[1]

[1] Jan Paweł II Geriatric Hospital, Katowice, Poland
jarek@derejczyk.com
[2] EMC Medical Institute S.A., Wrocław, Poland
[3] Silesian Technical University, Gliwice, Poland

Abstract. The early diagnosis of Alzheimer's disease and mild cognitive impairment (MCI) may help reduce disability and enhance quality of life through early introduce of multimodal therapy. This study evaluated the utility of the telephone interview which was compared in the evaluation of cognitive function in 55 patients of geriatric hospital, with the classical methods of neuropsychological diagnosis. Data demonstrates that TICS can be used as a simple and accessible method for the observation and estimation of cognitive status on early stage of dementia. Widespread tele-diagnostics of early phases of dementia and other functional deficits of the brain, connected with its aging, stands a chance of becoming a standard in the assessment of health condition of the elderly population, and of contributing to quicker provision of medical and community care for them.

Keywords: TICS, telephone interview for cognitive status, telehealth, aging, cognition.

1 Introduction

The quickly growing population of people in advanced senility generates the need for conducting screening examinations, which would allow to confirm or exclude the existence of a cognitive norm, on a wider scale than it presently happens. One of the methods for extending the diagnostic access to the assessment of cognitive skills at pre-clinical level is telephone interview method, having more than 20 years of tradition. Tele-diagnostic tests have been recognized by many researchers as a method having wide application possibilities as screening examination to detect disturbances in cognitive functions. The TICS (Telephone Interview for Cognitive Status) test, as well as its modifications – TICS-M – suggested by Brandt [1], Welsh [2], Hogervorst [3] are promising strategies, which

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_11, © Springer International Publishing Switzerland 2014

allow to identify groups of people with mild cognitive impairment, especially in large epidemiological studies. The neuropsychological studies applied at present, and considered a gold standard in detection of cognitive dysfunction, are a cost intense studies and, in reality, not available for a large population of patients, whereas the answer to the question whether TICS test may be treated as an alternative for a large group of senior citizens is not univocal at present.

2 Assumptions and Aim of the Study

The aim of the study is to verify the diagnostic usefulness of TICS test [4] for exclusion or confirmation of the presence of cognitive dysfunctions in people over 60 years of age, treated in a geriatric ward; and the assessment of its reliability in comparison with results of classical neuropsychological tests, which in this study were: Mini Mental State Examination (MMSE) [5] and Montreal Cognitive Assessment - MoCA [6]. The study was also performed in order to gain knowledge and experience in the assessment of usefulness of tools that enable screening tele-diagnostics of dementia in early phases of its development, as well as for the assessment of difficulties in its performance by psychologists who perform it.

3 Material and Methods

Telephone test of TICS class has been used for the assessment of cognitive skills in a tele-diagnostic system. In the test, the list of questions testing cognitive skills has been translated from the original version, and adapted for the needs of our study [4]. A group of 55 people has been tested, patients of the Jan Paweł II Geriatric Hospital in Katowice, Poland, between August and November 2013. The study group constituted of 55 patients, the average age was 81.2 yrs, in which 42 were women (average age 81.1 yrs) and 13 were men (average age 81.5 yrs). All the tested patients voluntarily gave consent for extending the routine neuropsychological assessment by an additional test, performed as a telephone interview. In inclusion criteria for the test were as follows: being over 75 years of age, absence of earlier diagnosis of dementia, no pro-cognitive drugs used by the patient before, absence of deep cognitive deficiencies on subjective examination, absence of behaviour disorders, absence of significant vision impairment (that would preclude the performance of some tasks in MMSE and MoCA tests), absence of substantial hearing loss (which would preclude the performance of TICS test). History taking included enquiry about the previous occurrence and duration of diabetes, incidence of Parkinson's disease, occurrence of depression syndrome, > 5 points in the 15-point GDS - Geriatric Depression Scale, stroke history [7]. Also the correlation of cognitive skills with selected laboratory parameters has been assessed, such as leucocytosis and C reactive protein (CRP) concentration. The examination was executed by performing three test in one diagnostic session with a neuropsychologist. The MMSE and MoCA tests were performed in standard manner – face to face. In case of the TICS test, the conversion was on the phone, during that the examined patient and the psychologist performing

1. Please tell me your full name. *1 point for first name; 1 point for last name.*
2. What is today's date (day, month, and year) – without looking at the calendar? *1 point for day; 1 point for month; 1 point for year.*
3. What is the day of the week? *1 point if correct.*
4. What season are we in? *1 point if correct.*
5. State your age and phone number. *1 point for age; 1 point for telephone number.*
6. Count backwards from 20 to 1. *2 points if correct on first trial; 1 point if correct on second trial.*
7. I'm going to read you a list of ten words. Please listen carefully and try to remember them. When I am done, tell me as many as you can, in any order. Ready? The words are: drum, curtain, bell, coffee, school, parents, moon, garden, hat, farmer. Now tell me all the words you can remember. *1 point for each correct word.*
8. One hundred minus seven equals what? And seven from that? (repeat "and seven from that" three times). *1 point for each correct subtraction. Stop at 5 subtractions.*
9. What do people usually use to cut paper? *1 point for correct response.*
10. How many things are in a dozen? *1 point for correct response.*
11. What do you call the green prickly plant that lives in a desert? *1 point for correct response.*
12. What animal does wool come from? *1 point for correct response.*
13. Please say this "no ifs, ands or buts". *1 point for each correct response on first trial.*
14. Please say this "absolute monarchy". *1 point for correct response.*
15. What is the name of the recent Pope? *2 points for full correct response; 1 point for partial response.*
16. What is the full name of the current Prime Minister of Poland? *1 point for correct first name; 1 point for correct last name.*
17. With your finger, tap five time on the part of the telephone you speak into. *2 points if 5 taps are heard; 1 point if more or less 5 taps are heard.*
18. I'm going to give you a word, and I want you to give me its opposite. For example, the opposite to hot is cold. What is the opposite to west? What is the opposite to generous? *1 point for each correct response.*
19. Remember the word list that you repeated before? Now tell me as many as you can. *1 point for each correct response.*

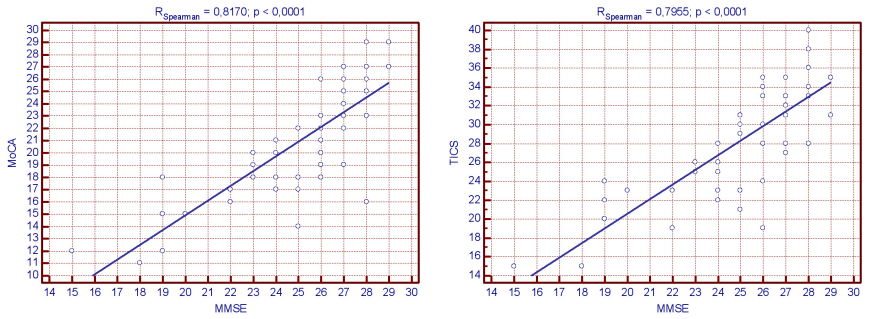
Fig. 1. Polish version of TICS test (Telephone Interview for Cognitive Status)

the examination were seated in separate rooms, and in case the elderly person tested clearly preferred the presence of testing person in the same room, they were seated back to back, with no eye contact. The first test to be solved was MMSE, which was followed by MoCA, and the last one was TICS. It has been assumed, on the basis of literature data, that in each of the first two tests the maximum score is 30 points [8,9]. The maximum result a person could reach in the TICS test, answering the questions and doing memory exercises is 49 points [10,11] (Fig. 1). The three tests solved differ slightly as to the range of cognitive functions investigated. All three of them allow to assess short term memory, attention, language functions, abstract thinking, as well as orientation, but not

in identical diagnostic range. The TICS test does not assess, in comparison with MMSE and MoCA, similarly to them the tasks concerning construction praxis and visual-spatial functions, as no exercises connected with drawing are performed there. In the TICS test, visual-spatial functions and abstract thinking are assessed by 2 questions, language functions by 8 questions, short-term memory by 3 questions, attention by 2 questions, and orientation by 4 questions. It has been assumed that results obtained by the patients examined, scored in points in MoCA and TICS tests, will be compared with results for MMSE, for which specific norms apply, defined for Polish population, which test has been considered the basic test, for the assessment of MoCA and TICS tests [12]. For the MMSE test, being the basis of norm for statistical analysis, it has been assumed that the score within norm is in the range of 28-30 points; the score of 24-27 points indicates mild cognitive problems; a score of 20-23 points corresponds to mild dementia. Patients with dementia who were found in the MMSE score of 20 points or less were excluded from the participation in the study. On the basis of such a qualification, the persons included in the study were patients without deficit of cognitive functions, patients with mild cognitive impairment, and patients with mild form of dementia. The patients who have not qualified for the study were those with fever, considerably dehydrated, as well as those taking drugs from benzodiazepine group and/or neuroleptics. In statistical analyses the results of MoCA and TICS tests have been assessed, which corresponded to the results of MMSE test at the level of 27 and 24 points, assuming that results of MMSE test below 23 [points] correspond to mild dementia, whereas scores above 27 [points] are within the norm. Statistical analysis has been performed on the basis of Spearman's rank correlation test, contingency coefficient test, Chi-square independence test, as well as tests for assessing the sensitivity and specificity, TD (Diagnostic Test), and ROC method (ROC curve analysis) [13].

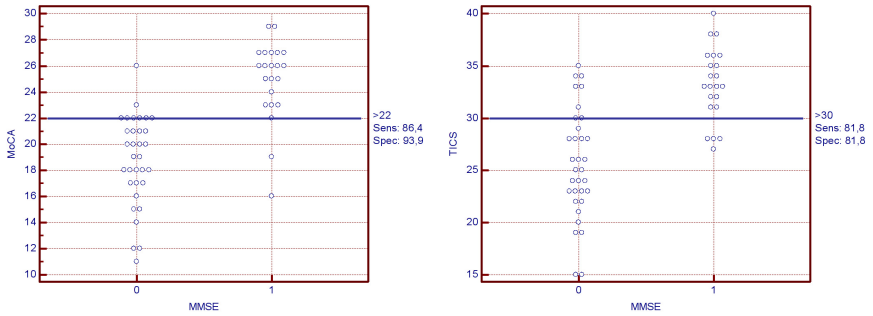
4 Results

A high correlation has been confirmed between the results of MMSE test and those of MoCA and TICS tests. Correlation between results of MMSE and MoCA tests was R = 0.8170 and correlation between results of MMSE and TICS test R = 0.7955, respectively (Fig. 2). The highest score obtained in the study by a patient from the study group of 55 persons in the TICS test was 40 points, the lowest was 15 points. It has been found that the boundary line (cat –off level) between full cognitive efficiency (MMSE from 27 to 30 points) and the level of mild cognitive impairments, which in the MMSE test was determined at the level lower 27 points, amounted to 22 points, for the MoCA test, compared with the MMSE test, and 30 points for the TICS test compared with MMSE. Fig. 3. Sensitivity and specificity of such cut-off level was, for the MoCA test: 0.8636 – 86.4% and 0.9394 - 93.9%, respectively, while for the TICS test: 0.8182 – 81.8% and 0.8182 – 81.8%, respectively. For the level being the borderline between mild cognitive impairment and mild dementia, which in MMSE was determined at 24 points [12], for the MoCA test 20 points were equivalent, while for the TICS test,



(a) (b)

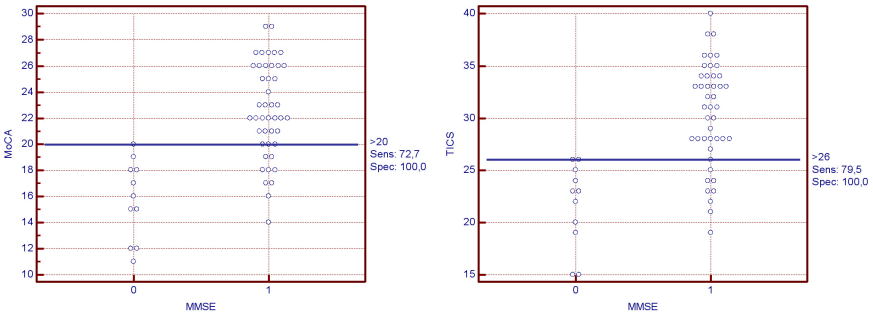
Fig. 2. Correlation between results of MMSE vs MoCA test (a) and MMSE vs TICS test (b)



(a) (b)

Fig. 3. Determination of the cut-off level for normal cognitive function between MMSE and MoCA test (a) and MMSE and TICS test (b) based on the level of 27 points in the MMSE test. (0 = MMSE from 0 to 26 points, and 1 = MMSE from 27 to 30 points)

compared with MMSE, the equivalent was 26 points (Fig. 4). The sensitivity and specificity for such cut-off level in case of MoCA test, amount to 0.7273 – 72.7% and 1.0000 - 100 %, respectively, while for the TICS test, they amount to 0.7955 – 79.5% and 0.8182 – 81.8% and 1.0000 – 100%, respectively. Using the analysis based on the Chi – square test, it has been found that correlation exists between the classical tests (MMSE and MoCA) and the TICS test for short term memory, attention, and orientation. Lower level of dependence has been found to exist between the assessment of language functions in MMSE and TICS tests, whereas in case of visual spatial skills and abstract thinking, no dependence has been found between TICS test and classical tests. Results are provided in Table 1. No correlation has been found, on the basis of Spearman rank analysis, between the tests performed on the study group of 55 patients and the age of patients, the number of years in their education history, the occurrence of diabetes, depression, and Parkinson's disease, as well as testing for leucocytosis



(a) (b)

Fig. 4. Determination of the cut-off level for mild dementia between MMSE and MoCA test (a) and MMSE and TICS test (b) based on the level of 24 points in the MMSE test. (0 = MMSE from 0 to 23 points, and 1 = MMSE from 24 to 30 points)

and CRP. The performance of the test, in the form used, was assessed as easy by the psychologists carrying them out. A certain difficulty for the subjects was the prolonged examination time, as the tests were performed in one go, and fatigue towards the end of tests has been noticed sporadically, in some of the subjects. The performance of all tests took some 65 minutes for one person, on the average, allowing for short breaks. The average time for performance of TICS test was 22 minutes. In the own assessment of psychologists, concerning their own reaction to the distinct form of performing the test, the description use was "adaptation of a methods, in which the performance of TICS test posed no barrier in the test execution". On the basis of opinions of psychologists who performed the test, the following advantages and disadvantages of TICS test were defined:

Table 1. Assessment of correlation of selected ranges of cognitive functions: orientation, attention, short term memory, language functions, as well as spatial functions and abstract thinking in classical tests (MMSE and MoCA)

Function assessed	Correlation between tests	N	R Spearman	t(N-2)	p level
Orientation	TICS vs MoCA	55	0.5199	4.4314	0.0000
	TICS vs MMSE	55	0.5172	4.3990	0.0001
Attention	TICS vs MoCA	55	0.5788	5.1678	0.0000
	TICS vs MMSE	55	0.6430	6.1128	0.0000
Short term memory	TICS vs MoCA	55	0.7551	8.3844	0.0000
	TICS vs MMSE	55	0.7085	7.3088	0.0000
Language functions	TICS vs MoCA	55	0.4557	3.7266	0.0005
	TICS vs MMSE	55	0.3289	2.5351	0.0142
Spatial and abstract functions	TICS vs MoCA	55	0.0321	0.2339	0.8160
	TICS vs MMSE	55	0.0094	0.0684	0.9458

- Advantages of the test:
  - it is easy to establish and maintain contact with the patient,
  - the test is concise and can be instantly interpreted,
  - it is more anonymous and acceptable by patients,
  - it is possible to be applied to patients with vision problems and movement difficulties (persons who, practically, would not have a chance to undergo psychological tests),
  - it is possible to apply the test for epidemiological studies,
  - it is easy to electronically register results.
- Disadvantages of the test:
  - it is possible that results will be falsified (calendar on the wall, handy notes, helpful spouse),
  - hearing problems may result in incorrect understanding of the questions asked,
  - the person conducting the test cannot see the patient's behaviour,
  - absence of follow up assessment (most of the tested people were white, well educated – one cannot refer theses results to a more heterogeneous population).

5 Discussion

A weakness of the analysis performed is a small group of subjects, not taking into account the Mungas [11] adjustment in the assessment of compliance with norm (reference to the number of education years and age), elimination of moderate and severe dementia in the assessment. The focus was on identification of commonly available and accepted diagnostic algorithms for the group of people in early stages of senility, who less frequently experienced education deficits, which would influence the understanding of tests performed. The confirmation of results should consist of extending them by observation of a larger group of people tested. A confirmation of the usefulness of applying TICS as a method for tele-diagnostics of disturbances of cognitive functions is the fact that population studies have been performed using this method in such well developed countries as USA, Canada, Australia, Germany, or the UK, for long-term population studies [14,15]. The conclusions from the studies conducted so far, which compare the reliability of tests using TICS vs. classical methods, such as MMSE or neuropsychological assessment, confirm the correctness of that method, and indicate its cost and time efficiency in performing diagnostics of dementia, as well as enable testing patients who – due to locomotor system disabilities and other limitations – would not be able to get to a specialized centre for a standard memory test. Widespread tele-diagnostics of early phases of dementia and other functional deficits of the brain, connected with its aging, stands a chance of becoming a standard in the assessment of health condition of the elderly population, and of contributing to quicker provision of medical and community care for them. A barrier for contemporary senior citizens may be the fact that memory assessment by phone deprives them of direct eye contact with the researched,

and having to use the phone during the test may arise certain reluctance. However, in the age of computers and digital world that problem will lose importance every year, and in a couple of dozen years will disappear completely. The finding of the study, that no correlation exists between the tests carried out in the study group and the age of patients, education history in years, presence of diabetes, depression, and Parkinson's disease, as well as leucocytosis and CRP results, may result from the small size of the study group, as well as the incidence of somatic diseases, which are present in the group of patients admitted for treatment in a geriatric ward. The results justify the expediency of further studies on the usefulness and suitability of tele-diagnostic tools for the assessment of cognitive deficits, in the groups of early as well as advanced senility in people. It is also important in the aspect of increasing role of early multimodal therapy, joined with pharmacological as well as education and behavioural interventions, as efficient way of postponing the progress of dementia process, in its various stages [15,16]. As concerns the assessment of learning ability, TICS is a more difficult test than MMSE and MoCA, due to the fact that repetition and memorizing of 10 words is required, whereas in MMSE - of 3 words, and in MoCA – of 5 words. The suggestion seems expedient to introduce the possibility of repeating the 10 words by the researcher a few times, which gives the subject a better chance to repeat them and better investigates the learning process [10].

6 Conclusions

1. The application of TICS test allows to exclude, with high level of probability, the existence of disturbances of cognitive functions. It is a promising and inexpensive strategy of population diagnostics in clinical and epidemiological studies.
2. To confirm the correct level of cognitive functions, on the basis of result of the test conducted in accordance with the suggested protocol, the diagnostic value is over 30 points.
3. For suspected presence of mild cognitive impairment (MCI) and mild dementia, on the basis of the suggested protocol, the diagnostic reference value for both conditions is - in TICS test - the score between 26 and 29 points for MCI, and below 25 points for mild dementia, respectively.
4. The TICS test may be helpful as a common diagnostic tool, for differentiation between cognitive norm on the one hand, and MCI mild dementia syndrome on the other hand.
5. A single screening examination with the use of TICS cannot be relied on for diagnosing MCI and dementia syndrome, as those conditions require additional and follow-up examinations, repeated over time.
6. A need exists for further search for specific and sensitive tools for assessing attenuation of functions and cognitive deficits in elderly people, at various stages of ageing.
7. The results obtained, pertaining to TICS test, confirm the expediency of further studies on its usefulness on a larger group of patients, which are planned in repeated sessions, over a longer time.

References

1. Brandt, J., Spencer, M., Folstein, M.: The telephone interview for cognitive status. Neuropsychiatry: Neuropsychol. Beh. Neurol. 1, 111–117 (1988)
2. Welsh, K.A., Breitner, J.C., Magruder-Habib, K.M.: Detection of dementia in the elderly using telephone screening of cognitive status. Neuropsychiatry Neuropsychol. Behav. Neurol. 6, 103–110 (1993)
3. Hogervorst, E., Bandelow, S., Hart Jr., J., Henderson, V.W.: Telephone word-list recall tested in the Rural Aging and Memory Study: two parallel versions for the TICS-M. Int. J. Geriatr. Psychiatry 19, 875–880 (2004)
4. Baker, A.T., Byles, J.E., Loxton, D.J., McLaughlin, D., Graves, A., Dobson, A.: Utility and Acceptability of the Modified Telephone Interview for Cognitive Status in a Longitudinal Study of Australian Women Aged 85 to 90. JAGS 15, 1217–1220 (2013)
5. Folstein, M.F., Folstein, S.E., McHugh, P.R.: Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. J. Psychiatr. Res. 12, 189–198 (1975)
6. Nasreddine, Z.S., Phillips, N.A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L., Chertkow, H.: The Montreal Cognitive Assessment (MoCA): A Brief Screening Tool For Mild Cognitive Impairment. Journal of the Amer. Ger. Soc. 53, 695–699 (2005)
7. Yesavage, J.A., Brink, T.L., Rose, T.L., Lum, O., Huang, V., Adey, M.B., Leirer, V.O.: Development and validation of a geriatric depression screening scale: A preliminary report. J. of Psych. Res. 17, 37–49 (1983)
8. Diagnostyka i leczenie otępień. Rekomendacje zespołu ekspertów Polskiego Towarzystwa Alzheimerowskiego. Medisfera, Otwock (2012)
9. MoCA test instructions, `http://www.mocatest.org/pdf_files/instructions/MoCA-Instructions-Polish.pdf`
10. Knopmann, D., Roberts, R.O., Geda, Y.E., Pankratz, V.S., Christianson, T.J.H., Petersen, R.C., Rocca, W.A.: Validation of the Telephone Interview for Cognitive Status-modified in Subjects with Normal Cognition, Mild Cognitive Impairment, or Dementia. Neuroepidemiology 34, 34–42 (2010)
11. Gallo, J.J., Breitner, J.C.: Alzheimer's disease in the NAS-NRC Registry of aging twin veterans.4. Performance characteristics of a twostage telephone screening procedure for Alzheimer's dementia. Psychol. Med. 25, 1211–1219 (1995)
12. Klich- Rączka, A., Siuda, J., Piotrowicz, K., Boczarska-Jedynak, M., Skalska, A., Krzystanek, E., Wizner, B., Świat, M., Skrzypek, M., Opala, G., Grodzicki, T.: Aspekty medyczne,psychologiczne, socjologiczne i ekonomiczne starzenia się ludzi w Polsce. In: Mossakowska, M., Więcek, A., Błędowski, P. (eds.) Termedia 2012, Poznań, pp. 109–121 (2012)
13. Zweig, X.H., Campbell, G.: Receiver - operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin. Chem. 39, 561–577 (1993)
14. Roccaforte, W.H., Burke, W.J., Bayer, B.L., Wengel, S.P.: Validation of a telephone version of the mini-mental state examination. J. Am. Geriatr. Soc. 40, 697–702 (1992)
15. Ruthirakuhan, M., Luedke, A.C., Tam, A., Goel, A., Kurij, A., Garcia, A.: Use of physical and intellectual activities and socialization in the management of cognitive decline of aging and dementia: a review. J. Aging Res. (2012), `http://www.hindawi.com/journals/jar/2012/384875/`
16. Jager, C.A., Budge, M.M., Clarke, R.: Utility of TICS-M for the assessment of cognitive function in older adults. Int. J. Geriatr. Psychiatry 18, 318–324 (2003)

Activity Monitoring of the Elderly for Telecare Systems – Review

Andrzej W. Mitas, Marcin Rudzki, Maria Skotnicka, and Paula Lubina

Silesian University of Technology, Faculty of Biomedical Engineering,
Zabrze, Poland
{andrzej.mitas,marcin.rudzki,maria.j.skotnicka}@polsl.pl,
paula.lubina@gmail.com

Abstract. Aging society in developed countries forms new challenges for medical care. Ensuring well-being of the elderly and provide them support when required is of special concern. Nowadays, solutions in form of telecare centers and monitoring systems are able to estimate condition of the patient and provide necessary help accordingly. This paper addresses the problem of reviewing current state of the art by describing systems available commercially and presented in scientific papers as well as inference methods for condition assessment. Conclusions and problems to be taken into account while designing such systems are also pointed out.

Keywords: assisted living, telecare, activity monitoring, sensor networks, review.

1 Introduction

Due to aging of the society in developed countries [31,32] health care faces the problem of ensuring safety and well-being of the elderly. Main aspect to be dealt with is to keep the elderly in good physical and mental condition as well as provide quick and professional support in case of emergency. Most of the elderly that are still self-reliant do not require continuous support provided for example by Assisted Living Facilities (ALF) and prefer to live in their own home environment. However, a fast help may always be required. From the literature it stems that there are many so called Ambient Assisted Living (AAL) systems being developed in order to address such issues. Some of them are reported as research projects, others are available on the market. The goal of this paper is to review these systems, compare their architecture and reported capabilities.

The paper is organized as follows: next section focuses on systems presented in the literature, as well as those available commercially, that are being developed for assisting the elderly in their daily life. Common architecture and methods used for activity detection and recognition are described in subsequent sections. The paper ends with a summary in the last section.

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_12, © Springer International Publishing Switzerland 2014

2 Services and Sensor Types

Nowadays AAL systems utilize easily accessible communication networks (Internet, GSM, GPRS) to form telecare centers providing support for many people at the same time [15,17,26] or are specially developed for hospitals or care units [11,29] to simplify monitoring of patients and allowing quicker response to an emergency. Services provided by telecare systems are diverse and can be broadly categorized into three groups:

- emergency calls,
- assessment of patient's condition,
- patient monitoring and providing safety for the entire household.

2.1 Emergency Calls

Those systems include a button-equipped device worn by a user, and a central station that asserts a warning signal to appropriate medical centers. Fast response can decrease health damage (i.e. caused by strokes) and even save ones life. The device in the simplest form consists of a device worn on the neck or wrist and a central station (Alert1 [3]). The wearable device can be additionally equipped in automatic fall detector and alarming mode (Auto Fall Guard Wireless Fall Detector [44]). More sophisticated systems can additionally sense if the patient left certain terrain and also constantly analyze all behaviors and inform about any anomalies, i.e. getting up in the night (Vivago CARE 8005 [55]).

This group is useful for elderly who stay home alone. Elderly person can quickly call for help if necessary. In case of emergency caused by fall and inability to personally call for help, a fall detector can automatically perform such action. Localization of the patient within a building or open space is then of vital importance as it is required to provide support [29,45]. Open space localization is done using GPS modules [41,14] while in-building location employs various techniques: using movement detectors [11,26], or radio-frequency/ultrasound [29]. A possibility of direct voice communication gives additional services to be provided like scheduling medical visits and planning free time [15].

2.2 Condition Assessment

This group can be considered the largest one as condition assessment can be performed by simple periodic questionnaires sent by mobile devices [6] or visual monitoring also capable of detecting falls and prolonged inactivity [37,59,60]. A lot of patient's dysfunctions can be detected while observing his physical activity. By analyzing the daily routine patient's condition can be reported.

First approach is to recognize activities by sensors placed within the environment (apartment or whole building) [8,24,50]. Falls, presence in forbidden zone, prolonged activity duration (e.g. lying in bed), unexpected presence and lack of activity in parts of a building (i.e. corridors) can be detected. In case of home installed systems information about typical and repeatable activities of daily

living (ADLs) like sleeping, preparing food, watching TV, toileting, etc. can be derived and well-being can be estimated. Recognized activity patterns are used to derive a model the current behavior is being compared to. Deviations from the model indicate change of patient's condition. Such approaches also provide alerts and localization data especially after detecting abnormal behavior or an emergency situation [29].

Such systems commonly use easily mountable environment sensors like PIR movement detectors [11,58], cameras [15,37,47,45], radio wave or ultrasound beacons [29], pressure (weight) [51], power consumption [50,58], door opening indicators [8,24].

These systems feature two main drawbacks. First, it is assumed that the person being monitored lives solitary (as person identification without wearable tag is limited). Then, sensors in all rooms are required. Such sensors also cannot provide detailed information about physical condition i.e. movement abnormalities, as they infer behavior in terms of generic ADLs.

In the second approach, physical activity monitoring is considered to be informative enough to estimate patient's condition [13,15,17]. Parameters like quantity of steps, traveled distance and motion speed allow ones physical condition to be assessed. There are devices used to monitor athletes (to improve their scores) as well as the elderly (to monitor their condition). The main component of such mobile devices is usually an inertial sensor. Gathered data is analyzed regarding patient's position (horizontal or vertical), activity type, energy expenditure, quantity of steps and motion speed. From sensor data activity patterns like walking, sitting, lying etc. are obtained and used for condition estimation as well as detection of falls is performed [13,47]. In case of emergency the device is capable of alerting the telecare center and provide voice communication [11,17]. Analysis of the location, traveling speed or elapsed time from leaving home allows condition decline (mobiCare [34]) to be detected. Frequently used sensors in the mobile device include: accelerometers, gyroscopes, temperature, pressure (altitude), GPS, "panic" button.

The devices can be worn on a belt (Activity Monitor [49]), on the thigh (ActivPal [42]), ankle (StepWatch [39]) or calf (OssurPAM [40]). ActiGraph device can be worn also on the wrist, thigh, knee or waist [1]. There is also a device with several sensors located on dedicated body parts: calf, wrist, feet, sternum (IDEEA [33]). Those devices are small and simple in use what is important for the users' comfort. Specially built sensors are relatively small, easily to be installed and more and more devices are being produced specially for this purpose. Inertial sensors are also used in motion capture systems that allow precise analysis of human gait [57]. Detection of abnormalities in walking patterns can be considered a premise that the person is more prone to falls and may require assistance thus a pro-active approach to fall risk estimation can be achieved [18,46].

Monitoring systems can also be accommodated to certain other tasks, like child care, diabetics, preventive ECG monitoring, or remote diagnosis. Dependent on patient's mobility mobile or stationary devices can be applied. Stationary

ones can process, analyze and transfer the data to a medical center where an action can be undertaken. Those devices are connected with a series of sensor chosen that fulfill certain tasks (type of telecare), i.e. ECG, sphygmomanometer, peak flow meter, pulse oximeter, spirometer, tonometer, weighing scale, thermometer, or glucometer.

Mobile devices are limited to ECG and respiratory movements detector. Analysis can be done in real time and any abnormalities of heartbeat can be compared to the patient's diary to determine the context of the event (i.e. increased frequency of heart beat while stressful meeting). Mobile version is also equipped with GPS, for localization of the patient in case of emergency (e-Care Platform [14]).

Systems performing additionally periodical measurements of cardiological signals like ECG, pulse or SpO2 are also available [12,27,53,47]. Detection of abnormalities and diagnosis is then provided.

2.3 Complex Systems

The final group consists of advanced monitoring systems, which enable to monitor patient's condition, improve life quality as well as improve household safety. Functionality of systems from previous groups is extended by functions which indicate when the medication should be taken, detectors of smoke, heat, motion, gas, humidity and also door sensor. Those functionalities bring safety not only for the human, but also for place of living [26,52]. Amber system, besides all of mentioned features can be connected with meters of blood pressure, pulse, sugar and cholesterol level, which enable to monitor the condition not only on the basis of activity analysis, but also by using biomedical measurements [54].

Simultaneous analysis of patient's activity and biomedical data allows comprehensive assessment of the elderly health status. Gas, smoke and humidity detectors can additionally increase the security of the entire home. Commercially available systems feature different capabilities. Some of these can precisely recognize the type of activity, but do not support biomedical measurements and interaction with medical centers. Other systems allow medical care, but limit the activity analysis just to determine whether the patient moves.

Services provided by AAL systems are summarized in Table 1 while sensors used are listed in Table 2.

3 Architecture

Common architecture consisting of four parts can be found in most AAL systems:

1. sensor network,
2. data acquisition and processing unit(s),
3. communication network gateway (Internet, GSM, GPRS),
4. telecare center.

Table 1. Functions provided by monitoring systems

activity recognition	[1,33,39,40,42,49,29,45,17,15,13,10,60,55,34]
alerts on demand	[3,34,52,54,55,17,44]
behavioral analysis	[14,50,26,8,11,58,24]
biomedical measurements	[14,34,12,51,47,27,53]
fall detection	[3,34,52,54,55,17,15,11,47,59,44]
home security	[34,52,54]
localization	[29,45,11,47,41,55,34,14]
voice communication	[17,26,15,3,44,34,54]

Table 2. Sensors used in monitoring devices

altitude	[13]
cameras	[45,15,47,37,60,59]
cardiac	[45,12,51,51,47,27,53]
GPS	[17,15,41,34,14]
household usage	[50,8,58,24,54]
inertial (acc., gyro.)	[29,45,17,2,15,13,11,49,42,1,39,40,33,54]
light	[15]
movement detectors	[26,8,11,47,58,24,54]
"panic" button	[29,50,26,15,3,44,52,34,54]
pressure/ weight	[8,51,24]
temperature	[45,13,53]

The task of the first part, the sensor network, is to acquire various signals from persons being monitored and/or their environment. A wide variety of sensors includes: IR movement detectors, cameras, temperature, pressure/weight, household appliances power consumption, inertial (accelerometers, gyroscopes), ECG/pulse, GPS, panic button and others. Signals ought to provide information about physical condition of the person - if and how he/she does household chores and if is physically active. Detection of falls and long inactivity is of special importance as rapid help from healthcare units may save life. With the availability of wireless communication modules (like BlueTooth, ZigBee) it is considered a must that the sensors form a wireless network. This requirement is valid not only for body worn sensors (as cable connections are troublesome and impair movements) [2,17] but also for sensor networks installed in the building simplifying their deployment cite1-s2.0-S0140366407005609.

Second part, the data acquisition and processing module, is responsible for acquisition and data processing in order to estimate physical condition from the raw sensor data. Algorithms and methods used depend heavily on sensor types, however commonly used approach is a comparison between actual behavior and model derived from previous measurements [35,50] and detection of specific dangerous situations [45]. Depending on chosen algorithms data processing may be as simple generating alerts when thresholds are exceeded [45]. As frameworks for distributed computing becomes available for sensor networks,

multi-agent approaches are also implemented [2]. Overview of the methods used are discussed in Section 4.

Third part is responsible for forwarding acquired or processed data and generating alerts that are sent either to the telecare center or person's relatives. It is usually implemented by a computer with Internet access, an application specific set top box or, in case of mobile-only approaches, using GSM/GPRS network [17]. Data to be transmitted has to be protected from unauthorized access [45] even at the telecare center, where localization of the patient is shown only after detecting an emergency situation [47].

Fourth part, the telecare center, gathers data from many persons and provides constant availability of support. IT systems do long term data analysis (statistics, data mining) and estimation of patient's condition. Staff is responsible for reviewing patient's data (especially after an alert) and responding adequately. Many systems also provide possibility of direct voice communication between subjects and staff [17,26].

4 Activity Recognition Methods

Activity detection, classification and pattern recognition in detecting abnormalities in movement patterns provides information about physical and/or psychological condition of the monitored person. Several activity recognition methods have been introduced in the last twenty years. One of the most popular tools to obtain the information about user's position, posture and movement is the accelerometer. Unlike the pedometer it allows not only the steps to be counted, but also measures the intensity of performed activity. Devices using more than one accelerometer have been shown to be also more precise than pedometers at slow walking speeds [20].

The development of the microelectromechanical systems (MEMS) technology have made it possible to reduce the size and the costs of accelerometers [30]. The MEMS technology encouraged also numerous attempts to recognize more specific activities such as Kung Fu moves, keyboard typing, writing on a whiteboard, shaking hands, looking at the watch, workshop activities including sawing, hammering, drilling and filing, gestures for American Sign Language and even gestures used to instruct pilots after landing [22]. One of the most successful researchers in this field - Ling Bao focused on more casual movements, such as vacuuming, folding laundry, watching TV or brushing teeth [5]. As can be seen, accelerometers have a wide range of application possibilities. In the monitoring of the elderly the most important activities to be recognized are sitting, standing, walking, ascending stairs, descending stairs and the transitions between standing and sitting. The time needed to perform those movements can be also used to predict the risk of falling [21].

Sensors such as accelerometers provide only raw data that has to be processed and analyzed. There are several approaches to activity recognition, most of them are using one of the following models or classifiers:

- decision trees,
- hidden Markov models,
- artificial neural networks,
- Bayesian classifiers,
- wavelet transform,
- thresholds.

4.1 Decision Trees

Decision trees are hierarchical binary structures where broad classifications are made in the top levels of the tree and more detailed subclassifications are made at lower levels. They allow to divide the classifications into modules and to test the algorithms independently. Decision trees were used by Bao and Intille [5], Karantonis et al. [23] and Bonomi et al. [7].

The laboratory-based trial involving the presented tree has been undertaken on six subjects using only one triaxial accelerometer [23]. Results indicate an overall accuracy of 90.8% across a series of 12 tasks (283 tests) involving a variety of movements related to normal daily activities. Distinction between activity and rest has been performed without error.

Decision trees are one of the simplest methods to obtain satisfying results even in real-time classifiers. They are slow to train but quick to run. The vast majority of signal processing can be performed on board the wearable unit using embedded intelligence. In the work of Bao et al. decision trees classifiers show the best performance recognizing twenty everyday activities with an overall accuracy rate of over 84% using 5 accelerometers [5]. Bonomi et al. have used one accelerometer to identify types of physical activity based on daily energy expenditure. The standard daily metabolic value (MET) has been obtained from a compendium of physical activity. The features selected for the classification are: the standard deviation of the acceleration in the vertical and medio-lateral directions of the body, the average acceleration in the vertical direction of the body, the peak-to-peak distance of the acceleration measured in the medio-lateral and antero-posterior directions of the body, and the frequency peak of the power spectral density of the acceleration measured in the vertical direction of the body. The activities including lying, sitting or standing, active standing, walking, running and cycling have been identified with overall accuracy of more than 93% (only active standing recognition caused some problems and resulted in the accuracy of 81%).

4.2 Hidden Markov Models

Another promising approaches are based on hidden Markov models (HMM), a great statistical tool which allows representing probability distributions over

sequences of observations. HMMs are used for arm gesture spotting module, which is a part of a big system designed to monitor the overall physical activity level of a person [22].

Wang et al. [56] has used a hidden Markov model to recognize six human daily activities from sensor signals collected from a single waist-worn tri-axial accelerometer. All training signals from the same activity class are modeled as generated by a HMM, while a Gaussian Mixture Model (GMM) is used to model the continuous observation for each hidden state. The tested signal is classified to the activity class corresponding to the HMM with the highest likelihood. The protocol includes six daily activities: walking, standing, running, jumping, sitting-down, and falling-down. The accuracy of 94.8% has been achieved. Not only the satisfying results but also the robustness and computational simplicity of this method make it worth of interest.

4.3 Artificial Neural Networks

Artificial neural networks are the basis of many activity recognition systems. They are modeled to resemble the human neural network and solve problems, with the purpose of self-directed information processing. A hypersphere clustering with an incremental neural network has been used by Fourty et al. [17].

This ANN is part of a big embedded system called HEBE [17]. It has several advantages:

- it can be easily implemented on a microcontroller,
- it is an incremental neural network, a new configuration can be learned without the need of learning process using the whole set of examples,
- it avoids the consideration of statistics, which provides a learning phase less sensitive to the learning set, and the rare events can be well identified
- unexplored spaces provide unknown responses and avoid misclassification.

A test on 5 subjects in a laboratory during 8h has shown a 20% rate of misclassification due to differences between subjects. A personalized learning process has been suggested for improvement.

Pärkä et al. [43] have automatically classified everyday activities like walking, running, and cycling. Sixteen persons have taken part in the experiment acquiring 31-hour data. Three classifiers including custom decision tree, automatically generated decision tree and ANN have been compared yelding the accuracy of 82%, 86% and 82%, respectively. The authors suggest that an ANN might be too prone to the noise of everyday data. The experiments conducted by Studenmayer [48] and Khan [25] in laboratory conditions using ANN, resulted in 88.8% and 97.65% of recognition accuracy, respectively.

4.4 Bayesian Classifier

Naïve Bayesian classifier is based on Bayes theorem. It relies on a probabilistic model of inertial data features, classes, and their relation. It is easy to build with

simple iterative parameters and can be used for very large datasets [4,5,16,36]. Muscillo [36] has classifyied different locomotor activities during real life condition, such as walking at different speeds, and with different slopes, stair descending and ascending. His naïve 2D-Bayes classifier has been updated online through the history of the estimated activities, in a Kalman-based scheme achieved correct classification rates higher than 90% for young and higher than 92% for the elderly. Atallah [4] compares a K-nearest neighbour classifier with different values of k to assess the effect of outlier points and a Bayesian classifier, where Gaussian distributions are used to model the priors of classes and the posterior probability. His goal was to find the best position for a sensor for different types of activities. The results for both methods were similarly high, but in some cases the Bayesian classifier turned out to be even more precise. Dobkin [16] uses a naïve Bayesian classifier to recognize walking, exercise, and cycling and to measure the speed of walking after stroke. The Bayesian classifiers give satisfying results in real life environment.

4.5 Wavelet Transform

Wavelet transforms are one of the most popular time-frequency-transforms. They have been often used in activity recognition based on gait measurements. Nyan [38] has described the classification of gait patterns among descending stairs, ascending stairs and level walking activities using a single accelerometer. The classification rates have been 97.72%, 93.18% and 93.93%, respectively. Godfrey [19] has also reduced the number of sensors to one and has classified activities such as sitting, standing, walking, and numerous postural transitions performed under supervised conditions by young and elderly subjects. He used a scalar product technique and vertical velocity estimates on a single tri-axial accelerometer and then compared it to a proven discrete wavelet transform method that incorporated accelerometers and gyroscopes. The activities were successfully detected with a sensitivity of 83% and specificity of 89% for elderly healthy subjects in home environment using the simplified algorithm. A complex wavelet approximation algorithm with only one sensor results in power consumption of 4 mA, which limits the recording period to 12h. These findings show that the wavelet transform might be very accurate but not suitable for long-term monitoring.

4.6 Thresholds

The use of thresholds is one of the simplest approaches. The simplicity is a very important feature in real-time monitoring systems. It allows the results to be obtained quickly without high battery consumption. Lyons [28] has introduced an accelerometer-based system that distinguishes between static and dynamic activities and detects the basic postures of sitting, standing and lying. He has used both, mid-point and "best suitable value". The first threshold results in accuracy of only 75%. Yet the application of the "best estimate" thresholding has improved the classification rate to 93% in home environment. Previous findings have shown that classificatin rate of 98% for the mid-point posture threshold

approach, but the data acquisition was based on protocols in laboratory environments. Bussmann [9] has introduced the Activity Monitor (AM) using a pre-set fixed threshold method for distinguishing postures and activities. While Bussmann needed five accelerometers to obtain the accuracy of more than 80%, three years later Lyons used only two sensors (one on the trunk and one on the thigh). Not only was the accuracy improved but also the comfort of the user, that should not be forgotten while designing a telecare system.

5 Summary

Various attempts are being made to develop AAL systems. Their architecture is usually similar while sensors used differ. For environment sensors movement detectors are widely used, mainly due to their availability in commercial security systems. Person's privacy is kept in a higher degree than with the use of cameras, yet the person recognition is limited. Forming and using a behavioral model from uncertain data (when several people may be present in an apartment) leads to improper reasoning. To overcome this issue person identification has to be performed using wearable tag. However, while developing a tag device, one can also equip it with "panic" button, sensors, GSM and GPS modules to develop an all-in-one device. Assuming that a wearable device is used daily by the same person forming and using behavioral models to detect condition changes is possible. Such a single device to be useful has to be properly designed to be ergonomic for the elderly. Size, shape, color and especially interactive elements (buttons, microphone, etc.) have to be optimized.

Acknowledgement. This work has been supported by the European Union under the 7th Frame Programme grant nr POIG.01.03.01-24-061/12.

References

1. ActiGraph: ActiGraph, `http://www.actigraphcorp.com/` (accessed January 11, 2014)
2. Aiello, F., Bellifemine, F., Fortino, G., Galzarano, S., Gravina, R.: An agent-based signal processing in-node environment for real-time human activity monitoring based on wireless body sensor networks. Engineering Applications of Artificial Intelligence 24(7), 1147–1161 (2011)
3. Alert1: Medical Alert System, `http://www.alert-1.com/` (accessed January 11, 2014)
4. Atallah, L., Lo, B., King, R., Yang, G.Z.: Sensor placement for activity detection using wearable accelerometers. In: 2010 International Conference on Body Sensor Networks (BSN), pp. 24–29 (2010)

5. Bao, L., Intille, S.S.: Activity Recognition from User-Annotated Acceleration Data. In: Ferscha, A., Mattern, F. (eds.) PERVASIVE 2004. LNCS, vol. 3001, pp. 1–17. Springer, Heidelberg (2004)
6. Bielli, E., Carminati, F., La Capra, S., Lina, M., Brunelli, C., Tamburini, M.: A wireless health outcomes monitoring system (whoms): development and field testing with cancer patients using mobile phones. BMC Medical Informatics and Decision Making 4(1), 1–13 (2004)
7. Bonomi, A., Plasqui, G., Goris, A., Westerterp, K.: Improving assessment of daily energy expenditure by identifying types of physical activity with a single accelerometer. Applied Physiology 107, 655–661 (2009)
8. Botia, J.A., Villa, A., Palma, J.: Ambient assisted living system for in-home monitoring of healthy independent elders. Expert Systems with Applications 39(9), 8136–8148 (2012)
9. Bussmann, J., Martens, W., Tulen, J., Schasfoort, F., van den Berg-Emons, H., Stam, H.: Measuring daily behavior using ambulatory accelerometry: The activity monitor. Behavior Research Methods 33, 349–356 (2001)
10. Chakraborty, S., Ghosh, S.K., Jamthe, A., Agrawal, D.P.: Detecting mobility for monitoring patients with parkinson's disease at home using {RSSI} in a wireless sensor network. Procedia Computer Science 19, 956–961 (2013)
11. Charlon, Y., Fourty, N., Bourennane, W., Campo, E.: Design and evaluation of a device worn for fall detection and localization: Application for the continuous monitoring of risks incurred by dependents in an Alzheimer's care unit. Expert Systems with Applications 40(18), 7316–7330 (2013)
12. Chen, C.M.: Web-based remote human pulse monitoring system with intelligent data analysis for home health care. Expert Systems with Applications 38(3), 2011–2019 (2011)
13. Chernbumroong, S., Cang, S., Atkins, A., Yu, H.: Elderly activities recognition and classification for applications in assisted living. Expert Systems with Applications 40(5), 1662–1674 (2013)
14. COMARCH: Platforma e-Care, `http://www.comarch.pl/e-zdrowie/` (accessed January 11, 2014)
15. Costa, A., Castillo, J.C., Novais, P., Fernández-Caballero, A., Simoes, R.: Sensor-driven agenda for intelligent home care of the elderly. Expert Systems with Applications 39(15), 12192–12204 (2012)
16. Dobkin, B., Xu, X., Batalin, M., Thomas, S., Kaider, W.: Reliability and validity of bilateral ankle accelerometer algorithms for activity recognition and walking speed after stroke. Stroke: Journal of the American Heart Association (2011)
17. Fourty, N., Guiraud, D., Fraisse, P., Perolle, G., Etxeberria, I., Val, T.: Embedded system used for classifying motor activities of elderly and disabled people. Computers & Industrial Engineering 57(1), 419–432 (2009)
18. Giansanti, D., Maccioni, G., Cesinaro, S., Benvenuti, F., Macellari, V.: Assessment of fall-risk by means of a neural network based on parameters assessed by a wearable device during posturography. Medical Engineering & Physics 30(3), 367–372 (2008)
19. Godfrey, A., Bourke, A., Ólaighin, G., van de Ven, P., Nelson, J.: Activity classification using a single chest mounted tri-axial accelerometer. Medical Engineering and Physics 33, 1127–1135 (2011)
20. Grant, P., Dall, P., Mitchell, S., Granat, M.: Activity-monitor accuracy in measuring step number and cadence in community-dwelling older adults. Journal of Aging and Physical Activity 16, 201–214 (2008)

21. Higashi, Y., Yamakoshi, K., Fujimoto, T., Sekine, M., Tamura, T.: Quantitive evaluation of movement using the timed up-and-go test. IEEE Engineering in Medicine and Biology Magazine (2008)
22. Junker, H., Amft, O., Lukowicz, P., Tröster, G.: Gesture spotting with body-worn inertial sensors to detect user activities. Pattern Recognition 41, 2010–2024 (2008)
23. Karantonis, D., Narayanan, M., Mathie, M., Lovell, N., Celler, B.: Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. IEEE Transactions on Information Technology in Biomedicine 10(1) (2006)
24. Kasteren, T., Englebienne, G., Kröse, B.: An activity monitoring system for elderly care using generative and discriminative models. Personal and Ubiquitous Computing 14(6), 489–498 (2010)
25. Khan, A., Lee, Y.K., Lee, S., Kim, T.S.: A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. IEEE Transactions on Information Technology in Biomedicine 14(5) (2010)
26. Kwon, O., Shim, J.M., Lim, G.: Single activity sensor-based ensemble analysis for health monitoring of solitary elderly people. Expert Systems with Applications 39(5), 5774–5783 (2012)
27. Lorenz, A., Oppermann, R.: Mobile health monitoring for the elderly: Designing for diversity. Pervasive and Mobile Computing 5(5), 478–495 (2009)
28. Lyons, G., Culhanea, K., Hilton, D., Gracec, P., Lyon, D.: A description of an accelerometer-based mobility monitoring technique. Medical Engineering and Physics 27, 497–504 (2005)
29. Marco, A., Casas, R., Falco, J., Gracia, H., Artigas, J., Roy, A.: Location-based services for elderly and disabled people. Computer Communications 31(6), 1055–1066 (2008)
30. Mathie, M., Coster, A., Lovell, N., Celler, B.: Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. Physiological Measurements (2004)
31. McDougall, R.: Ageing populations and hidden unemployment, http://www.populationmatters.org/issues-solutions/population/ageing/ (accessed January 11, 2014)
32. McDougall, R.: Too many people: Europe's population problem, http://www.populationmatters.org/issues-solutions/population/ (accessed January 11, 2014)
33. MiniSun: IDEEA - Physical Activity and Energy Expenditure Monitor, http://www.minisun.com/ (accessed January 11, 2014)
34. MobiCare SA: MobiCare, http://www.mobicare.com.pl/ (accessed January 11, 2014)
35. Mori, T., Takada, A., Noguchi, H., Harada, T., Sato, T.: Behavior prediction based on daily-life record database in distributed sensing space. In: 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005), pp. 1703–1709 (2005)
36. Muscillo, R., Schmid, M., Conforto, S., D'Alessio, T.: An adaptive kalman-based bayes estimation technique to classify locomotor activities in young and elderly adults through accelerometers. Medical Engineering and Physics 32, 849–859 (2010)
37. Nagai, N., Fujii, Y., Ueda, H., Maru, K., Kumakura, S.: Pc-based monitoring system for efficient use of day care centers for elderly people. Procedia - Social and Behavioral Sciences 2(1), 204–208 (2010)
38. Nyan, M., Tay, F., Seah, K., Sitoh, Y.: Classification of gait patterns in the time-frequency domain. Journal of Biomechanics 39 (2006)

39. Orthocare Innovations: StepWatch, http://orthocareinnovations.com/ (accessed January 11, 2014)
40. Ossur Life Without Limitations: Ossur PAM, http://ossur.com/ (accessed January 11, 2014)
41. Paiva, S., Abreu, C.: Low cost {GPS} tracking for the elderly and alzheimer patients. Procedia Technology 5, 793–802 (2012)
42. PAL Technologies Ltd: activPAL, http://www.paltechnologies.com/ (accessed January 11, 2014)
43. Pärkkä, J., Ermes, M., Korpipää, P., Mäntyjärvi, J., Peltola, J., Korhonen, I.: Activity classification using realistic data from wearable sensors. IEEE Transactions on Information Technology in Biomedicine 10(1) (January 2006)
44. Pioneer Emergency Medical Alarms: Auto Fall Guard Wireless Fall Detector, http://www.pioneeremergency.com/product/auto-fall-guard/ (accessed January 11, 2014)
45. Rajasekaran, M.P., Radhakrishnan, S., Subbaraj, P.: Sensor grid applications in patient monitoring. Future Generation Computer Systems 26(4), 569–575 (2010)
46. Riva, F., Toebes, M., Pijnappels, M., Stagni, R., van Dieën, J.: Estimating fall risk with inertial sensors using gait stability measures that do not require step detection. Gait & Posture 38(2), 170–174 (2013)
47. Rocha, A., Martins, A., Junior, J.C.F., Boulos, M.N.K., Vicente, M.E., Feld, R., van de Ven, P., Nelson, J., Bourke, A., ÓLaighin, G., Sdogati, C., Jobes, A., Narvaiza, L., Rodríguez-Molinero, A.: Innovations in health care services: The {CAALYX} system. International Journal of Medical Informatics 82(11), e307–e320 (2013)
48. Staudenmayer, J., Pober, D., Coruter, S., Bassett, D., Freedson, P.: An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. Journal of Applied Physiology 107, 1300–1307 (2009)
49. stayhealthy: Research Activity Monitor, http://www.stayhealthy.com/ (accessed January 11, 2014)
50. Suryadevara, N., Gaddam, A., Rayudu, R., Mukhopadhyay, S.: Wireless sensors network based safe home to care elderly people: Behaviour detection. Procedia Engineering 25, 96–99 (2011)
51. Tamura, T., Togawa, T., Ogawa, M., Yoda, M.: Fully automated health monitoring system in the home. Medical Engineering & Physics 20(8), 573–579 (1998)
52. Telecare Systems: Emergency Alarm Systems, http://www.telecaresystems.at/ (accessed January 11, 2014)
53. Vaidehi, V., Vardhini, M., Yogeshwaran, H., Inbasagar, G., Bhargavi, R., Hemalatha, C.: Agent based health monitoring of elderly people in indoor environments using wireless sensor networks. Procedia Computer Science 19, 64–71 (2013)
54. Visionic: Wireless Home Security, http://www.visonic.com/ (accessed January 11, 2014)
55. Vivago: Retirement Home Solutions, http://www.vivago.com/ (accessed January 11, 2014)
56. Wang, J., Chen, R., Sun, X., She, M., Wu, Y.: Recognizing human daily activities from accelerometer signal. Procedia Engineering 15, 1780–1786 (2011)
57. Xsens: MVN Awinda, http://www.xsens.com/en/mvn-awinda (accessed January 11, 2014)

58. Yang, C.C., Hsu, Y.L.: Remote monitoring and assessment of daily activities in the home environment. Journal of Clinical Gerontology and Geriatrics 3(3), 97–104 (2012)
59. Yu, M., Rhuma, A., Naqvi, S., Wang, L., Chambers, J.: A posture recognition-based fall detection system for monitoring an elderly person in a smart home environment. IEEE Transactions on Information Technology in Biomedicine 16(6), 1274–1286 (2012)
60. Zhou, Z., Dai, W., Eggert, J., Giger, J., Keller, J., Rantz, M., He, Z.: A real-time system for in-home activity monitoring of elders. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2009, pp. 6115–6118 (2009)

Device for Accelerometer and Gyroscope Measurements

Stefan Borik, Branko Babusiak, and Ivo Cap

University of Zilina, Faculty of Electrical Engineering,
Dept. of Electromagnetic and Biomedical Engineering, Zilina 010 26, Slovakia
stefan.borik@fel.uniza.sk
http://fel.uniza.sk/ktebi

Abstract. The article deals with the accelerometer and gyroscope measurements and gravity vector components compensation. A part of work involves also design and realization of device serving for acceleration and angle rotation measurement. Within the work algorithm for gravity compensation was tested.

Keywords: accelerometer, gravity, gyroscope, rotation, quaternions.

1 Introduction

Nowadays there are many solutions which serve for position detection or translation measurements. In these solutions various physical principles are used. For example we can mention of hand computer mouse which is well known pointing device. When we need to work in 2D it is without problems but working in 3D space with conventional computer mouse based on optical principles is not possible. For solving this challenge the inertial device consisting of accelerometer and gyroscope could be used. The data from accelerometer can provide information about position of object and the gyroscope data could be useful for estimation of rotation angle of the selected object. These measurements could be performed by using integrated sensors available on the current market [1,2,3,4,5].

2 Device for Acceleration and Rotation Angle Measurement

We used for our measurements LSM330DL device (Fig. 1). It is linear sensor module with 3D accelerometer sensor and 3D gyroscope sensor fusion. Its measurement principle is based on application of MEMS elements (Micro-Electro-Mechanical Systems). This device can be controlled by microcontroller through standard data protocols (I2C and SPI). We selected I2C communication protocol because we needed microcontroller SPI pins being free for its programming. We used microcontroller ATMEL - Atmega328P which is a low power 8-bit single chip device.

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_13, © Springer International Publishing Switzerland 2014

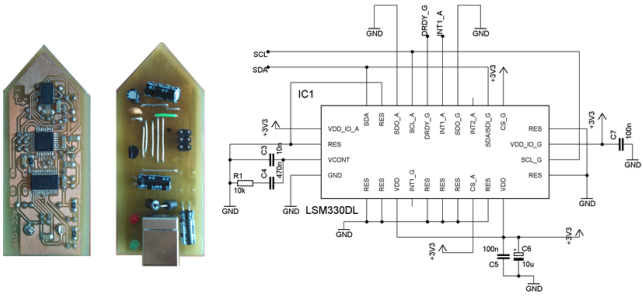


Fig. 1. Realized device and LSM330DL circuit

2.1 I2C Communication

ATMEL describes its I2C protocol such the 2-wire Serial Interface (TWI). TWI is ideally suited for typical microcontroller applications. The TWI protocol allows systems designer to interconnect up to 128 different devices using only two bi-directional bus lines, one for clock (SCL) and one for data (SDA). All devices connected to the bus have individual addresses, and mechanisms for resolving bus contention are inherent in the TWI protocol.

2.2 Data Acquisition

As described in the text above we used microcontroller for control of inertial device. It is used also for collecting the data from accelerometer and gyroscope. The acquired data are organized in frames as shown in the Fig. 2.

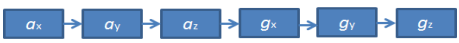


Fig. 2. Data frame

Then the data are sent via serial interface (using FT232RL which is serial to USB converter) to computer where they are further processed. For the better image how it works the program flowchart is shown on the Fig. 3.

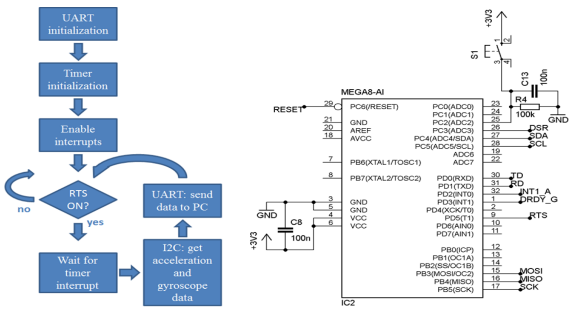


Fig. 3. MCU configuration

3 Accelerometer and Gyroscope Data Fusion

Positioning using inertial device such as accelerometer is based on the principle of acceleration integration. When we calculate the first integral of the accelerometer data we get velocity and after the second integration the result is actual position of an object which location we want to know. This basic principle can be described by the next equations:

$$v = \int a dt \tag{1}$$

$$s = \int v dt \tag{2}$$

Measurement errors influence the accuracy of integration results. For that reason we used high-pass filtering of the measured data for correction of these errors (see Section 4).

3.1 Earth Gravity Problem

When we want to calculate first and second integral from measured acceleration we encounter the Earth gravity problem. If the device is held in the standard position (see Fig. 4) gravity vector is placed antiparallel with z-axis.

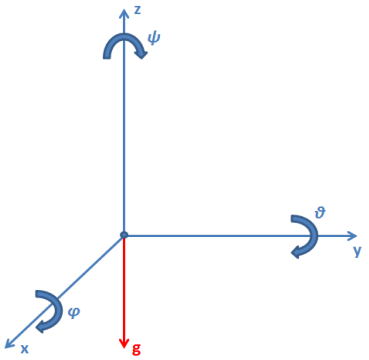


Fig. 4. Default position of device

We could simply subtract gravity vector [0 0 g] from the accelerometer data but we want to move and to rotate the device in any direction. The device rotation causes gravity acceleration decomposition to other axes and the gravity cannot be simple subtracted. It is necessary to predict how the gravity acceleration is redistributed to other axes. We used the algorithm described in the next section [2,3,4,5].

3.2 Quaternions

For simple gravity subtraction using gravity vector [0 0 g] we have to rotate the device to the default position (see Fig. 4.) We can do this operation when we know all rotation angles ϕ, θ and Ψ (see Fig. 4.) [1,2,3,4,5]. For estimation of these rotation angles the gyroscope was used. Single angles can be calculated by integration of gyroscope data (they are measured in [°/s]). For data acquisition we used sampling frequency 250 Hz and the gyroscope angular rate was set to 250 °/s which correspond to the angular sensitivity 8.75 mdps/digit. The rotation angle can be calculated:

$$\phi = \int_{t_1}^{t_2} \omega_\phi dt \tag{3}$$

where s_g is angular sensitivity, ω_ϕ is angular speed for rotation angle ϕ. The same equation holds generally also for angles θ and ψ.

By using equation (3) we know the actual device rotation angle in all the axes and we can rotate acceleration vector by using quaternions. The quaternion is defined as four dimensional vector [1,2,3,4,5]:

$$q = s + a_x i + a_y j + a_z k \tag{4}$$

where s, a_x, a_y, $a_z \in$ R, s denotes a 'degree' of rotation, a_x, a_y, a_z are acceleration vector components and i, j, k are the unit vectors of single axes of the given coordinate system. Rotation using quaternion can be expressed by following equation:

$$r = hq\bar{h} \tag{5}$$

where h is the unit quaternion with vector part characterizes axis around which we want to rotate vector expressed by quaternion q. The quaternion h and its conjugate $\bar{h}$ are described:

$$h = \left[cos\left(\frac{\phi}{2}\right) \quad sin\left(\frac{\phi}{2}\right) x_n \quad sin\left(\frac{\phi}{2}\right) y_n \quad sin\left(\frac{\phi}{2}\right) z_n \right] \tag{6}$$

$$\bar{h} = \left[cos\left(\frac{\phi}{2}\right) \quad -sin\left(\frac{\phi}{2}\right) x_n \quad -sin\left(\frac{\phi}{2}\right) y_n \quad -sin\left(\frac{\phi}{2}\right) z_n \right] \tag{7}$$

where ϕ is the rotation angle and x_n, y_n, z_n are rotation axis normalized components around which we rotate:

$$x_n = \frac{x}{\sqrt{x^2+y^2+z^2}}, \quad y_n = \frac{y}{\sqrt{x^2+y^2+z^2}}, \quad z_n = \frac{z}{\sqrt{x^2+y^2+z^2}} \tag{8}$$

Multiplication of two quaternions $q_1 = s_1 + x_1 i + y_1 j + z_1 k$ and $q_2 = s_2 + x_2 i + y_2 j + z_2 k$ can be calculated by using following matrix:

$$q_1 q_2 = \begin{bmatrix} s_1 s_2 - x_1 x_2 - y_1 y_2 - z_1 z_2 \\ s_1 x_2 + x_1 s_2 + y_1 z_2 - z_1 y_2 \\ s_1 y_2 - x_1 z_2 + y_1 s_2 + z_1 x_2 \\ s_1 z_2 + x_1 y_2 - y_1 x_2 + z_1 s_2 \end{bmatrix} \tag{9}$$

For vector rotation we can use also rotation matrix:

$$R = \begin{bmatrix} s^2 + x^2 - y^2 - z^2 & 2xy - 2sz & 2xz + 2sy \\ 2xy + 2sz & s^2 - x^2 + y^2 - z^2 & 2yz - 2sx \\ 2xz - 2sy & 2yz + 2sx & s^2 - x^2 - y^2 + z^2 \end{bmatrix} \tag{10}$$

The quaternion components s, x, y and z are calculated from angles ϕ, θ and ψ measured by gyroscope:

$$\begin{bmatrix} s \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} cos\left(\frac{\phi}{2}\right) cos\left(\frac{\theta}{2}\right) cos\left(\frac{\psi}{2}\right) + cos\left(\frac{\phi}{2}\right) cos\left(\frac{\theta}{2}\right) cos\left(\frac{\psi}{2}\right) \\ sin\left(\frac{\phi}{2}\right) cos\left(\frac{\theta}{2}\right) cos\left(\frac{\psi}{2}\right) - cos\left(\frac{\phi}{2}\right) sin\left(\frac{\theta}{2}\right) sin\left(\frac{\psi}{2}\right) \\ cos\left(\frac{\phi}{2}\right) sin\left(\frac{\theta}{2}\right) cos\left(\frac{\psi}{2}\right) + sin\left(\frac{\phi}{2}\right) cos\left(\frac{\theta}{2}\right) sin\left(\frac{\psi}{2}\right) \\ cos\left(\frac{\phi}{2}\right) cos\left(\frac{\theta}{2}\right) sin\left(\frac{\psi}{2}\right) - sin\left(\frac{\phi}{2}\right) sin\left(\frac{\theta}{2}\right) cos\left(\frac{\psi}{2}\right) \end{bmatrix} \tag{11}$$

The equation (10) serves for gravity removing by rotating actual vector to inertial (default) position when gravity vector is [0 0 g]:

$$a_i = Ra_m^T + [0\,0\,g]^T \tag{12}$$

where a_m is measured acceleration vector [1,2,3,4,5]. Gravity removing algorithm is described in the next block diagram:

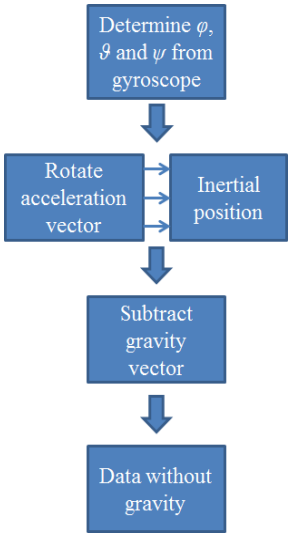
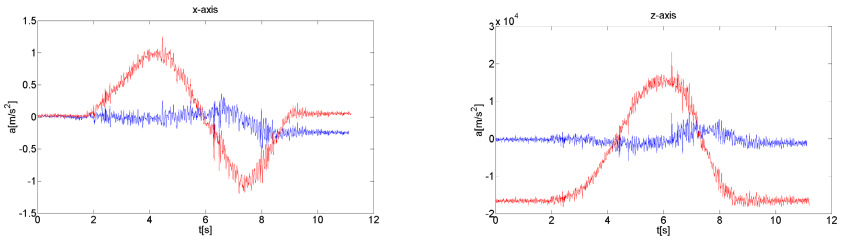


Fig. 5. Flowchart of gravity removing algorithm

4 Measurement Results

4.1 Gravity Components Removing

For designed algorithm testing we selected following rotation sequence. We rotated realized device (see Fig. 1) around its y-axis from 0 to 360 degrees. This movement causes gravity vector decomposition to x and z-axes. We can see measured signal in the Fig. 6(a) and Fig. 6(b).



(a) measured from x-axis (b) measured from z-axis

Fig. 6. Acceleration signal

Red curves characterize acceleration in x-axis (Fig. 6(a)) and z-axis (Fig. 6(b)). Blue curves are acceleration data after application algorithm for removing Earth gravity. We can see that blue curves contain only dynamic part of acceleration.

4.2 Position Measurement Testing

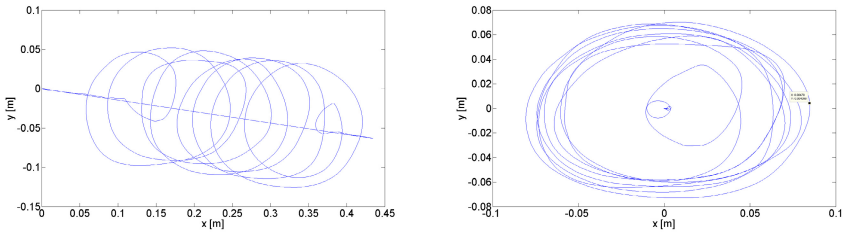
To test functionality of used algorithm we have selected circular trajectory with diameter d = 0.2 m in x-y plane. We tested the double integration of acceleration data with quaternion correction. The measurements were performed by using high-pass filter and without it. The cutoff frequency of used high-pass filter was 0.6 Hz. In this way removing of static gravity was achieved The experiment duration was about 10 s and the magnitude of acceleration was around 1 g. The next figures show measurement results:

It results from the Figures 7(a) and 7(b) that better measurement accuracy can be achieved by filtering accelerometer data. Measured diameter of circle was approximately 0.18 m.

4.3 Device Applications

Realized device can be used in many application areas such as listed in the LSM330DL data-sheet:

- GPS navigation systems
- Impact recognition and logging



(a) without application of high-pass filter (b) with application of high-pass filter

Fig. 7. Position measurement

- Gaming and virtual reality input devices
- Motion-activated functions
- Intelligent power saving for handheld devices
- Vibration monitoring and compensation
- Free-fall detection
- 6D-orientation detection

In biomedical engineering area we see appropriate application solution in analysis of spinal curvature because the spinal curvature is one of the most important parameters in quantitative anatomical spine evaluation [6]. The device could find its application also directly in hospitals for example in the intensive care unit departments where it could serve as the auxiliary device connected with recording system optimized for long term measurement of bed headrest elevation angle described in [7] - our solution would help to detect changes of headrest position as well as would be helpful at detection of artifacts.

5 Conclusion and Future Work

Designed device can be used for acceleration and rotation angle measurements. Used signal processing algorithm can be used for Earth gravity compensation in accelerometer and gyroscope measurements. This compensation is necessary for removing static gravity component from measured data because at integrating of accelerometer signal it will cause calculation errors and the position measurement it will be not possible. Our future work will be focused on data processing algorithm improvements and on design of algorithm for actual object position estimating and spinal curvature analysis.

References

1. Zhou, F., Wu, Y., Fan, N.: Improved accelerometer and magnetometer-based solution to solve quaternion. In: 2012 12th International Conference on Control, Automation and Systems (ICCAS), pp. 2093–2096. IEEE (2012)

2. Zhang, W., Ghogho, M., Yuan, B.: Mathematical model and matlab simulation of strapdown inertial navigation system. Modelling and Simulation in Engineering 2012, 1 (2012)
3. Baerveldt, A.J., Klang, R.: A low-cost and low-weight attitude estimation system for an autonomous helicopter. In: Proceedings of the IEEE International Conference on Intelligent Engineering Systems, INES 1997, pp. 391–395. IEEE (1997)
4. Tan, C.W., Park, S.: Design of accelerometer-based inertial navigation systems. IEEE Transactions on Instrumentation and Measurement 54(6), 2520–2530 (2005)
5. Ang, W.T., Khosla, P.K., Riviere, C.N.: Design of all-accelerometer inertial measurement unit for tremor sensing in hand-held microsurgical instrument. In: Proceedings of the IEEE International Conference on Robotics and Automation, ICRA 2003, vol. 2, pp. 1781–1786. IEEE (2003)
6. Barabas, J., Babusiak, B., Gala, M., Capka, M., Radil, R.: Computer-assisted analysis of spinal curvature parameters from ct images. In: 2012 35th International Conference on Telecommunications and Signal Processing (TSP), pp. 574–577 (2012)
7. Krefft, M., Zamaro-Michalska, A., Zabołotny, W.M., Zaworski, W., Grzanka, A., Łazowski, T., Tavola, M., Siewiera, J., Mikaszewska-Sokolewicz, M.: Head of the bed elevation angle recorder for intensive care unit. In: Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2013. International Society for Optics and Photonics, p. 89031A (2013)

Wearable System for Activity Monitoring of the Elderly

Andrzej W. Mitas, Marcin Rudzki, Wojciech Wieclawek, Piotr Zarychta, and Seweryn Piwowarski

Silesian University of Technology, Faculty of Biomedical Engineering, Zabrze, Poland
{andrzej.mitas,marcin.rudzki,wojciech.wieclawek, piotr.zarychta,seweryn.piwowarski}@polsl.pl

Abstract. The paper presents a concept of an Ambient Assisted Living system for daily monitoring of the elderly. Description and testing of the prototype of a mobile data acquisition device for physical activity detection using inertial sensors is made. Testing has been performed on a mechanical test stand and a group of healthy individuals during normal walking. The purpose has been to assess the repeatability of acquired data. Obtained results are satisfactory and further development takes place.

Keywords: assisted living, activity detection, inertial sensors.

1 Introduction

The contemporary world is growing mature. As World Health Organization data shows[1], life expectancy for children born in developed countries in 2011 countries reaches 75-90 years. With birth rate already lower than death rate in some regions[2] it is expected, that a society will soon face increasing number of age-related problems. One of them is how to effectively care, diagnose and treat elder people keeping them as active members of society as long as possible.

It is a complex problem and many different issues need to be considered before effective solutions will wildly be accepted:

- should medical procedures be identical for elder and young people?
- how to prevent age-related diseases?
- how to aid elder people in overcoming life-difficulties related to age?
- etc.

[1] Life expectancy at birth: female (http://gamapserver.who.int/mapLibrary/Files/Maps/Global_LifeExpectancy_females_2011.png and male http://gamapserver.who.int/mapLibrary/Files/Maps/Global_LifeExpectancy_males_2011.png, 2011.

[2] Crude birth and death rate: http://apps.who.int/gho/data/node.main.CBDR107?lang=en

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_14, © Springer International Publishing Switzerland 2014

Answers may be given by medical, sociological and engineering experts. The members of the former groups will likely focus more on health-related aspects and prevention, while latter provide various applications of modern technologies suited to elderly people's needs.

Personal monitoring as a tool created to both keep elderly people independent and detect medically relevant symptoms of incoming breakdowns fits all categories. It presents some weaknesses, though.

A shortcoming of personal electronic monitoring, with respect to the presence of a real person, expresses itself by limited nature of continuous information flow from the patient. Traditional, multi-sensoric contact with other person includes, among others, sight, hear, touch and extra-verbal communication. In an inheritable manner it allows fuzzy reasoning while determining interventions. System developed by engineers, by using sensors and wearable processing units is nowadays hardly able to copy this behavior and provide similar level of protection. Nevertheless some attempts are made and Ambient Assisted Living (AAL) systems are constructed.

State-of-the-art systems handle some tasks typically performed by a human assistant. Information about current condition or actions (various kinds of sensors) of the monitored subject is processed to detect known (less or more severe) threats. The specific objective is the use of diagnostic information to distinguish the state of subjectively perceived discomfort from the objective state of emergency and protect human in the latter case.

To meet growing expectations, Ambient Assisted Living systems are subjects of ongoing research. Their application in general is to support the elderly and provide better care by medical units. However many variants have been developed. Starting from systems designed specifically for hospitals [12] and care units [11] providing not only activity monitoring but also localization within the area. Requirements imposed onto such systems assume that the staff is always nearby and ready to react in an emergency. This is quite different in systems devoted to home monitoring of solitary people, where not only data is acquired from the person but also from his environment [13,2]. Incorporating data from person's environment may enhance detection of abnormal activities. As an additional feature one may consider the functionality of home security from gas or smoke [10]. Some systems are built with the use of PIR movement detectors [8] or cameras [14]. However, as providing most useful information and at the same time not violating subject's discretion are systems utilizing wearable movement detectors [5,6]. Various activities can be detected and analyzed, especially fall detection is considered as the most important. Long term data analysis may discover trends in changes of the behavior that might indicate that the person may require continuous support by a human caregiver. The system should also provide possibility of voice communication [5].

In this paper a concept of an ALL system and initial tests of a monitoring device are presented. The paper is organized as follows: the next section describes the system, then Section 3 presents the prototype of the monitoring device,

followed by testing and data analysis in Section 4. Discussion on obtained results conclude the paper in the last section.

2 Concept of the System

In this study a new Ambient Assisted Living (AAL) system is introduced. The general aim of the system is to ensure well-being of the elderly people (patient) who are living (possibly alone) in their homes. The basic assumption is, that although patient's medical condition or age places him/her in higher risk group for downfalls, fainting etc, the subject can still live normal life as long as possible threats or dangerous changes in condition will be detected (or even predicted) in time to provide suitable help.

Five specific goals have been defined:

- detect present incidents and threats and automatically raise an alarm,
- estimate short-term risk of a dangerous situation possible to occur during the ongoing day to give a caregiver chance to react before the incident happens,
- forecast long-term degradation of patient's condition that might indispose independent living due to unacceptable high risk of dangerous situation (fall, lack of consciousness, hypokinesy, etc.),
- establish continuous possibility of providing help on demand and reacting to emergency situations (e.g. alarm button, voice communication etc.)
- gathered data with diagnostic significance should be archived to permit geriatrics (medical) evaluation.

Some of them can be accomplished only by an on-line data processing (e.g. detection of present incidents). Other goals put various constraints on the system (analysis of huge amount of long-term data should be possible, system availability is crucial for safety etc.) Detailed analysis allowed a hierarchical architecture to be proposed (Fig. 1).

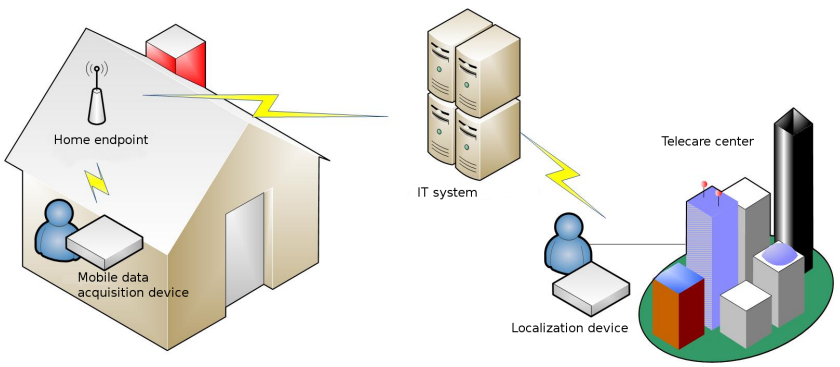


Fig. 1. System architecture

In the proposed solution the patient is equipped with a mobile data acquisition device (MDAD) collecting data from various distributed sensors. MDAD preprocesses acquired data and forwards it to a home endpoint (HE). HE may perform further data processing but is also a gateway to a central system (CS) – a heart of the whole AAD system. Part of the CS – the telecare (TC), provide continuous support for all monitored patients.

In this architecture, real time data processing required to generate urgent notifications or alarms can be performed both in MDAD (less demanding algorithms) or HE (more advanced computations) even without being connected to the CS. In such case, HE may also serve as a buffer for data sent from MDAD.

Long-term forecast generation and data gathering is a task of CS. The CS perform analysis of long-term trends to detect alarming evolution of patient's state and provide diagnostic data and other relevant information to geriatric physician, physiotherapist etc.

In other words, the CS center is equipped with IT system capable of:

- archiving large amounts of data,
- providing long-term data analysis using statistical and datamining methods,
- providing two-way voice (or video) communication with the patient (TC),
- enabling access by qualified experts to perform diagnosis based on gathered and processed data.

The proposed architecture conforms to systems described in the literature. The patient monitoring is based on the analysis of physical activities. This includes: (1) overall activity level, typical (2) daily behavior patterns, and (3) movement analysis form main data stream. Currently accelerometric data is used as most discriminative. Other MDAD provided signals[3] (blood pressure, pulse, temperature) are still considered as supplementary but not used at present. Dangerous situations detected by the monitoring system, like increasing number of falls, slips, loss of balance or, on the contrary, prolonged periods of inactivity give information about degradation of patient's condition.

In other words, in the proposed AAL system, the main source of data is the mobile data acquisition device. Analysis of MDAD-provided parameters permit the current patient state and its changes to be assessed. When an expert knowledge is provided, based on the same data more advanced reasoning is also possible and a forecast can be generated. MDAD must therefore be the most reliable and robust element of the system.

This paper focuses on the MDAD design and testing.

3 Mobile Data Acquisition Device

Mobile Data Acquisition Device is a set of electronic devices worn by a patient. The MDAD consists of the following parts (Fig. 2):

[3] Incorporation of environmental sensors like atmospheric pressure, humidity or ambient temperature, intensively perceived by some people and influencing their state is also under research.

– inertial sensors equipped with 3-axial accelerometer, 3-axial gyroscope and a magnetometer,
– integrating device controlling the sensor modules and performing signal preprocessing,
– additional instruments (i.e. microphone, pulsometer).

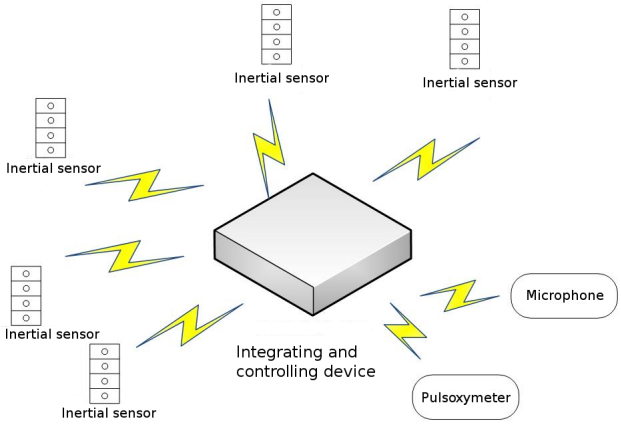


Fig. 2. Mobile Data Acquisition Device

Initially, 5 inertial sensors are to be placed in the following manner:

– above ankle of each leg (2),
– on waist laterally (2),
– on person's back at T6 vertebra (1).

Such sensor placement enables not only measuring activity level or performing recognition of activity type but also detecting abnormal behavior [1].

The MDAD prototype has been developed that conforms to the above architecture and is composed of:

– 5 MPU6050 sensor modules,
– Raspberry PI model B as the integrating and controlling device,

and is controlled by custom software developed in C++ language. Acquired data are available through a wireless link.

4 MDAD Testing

The prototype has been subjected to testing mainly to assess repeatability and similarity of signals acquired by all 5 sensors. Two variants of testing procedure have been performed. First, a stable arm has been employed in order to evaluate the repeatability of measurements and comparability of sensors. Then, the signals acquired by sensors attached to a person back have been registered in order to assess the intra- and inter-trial variability.

4.1 Measurement Procedures

In the first attempt a robotic arm has been employed (ROBOTICS SOLUTIONS[4] Fig. 3). The MDAD prototype has been attached to the robot arm then, the robot has performed predefined movements and the signals from five sensors have been recorded. Only the signals recorded by the sensors during the execution of movements in three directions: X, Y, Z have been analyzed. The presence of noise caused by the arm vibrations excluded the signal from a further analysis.

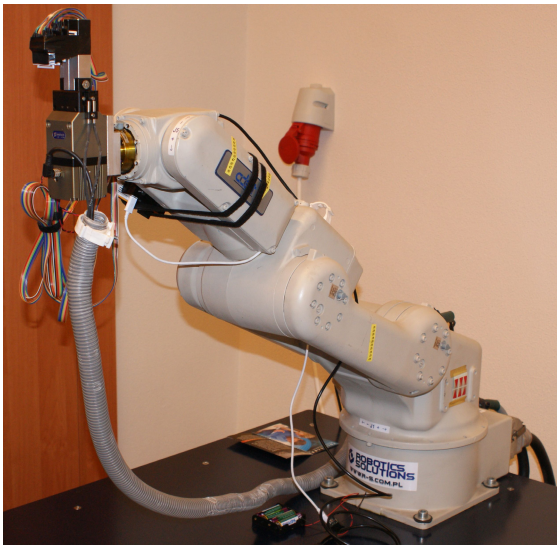


Fig. 3. Testing with the use of robotic arm

Second MDAD evaluation attempt has been made using mechanical arm (Fig. 4) with only one degree of freedom (rotation along vertical axis). All sensors have been firmly attached and 10 trials have been performed. Each trial starts with releasing of the arm, followed by angular acceleration and ending in an impact. Thus two of three sensing axes (X and Z) could be measured.

Final evaluation procedure have focused on repeatability of signals during normal walking. Tests have been performed on a randomly selected group of healthy people aged between 26 and 38 years. During this testing phase subjects have been wearing all 5 sensors attached to their backs (Fig. 5) and have been walking along a straight path. Other than sensor-dependent errors can be caused by movements of the patient specific anatomical structures, such as chest movement during respiration. Analysis of the available literature [7,9] states, however, that there is no need to eliminate these interferences due to introduced vanishingly small noise. The aim of those tests has been to assess the similarity

[4] www.R-S.com.pl

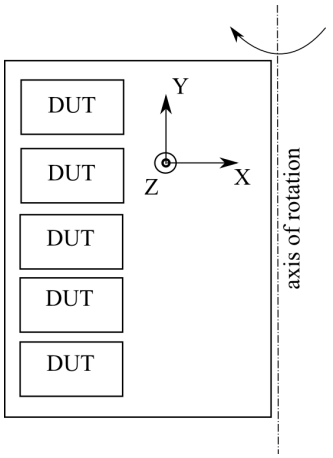


Fig. 4. Testing using mechanical arm

between signals acquired in conditions similar to the normal use and compare them to those obtained in controlled conditions (mechanical excitation). In all testing scenarios the sensors have been working at sampling frequency of 100Hz.

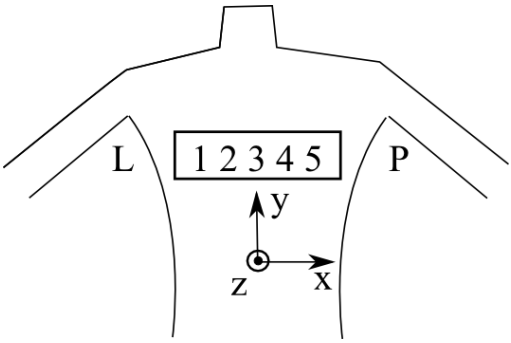


Fig. 5. Testing during walking

4.2 Data Analysis

Before the analysis all acquired waveforms have been subject to low pass median filtering. Then, two aspects of repeatability have been analyzed:

- inter-sensor cross correlation for each trial (intra-trial),
- inter-sensor cross correlation between trials (inter-trial).

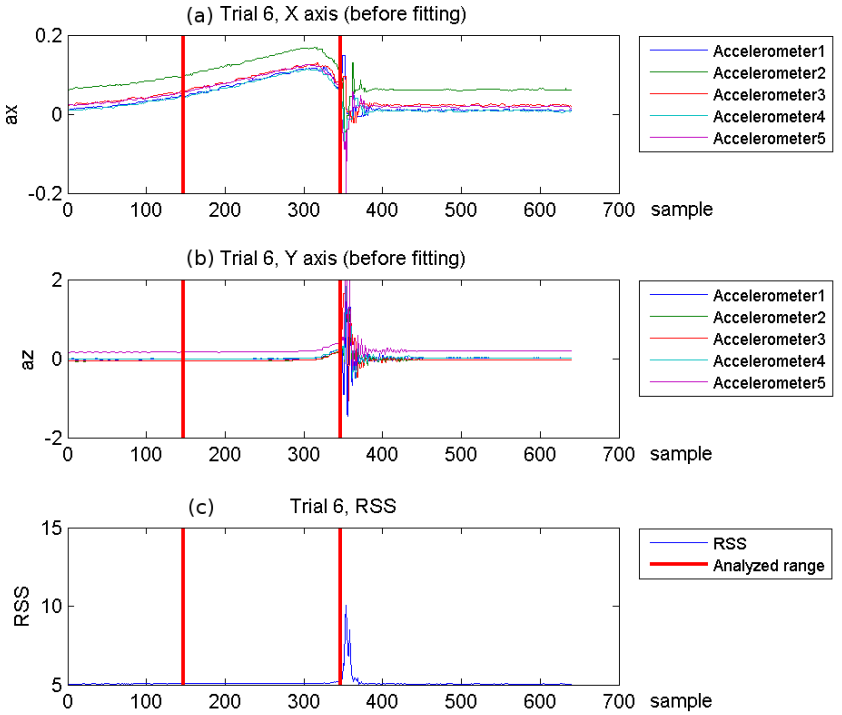


Fig. 6. Waveforms acquired by all accelerometers during trial 6 of mechanical excitation

The iter- and intra-trial accuracy has been found by correlating the signals within a predefined time window. A fiducial point that gates the signals is searched on a root-sum-of-squares (RSS, [3,4]) waveform (Fig. 6)

$$RSS = \sqrt{\sum_{id=1}^{id=5} [x_{id}^2 + y_{id}^2 + z_{id}^2]}, \tag{1}$$

where id indicates the sensor and x, y, z are values registered by them in each direction. Sample for which RSS increase exceeds value 0.5 is referred to as a fiducial point. The left edge of the window has been marked by selecting 200 samples registered before the fiducial point. Vertical red lines in Fig. 6 represent the obtained time window. Waveforms shown in Fig. 7 were subject to windowing operation. The waveforms within the obtained time window are shown in Fig. 8.

Next, correlation coefficients (CCs) between signals acquired during each trial have been calculated within the selected window. The analysis has been done separately for each sensing axis and trial giving in total 20 matrices 5x5. The p-values have been also calculated. Exemplary results are presented in Table 1. Minimal, average and maximal values of correlation coefficients between sensors for a specific trial are shown in Table 2 and in Table 3 – for a given sensor between trials.

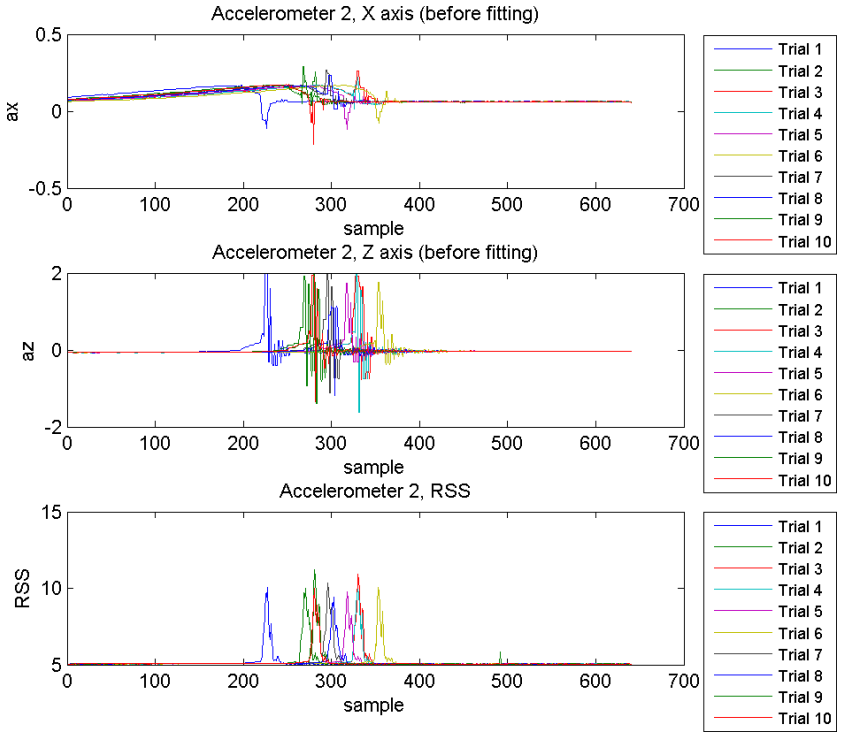
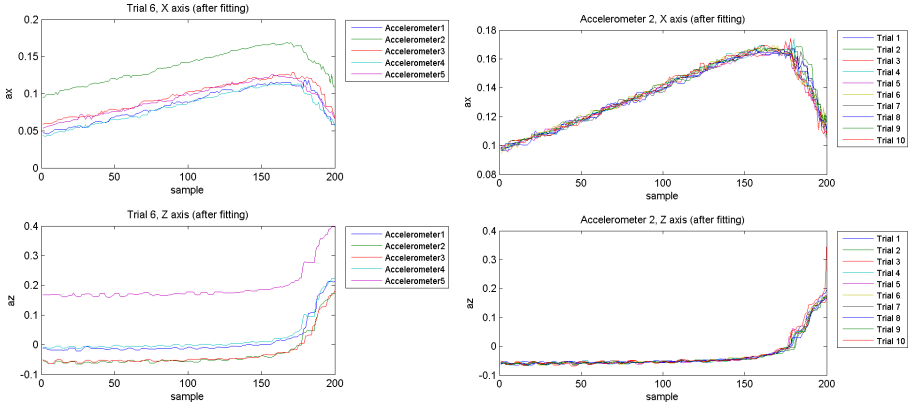


Fig. 7. Waveforms acquired during all trials by accelerometer ID=2 during mechanical excitation



(a) (b)

Fig. 8. Waveforms time-windowing: (a) single trial, all sensors, (b) single sensor, all trials

Table 1. Exemplary matrix of CCs and p–Values between sensors for a single (6th) trial

(a) CCs, x axis

	Accel1	Accel2	Accel3	Accel4	Accel5
Accel1	1.0000	0.9807	0.9905	0.9831	0.9818
Accel2	0.9807	1.0000	0.9806	0.9953	0.9950
Accel3	0.9905	0.9806	1.0000	0.9835	0.9801
Accel4	0.9831	0.9953	0.9835	1.0000	0.9941
Accel5	0.9818	0.9950	0.9801	0.9941	1.0000

(b) CCs, z axis

	Accel1	Accel2	Accel3	Accel4	Accel5
Accel1	1.0000	0.9926	0.9935	0.9909	0.9915
Accel2	0.9926	1.0000	0.9938	0.9985	0.9973
Accel3	0.9935	0.9938	1.0000	0.9930	0.9919
Accel4	0.9909	0.9985	0.9930	1.0000	0.9979
Accel5	0.9915	0.9973	0.9919	0.9979	1.0000

(c) p-values, x axis

	Accel1	Accel2	Accel3	Accel4	Accel5
Accel1	1.00e+00	2.96e-142	1.94e-172	5.48e-148	8.07e-145
Accel2	2.96e-142	1.00e+00	5.29e-142	6.53e-203	4.86e-200
Accel3	1.94e-172	5.29e-142	1.00e+00	5.44e-149	5.62e-141
Accel4	5.48e-148	6.53e-203	5.44e-149	1.00e+00	4.71e-193
Accel5	8.07e-145	4.86e-200	5.62e-141	4.71e-193	1.00e+00

(d) p-values, z axis

	Accel1	Accel2	Accel3	Accel4	Accel5
Accel1	1.00e+00	2.38e-183	4.37e-189	1.99e-174	2.56e-177
Accel2	2.38e-183	1.00e+00	1.44e-190	1.00e-251	5.06e-227
Accel3	4.37e-189	1.44e-190	1.00e+00	1.49e-185	1.71e-179
Accel4	1.99e-174	1.00e-251	1.49e-185	1.00e+00	6.29e-238
Accel5	2.56e-177	5.06e-227	1.71e-179	6.29e-238	1.00e+00

The inter- and intra-trial waveforms are shown in Fig. 8. The presentation of all sensor measures superimposed within one trial is followed by a superposition of results given by one randomly selected sensor during 10 independent trials.

Similar analysis has been performed on data acquired during the walking test. After median filtering, the waveforms have been time-windowed to 10 gait periods and analyzed as previously. Fig. 9 presents the waveforms after time windowing, showing waveforms from all sensors during one trial, and waveforms from sensor 5 during all trials.

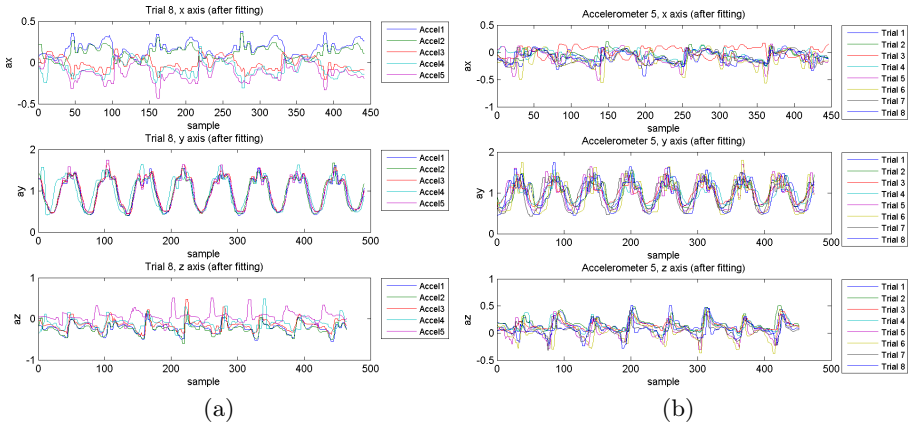
Similarly, the CCs have been computed and summarized in a per-sensor (Table 4) and per-trial (Table 5) manner.

Table 2. Summary of CCs between sensors for a specific trial (in all cases $p < 0.05$)

Trial	x axis			z axis		
	min	max	avg	min	max	avg
1	0.9791	0.9952	0.9888	0.9924	0.9966	0.9946
2	0.9802	0.9942	0.9882	0.9820	0.9969	0.9904
3	0.9784	0.9930	0.9893	0.9844	0.9960	0.9914
4	0.9789	0.9933	0.9878	0.9911	0.9959	0.9928
5	0.9733	0.9931	0.9853	0.9833	0.9956	0.9891
6	0.9801	0.9953	0.9865	0.9909	0.9985	0.9941
7	0.9889	0.9945	0.9925	0.9904	0.9969	0.9932
8	0.9766	0.9947	0.9860	0.9807	0.9963	0.9900
9	0.9857	0.9955	0.9891	0.9829	0.9976	0.9904
10	0.9800	0.9919	0.9873	0.9772	0.9964	0.9875
Worst case:	0.9733	0.9955	0.9881	0.9772	0.9985	0.9914

Table 3. Summary of CCs between trials for a specific sensor (in all cases $p < 0.05$)

Sensor	x axis			z axis		
	min	max	avg	min	max	avg
1	0.9625	0.9962	0.9857	0.9792	0.9967	0.9902
2	0.9707	0.9954	0.9876	0.9760	0.9965	0.9894
3	0.9584	0.9943	0.9861	0.9778	0.9975	0.9890
4	0.9694	0.9954	0.9887	0.9793	0.9969	0.9899
5	0.9731	0.9945	0.9886	0.9878	0.9965	0.9926
Worst case:	0.9584	0.9962	0.9873	0.9760	0.9975	0.9902



(a) (b)

Fig. 9. Waveforms after time-windowing: (a) single trial, all sensors, (b) single sensor, all trials

Table 4. Summary of CCs between sensors for a specific walking trial (in all cases $p < 0.05$)

Trial	x axis			y axis			z axis		
	min	max	avg	min	max	avg	min	max	avg
1	-0.6548	0.9402	-0.0311	0.7431	0.9785	0.8847	-0.4834	0.9485	0.3350
2	-0.7030	0.9429	-0.0069	0.5563	0.9574	0.8124	0.6683	0.9431	0.8113
3	-0.6488	0.9183	0.2016	0.7336	0.9845	0.8872	0.6549	0.9549	0.8145
4	0.5948	0.9519	0.7764	0.5734	0.9770	0.8178	-0.4739	0.9436	0.3335
5	0.3915	0.9216	0.6480	0.6331	0.9676	0.8366	-0.3825	0.9396	0.3493
6	0.4431	0.9259	0.6710	0.6618	0.9744	0.8468	0.5590	0.9533	0.7795
7	-0.6985	0.9243	-0.0446	0.8418	0.9903	0.9325	-0.3122	0.9088	0.3267
8	-0.8017	0.8910	-0.1348	0.8235	0.9905	0.9162	-0.2303	0.8877	0.3243
Worst case:	-0.8017	0.9519	0.2600	0.5563	0.9905	0.8668	-0.4834	0.9549	0.5093

Table 5. Summary of CCs between trials for a specific sensor (in all cases $p < 0.05$)

Sensor	x axis			y axis			z axis		
	min	max	avg	min	max	avg	min	max	avg
1	-0.6209	0.8337	0.2877	0.5794	0.9182	0.7513	0.2872	0.8991	0.5260
2	-0.5067	0.8445	0.3675	0.5440	0.9785	0.7535	0.2978	0.8867	0.5712
3	0.4165	0.8564	0.6172	0.5636	0.9174	0.7356	-0.7209	0.8557	0.3214
4	0.4316	0.8208	0.6118	0.5849	0.9059	0.7547	-0.5523	0.9057	0.3678
5	-0.5320	0.8408	0.3148	0.5893	0.9284	0.7348	0.3447	0.8784	0.6166
Worst case:	-0.6209	0.8564	0.4398	0.5440	0.9785	0.7460	-0.7209	0.9057	0.4806

5 Conclusions

Results obtained form mechanical excitation tests show statistically significant high correlation between acquired data. However, constant value offsets can be observed and require correction if velocity or displacement calculation and analysis is performed.

Data registered during the walking test show that waveforms on x and z axis are in some trials off-phase. This result could be expected due to sensor placement and considering individual walking patterns. When walking the trunk performs rotation along vertical axis, thus sensors placed laterally from spine are in opposite phase. However, not in all individuals this could be observed, but has to be taken into consideration in further analysis. Another aspect is sensor mounting. Currently the sensors are being fixed to the body using velcro strips, what causes additional variation of acquired signals due to hindered mounting in exactly the same place on the body.

Despite the noticed problems, it can be concluded that the MDAD is useful as movement data acquisition device and is being constantly improved. Further research focuses on improving reliability and incorporating other sensing modalities.

Acknowledgement. This work has been supported by the European Union under the 7th Frame Programme grant nr POIG.01.03.01-24-061/12.

References

1. Bao, L., Intille, S.S.: Activity recognition from user-annotated acceleration data. In: Ferscha, A., Mattern, F. (eds.) PERVASIVE 2004. LNCS, vol. 3001, pp. 1–17. Springer, Heidelberg (2004)
2. Botia, J.A., Villa, A., Palma, J.: Ambient assisted living system for in-home monitoring of healthy independent elders. Expert Systems with Applications 39(9), 8136–8148 (2012)
3. Bourke, A., van de Ven, P., Gamble, M., O'Connor, R., Murphy, K., Bogan, E., McQuade, E., Finucane, P., ÓLaighin, G., Nelson, J.: Evaluation of waist-mounted tri-axial accelerometer based fall-detection algorithms during scripted and continuous unscripted activities. Journal of Biomechanics 43(15), 3051–3057 (2010)
4. Charlon, Y., Fourty, N., Bourennane, W., Campo, E.: Design and evaluation of a device worn for fall detection and localization: Application for the continuous monitoring of risks incurred by dependents in an alzheimer's care unit. Expert Systems with Applications 40(18), 7316–7330 (2013)
5. Costa, A., Castillo, J.C., Novais, P., Fernández-Caballero, A., Simoes, R.: Sensor-driven agenda for intelligent home care of the elderly. Expert Systems with Applications 39(15), 12192–12204 (2012)
6. Fourty, N., Guiraud, D., Fraisse, P., Perolle, G., Etxeberria, I., Val, T.: Embedded system used for classifying motor activities of elderly and disabled people. Computers & Industrial Engineering 57(1), 419–432 (2009)
7. Godfrey, A., Bourke, A., Ólaighin, G., van de Ven, P., Nelson, J.: Activity classification using a single chest mounted tri-axial accelerometer. Medical Engineering & Physics 33(9), 1127–1135 (2011)
8. Kasteren, T., Englebienne, G., Kröse, B.: An activity monitoring system for elderly care using generative and discriminative models. Personal and Ubiquitous Computing 14(6), 489–498 (2010)
9. Khan, A., Lee, Y.K., Lee, S., Kim, T.S.: A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. IEEE Transactions on Information Technology in Biomedicine 14(5), 1166–1172 (2010)
10. Kwon, O., Shim, J.M., Lim, G.: Single activity sensor-based ensemble analysis for health monitoring of solitary elderly people. Expert Systems with Applications 39(5), 5774–5783 (2012)
11. Marco, A., Casas, R., Falco, J., Gracia, H., Artigas, J., Roy, A.: Location-based services for elderly and disabled people. Computer Communications 31(6), 1055–1066 (2008)

12. Rajasekaran, M.P., Radhakrishnan, S., Subbaraj, P.: Sensor grid applications in patient monitoring. Future Generation Computer Systems 26(4), 569–575 (2010)
13. Suryadevara, N., Gaddam, A., Rayudu, R., Mukhopadhyay, S.: Wireless sensors network based safe home to care elderly people: Behaviour detection. Procedia Engineering 25, 96–99 (2011)
14. Zhou, Z., Dai, W., Eggert, J., Giger, J., Keller, J., Rantz, M., He, Z.: A real-time system for in-home activity monitoring of elders. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2009, pp. 6115–6118 (2009)

Part IV

Experimental Bioengineering

Hetero- and Homogeneous Multiclassifier Systems Based on Competence Measure Applied to the Recognition of Hand Grasping Movements

Marek Kurzynski and Andrzej Wolczowski

Wroclaw University of Technology, Department of Systems and Computer Networks
Wyb.Wyspianskiego 27, 50-370 Wroclaw, Poland,
marek.kurzynski@pwr.wroc.pl

Abstract. The paper presents an advanced method of recognition of patient's intention to move of multijoint hand prosthesis during the grasping and manipulating objects in a dexterous manner. The proposed method is based on a two-level multiclassifier system (MCS) with heterogeneous and homogeneous base classifiers dedicated to EMG and MMG biosignals and with combining mechanism using a dynamic ensemble selection scheme and probabilistic competence function. The performances of two MCSs with the proposed competence function and combining procedure were experimetally compared against three benchmark MCSs using real data concerning the recognition of six types of grasping movements. The systems developed achieved the highest classification accuracies demonstrating the potential of multiple classifier systems with multimodal biosignals for the control of bioprosthetic hand.

Keywords: Multiclassifier system, Competence measure, Hand grasping movements.

1 Introduction

Nowadays, many researchers focus on Multiple Classifier Systems (MCS) because this approach has been shown to outperform single classifiers for a wide range of classification problems. Two main approaches used for the combination of classifiers in the ensemble, are classifier fusion and classifier selection [7]. In the first approach, all classifiers in the ensemble contribute to the decision of the MCS, e.g. through sum or majority voting. In the second approach, a single classifier is selected from the ensemble for each test example and its decision is used as the decision of the MCS. The selection of a classifier can be either static or dynamic. In static classifier selection, a region of competence in the feature space is assigned for each classifier during the training phase and classification is made by the classifier assigned to the competence region that contains the test example. In dynamic classifier selection, competences of classifiers are calculated during the classification phase, i.e. at the time when the test example is presented. The classifier with the highest value of competence is used for the classification of the test example.

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_15, © Springer International Publishing Switzerland 2014

Recently, dynamic ensemble selection (DES) methods are intensively developed as an effective approach to the construction of multiple classifier systems ([12,23,24]). In these methods, first an ensemble of base classifiers is dynamically selected and then the selected classifiers are combined by majority voting. The most DES schemes use the concept of classifier competence on a defined neighbourhood or region, such as the local accuracy estimation, Bayes confidence measure, multiple classifier behaviour or probabilistic model, among others.

In this study the multiclassifier system was applied to the recognition of hand grasping movements, which is a fundamental problem in the control of the dexterous bioprosthetic hand [16]. The proposed method is based on a two-level multiclassifier system (MCS) with heterogeneous and homogeneous base classifiers dedicated to EMG and MMG biosignals and with combining mechanism using a dynamic ensemble selection scheme and original probabilistic competence function. This paper is a sequel to the authors' earlier publications [8,9,10] and it provides an extension of the results included therein.

The paper arrangement is as follows. Chapter 2 includes the concept of prosthesis control system based on the recognition of patient intent and provides an insight into steps of the whole recognition procedure. Chapter 3 presents the key recognition algorithm based on the multiclassifier system with the dynamic ensemble classifier selection strategy. Chapter 4 presents experimental results confirming adopted solution and chapter 5 concludes the paper.

2 Recognition of Hand Movements as a Tool for Bioprosthesis Decison Control

Existing active prostheses of hand are generally controlled on myoelectric way – they react to electrical signals that accompany the muscle activity (called electromyography signals – EMG signals). The control is feasible since after the amputation of the hand, there remain a significant number of the muscles in the arm stump that normally controlled the finger action. The tensing of these muscles still depends on the patient will and may express her/his intentions as to the workings of her/his prosthesis [13,26].

Nevertheless, reliable recognition of intended movement using only the EMG signals analysis is a hard problem. A recognition error increases along with the cardinality of movement repertoire (i.e. with prosthesis dexterity). A natural solution to overcome this error and increasing the efficiency of the recognition stage may be achieved through the following activities:

1. by introducing the concept of simultaneous analysis of two different types of biosignals, which are the carrier of information about the performed hand movement – authors studied the fusion of EMG signals and the mechanomyography signals (MMG signals)[9];
2. through improving the recognition method – authors proposed to use the multiclassifier system with heterogeneous and homogeneous base classifiers dedicated to particular registered biosignals;

3. by the appropriate choice of feature extraction methods (biosignals parameterization) justified by the experimental results of comparative analysis.

The above analysis shows, that – according to authors' proposition – the bioprosthesis control is performed by recognizing its intended movement on the base of classification of EMG and MMG signals from the user arm stump. This requires the development of three stages:

1. acquisition of signals;
2. reduction of dimensionality of their representation;
3. classification of signals.

The acquisition must take into account the nature of the measured signals and their measurement conditions. A quality of the obtaining information depends essentially on the ratio of the measured signal power to the interfering signal power, defined as SNR (Signal to Noise Ratio). For the non-invasive methods of measurements carried out on the surface of the patient's body, it is difficult to obtain a satisfactory SNR [2]. The noise amplitude usually exceeds many times the amplitude of the measured signal. For the EMG signals the amplitude of voltages induced on the patient body as a result of the influence of external electric fields, may exceed more than 1000 times, the value of useful signals. To overcome this difficulty a differential measurement system was applied. The system encompasses two signal electrodes placed above the examined muscle and an reference electrode placed as far as possible above electrically neutral tissue (above a bone or a joint). Signals obtained from signal electrodes are subtracted from each other and amplified. The common components, including surrounding noise, are thus excluded and the useful signal is amplified.

The MMG signals are mechanical vibrations propagating in the limb tissue as the muscle contracts. They have low frequency (up to 200 Hz) and small amplitude and can be registered as a "muscle sound" on the surface of the skin using microphones [14,20]. This sound carries essential information about individual muscle group excitation. The basic problem when designing the MMG sensor is to isolate the microphone from the external sound sources along with the best acquisition of the sound propagating in the patient's tissue.

After the acquisition stage, the recorded signals have the form of strings of discrete samples. Their size is the product of measurement time and sampling frequency. For a typical motion, that gives a record of size between 3 and 5 thousand of samples (time of the order of 3-5 s, and the sampling of the order of 1 kHz). This "primary" representation of the signals hinders the effective classification and requires the reduction of dimensionality. This reduction leads to a representation in the form of a signal feature vector. To determine the algorithm of features extraction, the database records were analyzed in time and frequency using Short Time Fourier Transform (STFT). Fig. 1 shows the exemplary results.

As we can see, the MMG histogram has two amplitude peaks: at the beginning and at the end of the movement, and relatively low amplitude in the middle while the EMG histogram shows a peak in the middle of the movement time

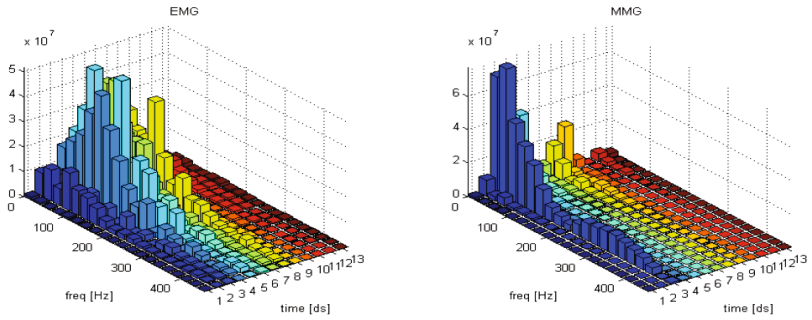


Fig. 1. Exemplary histograms for EMG signal (left) and for MMG signal (right)

span. The analyses of histograms for the tested movements allowed selecting the localization of the best signal features (the best points in time and frequency) securing the best differentiation of the movements.

The resulting algorithm has the following form:

Step 1. Extract from the recorded signal, the signal segments representing the specified movements (using video information). Each extracted segment has new time span ($t \in [0, T]$);
Step 2. Apply the STFT to each segment;
Step 3. Choose as signal features the values from the STFT product corresponding to the k (most representative) time slices;
Step 4. Repeat steps 2 and 3 for every channel;
Step 5. Use all the obtained (in steps 2 and 3) values as elements of the feature vector representing the analyzed signal segment.

This procedure allows creating input vectors with an adjustable size. The structure of this feature vector used as an input in the classifier is given by:

$$(A_{t_1}^{CH_i}, A_{t_2}^{CH_i}, \ldots, A_{t_k}^{CH_i})_{i=1,2,\ldots,n}, \tag{1}$$

where k is the number of time slices and n denotes the number of signal channels.

Although different methodological paradigms can be used as a classifier construction, we suggest using multiclassifier systems (MCS) with base classifier dedicated to particular registered biosignals and with the dynamic ensemble selection method using original procedure of fusion/selection based on competence measure.

3 Multiclassifier System

3.1 Preliminaries

In the multiclassifier (MC) system we assume that a set of trained classifiers $\Psi = \{\psi_1, \psi_2, \ldots, \psi_L\}$ called base classifiers is given. A classifier ψ_l is a function

$\psi_l : \mathcal{X} \rightarrow \mathcal{M}$ from a feature space to a set of class labels $\mathcal{M} = \{1, 2, \ldots, M\}$. Classification is made according to the maximum rule

$$\psi_l(x) = i \Leftrightarrow d_{li}(x) = \max_{j \in \mathcal{M}} d_{lj}(x), \tag{2}$$

where $[d_{l1}(x), d_{l2}(x), \ldots, d_{lM}(x)]$ is a vector of class supports (classifying function) produced by ψ_l. Without loss of generality we assume that $d_{lj}(x) \geq 0$ and $\sum_j d_{lj}(x) = 1$.

The ensemble Ψ is used for classification through a combination function which, for example, can select a single classifier or a subset of classifiers from the ensemble, it can be independent or dependent on the feature vector x (in the latter case the function is said to be dynamic), and it can be non-trainable or trainable [7]. The proposed multiclassifier system uses dynamic ensemble selection (DES) strategy with trainable selection/fusion algorithm. The basis for dynamic selection of classifiers from the pool is a competence measure $c(\psi_l|x)$ of each base classifier ($l = 1, 2, \ldots, L$), which evaluates the competence of classifier ψ_l, i.e. its capability to correct activity (correct classification) at a point $x \in \mathcal{X}$. For the training of competence it is assumed that a validation set

$$\mathcal{V} = \{(x_1, j_1), (x_2, j_2), \ldots, (x_N, j_N)\}; \quad x_k \in \mathcal{X}, \ j_k \in \mathcal{M} \tag{3}$$

containing pairs of feature vectors and their corresponding class labels is available.

The construction of the competence measure consists of the two following steps. In the first step, a hypothetical classifier called a randomized reference classifier (RRC) is constructed. The RRC can be considered to be equivalent to the classifier ψ_l and its probability of correct classification $Pc^{(RRC)}(x_k)$ can be used as the competence $C(\psi_l|x_k)$ of that classifier. In the second step, the competences $C(\psi_l|x_k), \ x_k \in \mathcal{V}$ are used to construct the competence function $c(\psi_l|x)$. The construction is based on extending (generalizing) the competences $C(\psi_l|x_k)$ to the entire feature space $\mathcal{X}$. The next two subsections describe the steps of the method in detail.

3.2 Randomized Reference Classifier

The RRC is a stochastic classifier and therefore it is defined using a probability distribution over the set of class labels $\mathcal{M}$ or, assuming the canonical model of classification, over the product of class supports $[0, 1]^M$. In other words, the RRC uses the maximum rule and a vector of class supports $[\delta_1(x), \delta_2(x), \ldots, \delta_M(x)]$ for the classification of the feature vector x, where the j-th support is a realization of a random variable (rv) $\Delta_j(x)$. The probability distributions of the rvs are chosen in such a way that the following conditions are satisfied (throughout this description, the index l of the classifier ψ_l and its class supports is dropped for clarity):

(1) $\Delta_j(x) \in [0, 1]$;
(2) $E[\Delta_j(x)] = d_j(x), \ j = 1, 2, \ldots, M$;
(3) $\sum_{j=1,2,\ldots,M} \Delta_j(x) = 1$,

where E is the expected value operator. The above definition denotes that the RRC can be considered to be equivalent to the classifier ψ for the feature vector x since it produces, on average, the same vector of class supports as the modeled classifier.

Since the RRC performs classification in a stochastic manner, it is possible to calculate the probability of classification of an object x to the i-th class:

$$P^{(RRC)}(i|x) = Pr[\forall_{k=1,\ldots,M,\ k\neq i}\ \Delta_i(x) > \Delta_k(x)]. \tag{4}$$

In particular, if the object x belongs to the i-th class, from (4) we simply get the conditional probability of correct classification $Pc^{(RRC)}(x)$.

The key element in the modeling presented above is the choice of probability distributions for the rvs $\Delta_j(x), j \in \mathcal{M}$ so that the conditions 1-3 are satisfied. In this paper beta probability distributions are used with the parameters $\alpha_j(x)$ and $\beta_j(x)$ $(j \in \mathcal{M})$. The justification of the choice of the beta distribution, resulting from the theory of order statistics, can be found in [23].

Applying the RRC to a validation point x_k and putting in (4) $i = j_k$, we get the probability of correct classification of RRC at a point $x_k \in V$:

$$Pc^{(RRC)}(x) = \int_0^1 b(u, \alpha_1(x_k), \beta_1(x_k)) \left[\prod_{j=2}^{M} B(u, \alpha_1(x_k), \beta_1(x_k))\right] du, \tag{5}$$

where $B(\dot{})$ is a beta cumulative distribution function. The MATLAB code for calculating probabilities (5) was developed and it is freely available for download [25].

3.3 Measure of Classifier Competence

Since the RRC can be considered equivalent to the modeled base classifier $\psi_l \in \Psi$, it is justified to use the probability (5) as the competence of the classifier ψ_l at the learning point $x_k \in \mathcal{S}$, i.e.

$$C(\psi_l|x_k) = Pc^{(RRC)}(x_k). \tag{6}$$

The competence values for the validation objects $x_k \in \mathcal{V}$ can be then extended to the entire feature space $\mathcal{X}$. To this purpose the following normalized Gaussian potential function model was used ([22]):

$$c(\psi_l|x) = \frac{\sum_{x_k \in \mathcal{V}} C(\psi_l|x_k) exp(-dist(x, x_k)^2)}{\max_{x \in \mathcal{X}} \sum_{x_k \in \mathcal{V}} C(\psi_l|x_k) exp(-dist(x, x_k)^2)}, \tag{7}$$

where $dist(x, y)$ is the Euclidean distance between two objects x and y.

3.4 Dynamic Ensemble Selection System

Since recognition of the patient's intent is made on the basis of analysis of two different biosignals (EMG and MMG), the multiple classifier system – according to the proposed concept of the recognition method – consisits of two submulti-classifiers, each of them dedicated to particular types of data. It leads to the two level structure of MC system presented in Fig. 2, in which the DES method is realized at the first level, whereas the combining procedure at the second level is consistent with the continuous-valued dynamic fusion scheme.

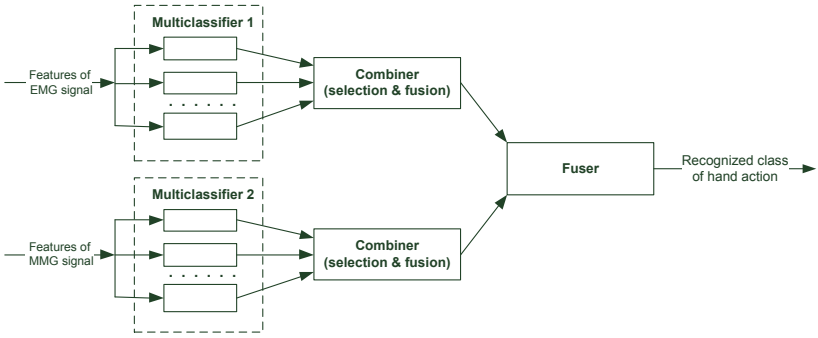


Fig. 2. Block diagram of the proposed multiclassifier system

DES Systems at the First Level. Let Ψ_1 and Ψ_2 denote sets (ensembles) of base classifiers dedicated to the EMG and MMG signals, respectively. The DES system for the ensemble Ψ_i $(i = 1, 2)$ is constructed using the developed measure of competence and classifies the feature vector $x^{(i)}$ ($x^{(1)}$ and $x^{(2)}$ denote the vector of features obtained from the EMG and MMG signal, respectively) in the following manner.

First, the competence function $c(\psi_k^{(i)}|x)$ $(k = 1, 2, ..., L_i)$ are constructed for each classifier in the ensemble. Then, a subset $\Psi_i^*(x)$ of base classifiers with the competences greater than the probability of random classification is selected. This step eliminates inaccurate classifiers and keeps the ensemble relatively diverse [11]. The selected classifiers are combined on the continuous-valued level [7], i.e. class suports are calculated as the weighted sum of supports given by base classifiers from $\Psi_i^*(x)$,viz.

$$d_j^{(i)}(x) = \sum_{\psi_k^{(i)} \in \Psi_i^*(x)} c(\psi_k^{(i)}|x) d_{k,j}^{(i)}(x). \tag{8}$$

Fusion Procedure at the Second Level. At the second level of MC, supports (8) are combined by the weighted sum:

$$d_j(x) = \sum_{i=1,2} c^{(i)}(x) d_j^{(i)}(x), \tag{9}$$

where weight coefficients $(i = 1, 2)$

$$c^{(i)}(x) = \frac{1}{|\Psi_i^*(x)|} \sum_{\psi_k^{(i)} \in \Psi_i^*(x)} c(\psi_k^{(i)}|x). \tag{10}$$

denote mean competence of base classifiers from $\Psi_i^*(x)$.

Finally, the MC system classifies $x = (x^{(1)}, x^{(2)}, x^{(3)})$ using the maximum rule:

$$\psi_{MC}(x) = i \;\Leftrightarrow\; d_i(x) = \max_{j \in \mathcal{M}} d_j(x). \tag{11}$$

4 Experiments

4.1 Experimental Setup

In order to study the performance of the proposed method of EMG and MMG signals recognition, some computer experiments were made. The experiments were conducted in MATLAB using PRTools 4.1 [5] and Signal Processing Toolbox. In the recognition process of grasping movements, 6 types of objects (a pen, a credit card (standing in a container), a computer mouse, a cell phone (laying on the table), a kettle and a tube (standing on the table)) were considered. Our choice is deliberate and results from the fact that the control functions of simple bioprosthesis are hand closing/opening and wrist pronantion/supination, however for the dexterous hand these functions differ depending on grasped object [16].

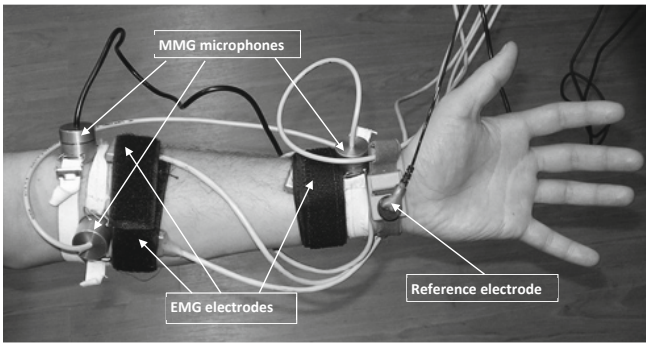


Fig. 3. The layout of the EMG electrodes and the MMG microphones on the forearm

The dataset used to test the proposed classification methods consisted of 400 measurements, i.e. pairs "EMG and MMG signals segment/movement class". Each measurement lasted 6 s and was preceded by a 10-second break. The values from the STFT product (1) corresponding to the $k = 3, 4, 5$ most representative time slices were considered as feature vector. Consequently, we got 3 datasets each containing 400 objects desribed by a different number of features.

The training and testing sets were extracted from each dataset using two-fold cross-validation. One half of the objects from the training dataset was used as a validation dataset and the other half was used for the training of base classifiers. Three experiments were performed which differed in the biosignals used for classification (EMG signals, MMG signals, both EMG and MMG signals).

The experiments were conducted using two types of ensembles Ψ_1 and Ψ_2: homogeneous and heterogeneous. The homogeneous ensemble consisted of 20 feed-forward backpropagation neural network classifiers with one hidden layer (containing 10 neurons) and the number of learning epochs set to 100. The heterogeneous ensemble consisted of the following 10 classifiers [4]: (1, 2) linear (quadratic) classifier based on normal distributions with the same (different) covariance matrix for each class, (3) nearest mean classifier, (4 – 6) k-nearest neighbours classifiers with $k = 1, 5, 15$, (7, 8) Parzen density based classifier with the Gaussian kernel and the optimal smoothing parameter h_{opt} (and the smoothing parameter $h_{opt}/2$), (9) pruned decision tree classifier with Gini splitting criterion, (10) support vector machine classifier with radial basis kernel function. For both ensemble types, classifiers were trained using bootstrapping of the training set.

The performances of the systems constructed (MC_{Hetero} and MC_{Homo} (with hetero-and homogeneous base classifiers,respectively) were compared against the following three multiple classifier systems: (SB) – The single best classifier in the ensemble [7]; (MV) – Majority voting (MV) of all classifiers in the ensemble [7]; (LA) – DCS-local accuracy (LA) system: this system classifies x using selected classifier with the highest local competence (the competence is estimated using k nearest neighbours of x taken from the validation set [15].

4.2 Results and Discussion

Classification accuracies (i.e. the percentage of correctly classified objects) for methods tested are listed in Table 1 (k denotes the number of time slices per signal channel). The accuracies are average values obtained over 10 runs (5 replications of two-fold cross validation). Statistical differences between the performances of the MC_{Hetero} and MC_{Homo} systems and the three MCS's were evaluated using Dietterich's 5x2cv test [3]. The level of $p < 0.05$ was considered statistically significant. In Table 1, statistically significant differences are given under the classification accuracies as indices of the method evaluated, e.g. for the dataset with $k = 3$ and EMG signals the MC_{Homo} system produced statistically different classification accuracies from the SB and MV methods.

These results imply the following conclusions:

1. The both MC_{Hetero} and MC_{Homo} systems produced statistically significant higher scores in 37 out of 54 cases (9 datasets × 3 classifiers × 2 systems developed);
2. There are no statistically significant differences between scores of MC_{Hetero} and MC_{Homo} systems.
3. The multiclassifier systems using both EMG and MMG signals achieved the highest classification accuracy for all datasets.

Table 1. Classification accuracies of MSCs compared in the experiment (description in the text). The best score for each dataset is highlighted

	Classifier / Mean (SD) accuracy [%]				
k	SB (1)	MV (2)	LA (3)	MC_{Hetero} (4)	MC_{Homo} (5)
	EMG signals				
3	77.2/2.3	74.5/1.5	78.3/1.6	78.5/2.3 1,2	**78.7**/1.9 1,2
4	**85.7**/1.9	83.2/1.3	85.1/1.8	85.4/2.5 2	85.1/1.9 2
5	90.5/2.2	92.6/1.8	91.8/1.7	93.1/2.3 1,3	**93.5**/1.6 1,2,3
	MMG signals				
3	**47.8**/1.1	43.5/1.5	46.8/1.6	45.9/1.3 2	44.8/1.1 2
4	52.4/1.3	51.2/1.2	50.6/0.8	**54.2**/0.9 1,2,3	53.8/1.2 1,2,3
5	65.8/1.1	63.9/0.7	65.4/0.9	**67.2**/1.3 1,2,3	66.8/0.9 2,3
	MMG and EMG signals				
3	82.5/2.1	81.8/1.5	83.1/1.6	**84.3**/1.5 1,2	83.9/1.8 1,2
4	92.7/1.7	92.1/1.3	91.9/2.0	**93.8**/2.1 2,3	93.0/1.5 2,3
5	95.9/1.3	95.1/0.7	94.7/0.9	96.8/1.1 2,3	**97.0**/1.2 1,2,3

5 Conclusion

Experimental results indicate that the proposed methods of grasping movement recognition based on the dynamic ensemble selection with probabilistic model of competence function, produced – regardless of the type of base classifiers – accurate and reliable decisions, especially in the cases with features coming from both the EMG and MMG biosignals.

The problem of deliberate human impact on the mechanical device using natural biological signals generated in the body can be considered generally as a matter of "human – machine interface". The results presented in this paper significantly affect the development of this field and the overall discipline of biosignal recognition, thereby contributing to the comprehensive development of biocybernetics and bioengineering. But more importantly, these results will also find practical application in designing a dexterous prosthetic hand – in the synthesis of control algorithms for these devices, as well as development of computer systems for learning motor coordination, dedicated to individuals preparing for a prosthesis or waiting for hand transplantation [21].

Acknowledgement. This work was financed from the National Science Center resources in 2012-2014 years as a research project No ST6/06168.

References

1. Boostanl, B., Moradi, M.: Evaluation of the forearm EMG signal features for the control of a prosthetic hand. Physiolog. Measurement 24, 309–319 (2003)
2. De Luca, C.: Electromyography. In: Webster, J.G. (ed.) Encyclopedia of Medical Devices and Instrumentation, pp. 98–109. John Wiley Publisher (2006)
3. Dietterich, T.: Approximate statistical tests for comparing supervised classification learning algorithms. Neural Computation 10, 1895–1923 (1998)
4. Duda, R., Hart, P., Stork, D.: Pattern Classification. John Wiley and Sons, New York (2000)
5. Duin, R., Juszczak, P., et al.: PRTools4. A Matlab Toolbox for Pattern Recognition. Delft University of Technology (2007)
6. Englehart, K.: Signal representation for classification of the transient myoelectric signal. Ph.D. Thesis, University of New Brunswick, Fredericton, New Brunswick (1998)
7. Kuncheva, I.: Combining Pattern Classifiers: Methods and Algorithms. Wiley-Interscience (2004)
8. Kurzynski, M., Wolczowski, A.: Dynamic selection of classifier ensemble applied to the recognition of EMG signal for the control of bioprosthetic hand. In: Proc. 11 th Int. Conf. on Control, Automation and Systems, Seoul, pp. 175–182 (2011)
9. Kurzynski, M., Wolczowski, A., Tito, A.: Control of bioprosthetic hand based on EMG and MMG signals recognition using multiclassifier system with feedback from the prosthesis sensors. In: Proc. 2nd Int. Conf. on Systems and Control, Marrakech, Morocco, pp. 280–285 (2012)
10. Kurzynski, M., Wolczowski, A.: Control of bioprosthetic hand based on multimodal biosignals recognition and feedback signals from the prosthesis sensors. In: Int. Conf. Biological Engineering and Natural Science, Bangkok, pp. 315–322 (2013)
11. Lysiak, R., Kurzynski, M., Woloszynski, T.: Probabilistic approach to the dynamic ensemble selection using measures of competence and diversity of base classifiers. In: Corchado, E., Kurzyński, M., Woźniak, M. (eds.) HAIS 2011, Part II. LNCS, vol. 6679, pp. 229–236. Springer, Heidelberg (2011)
12. Lusiak, R., Kurzynski, M., Woloszynski, T.: Optimal selection of ensemble classifiers using measures of competence and diversity of base classifiers. Neurocomputing 126, 29–35 (2014)
13. Nishikawa, D.: Studies on Electromyogram to Motion Classifier. Ph.D. Thesis, Graduate School of Engineering, Hokkaido University, Sapporo (2001)
14. Orizio, C.: Muscle sound: basis for the introduction of a mechanomyographic signal in muscle studies. Critical Reviews in Biomedical Engineering 21, 201–243 (1993)
15. Smits, P.: Multiple classifier systems for supervised remote sensing image classification based on dynamic classifier selection. IEEE Trans. on Geoscience and Remote Sensing 40, 717–725 (2002)
16. Wolczowski, A., Kurzynski, M.: Human – machine interface in bio-prosthesis control using EMG signal classification. Expert Systems 27, 53–70 (2010)

17. Wolczowski, A., Kurzynski, M.: Control of artificial hand via recognition of EMG Signals. In: Barreiro, J.M., Martín-Sánchez, F., Maojo, V., Sanz, F. (eds.) ISBMDA 2004. LNCS, vol. 3337, pp. 356–367. Springer, Heidelberg (2004)
18. Wolczowski, A.: Smart hand: The concept of sensor based control. In: Proc. of 7th IEEE Int. Symposium MMAR, Miedzyzdroje, pp. 783–790 (2001)
19. Wolczowski, A., Krysztoforski, K.: Artificial hand control via EMG signal classification - experimental investigation of alghoritms. In: Tchon, K. (ed.) Progress in Robotics, pp. 97–122. WKL, Warszawa (2008)
20. Wolczowski, A., Suchodolski, T.: Bioprosthesis control: human-machine interaction problem. In: Eizmendi, G., Azkoitia, J.M., Craddock, G. (eds.) Challenges for Assistive Technology, pp. 558–560. IOS Press, Amsterdam (2007)
21. Wolczowski, A., Kurzynski, M., Zaplotny, P.: Concept of a system for training of bioprosthetic hand control in one side handless humans using virtual reality and visual and sensory biofeedback. Medical Information Technologies 18, 85–91 (2011)
22. Woloszynski, T., Kurzynski, M.: A measure of competence based on randomized reference classifier for dynamic ensemble selection. In: 20th Int. Conf. on Pattern Recognition, pp. 4194–4197. IEEE Computer Press, Istanbul (2010)
23. Woloszynski, T., Kurzynski, M.: A probabilistic model of classifier competence for dynamic ensemble selection. Pattern Recognition 44, 2656–2668 (2011)
24. Woloszynski, T., Kurzynski, M., Podsiadlo, P., Stachowiak, G.: A measure of competence based on random classification for dynamic ensemble selection. Information Fusion 13, 207–213 (2012)
25. Woloszynski, T.: Matlab Central File Enchange (2010), `http://www.mathwork.com/matlabcentral/fileenchange/28391-classifier-competence-based-on-probabilistic-modeling`
26. Zecca, M., Micera, S., Carrozza, M., Dario, P.: Control of Multifunctional Prosthetic Hands by Processing the Electromyographic Signal. Critical Reviews in Biomedical Engineering 30, 459–485 (2002)

Prototype of the Device Registering Encircling Images

Mirosław Dziewoński[1], Ewa Majchrzak[1],
Mariusz Ciesielski[2], and Sebastian Freus[2]

[1] Silesian University of Technology, Institute of Computational Mechanics and Engineering, Konarskiego 18a, 44-100 Gliwice, Poland
{miroslaw.dziewonski,ewa.majchrzak}@polsl.pl
[2] Czestochowa University of Technology, Institute of Computer and Information Sciences, Dąbrowskiego 69, 42-200 Częstochowa, Poland
{mariusz.ciesielski,sebastian.freus}@icis.pcz.pl

Abstract. Within the project entitled "The diagnostic system supporting the healing process of burn and chronic wounds" funded by The Polish National Centre of Research and Development the diagnostic system and device for recording thermal images/videos was created. The paper concerns the description of the prototype of the device registering encircling thermograms and video images.

The novelty consists in using 3D thermal imaging techniques in the analysis of circumferential burns, for which 2D thermography is insufficient. Developed instrument records thermograms, both radial and longitudinal, alongside the scanned object, which allows recording of temperature field on large body surfaces with improved measuring accuracy as compared with classic 2D thermograms.

1 Introduction

Wound healing is a complex, slow and time consuming process, where treatment consists in coordination of events on cellular, physiological and biochemical level. It also depends on the type of injury (mechanical, chemical, thermal).

System IRDiagnostics offering comprehensive support of the diagnostics and treatment of burn wounds and chronic wounds, has been developed collectively by a team of specialists from Institute of Computational Mechanics and Engineering (IMIO), Biomedical Engineering Centre (CIB) of the Silesian University of Technology and Institute of Medical Technologies and Instruments (ITAM). The system employs imaging in visible and infrared spectrum. It also allows evaluation of parameters related to treatment process as early as in the initial phase of wound healing and allows implementing targeted therapy.

The impulse for development of the invention came from Burns Treatment Centre (CLO) in Siemianowice Śląskie, Poland's best centre for burns therapy. Phases of treatment in which designed system can support other therapeutic

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_16, © Springer International Publishing Switzerland 2014

methods were determined in collaboration with CLO [1,2,3,4]. The system can assist in:

- classification of wounds,
- monitoring of wound infection progress,
- diagnostics of skin grafts integration,
- determination of healing degree of burns and hard-to-heal wounds,
- determination of healing degree of wounds treated with hyperbaric oxygen.

In all the above-mentioned cases recording flat thermal images (thermograms) allows only for a qualitative assessment, while measuring the surface area of wounds (usually extensive ones) is almost impossible. During the research on the diagnostic system a thermal encircling image recording device was designed allowing to improve the qualitative and providing accurate quantitative assessment.

In addition to described application in wound therapy, the system can be used in monitoring the progress of chemotherapy and radiotherapy of skin tumours and in changes recording in therapy of other dermatological diseases.

2 System Components

Major system components are thermal imaging camera and video camera installed on the instrument recording three-dimensional thermograms and software

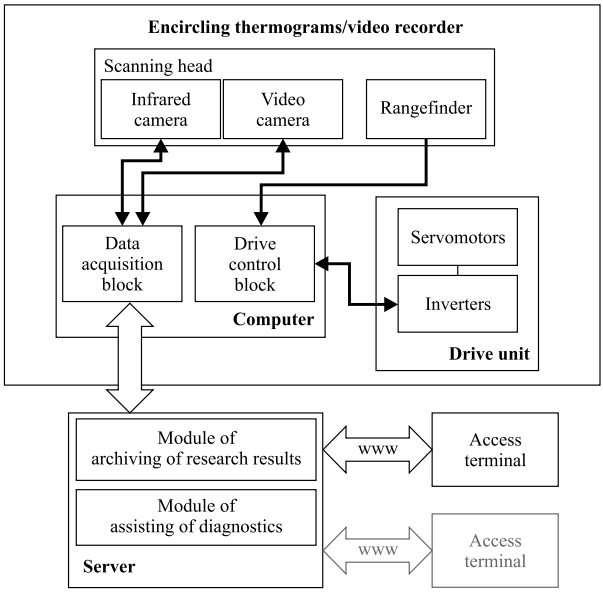


Fig. 1. Structure of the IRDiagnostic system

comprising database, image archiving module and a module of images processing for assisting of diagnostics [5] (Fig. 1).

The recorder is controlled locally with a computer system, which during the examination sends the thermal images/videos to the central database that allows users to access them supporting the diagnostic processes. Considering this, the system is multi-user and allow users to interactively manage the available diagnostic applications from the level of work stations and access terminals.

Developed applications (included in the system) allow the determination and measurement of wound surface area, both on the thermograms and the images in the visible range. They also allow three-dimensional visualization of the wound and their comparison in the various stages of the treatment.

3 Device for Recording Encircling Thermal and Video Images

The device registering encircling images consists of, among others, a moveable outrigger with the device head (inside the thermal and video cameras, rangefinders) and image recording device (Fig. 2).

It should be emphasised that the device will provide a semi-automatic recording of thermal images both radially and along the examined object, which will allow us to record temperature distribution on extensive body surface areas providing a considerably higher measurement accuracy than in the case of "classic" flat thermal images.

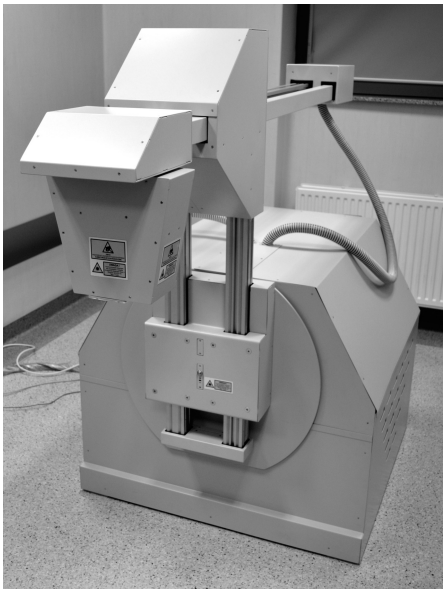


Fig. 2. Encircling thermograms/video recorder

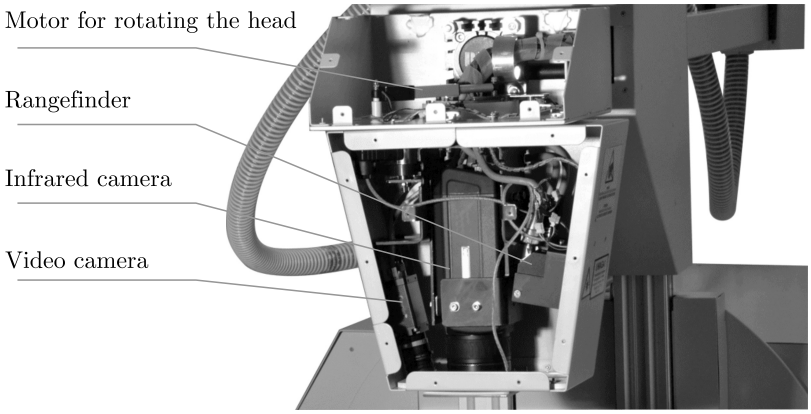


Fig. 3. Head of the recording device

Rangefinders are coupled with the camera and allow ones to measure the distance to the patient's body in order to ensure a good dioptre adjustment for preservation of the image sharpness (Fig. 3).

The idea of recording the encircling thermograms lies in the fact that the infrared camera records the temperature of points on the body within the strip more than ten pixels wide and the preset length L, while the recording will be conducted continuously at a specified frequency, when the camera moves around at a constant angular velocity. After recording the strip with a width L in this way, the camera moves linearly by the same section with a necessary overlap and the system records another body strip (Fig. 4).

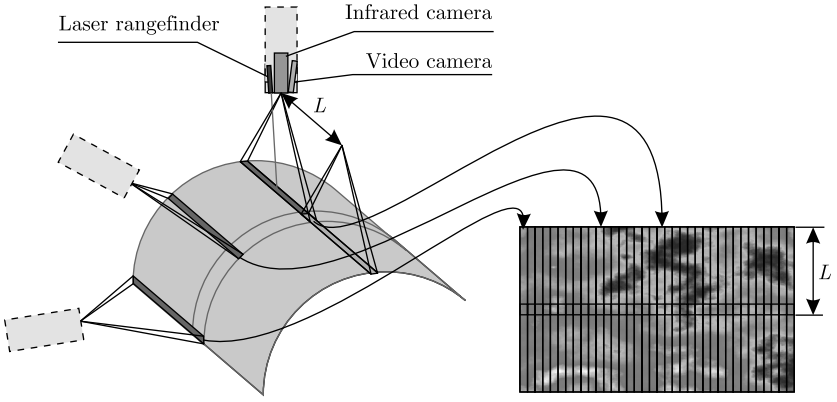


Fig. 4. Visualization of the idea of recording encircling thermograms

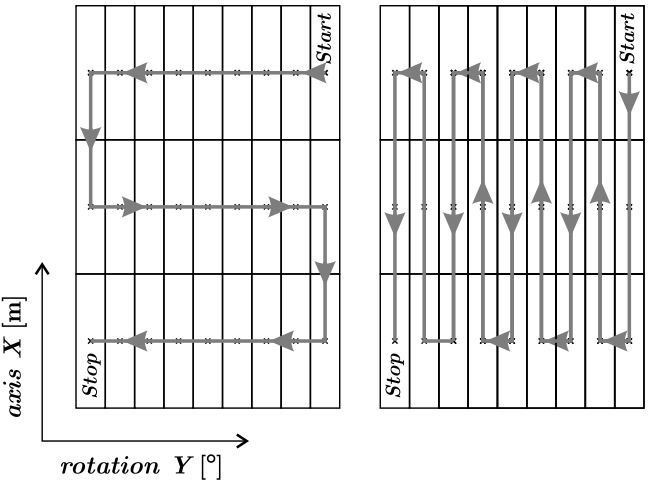


Fig. 5. Motion trajectory of the head: rotational-linear and linear-rotational

The simultaneous band recording of thermal image and video together with measuring the distance from the area under examination makes it possible to project the recorded image at a high precision onto a plane and measure the surface area of the wound.

Both the angle of camera movement and the total length of linear displacement are adjustable.

During the scanning process operator selects the scan-starting point on the patient's skin surface by moving the head using the buttons on the system interface. The control is done in a manual or automatic way. After the start of the scan, the head moves in a rotational motion until the end point is indicated by the operator (the scan of one strip) and then the head moves linearly with a fixed distance to scan another strip at the patient's body in the opposite direction. These steps are repeated until the end of the scan is indicated by the operator. Depending on the scanning mode, the movement of the head device can be done in reverse order, i.e. the first has linear movement, rotational then. Described motion trajectory of the head is in a schematic way shown in Fig. 5.

4 Making of the Encircling Thermograms

To obtain the thermograms of the surface temperature of large wounds, an algorithm of image merging is proposed.

During scanning a selected area of the patient's body surface with burn wound, infrared and video images in the device are recorded. Images are recorded in certain discrete positions of the device head whose have been determined by the scanning algorithm. It is assumed that the step length of position head is fixed for both axes (Δx [m] for axis X and Δy [°] for axis Y) - see Fig. 6.

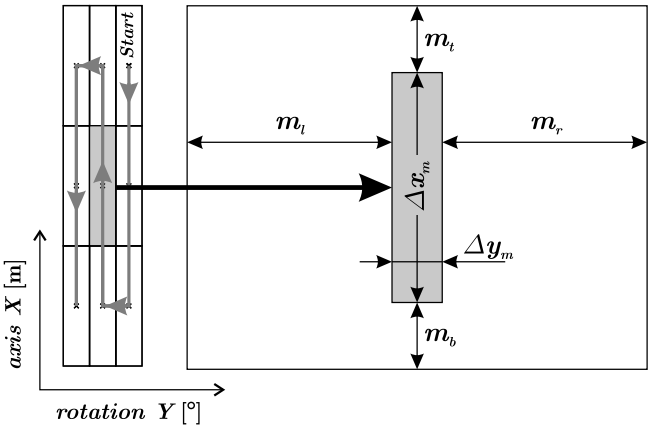


Fig. 6. Cropping of the scanned images

Before scanning, the following major data should be set:

- the scan mode: rotational-linear or linear-rotational,
- the position of the head of device at the selected initial scanning point,
- number of scanned steps (N_x and N_y) in each direction (in order to determine the scanned image number).

During the recording of video and thermal images in the device, for each recorded image (with indexes i and j), the following information are stored:

- the current head position (three coordinates): $x_{i,j}$, $y_{i,j}$, $z_{i,j}$,
- the current distance from the rangefinder to the body of the patient: $dk_{i,j}$,
- scale parameters for image,
- image data in the FPF file format (FLIR System) - in the case of the infrared camera,
- image data in the standard BMP file format - in the case of the video camera.

One can note that the distance $\overline{dk}$ of the head to the body of the patient during scanning is not always constant - because it is associated with a continuous motion of the head and insufficiently fast correction of the head position in the z-coordinate. Deviations from the intended distance were low and the images were recorded at a depth of field resulting from the optical properties of both cameras. In a similar way, this problem occurs when the images are acquired from both cameras at the determined position by the algorithm. For this purpose, some geometric corrections during merging images must be taken into consideration. Schemas of head positions in rotational and linear motions are presented in Fig. 7.

After scanning, the obtained data (set of thermal and video images) are merged in the computer system. In order to merge images, the following algorithm is proposed:

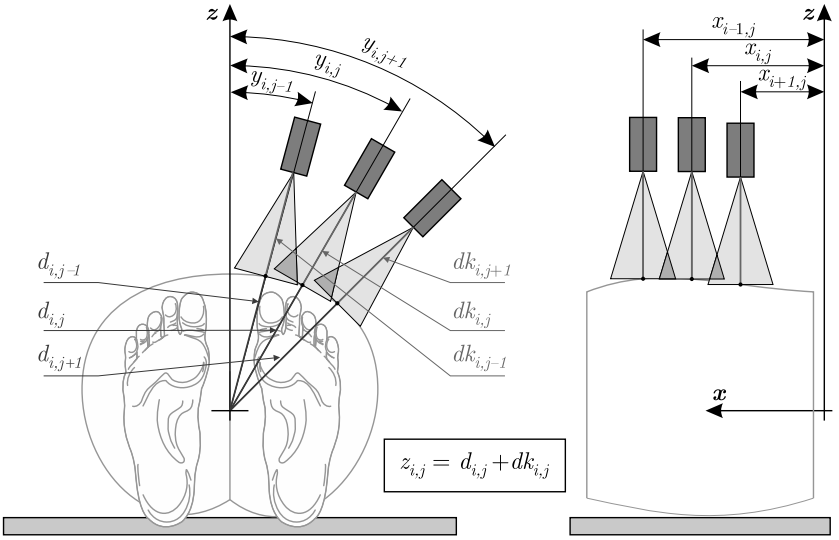


Fig. 7. Schemas of head positions in rotational and linear motions

1. Determine size of the field of view of both cameras on the base of the distance from rangefinder:

$$\Delta x_{i,j}^{IR} = 480/\text{Scale}^{IR}\,(dk_{i,j}) \quad [m]\ , \tag{1}$$

$$\Delta y_{i,j}^{IR} = 640/\text{Scale}^{IR}\,(dk_{i,j}) \quad [m]\ , \tag{2}$$

$$\Delta x_{i,j}^{V} = 480/\text{Scale}^{V}\,(dk_{i,j}) \quad [m]\ , \tag{3}$$

$$\Delta y_{i,j}^{V} = 640/\text{Scale}^{V}\,(dk_{i,j}) \quad [m] \tag{4}$$

   where Scale( ) is the function used to determine the number of image pixels per 1 meter for the rangefinder distance.
2. Determine the average size of the image fragments:

$$\Delta x_{m,pix}^{IR} = \Delta x_m\ \text{Scale}^{IR}\left(\overline{dk}\right) + 0.5 \quad [pix]\ , \tag{5}$$

$$\Delta y_{m,pix}^{IR} = \Delta y_m\ \text{Scale}^{IR}\left(\overline{dk}\right) + 0.5 \quad [pix]\ , \tag{6}$$

$$\Delta x_{m,pix}^{V} = \Delta x_m\ \text{Scale}^{V}\left(\overline{dk}\right) + 0.5 \quad [pix]\ , \tag{7}$$

$$\Delta y_{m,pix}^{V} = \Delta y_m\ \text{Scale}^{V}\left(\overline{dk}\right) + 0.5 \quad [pix] \tag{8}$$

   where

$$\Delta x_m = \Delta x\ , \tag{9}$$

$$\Delta y_m = 2\ d_m \tan\!\left(\frac{\Delta y}{2}\,\frac{\pi}{180}\right)\ , \tag{10}$$

$$d_m = \frac{1}{N_x N_y}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y} d_{i,j}\ . \tag{11}$$

3. Calculate the crop margins (m_l, m_r, m_t, m_b) [pix] (see Figure 6) of each recorded image on the base of mathematical formulas which take into account the scale and offset correction for position of the head in the registration point:

$$m_l^{IR} = \left(\frac{\Delta y_{i,j}^{IR}}{2} - d_{i,j} \tan\left(\frac{\Delta y}{2} \frac{\pi}{180} \right) + y_{offset} \right) \text{Scale}^{IR}\left(dk_{i,j}\right) + 0.5 \ , \quad (12)$$

$$m_r^{IR} = \left(\frac{\Delta y_{i,j}^{IR}}{2} - d_{i,j} \tan\left(\frac{\Delta y}{2} \frac{\pi}{180} \right) - y_{offset} \right) \text{Scale}^{IR}\left(dk_{i,j}\right) + 0.5 \ , \quad (13)$$

$$m_t^{IR} = \left(\frac{\Delta x_{i,j}^{IR}}{2} - \frac{\Delta x_m}{2} - x_{offset} \right) \text{Scale}^{IR}\left(dk_{i,j}\right) + 0.5 \ , \quad (14)$$

$$m_b^{IR} = \left(\frac{\Delta x_{i,j}^{IR}}{2} - \frac{\Delta x_m}{2} + x_{offset} \right) \text{Scale}^{IR}\left(dk_{i,j}\right) + 0.5 \ . \quad (15)$$

4. Cut out a necessary part of every image and next merge those pieces together (using stretching - if needed) into a single thermovision image and a single video image.
5. Approximate the shape of the scanned area of the patient and visualization 3D of scan images (optionally).

As a result of this algorithm, the temperature distribution over a large area of the body and the visual image on a two dimensional plane are obtained.

5 Planimetry of the Encircling Thermograms and Pictures

While measuring the surface of the areas marked in the pictures, commonly known algorithms are used. The same algorithms are also applied for example in geodesy.

The first stage of wound measurement and calculation process is point digitalization which involves marking the given points with a cursor by an operator. When matched together, the points make a contour, which is the boundary of the burn area (Fig. 8).

The digitized coordinates of the points identified are recorded and on this basis the surface area P is calculated using the following formulas:

$$2P = \sum_{i=1}^{n} \xi_i \left(\eta_{i+1} - \eta_{i-1} \right) \quad (16)$$

where n is the number of vertices, $\eta_{n+1} = \eta_1, \eta_0 = \eta_n$ and

$$2P = \sum_{i=1}^{n} \eta_i \left(\xi_{i+1} - \xi_{i-1} \right) \quad (17)$$

where $\xi_{n+1} = \xi_1, \xi_0 = \xi_n$.

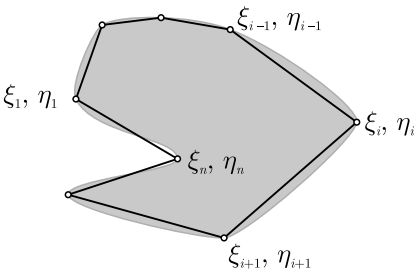


Fig. 8. Point digitizing

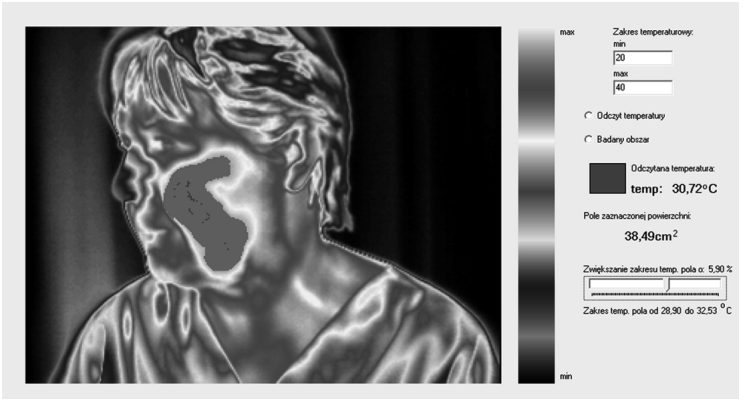


Fig. 9. An example of measurement of the area on the thermogram

Received values of the marked surface area should be calculated basing on an appropriate scale S which can be determined by setting the frame size based on the optics geometry of a camera and the distance from the object.

Thermographic pictures on average do not have tight edges allowing for an identification of borders between the areas with pathological changes and the undamaged tissue. Using point digitalization in analyzing thermograms is difficult because, as it has already been mentioned, the transition between the areas of various temperatures is vague, without clear boundaries. Therefore, hand-held attempts of indicating the areas are strenuous and prone to serious mistakes [6]. Having consulted a group of doctors, it was assumed that in case of thermograms, the parameter contributing to the assessment of wound healing progress can be the measurement of the surface area of similar temperatures in the subareas suspected of resulting in pathological changes. A thermogram is a dispersion of radiation intensity measured by the detector's particular cells (e.g. microbolometers) and registered in two dimensional arrays having the measurements consistent with the detector's resolution (e.g. 640x480). Each pixel (an element of the array) can be presented either in an arithmetic form, for example showing the existing temperatures of an object, or in a binary form. In the

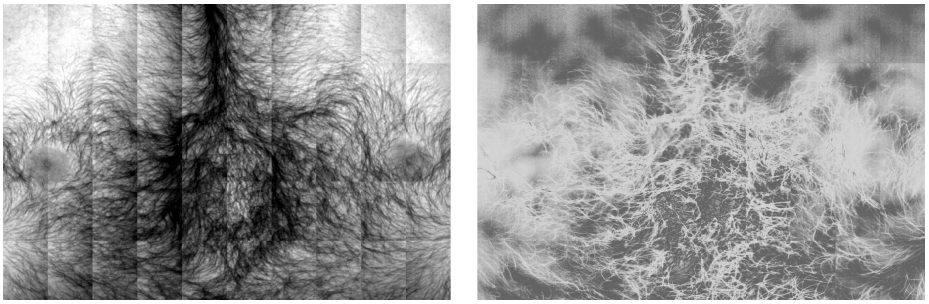


Fig. 10. Recorded field of the video image (left side) and the temperature distribution (right side) of the chest on a two dimensional plane

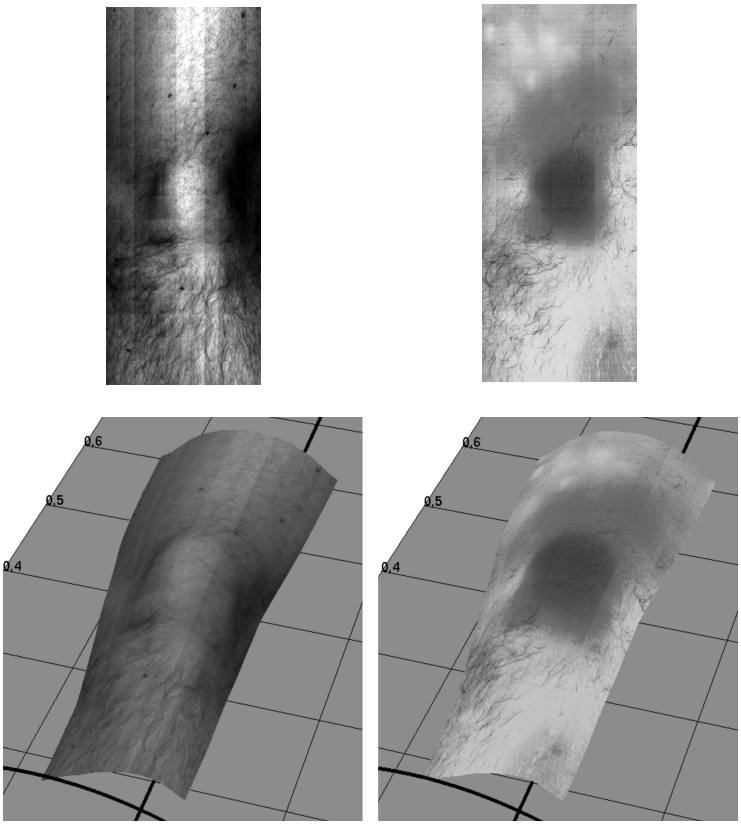


Fig. 11. Recorded field of the video image (left side) and the temperature distribution (right side) of the knee

latter case, in order to read the registered temperatures it is essential to know the temperature range set in the camera during the measurement.

As in the case of the surface area measurement, it is also necessary to know the actual frame size in order to set an appropriate scale. If we indicate the actual frame size, we can easily calculate the area of each pixel comprising a thermogram. Counting the pixels of a given temperature (or from the given temperature range) we will calculate the surface area of the wound fragment in question. The areas of skin necrosis have the temperature significantly lower than the temperature of a healthy tissue (being the point of reference) but the inflammatory condition developing in the wound increases it. An example measurement of a wound surface area is shown in the Fig. 9.

6 Examples of Registered Images

To verify the correctness of the proposed algorithm, a torso (chest) and the knee of a healthy male without burn wounds has been scanned.

During the first scan, 50 (10x5) images have been recorded. Both cameras acquire single image with resolution 640x480 pixels. As a result of the merging algorithm, two bitmap images were obtained. The resolution of the IR and video image is 1460x965 pixels. In Fig. 10 these images are shown.

During the next scan, 176 (11x16) images have been recorded. The resolution of the IR and video image is 627x1536 pixels. The images and their spatial visualizations (3D) on the base of the geometrical data are shown in Fig. 11.

7 Conclusion

The diagnostic system developed by authors consisting of a device for recording encircling thermal and video images and a computer system will constitute a completely new tool intended to assist the diagnostics and the treatment process of burn and chronic wounds. A significant increase in the accuracy of measurement of the temperature field in comparison to the classic flat thermograms is observed. Currently, the clinical utility of the system is verifying.

Acknowledgement. The patent application covering the device presented in the article has been filed in the Polish Patent Office under the number of P.397895 The article and research were financed within the project N R13 0124 10 sponsored by Polish National Centre for Research and Development.

References

1. Mani, R., Romanelli, M., Shukla, V.: Measurements in Wound Healing: Science and Practice. Spinger, London (2012)
2. Monstrey, S., Hoeksema, H., Verbelen, J., Pirayesh, A., Blondeel, P.: Assessment of burn depth and burn wound healing potential. Burns 34, 761–769 (2008)
3. Więcek, B., De Mey, G.: Thermovision in infrared. Bases and applications. Wydawnictwo PAK, Warszawa (2010)

4. Kawecki, M., Glik, J., Knefel, G., Podlewski, B., Nowak, M.: Application of hyperbaric oxygenation and vacuum therapy in the treatment of necrotic soft tissue infections. Leczenie Ran 6 (2009)
5. Majchrzak, E., Dziewoński, M., Nowak, M., Kawecki, M., Bachorz, M., Kowalski, P.: The design of a system for assisting burn and chronic wound diagnosis. In: Piętka, E., Kawa, J. (eds.) ITIB 2012. LNCS, vol. 7339, pp. 110–117. Springer, Heidelberg (2012)
6. Dziewoński, M.: Planimetry of thermograms in diagnosis of burn wounds. Scentific Research of the Institute of Mathematics and Computer Science 1(8), 33–38 (2009)

Preliminary Study of Clinial Workflow Integration of Spatial Thermograms Application for the Assessment of Burn Wound Healing Process

Dominik Spinczyk

Faculty of Biomedical Engineering, Silesian University of Technology, Zabrze, Poland
dspinczyk@polsl.pl
http://ib.polsl.pl

Abstract. This paper presents a preliminary study of prototype application supporting monitoring of healing process of burn and chronic wounds. Application has been designed in the client-server architecture that allows you to: adding new patients, looking for patients according their data, presentation of infrared burn wound map section. Course of patient's healing process is represented in the form of history of patient's medical visits which consists of: infrared images of burnt places, desciription of Computed Tomography (CT) screening i ultrasound screening, and results of microbiological and histopatological test. The presentation of a burn wound section includes: histogram presentation, selection of regions of interest by thresholding, imposition of colour masks on a thresholded section, manually selecting the region of interest by imposing a polygonal curve, slice and regions of interest descriptions. A prototype has been tested on a synthetic test images and anonymized pictures of pilot sample of test patients.

Keywords: burn wounds, chronic wounds, burn wounds thermograms, healing process of burn wounds.

1 Introduction

Today we are developing methods and technical means by which computer-aided diagnosis and therapy is entering new areas. One of them is the process of healing of burn wounds. This process is complex and time-consuming, can be monitored with a thermal image, which represent wounds' surface temperature [5]. Wound's temperature informs whether it is developed to necrosis or inflammation. Human tissue temperature is largely related to their blood supply. Level of blood supply tells us whether injured tissue is being healed. Subcutaneous tissue damage invisible to the naked eye are observed using infrared, it allows you to capture heat emitted by the human body. If body temperature at burns is less than normal tissue temperature, there is a high probability of necrosis to take place in a given

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_17, © Springer International Publishing Switzerland 2014

place. However, if temperature is higher, it can be concluded that inflammatory condition come into being in a given place [4]. Undoubtful specification of healing process progress is not easy, because in this type of wounds their appearance between individual changes of dressings or therapeutic treatments varies only slightly or effects are not visible at all. The purpose of the current approach is objectification of burn wounds healing process which uses spatial thermograms of burn wounds. The literature summarizing the positions of infrared work on issues relating to the treatment of burn wounds could be found in [1].

2 Material and Methods

Assessment of burn wounds healing process shall be carried out on the basis of infrared photos performed in successive stages of healing process. In order to make an evaluation of the progress of burn wound healing process more objective, measurement of wound's surface is applied additionally. Quantitative assessment, linked to the measurement of surface area, in the case of classic thermograms is very difficult, and in the case of extensive wounds is impossible [2]. In order to improve qualitative assessment and for exact quantitative assessment, belt thermograms are used in the project.

2.1 Equipment Which Registers Belt Thermograms

The main components of registering system are infrared cameras and video capture device installed on the arm of equipment which allows to register belt thermograms. Equipment registers semi-automatically thermograms both radially and along the object, which allows you to record temperature field on the large body areas, with significant increase of accuracy comparing to the "classic" flat thermograms (Fig. 1). Simultaneous thermogram and videogram belt recording together with distance measurment from the nearest investigated area allows for high accuracy in transforming registered image on a plane, and then calculate wound's surface (Fig. 2) [3]. Camera captures the temperature of body surface points in a belt several pixels wide and of given length. Registration is carried out continuously with the specified frequency, while the camera moves in a circular path at a constant angular speed. Once belt with previously specified length is registered linear displacement of camera takes place for the same distance and registration of next body belt [3].

Putting acquired images together is done by computer application. Three types of files are generated, which are used in the further stage of data processing.

- photo and infrared photo
- the file that contains surface area
- file containing temperatures of given areas

Designed diagnostic system (Fig. 4) consists of two essential parts: equipment for acquisition of belt thermograms and videograms (Fig. 3) and computer system which archives tests results and supports diagnostics and process of healing of wounds.

Data acquisition equipment is controlled locally by means of a computer system, which at the time of the study simultaneously transmits recorded data to the main database shared with users for further diagnosis.

2.2 System Architecture and Data Entry

The application requirements have been gathered for the management and viewing of data in the diagnosis and monitoring of burn wound treatment. On the basis of gathered requirements and experience in designing and deploying computer assisted medical diagnosis systems [6,7] prototype database and prototype application in a client-server architecture [8] has been implemented. Basic layer of the created application is the data repository management which involves storing the data included in the description of a clinical case: wound photographic images, flat surface maps of wounds, parts of flat surface maps of wounds in the form of image files. The designed prototype database and related to it data entering interface allows you to enter new patients, keep medical appointments chronology, associated with the diagnosis and treatment of burn wounds. In addition, the system collects information on other results of monitored patient including: description of CT test, ultrasound examination, microbiological and histopathological study results. Patient visit view is shown in Fig. 5. Interface available for patient search is shown in Fig. 6. Navigation through the application is facilitated by navigation bar, as shown in Fig. 7.

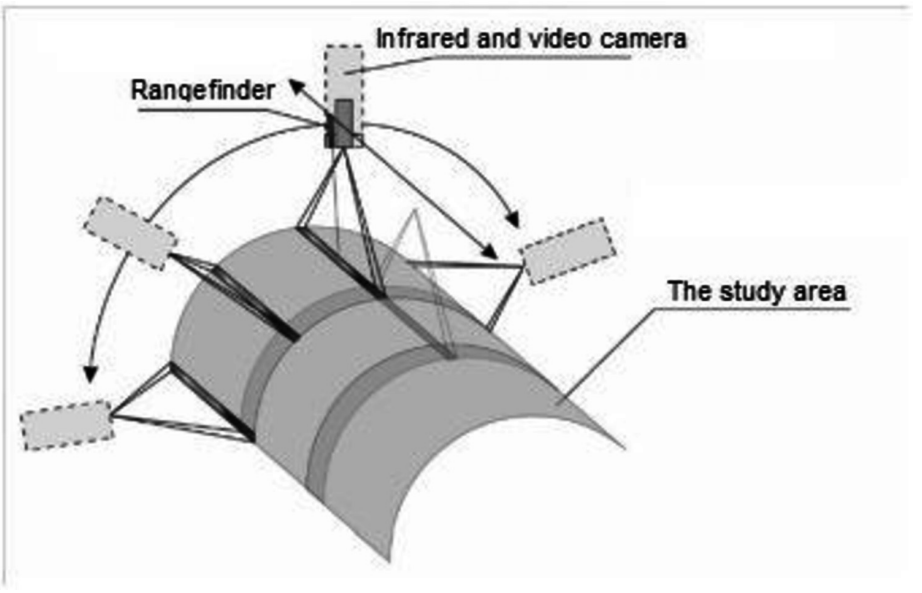


Fig. 1. Recorder of belt thermograms operating diagram [3]

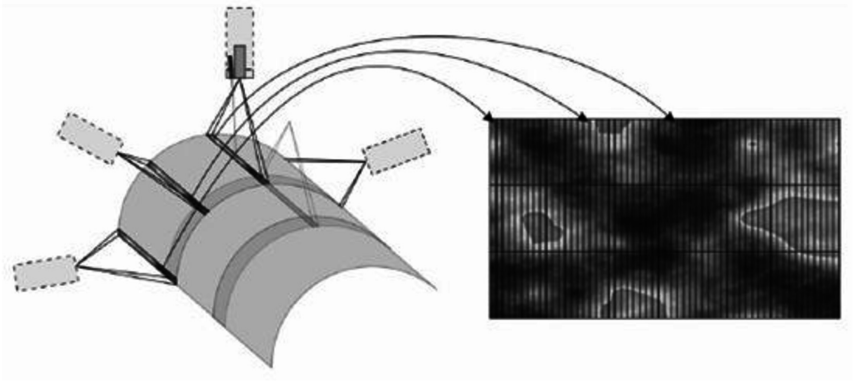


Fig. 2. Belt thermograms and videograms creation [3]

Security of stored data is assured by system of rights based on application predefined roles. Following roles were defined: administrator, physician, student. Administrator has the right to set up user accounts, and to assign rights. Each procedure with patient is in the system registered as a new visit for the given patient, owner of which is physician who carry out the visit. Only the physician who carries out the visit has the right to modify data stored on the visit. Other physicians may only view visits of which they are not the owners. Student role is to view data in read-only mode, without editing rights. Data about the patient and his visits are stored in relational database, the images are stored in a separate repository. In order to ensure the safety any use of any image in the system generates a random id for the given image, and random provisional place of storage, which after user's session is destroyed. Integration of the equipment for registering histograms with computer aided assessment of wound healing process is provided by the layer of services implemented in the system. Its main task is to track of new examinations for registered and new patients and to provide the information to the application for monitoring of wound's healing process. This kind of solution release personnel operating the system from having to remember patient ids, history of examinations and updating cases, which have already been analyzed. The service automatically remembers which tests have been already analyzed and prevents accidental changes of results' interpretation.

2.3 Method of Assessment of Wound Healing Process

The presentation of a burn wound section includes: histogram presentation of thermograpic image of burned area, selection of regions of interest by thresholding of this histogram, imposition of colour masks on a thresholded section, manually selecting the region of interest by imposing a polygonal curve, slice and regions of interest descriptions. Designed graphical interface of wound healing process is shown in Fig. 8.

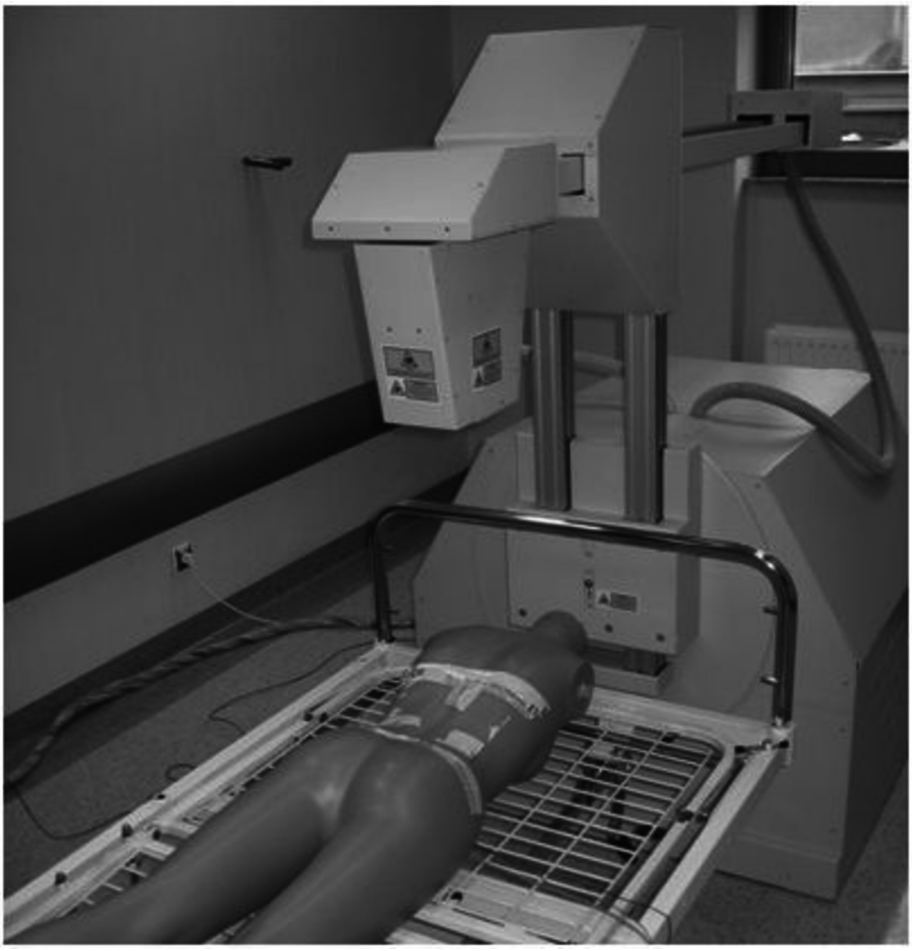


Fig. 3. Equipment for acquisition of thermograms and videograms

'Thermo-Vision' tab allows to add regions of interests. Changing temperature range on histogram makes region of interest apparent (background is darkened). Boundary values can be entered manulally or by moving on a histogram, for the given region surface area and area percentage of the image is calculated. In the main window thermographic image of burnt place is shown. Presentation can be switched between photo and infrared image, or impose the two images on each other semitransparently - there is an option which allows to set up transparency coefficient. This is a characteristic wanted by physicians, which enables visual correlation between wound photo and its infrared photo. On the right-hand side temperature histograms of burn wound are presented: global - upper histogram and detailed - lower histogram. The interface allows you to add region of interest to the analyzed image, which is shown in the table of regions of interest in the

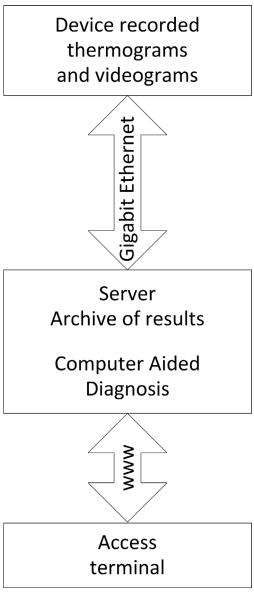


Fig. 4. Diagnostic system structure

Fig. 5. Visits view

upper right-hand corner of the view. Histograms show information for the currently selected region of interest presenting temperature boundary values, which fall within its composition. Temperature boundary values can be changed using the sliders for histograms, where the current threshold value corresponding to the sliderposition is displayed during its move, or manually by entering arbitrary temperature values in the regions of interests table header. For the currently selected thresholds area surface of the region and percentage of the total area of the presented section of the wound in the selected region is calculated. There is also possibility of presentation of a region mask in the form of any color, as indicated by the user, which is also imposed on presented images. In the window two

Fig. 6. View Patient search

Fig. 7. Application navigation bar

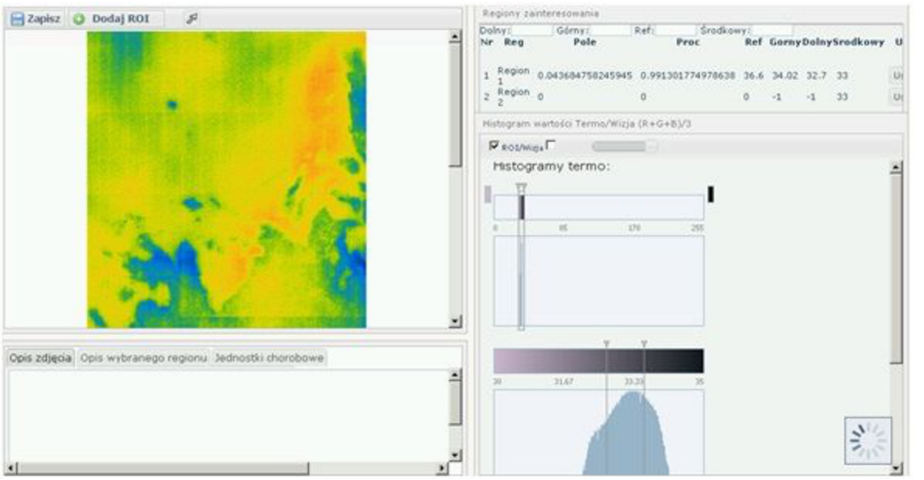


Fig. 8. Thermo-Vision view

histograms: global and detailed are presented in order to allow more precise selection of boundary temperatures for the regions of interest. Detailed histogram range are dependent on boundaries set in global histogram. While a subset of temperatures set on a detailed histogram decides on temperature boundaries for the current region of interest, in the lower part of the interface description of the image, description of the selected region, disease entities, list of burnt places can be added. "Planimetria" tab (Fig. 9) allows area surface calculation for the selected regions of interest. Selection of region of interest is made by boundary

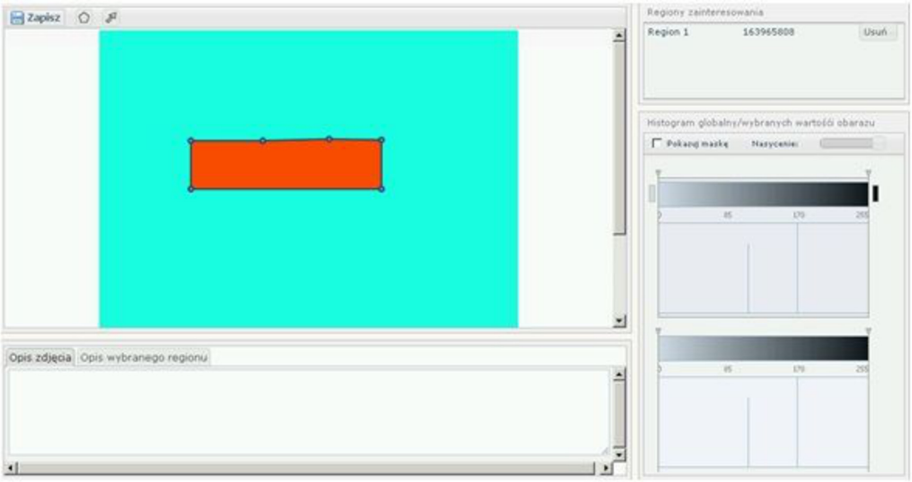


Fig. 9. Planimetria view

points which are set of points of polyline of region of interest. For calculation of surface area rectangle fill algorithm based on parity check has been used. Using histogram we can mask regions which are out of interest. In the lower part of description of the image and description of the selected region can be added.

3 Results

The aim of the study was to design and implement computer application aiding the assesment process of wound's healing process on the basis of infrared images of wound surface. Having regard to the collected requirements the application is designed and implemented in a client-server architecture with ASP .NET MVC technology [8][9]. A prototype has been implemented and tested on a synthetic test images and pilot sample of anonymized infrared photos made on patients in Centre for Burns Treatment in Siemianowice Śląskie. The project focuses on the use of structural (white-box) and functional (black-box) tests. Structural tests rely on testing applications by feeding the input data, in such way that the program has been applied for each path implemented in the code. These rules are defined by the cover criteria for each loop and all possible conditions [10]. Test uses artificially, especially prepared data in order to test program response precisely. Functional tests were conducted by persons associated with the project but who does not know the code they have been testing. Implemented system was approved by clinical physicians. At a later stage, database of clinical cases collection and detailed analysis of collected results is planned in order to develop quantitative criteria for quantitative evaluation of the healing of the burn wound.

Acknowledgement. The author would like to acknowledge financial support from the Polish National Centre for Research and Development - grant: 'Diagnostic system for treatment support of burn and chronic wounds'.

References

1. Nowakowski, A.: Rozwój diagnostyki termicznej metodami detekcji podczerwieni (ilościowa diagnostyka ran oparzeniowych i inne aplikacje). Akademicka Oficyna wydawnicza EXIT (2009)
2. Dziewoński, M.: Wspomaganie oceny postępu gojenia się ran trudno gojących się oparzeniowych na podstawie analizy zdjęć i termogramów. In: XVIII Krajowa Konferencja 'Biocybernetyka i Inżynieria Biomedyczna', vol. 79 (2012)
3. Dziewoński, M., Majchrzak, E., Bachorz, M., Kowalski, P.: Koncepcja urządzenia do rejestracji termogramów opasających. Model. Inż 12, 45–52 (2012)
4. Więcek, B., De Mey, G.: Termowizja w podczerwieni: podstawy i zastosowania. Wydawnictwo PAK (2010)
5. Nowakowski, A.: Postępy termografii: aplikacje medyczne. Wydawnictwo Gdańskie (2001)
6. Pietka, E., Kawa, J., Badura, P., Spinczyk, D.: Open architecture computer-aided diagnosis system. Expert Systems 27, 17–39 (2010)
7. Spinczyk, D.: Prototyp aplikacji do wprowadzania danych pacjenta oddziału oparzeniowego. Zeszyt streszczeń. Polskie Towarzystwo Mechaniki Teoretycznej i Stosowanej. Oddział Gliwice, Komitet Mechaniki Polskiej Akademii Nauk, Katedra Mechaniki Stosowanej Politechniki Śląskiej, pp. 244–245 (2013)
8. Rosenfeld, L., Morville, P.: Information Architecture for the World Wide Web. O'Reilly Media (2007)
9. Zabir, O.: Building a Web 2.0 Portal with ASP.Net 3.5. O'Reilly Media (2008)
10. Wiszniewski, B., Bereza-Jarociński, B.: Teoria i praktyka testowania programów. PWN (2009)

Active Dynamic Thermography Imaging of Wound Healing Processes in Cardio Surgery

Mateusz Moderhak[1], Antoni Nowakowski[1], Mariusz Kaczmarek[1], Piotr Siondalski[2], and Łukasz Jaworski[2]

[1] Gdańsk University of Technology, ul. Narutowicza 11/12, 80-233 Gdańsk, Poland
matmod@biomed.eti.pg.gda.pl
[2] Gdańsk University of Medicine, ul. Dębinki 7, 80-211, Gdańsk, Poland

Abstract. The surgery is a branch of medicine, that is integrally connected to wounds. Despite using sterile tools and compliance with aseptic rolls, some of the surgery wounds become infected. In clinical practice there is a lack of cheap, objective methods and tools for quantitative definition and estimation of the surgery wound healing progress. This paper presents preliminary results of Active Dynamic Thermography (ADT) parametric imaging of sternum surgery wounds. The study was performed in the Clinics of Cardio Surgery at the Medical University of Gdańsk. The results are presented in the form of the simplified exponential parameters approximation of examined objects' thermal response: dT_{norm} and t_{90_10} images.

Keywords: Active Dynamic Thermography (ADT), diagnostics, cardio surgery, wound healing, medical imaging.

1 Introduction

One of the hazards resulting from surgical injury is abnormal wound healing. Comparing to the extent and difficulty of the procedure, the problem of wound healing seems to be marginal and the least important element of the treatment. However, because of the duration of the operation, long-lasting intubation and respiratorotherapy, breakage of the anatomic barriers the number of complications related to the surgical wound healing is 1-5% of operated patients. It is often emphasized that treatment of patient's with complications is much more expensive and longer; their comfort of living is lower and death rate is higher. Authors at the end of 90s of 20th century announced 5% of death rate by serious surgery wound and mediastinum inflammation. Earlier the death rate was estimated even up to 46%.

There are two types of thermograpic cameras depending on the measured band of the thermal radiation: the middle infrared spectral range $3\mu m - 5\mu m$ (MWIR - Medium Wave InfraRed) and far-infrared band $8\mu m - 12\mu m$ (LWIR - Long Wave InfraRed). Due to the greater amount of radiation emitted by a human body at its temperature the LWIR cameras are used [1,2,3,4]. One of

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_18, © Springer International Publishing Switzerland 2014

the most sensitive symptoms of wound inflammation is local growth of temperature value. This observation was noticed in the ancient times and was a symptom helpful in a wound infection diagnosis. In this situation it is surprising that thermal diagnostics has not been used in this field. It is well known that increased metabolism is characteristic for properly proceeded treatment causes local increase of temperature; high temperature can be also a sign of prolonged inflammatory processes. Similarly, local decrease of temperature characteristic for ischemia or process of necrosis can be an important diagnostic sign, too.

These typically functional symptoms can be quite easily visualized by the infrared cameras using methods of Active Dynamic Thermography (ADT). In ADT heating or cooling of the object under test is performed. The diagnostic information is contained in the transitional characteristics of the object's thermal response.

Since the ADT is performed on living subjects there is a need to ensure that the study does not result in its tissue damage and does damage subject's health. This problem is reflected in the tissue stimulation methods. In the case of heat excitation there must be absolute confidence that no painful burns of examined body area are caused. On the other hand, while using cold excitation there is a possibility of chilblain changes as well as different infections. This shows that maintaining proper control of the stimulation source is essential.

To avoid the adverse effects of thermal stimulation we accepted to use the minimum and maximum surface temperature value of the body to which you can bring in the initial phase of the exam. That is 42°C in the case of heating and 28°C in the case of cooling. The most suitable means of thermal stimulation appear to be heating by hot air, radiant heat, and cooling by the stream of cold air or other working gases [5,6,7,8].

2 Methodology

In Fig. 1 typical idea of the ADT imaging unit is presented. It consists of: thermographic camera responsible for the recording of the temperature transient, RGB camera for visual image acquiring, excitation source responsible for inducing objects' local temperature decrease and a PC laptop with software for data

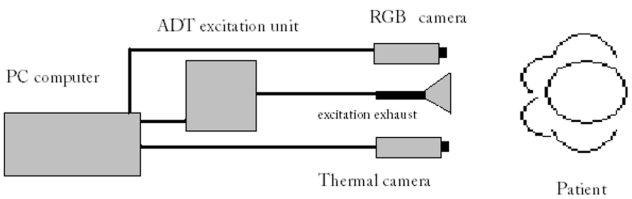


Fig. 1. Scheme of the ADT imaging unit

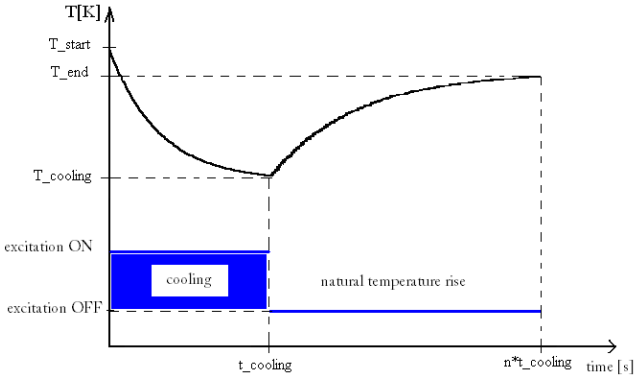


Fig. 2. Typical temperature transient during ADT exam

acquisition and exam control. The typical ADT transient is shown in Fig. 2. Two phases of the process can be distinguished. The first one is the thermal excitation phase (cooling phase) and the other is the natural temperature return phase to its equilibrium. In clinical practice those two phases last: cooling - 60s and the natural return phase - 180s. Temperature values at 180s are not equilibrium state temperatures. While performing ADT exam the best quality of thermographic sequence should be provided. This means that patient and ROI movements should be minimized, object should be centered in in the image as well as thermograms should have good focus. To obtain thermal parametric images further processing of raw thermal data is needed. Image and sequence preprocessing consists of image matching for ROI translations elimination, detection of the excitation range and time transient segmentation for determining the exact parameters of the temperature curve. The method of image matching is FFT spectra based Direct Phase Substitution (DPS) algorithm which is very well suited for thermal images that consist of smooth temperature gradients rather than distinctive features like edges and corners that feature based matching algorithms are based on [9].

Sequence processing diagram is presented in Fig. 3.

Such processed sequence of thermograms is passed on to the input of parametric image calculation algorithm. The idea of dTnorm, t90_10 (1),(2) images is explained in Fig. 4.

$$dT_{norm} = \frac{dT_{return}}{dT_{cooling}} \tag{1}$$

$$t_{90\_10} = t_{90} - t_{10} \tag{2}$$

In Fig. 4:

$T0$ - the initial value of the temperature of the object,
T_{stop} - the temperature value at the time of the experiment of termination,
$T_{cooling}$ - temperature value while the thermal excitation is switched off,

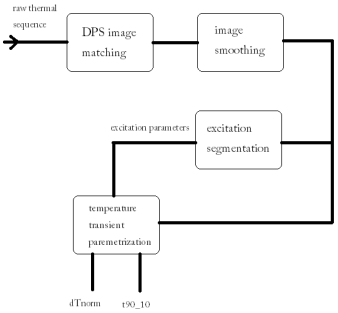


Fig. 3. Thermal sequence processing diagram; at the input there is raw thermal sequence and at the output there are parametric images

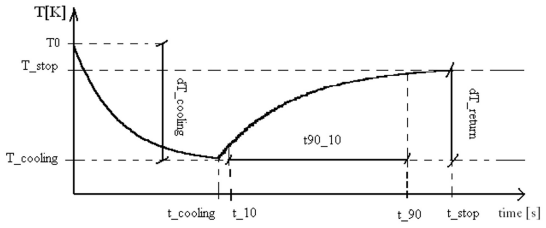


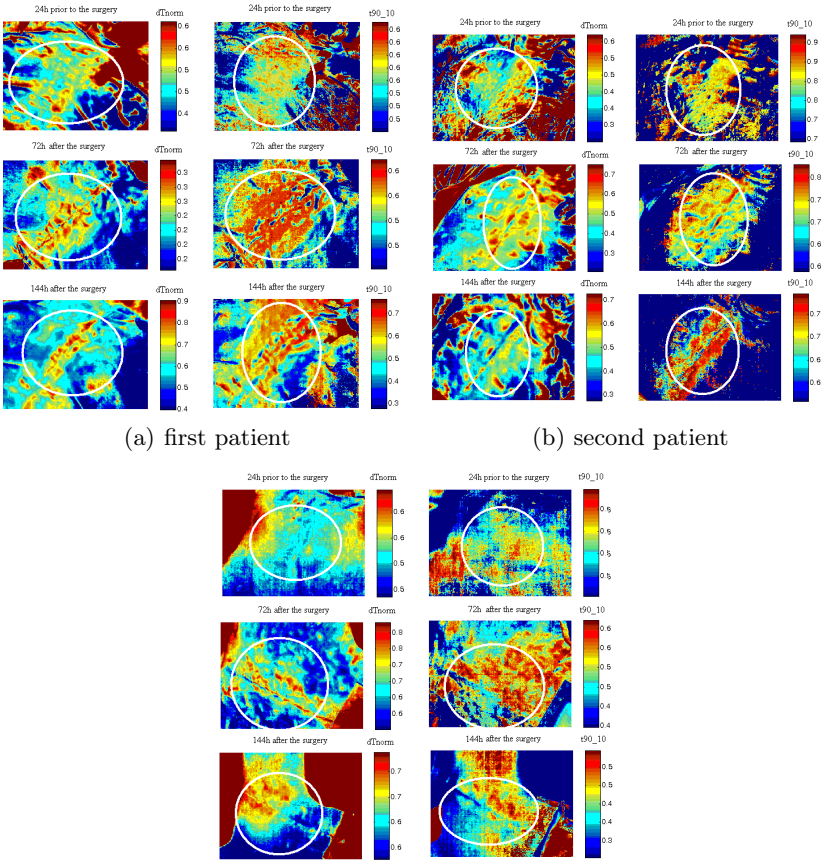
Fig. 4. The idea of thermal parametric imaging presented in this paper

$t_{cooling}$ - the duration of the thermal excitation,
t_{stop} - time moment of the end of the experiment,
$dT_{cooling}$ - total change of temperature value of the object while cooling,
dT_{return} - total rise of temperature value after switching off the excitation,
$t90\_10$ - a difference of time, after which the object temperature reaches 90% of final value dT_{return} and time to reach 10% of this value.

The dT_{norm} is a magnitudal parameter describing relative temperature rise in the phase of natural temperature return, on the other hand t_{90_10} is the temporal parameter that illustrates the speed of regaining the steady state temperature value by the examined object.

3 Results

Figure 5 present the results of sternum wound imaging of 3 patients. ADT exams were performed 24 hours prior to the surgery, 72 hours (3 days) after the surgery and 144 hours (6 days) after the surgery. High values of dT_{norm} parameter indicate the high rate of temperature return. That means that in the natural return phase the temperature reached the value from before the cooling phase. This is usually induced by high thermal activity that can be related with high metabolism, inflammation as well as blood perfusion within the excited volume.



(a) first patient (b) second patient

(c) third patient

Fig. 5. The results of thermal parametric imaging of surgical wound of three patients; the exams were performed 24h prior to the surgery; 72h after the surgery and 144h after the surgery; on the left dT_{norm} images are presented and t_{90_10} on the right; wound area is encircled in white

On the other hand low values of t_{90_10} parameter indicate rapid temperature return of the examined surface. In other words the final temperature at the end of the exam is regained very fast and the speed of the process is related with the same functional features as in the case of dTnorm images.

4 Conclusions

Presented in this paper results of thermal parametric wound imaging show that ADT with its functional imaging capabilities seems to be a good method of surgical wound diagnostics. However to prove its value in this field the results of

well healing as well as complicated wounds should be compared. Then the characterization of the specific features of dT_{norm} and t_{90_10} images aimed at the assessment of wound recovery will be possible. Even though, with ADT parametric imaging we are able to observe physiological processes that occur in wound after the surgery such as: increased blood flow and tissue inflammation. There are no non invasive, objective and fast methods enabling observation of physiological changes in tissues. This is considered to be the main ADT advantage. The diagnosis is based on doctors' experience and clinical guidelines.

References

1. Nowakowski, A. (red.): Postępy termografii – Aplikacje medyczne. Wydawnictwo Gdańskie (2001)
2. Więcek, B., De Mey, G.: Termowizja w podczerwieni. Podstawy i zastosowania. Wydawnictwo PAK (2011)
3. Minkina, W.: Pomiary termowizyjne: przyrządy i metody. Wydawnictwa Politechniki Częstochowskiej (2004)
4. Maldague, X.: Infrared methodology and technology. Gordon and Breach Science Publishers (1992)
5. Rumiński, J., Kaczmarek, M., Renkielska, A., Nowakowski, A.: Thermal parametric imaging in the evaluation of skin burn depth. IEEE Transaction on Biomedical Engineering 54(2), 303–312 (2007)
6. Renkielska, A., Nowakowski, A., Kaczmarek, M., Rumiński, J.: Burn depths evaluation based on active dynamic IR thermal imaging – A preliminary study. Burns 32, 867–875 (2006)
7. Renkielska, A., Nowakowski, A., Kaczmarek, M., Dobke, M., Grudziński, J., Karmoliński, A., Stojek, W.: Static thermography revisited – An adjunct method for determining the depth of the burn injury. Burns 31(6), 768–775 (2005)
8. Nowakowski, A., Kaczmarek, M., Bajorek, M., Moderhak, M., Suchowirski, M.: Advances of IR–Thermal diagnostics in medicine. In: Diagnosis of Processes and Systems, pp. 85–104. PWNT, Gdańsk (2009)
9. Moderhak, M.: FFT spectra based matching algorithm for active dynamic thermography. Qirt Journal 8(2), 239–242 (2011)

About the Measurement Methods in Music Therapy

Maria Skotnicka and Andrzej W. Mitas

Faculty of Biomedical Engineering, Silesian University of Technology, Zabrze, Poland
{maria.j.skotnicka,andrzej.mitas}@polsl.pl

Abstract. The chapter begins with a necessary introduction describing the fundamental importance of music in human life. Sounds of nature and typical music created by humans according to accepted rules (that are also nature philosophy derivatives) are considered here. Nowadays, music is used to intentional change the psychophysical state of a human being. It would be relatively easy to indicate that music stimulates activities, it is much more complicated to select such types of sound stimulation, which would be useful for recreation - music therapy is meant here. Currently, the choice is rather heuristic than algorithmic.

For a holistic analysis a snapshot of formal methods known in psychology is provided, such as Faces Pain Scale, State - Trait Anxiety Inventory (to assess anxiety), Rating Anxiety in Dementia (effective studying the influence of quiet music), Relax Rating Scale (a simple instrument to assess the relaxation degree), and techniques commonly known as biomedical measurement such as ECG , HR, BP , EEG, EEG or temperature measurement.

In the following section a carefully selected case study, performed to verify the results presented in the literature, was described. In the summary and conclusions the results of own works and the directions of further research are presented.

Keywords: biomedical measurements, music therapy, psychophysiological condition.

1 Music for Creating of Human Specified Reaction – Introduction

Music surrounds us every day, in every place and in every situation. It is the background for our emotional experiences in the cinema or in the theater, it spurs us to dance at the parties, it accompanies us at work as it is transmitted by radios.

Music also soothes our senses and fulfils a variety of functions such as serving as a component of the culture, it takes part in the daily life of man for ages. In the past, music was served not only to inform about the hazards or call to battle but also to induce state of relaxation. The first traces of discussions about music and its educational and social values can be found in the ancient history of Near and Far East [22].

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_19, © Springer International Publishing Switzerland 2014

Nowadays music is used to make travelling more enjoyable and it is a form of entertainment. It is also used to induce specific reactions (for instance making someone relax, enter into a reverie) increasingly.

The field that deals with the treatment of music is called music therapy. One of its tasks is to improve the mental health of humans via relaxation. Unfortunately, music therapy has an incommensurable nature, we do not know how music affects people. The division of music which is used in music therapy is not standardized, classical music is often used as well sounds coming from nature (such as the singing of birds, the noise of a stream or thunderstorms), but there are also new songs which are being created in order to induce the human relaxation state. Music therapists usually use their experience and intuition to observe the patient during treatment, and on that basis decide whether the selected compositions are appropriate or if it should be changed [12,13].

The fact that music affects our bodies is undeniable. Sounds that we receive by our hearing system or feel by the whole body are able to induce in us sadness, gladness or thoughtfulness. There are also more objective evidence of music impacts in the human body, such as changes in the rhythm of our heart rate, in the body temperature and in the breathing speed.

Influence of music on humans can be assessed subjectively - by the opinions of the listener or objectively - by measuring physiological parameters or concentrations of chemical substances (e.g. hormones) in the blood.

2 Standardized Psychological Tests – The Secondary Indication of Human Responses

The impact of music therapy can be evaluated using patients opinion about their frame of mind. There are being used some standardized psychological tests, to make the assesment more unambigous.

The **Faces Pain Scale (FPS)** allows children to evaluate their pain. The FPS is a self-report measure. It consists of 6 schematic faces which express different pain intensity (Fig. 1) [7].

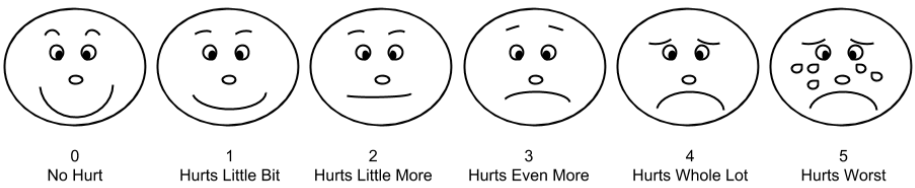


Fig. 1. Faces Pain Scale

This test has been used, among others, to assess the effects of interactive music therapy in 65 hospitalized children with cancer. Within 4 months in every child it was carried out from 1 to 3 trials of music therapy sessions lasting from 15 to 45 minutes. The sessions consisted of creating the songs, singing,

listening to recorded songs or songs intended for younger children. Children and parents participated in the selection of music and the instruments used during the session. Before and after each musical session children rate their feelings using the FACES scale. Analysis of the tests showed an improvement in children feelings experienced after music therapy [1].

Another test is the **State-Trait Anxiety Inventory (STAI)** which is used to assess anxiety. It measures two types of anxiety - the first one applies to state anxiety (temporary and situationally conditioned state) the second one - to trait anxiety (relatively stable personality trait). The STAI consists of 40 questions with 4 answers, it is a self-report test [28]. This survey has been used, among others, to assess the impact of music therapy on patients undergoing cataract surgery and it was conducted among 121 elderly people. Before the surgery, all patients completed a STAI questionnaire. During treatment everyone had to wear headphones and listened to one of the four recording from cassette tapes (depending on the affiliation to a particular group). One group auditioned the recordings of relaxing suggestions repeated at 1-2 minute intervals, the second group listened to white noise, the third one pre-recorded sounds from the operating room, and the fourth group was exposed to relaxing music which consisted of "classical composition accompaniment of soothing sounds of nature." After the surgery participants filled STAI again. Anxiety decreased after surgery in groups listening to relaxing music and relaxing sentences. Patients from relaxing music group "were more satisfied with the experience of operating and attendant music than those belonging to other groups and, what is more, vast majority of them (31 to 32 persons) declared that would have gone again to listen to music during surgery" [8].

In the same study ([8]) there was used the **Visual Analog Scale (VAS)** too. It is a measurement instrument for evaluating chronic pain intensity. Most patients cannot discretize pain intensity (e.g., describe it using integers from 0 to 10) for this reason the VAS is usually a horizontal line, 100 mm in length, with the extremes labelled ”least possible pain” and ”worst possible pain” (Fig. 2). The measurement lies in selecting by a patient a point on a line the most corresponding to pain intensity [3].

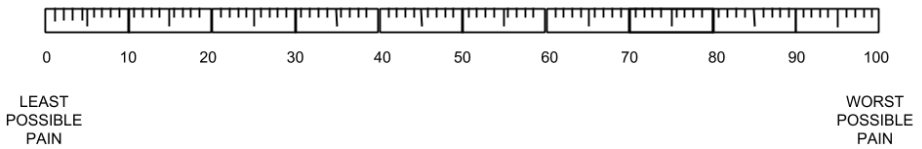


Fig. 2. Visual Analog Scale

The **RAID (Rating Anxiety in Dementia)** is an instrument developed to use with people with dementia. This questionnaire consists of 20 items, each scored on a 4-point scale. The questions concern e.g. the worry about physical health or finances, sleep disturbances and somatic symptoms. The clinician completes a survey based on observations, medical records and conversation with the patient [24]. This survey was used in the study of anxiety of the elderly (over

65 years) with dementia in care homes. Groups of 29 people for 6 weeks twice a week for half an hour were listening to relaxing music (during the continuation of the standard medical care). In the control group there were 23 individuals who were only on standard care. The relaxed music was selected on the basis of Music Preference Survey (this is a survey on the importance of music in human life, about favorite types of music, artists and songs). The most frequently chosen type was regional popular music of the teenage years of the patients. To assess the state of anxiety it was used the RAID before and after the treatment - survey was completed by doctors or nurses, who described the condition of patients. After 6 weeks of the music therapy there was a significant reduction of anxiety in the research group. In the control group reduction of anxiety was observed too, but it was not significant [25].

The **Relaxation Rating Scale (RRS)** is a simple self-reported instrument that is used to assess the degree of relaxation. In using the RRS the patient needs to select a point in the 10-point scale of the degree of relaxation. The beginning of the scale (value 1) is described as "not working" and the end (value 10) - "completely relaxed". This questionnaire was used to assess the effects of music therapy on women who suffer from arachnophobia. Each of the 60 women who participated in the study, filled RRS and then the subjects were divided into 3 groups. The first one was listening to own relaxing music with headphones for 10 minutes (the music was considered by participants to be the most relaxing), the second group in the same time was listening to classical music (2 fragments of Bach concert), while the third group (the control group) stayed at this time in silence. After listening to the recordings participants filled RRS again and then each of the ladies looked at the 30 realistic images of spiders. Images were chosen in such a way that each next picture influenced the examined person increasingly, this lasted about 5 minutes. In the next stage, the **Behavioral Avoidance Test (BAT)** was performed, it is used to measure anxiety in arachnofoby. This test consists of measuring the distance at which the test person will be closer to terrarium with tarantulas. At the end each of the ladies fulfilled the Experienced Anxiety Scale, which assessed the anxiety experienced during the behavioral test. The highest increase in the perceived level of relaxation (based on RRS) occurred in the group with their own music, in a group of classical music this increase was significantly lower. In the control group there was a decrease in the RRS [2] which means the relaxation decreased.

3 Biomedical Measurements as a Proposal of Objective Methods of Assessing the Music Impact on Human Body

For more objective evaluation of the music therapy impact on human body we should use measurements of parameters which are changing without our awareness. Completing the tests we may choose responses that are expected instead of those that are true. In this case, the assessment of effectiveness of the therapy will be incorrect. To avoid this (instead use tests) we can use biomeasurements

such as electrocardiography (ECG), heart rate (HR), blood pressure (BP), galvanic skin response (GSR), electromyography (EMG), electroencephalography (EEG) and temperature. All recorded signals are processed and then analyzed (Fig. 3).

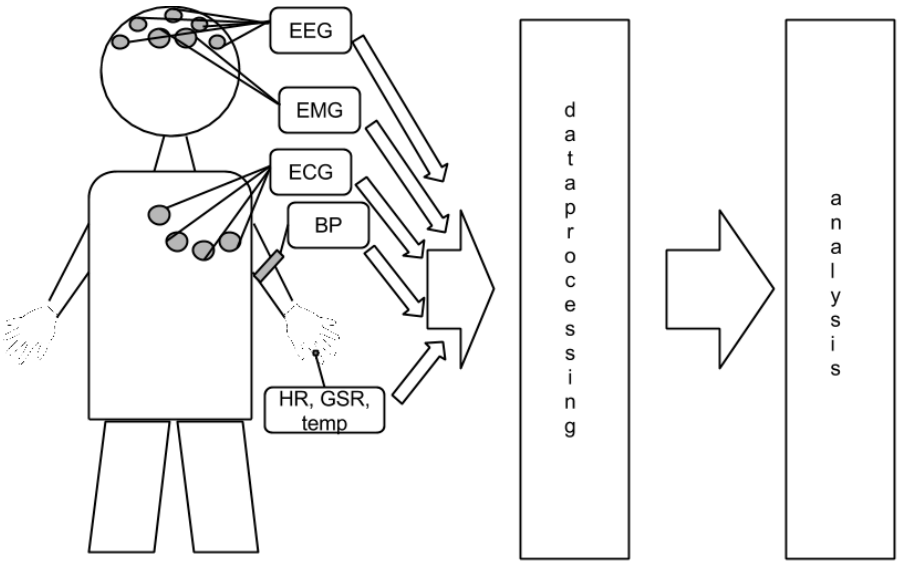


Fig. 3. Registration of biomedical measurements

Electrocardiography is the basic method of the heart activity measurements. Voltages which are derived from stimulated cardiac cells can be measured at the skin surface. The ECG signal shows the phases of the heart depolarization and repolarization. On the basis of the ECG amplitudes, time relationships of each wave and the length of sections lying between waves are determined. Changes in these values enable the evaluation of the electrical heart activity at a certain time [27]. ECG was used to evaluate the performance of the music in [11] wherein the autonomic activity were evaluated. The study was conducted on a group of 10 people, each of them was listening to New Age music, designer music and rock for 15 minutes and there was a two-day break between each session. The survey consisted of 15 minutes of monitoring ECG at rest then for 15 minutes the subjects listened to the particular type of music at constantly monitored ECG. The control group remained for 15 minutes in silence. During the designer music session, the total autonomic activity increased (the assessment was made, based on the analysis of ECG signal). For the two other types of music there were no changes in autonomic activity.

Heart systole generates pressure that allows the flow of blood through the bloodstream. Blood is pumped with a pressure of 90-140 mm Hg from left ventricle into the aorta [27]. **Blood pressure** was applied in [8] together with evaluating the music performance using questionnaires (described above). Systolic blood pressure increased during the first period of measurement on groups that listened to the relaxing suggestions, white noise, and relaxing music. In the group that auditioned the recording of sounds coming from the operating room systolic blood pressure decreased in 2 and 3 section of the test. Blood pressure as an indicator of the music action was also used by Chiu et al [6]. They studied a group of 68 hospitalized people waiting for surgery. It was based on the measurement of basic physiological parameters before and during music therapy, which was to lower the anxiety in patients. Each of the patients was lying in a dark room at operational bed to stabilize vital signs for 5 minutes. Their blood pressure was measured immediately after the rest. In the next stage of the study, patients were divided into two equal groups, one group listening to the sounds of nature for 5 minutes, and the second group remained in silence. After this stage blood pressure was measured again. The analysis showed that "blood pressure is a wrong indicator" because its value decreased in both groups.

Heart rate (number of systoles per minute) was another parameter used in [8] to assess the effectiveness of music therapy on patients during surgery. There were no differences in heart rate values for the groups stimulated with different types of sound. Measurement of heart rate was also used in [5] (study described below). In this case differences in HR values were noticed - in the group listening to classical music was observed slower rhythm than in the control group and listening to their own music.

Electroencephalography is a method to study the spontaneous electrical activity of brain neurons. EEG is recorded on the surface of the skull [27]. In [9] has been shown the impact of music (assessed on the basis of the EEG signal and quantity of tasks performed) on the work effectiveness of people of varying level of perceived anxiety. The study involved 25 people, where 15 people were classified into groups with normal levels of anxiety, and 10 people to a group of high anxiety. Musical stimulation were classical music (Tchaikovsky, Grieg and Mozart) and rock (The Rolling Stones) with variable intensity of impact. EEG was monitored during the task (recognition of distorted images) both in silence and while listening to music. Analysis of EEG allowed to conclude: "while working music causes stimulation of the brain (in a group of people with a normal degree of anxiety) or reductions in brain activity (in the group with a higher anxiety)".

Body **temperature** (in particular the internal temperature) is subject to self-regulation in warm-blooded organisms. The internal temperature of the body is the temperature in the central area of the head or torso. Body temperature can vary, it depends on the time of the day, physical activity, emotional state and environment temperature [27]. Attempting to use the temperature measurement to assess the impact of music therapy on the mental and physical condition of people exposed to stress factor has been described in [5]. The study involved 60

people, who were divided into four groups, which during the experiment listened to classical music, hard rock, music indicated as relaxing or remained in silence. At the beginning of the study all participants were informed about the coming stressful test (raising the level of arousal). Thermistor was dressed to the left index finger (for temperature measurement). For 10 minutes the temperature was monitored in silence, then continue measuring for 10 minutes while listening to a particular kind of music and for the last 10 minutes measurement in silence was repeated . There were no significant differences in the temperature measurement.

In [26] it was described the change of **skin galvanic response** while watching movies with different audio tracks. A group of 60 people was divided into 3 subgroups, one of which was the control group was watching a movie without music the other was watching the movie with "documentary soundtrack" which aimed at reducing the perceived emotions and the third one was watching a film with "horror soundtrack" (which was aimed to strengthen emotions). For the first 10 minutes of the test GSR was recorded in silence, for the next seven minutes the subjects were watching a blank screen, then for 12 minutes it was displayed a black-and- white film about industrial security depicting three accidents, finally re- recorded signal at the 7 minute observation blank screen. GSR signal was measured at all stages of the study. The biggest conductivity was registered in the group watching a movie with "horror soundtrack", lower in the control group, and the lowest conductivity of the skin has occurred in the group of people watching a video of "documentary soundtrack".

EMG (**electromyography**) was used to assess the effects of music impact on the human body for example in the study described in [10]. The study has used two kinds of music: the flute music and hard rock music. The frontal muscle activity was measured. In the first stage of the experiment there was EMG signal measurement in silence for 5 minutes , for a further 5 minutes (continuing the EMG registration) examined listened to the flute music (using headphones), then repeated 5 minute registration in silence, and at last 5 minutes a man listened to hard rock music. It was determined power spectral density (PSD) in EMG signal, which is the Fourier transform of the autocorrelation function of EMG signal (1).

$$S_{xx}(\omega) = \int_{-\infty}^{\infty} \gamma(\tau) e^{-i\omega\tau} d\tau \tag{1}$$

PSD decreased while listening to the sounds of the flute as compared to the value recorded in silence. While listening to hard rock music the power spectral density increased again.

4 Case Study – Rhythmic Stimulation

Most of the studies describe the effect of the whole composition to human. During such stimulation there are changes in many parameters of music, therefore it is difficult to clearly assess music impact on human body.

In the following study it is described the effects of decreasing music rate on heart rate and blood pressure, these parameters are simple to acquisition and previous studies have shown their variability.

In the five-minute composition the sound of a gong of release tempo was used. The gong sound was chosen because it has a distinctive sound that humans know from the beginning of time. The initial stimulation tempo has a value of 70, which corresponds to the normal heart rate. For 4 minutes of the composition the tempo decreases of 3 units per minute, to the value of 58. For last minute the gong resounds at the same tempo (Fig. 4).

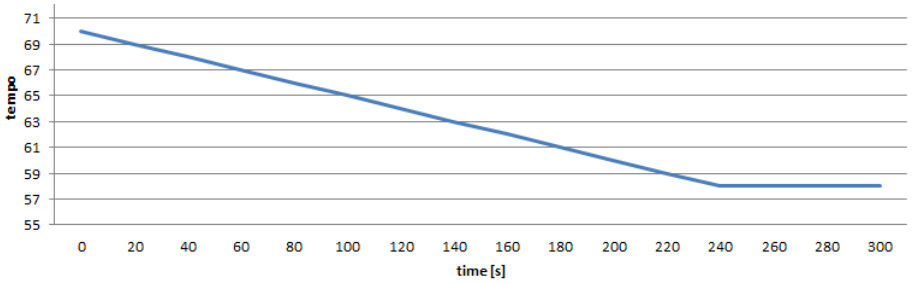


Fig. 4. Graph of tempo changes during the composition

The study was conducted on a group of 30 people aged between 20 and 24 years. The total study time was approximately 15 minutes (Fig. 5).

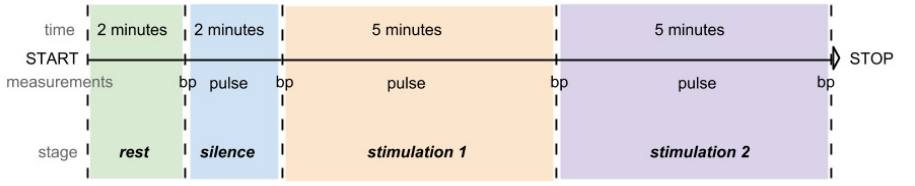


Fig. 5. Study stages

Patients were sitting on a chair in a comfortable position, they decided whether they wish to listen to recording with closed or open eyes. Plethysmograph sensor was placed on the index finger on the right hand. Sleeve pressure gauge was assumed on the left arm. The study begins with two minutes of rest, then the blood pressure is measured for the first time. For the next two minutes the patient's pulse is recorded in silence. The blood pressure is measured again. Further the patient is sitting in a comfortable position and is listening to the 5-minute composition, pulse is recording during this section, after stimulation blood pressure is measured. At the end the last two stages of the study are repeated (Fig. 6).

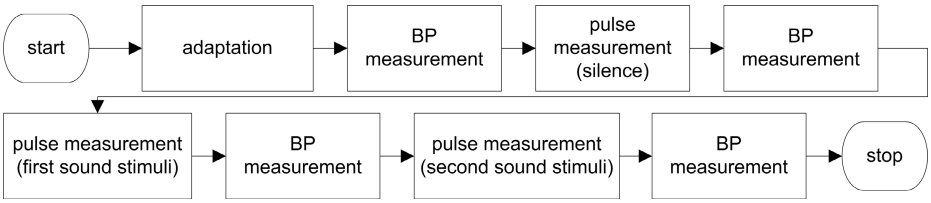
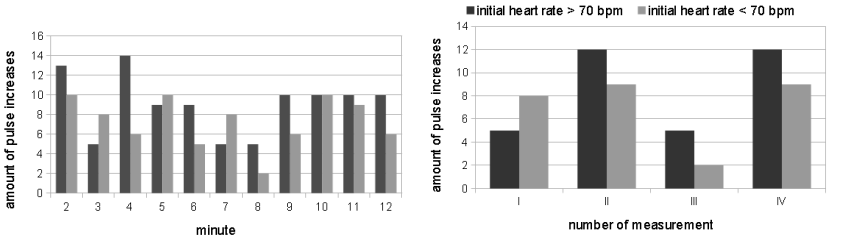


Fig. 6. Study design

After rejecting the extreme values of the initial pulse (3 cases) patients were divided into 2 groups - a group of initial pulse above 70 beats per minute (15 people) and a group of initial pulse below 70 bpm (12 people).

Figure 7(a) shows the changes of the pulse in the subsequent minutes of the test. Columns that are described as the second minute of the study show the number of people with increase in heart rate in the second minute of silence compared to the first minute. Third column describe changes in the first minute of the sound stimulation. Eighth column shows at how many people pulse increase in first minute of the second sound stimulation (70 bpm) compared to the last minute of previous stimulation (58 bpm).



(a) Changes in the subsequent minutes of the test

(b) Changes in the subsequent stages of the test

Fig. 7. Changes of the pulse

In the group with higher initial pulse the first minute of stimulation (gong sound with 70 bpm tempo) caused a decrease in heart rate and in the group with lower initial pulse - it grows. The first minute of the second stimulation (8th column; 70 bpm sound stimulation after 58 bpm stimulation) resulted in a decrease of the pulse in both groups, regardless the heart rate in the previous minute (higher or lower than 70 bpm).

There is no significant changes in pulse between each minute of the same stimulation. Figure 7(b) shows the amount of people whose heart rate increased in the first minute of each sound stimulation (first and third column) and during the whole stimulation (second column it is the number of pulse growth in 5th minute of the first stimulation compared to the 1st minute of this stimulation, fourth column is the same about second sound stimulation).

At the end of the first stimulation (second column) in both groups it was observed a pulse increase in comparison to the first minute of the stimulation (21 cases out of 27). In the last minute of the second stimulation (forth column) it was also observed a pulse increase relative to the first minute of the stimulation (in 21 cases out of 27 - it was a different group of people than the first one).

There was no significant changes in blood pressure during the entire study.

5 Conclusion and Future Work

The studies, which are briefly characterized in the chapter, were inspired by the eternal and till now unfulfilled need to know what is exactly the impact of sound (acoustic waves) on the human being. It is done a review of the basic techniques used for the evaluation of human responses to music. An important conclusion from the literature analysis, requiring further experimental research, is that music can cause the brain activity by people with normal levels of anxiety, just in opposite to them with greater anxiety, (reducing brain activity). In this context, also referring to well known references that the music tempo changes cause changes in heart rate, series of tests were done. They are described in details in the article.

Our case study was to examine the impact of tempo changes on heart rate and blood pressure. Gong was selected for this purpose because its sound is known for a very long time, and often it is used in many types of music therapy. During the experiments the frequency was reduced from 70 beats per minute to 58. The test time was limited to 15 minutes. A plethysmograph sensor and blood pressure were used after a two-minute rest and after the whole stimulation with music. The patients group were divided into two sections, depending on the significant reaction trend. The changes in heart rate were correlated with tempo changes, while the changes in pressure have no special significance. The results are shown in the corresponding figures.

The data analysis indicates that the sensitivity to tempo changes of sound stimulation is rather an individual feature. On the basis of the achieved results some detailed observations were done. They lead to the conclusion that the response to music stimulation may be classified as an important method of behavioural biometrics [23,18,17,4].

The topic for next researches could be formulated on the statement (published in the literature and indirectly confirmed by own experiments) that the greatest increase (subjectively perceived) level of relaxation occurs in a group of patients with their own relaxing music. This confirms the desirability of this research in the field of music therapy, aimed at the more objective selection of music, depending on the specific psychophysiological features of human being. This is an important indicator that the universal relaxation music rather does not exist and its matching should be supported by a detailed analysis of the human reaction.

A separate field of researches is the optimisation of the conditions for the driver [19,16,15,14,20,21]. Biocybernetic aspect of transport safety involves the

evolution of psychophysiological state of the driver also by appropriate stimulation (e.g. sound stimulation) to reach a positive impact on the driver.

References

1. Barrera, M.E., Rykov, M.H., Doyle, S.L.: The effects of interactive music therapy on hospitalized children with cancer: a pilot study. Psycho-Oncology 11(5), 379–388 (2002)
2. Bartoń, J., Blaut, A.: Efektywność muzyki relaksacyjnej (klasycznej i wybranej samodzielnie) w terapii behawioralnej kobiet z arachnofobią. Psychoterapia (4), 41–50 (2011)
3. Bijur, P.E., Silver, W., Gallagher, E.J.: Reliability of the visual analog scale for measurement of acute pain. Academic Emergency Medicine 8(12), 1153–1157 (2001)
4. Bugdol, M.D., Mitas, A.W.: Multimodal biometric system combining ECG and sound signals. Pattern Recognition Letters (38), 107–112 (2014) ISSN 0167-8655
5. Burns, J.L., Labbé, E., Arke, B., Capeless, K., Cooksey, B., Steadman, A., Gonzales, C.: The effects of different types of music on perceived and physiological measures of stress. J. Music Ther. 39(2), 101–116 (2002)
6. Chiu, H.W., Lin, L.S., Kuo, M.C., Chiang, H.S., Hsu, C.Y.: Using heart rate variability analysis to assess the effect of music therapy on anxiety reduction of patients 30, 469–472 (2003)
7. Hicks, C.L., von Baeyer, C.L., Spafford, P.A., van Korlaar I, G.B.: The faces pain scale-revised: toward a common metric in pediatric pain measurement. Pain 93(2), 173–183 (2001)
8. Cruise, C.J., Chung, F., Yogendran, S., Little, D.: Music increases satisfaction in elderly outpatients undergoing cataract surgery. Can. J. Anaesth. 44(1), 43–48 (1997)
9. Frolov, M.V., Milovanova, G.B., Mekhedova, A.Y.: Effect of musical accompaniment on the operation performance of subjects with different anxiety levels. Human Physiology 31(2), 164–170 (2005)
10. Kara, S., Ozel, P.: Determination of effects of different music on frontal muscle by using emg signal. In: Frontiers in the Convergence of Bioscience and Information Technologies, FBIT 2007, pp. 449–452 (2007)
11. Mccraty, R., Atkinson, M.I.K.E., Glen, R.E.I.N., Watkins, A.D.: Music Enhances the Effect of Positive Emotional States on Salivary IgA. Stress Medicine 12(3), 167–175 (1996)
12. Mitas, A.W.: About the possibilities of the informatics technologies application in music therapy. Journal of Medical Informatics & Technologies 12, 189–193 (2008)
13. Mitas, A.W.: Stimulation methods in music therapy - short discussion towards the bio-cybernetic aspect. Journal of Medical Informatics & Technologies 13, 255–258 (2009)
14. Mitas, A.W., Bugdol, M., Ryguła, A.: Computer aid assessment of driver's fatigue during driving based on eye movements analysis. Journal of Medical Informatics & Technologies 12, 195–200 (2008)
15. Mitas, A.W., Bugdol, M., Ryguła, A.: The psycho-physiological conditionings of driver's work under the aspect of traffic safety. Transport Problems 4, 87–94 (2009)
16. Mitas, A.W., Bugdol, M., Ryguła, A.: Simultaneous analysis of driver's physiological and behavioural parameters under the aspect of transport safety. Journal of Medical Informatics & Technologies 13, 241–247 (2009)

17. Mitas, A.W., Bugdol, M.D.: Strengthening a cryptographic system with behavioural biometric. In: Piętka, E., Kawa, J. (eds.) ITIB 2012. LNCS, vol. 7339, pp. 266–276. Springer, Heidelberg (2012)
18. Mitas, A.W., Bugdol, M.D., Konior, W., Ryguła, A.: Phase angles of sound as a biometric feature. In: Piętka, E., Kawa, J. (eds.) ITIB 2012. LNCS, vol. 7339, pp. 256–265. Springer, Heidelberg (2012)
19. Mitas, A.W., Ryguła, A.: Biomedically aided car driver safety system. In: Piętka, E., Kawa, J. (eds.) Information Technologies in Biomedicine. AISC, vol. 69, pp. 229–236. Springer, Heidelberg (2010)
20. Mitas, A.W., Ryguła, A., Pyciński, B.: Driver's biocybernetic monitoring and early warning system. In: Mikulski, J. (ed.) TST 2012. CCIS, vol. 329, pp. 52–59. Springer, Heidelberg (2012)
21. Mitas, A.W., Ryguła, A., Pyciński, B., Bugdol, M.D., Konior, W.: Driver biomedical support system. In: Piętka, E., Kawa, J. (eds.) ITIB 2012. LNCS, vol. 7339, pp. 277–285. Springer, Heidelberg (2012)
22. Mitas, A.W.: Spectral analysis in stimuli art predefining in music therapy. In: International Conference on Biometrics and Kansei Engineering, ICBAKE 2009, pp. 123–126 (2009)
23. Mitas, A.W.: Behaviorual and physiological biometrics advantages and limitations. In: Biometrics 2010, vol. 4, pp. 153–168 (2011)
24. Seignourel, P.J., Kunik, M.E., Snow, L., Wilson, N., Stanley, M.: Anxiety in dementia: a critical review. Clin. Psychol. Rev. 28(7), 1071–1082 (2008)
25. Sung, H.C., Chang, A.M., Lee, W.L.: A preferred music listening intervention to reduce anxiety in older adults with dementia in nursing homes. Journal of Clinical Nursing 19(7-8), 1056–1064 (2010)
26. Levenson, R.W., Thayer, J.F.: Effects of music on psychophysiological responses to a stressful film. Psychomusicology: A Journal of Research in Music Cognition 3(1), 44–52 (1983)
27. Torbicz, W., Filipczyński, L., Maniewski, R., Nałęcz, M., Stolarski, E.: Biocybernetyka i Inżynieria Biomedyczna 2000 Tom 2 Biopomiary. Akademicka Oficyna Wydawnicza Exit (2011)
28. Wrześniewski, K.: Inwentarz stanu i cechy lęku STAI: polska adaptacja STAI: podręcznik. Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego (2011)

Eye Tracking Method in Low Resolution Video Recordings Using Fixed Reference Objects

Branko Babusiak and Jan Barabas

University of Zilina, Slovakia,
branko.babusiak@fel.uniza.sk, jan.barabas@uniza.sk

Abstract. This article details our developed algorithm for eye tracking in low resolution video recordings, especially those with sub-VGA resolution and using lossy compression. The proposed method for eye tracking is an important step towards the successful detection of eye blink artifacts which result in false delta activity being recorded in EEG records. Once detected, we can automatically remove the erroneous delta activity during pre-processing of the EEG records and ultimately improve the precision of automated EEG analysis.

Keywords: eye tracking, low resolution videos, EEG analysis.

1 Introduction

Electroencephalography is a neurophysiological approach to measuring the electrical activity of brain using electrodes placed on the head surface or, in special cases, within the brain itself. Results are presented in the form of electroencephalograms (EEG), which are a graphical representation of recorded brain waves. EEG is used in many areas, especially in brain damage assessment, epilepsy and various other physiological problems. However, many types of artifacts are present in EEG recordings, significantly degrading the informational value thereof. Identification and removal of the said artifacts is crucial for a successful analysis. In view of the said, this article will deal with a method to detect eye blink artifacts, based on eye localization with respect to head position. Eye blinking introduces erroneous delta or theta waves, which are undesirable during automated EEG processing and require additional attention to pinpoint their source.

2 Eye Tracking Method

Most neurology departments are equipped with a camera monitoring system recording the patient session [1]. The aim of the recording is to monitor patient movement and state during EEG analysis. The recordings are archived for a certain amount of time and thus, in order to save disk space, the recordings are made with a lower resolution, frame rate and more aggressive lossy compression. Parameters chosen for the purpose of this article are shown in Fig. 1.

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_20, © Springer International Publishing Switzerland 2014

The recording consists of two camera feeds – a color feed visualizing the whole patient body and a gray scale feed of the patient face. The compression caused artifacts in the form of blocks, complicating and in certain cases also making it impossible to use effective methods to extract characteristic picture properties. Our eye tracking algorithm requires a 152x126px region of interest from the gray scale feed.

File Size	97 703 480 bytes
Number of Frames	11 597
Frames per Second	14.9914 s^{-1}
Time Length	12 min. 52 sec.
Width	352 pixels
Height	288 pixels
Image Type	Truecolor
Video Compresion	DivX50

Fig. 1. Recording sample (left) and parameters thereof (right)

Readily available methods for the detection of eye position, e.g. [2,3], were not possible due to the low quality of the recording. Thus a new and effective method was necessary. Our proposed algorithm uses three reference objects (Fig. 2) placed in the vertices of an equilateral triangle. Using our previously published method [4,5] we are able to calculate the center of reference objects. Based on their mutual position we are able to determine the rotation of the triangle in all three axes. The reference object set is affixed to the head of the patient and thus the triangle rotation is proportional to the rotation of the head. Upon determining the rotation of the head we can calculate the eye location.

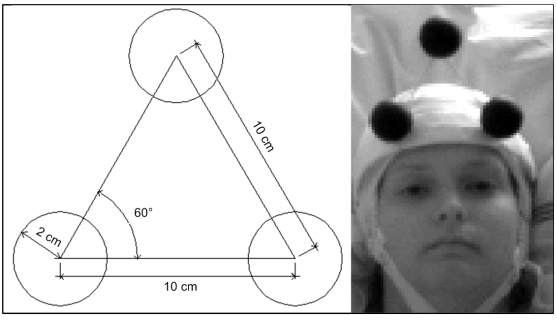


Fig. 2. Location of circular reference objects in the vertices of an equilateral triangle(left). Triangle position on the patient head (right)

2.1 Definition of the Three-Dimensional Model

In order to determine head rotation we must first define a three-dimensional model of the head including the reference triangle. Distances in the model are relative to the key points and respect the general anatomy of the head. Chosen key points include: eyes, centers of reference objects and center of axial system. The center of axial system is located approximately in the location of atlas and axis vertebrae (C1 and C2), allowing the rotation of the head (Fig. 3). Distances in the model are approximate and are calibrated based on individual patient anatomy obtained from reference frame.

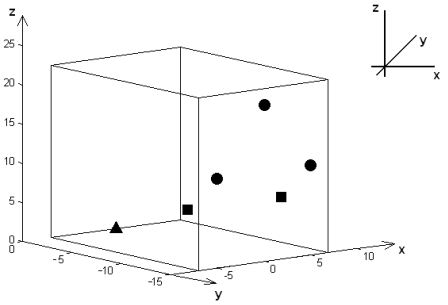


Fig. 3. General three-dimensional model showing the location of eyes (square marker), centers of reference objects (round markers) and center of the axial system (triangular marker). Distances are in centimeters

Rotation of the reference points and eye points around all three axes results in revised point locations (coordinates) in three-dimensional space. Point rotation is accomplished using three basic matrixes, each representing rotation around a given axis. The resulting point coordinates after rotation $[x', y', z']$ can be obtained by multiplication:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R(\varphi) \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{1}$$

Where R represents one of the rotational matrices based on the chosen rotation axis, φ is the rotation angle and $[x, y, z]$ are starting coordinates.

2.2 Parallel Projection

Points in the three-dimensional model can be transformed into two-dimensional space by means of parallel projection. Computer graphics usually utilize rectangular projection and the same applies to our algorithm. Fig. 4a shows the basic arrangement of parallel rectangular projection (axonometric projection).

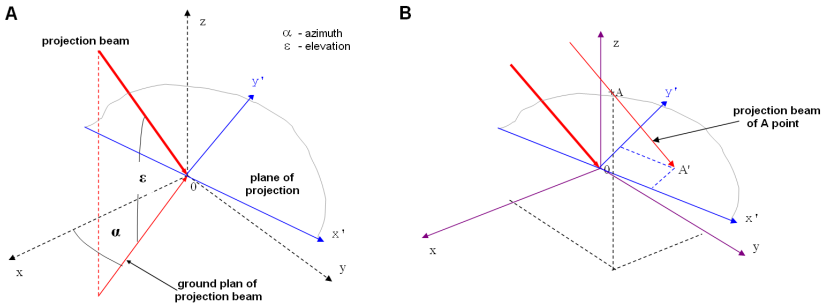


Fig. 4. Axonometric projection (A), projection of point A in parallel rectangular projection (B)

Fig. 4b shows parallel rectangular projection of point A. Coordinates $[x'_a, y'_a]$ of projected point A' can be calculated as follows:

$$x'_a = -x_a \sin(\alpha) + y_a \cos(\alpha) \tag{2}$$

$$y'_a = z_a \cos(\varepsilon) - (x_a \cos(\alpha) + y_a \sin(\alpha)) \sin(\varepsilon) \tag{3}$$

In our case we employed the X-Y projection plane. Thus the azimuth α is equal to -90^o and elevation ε equals 0. The azimuth has a negative value because positive values on the Y axis are oriented backwards (Fig. 3).

The previous paragraphs described the transformation from a 3-dimensional space into two-dimensions using parallel projection. However, the situation is opposite in real-life scenarios. The camera recording is the result of parallel projection of three dimensional space. We are able to detect the position of reference point centers using our method of circular object detection. The centers of reference objects represent the vertices of the triangle under arbitrary rotation. In order to successfully determine the eye location we must first determine the triangle rotation in three-dimensional space. Reconstruction of three-dimensional space from known two-dimensional space is no longer possible using inverse equations (2)(3) because we are missing the third space variable. Thus we must determine the rotation of the triangle in three-dimensional space based on the position of its vertices in two-dimensional space.

2.3 Rotation in Three-Dimensional Space

In order to determine the eye position we must first determine the rotation angle of the head in three-dimensional space using the reference triangle affixed to the patient head. To determine the angle of rotation we must recalculate all coordinates in the model. Dimensional coordinates in Fig. 3 are expressed in

centimeters; however these are converted to pixels for use in subsequent calculations. The calculation will first be limited to the set of reference objects. Recalculation requires the definition of the conversion ratio which defined by the screen distance between reference points in horizontal plane (Fig. 5). This distance must be determined at zero angular rotation around the Z and Y axis, whereby the distance diminishes during rotation around the Z axis. At zero rotation around the Y axis the two lower reference points represent a horizontal line. Thus the frame with zero rotation around the Z and Y axis is automatically determined or we can ask the patient to look directly at the camera prior to the recording.

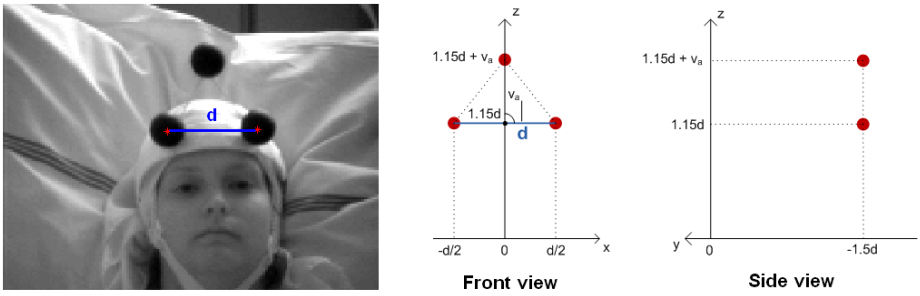


Fig. 5. Expression of reference object coordinates in picture points (pixels)

We place the triangle obtained by parallel projection of the rotated 3-D model into a new axial system with the center located in reference point 2 (Fig. 6). The coordinates of translated reference points are:

$$\begin{bmatrix} x_1' & x_2' & x_3' \\ y_1' & y_2' & y_3' \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix} + \begin{bmatrix} -x_2 & -x_2 & -x_2 \\ -y_2 & -y_2 & -y_2 \end{bmatrix} \tag{4}$$

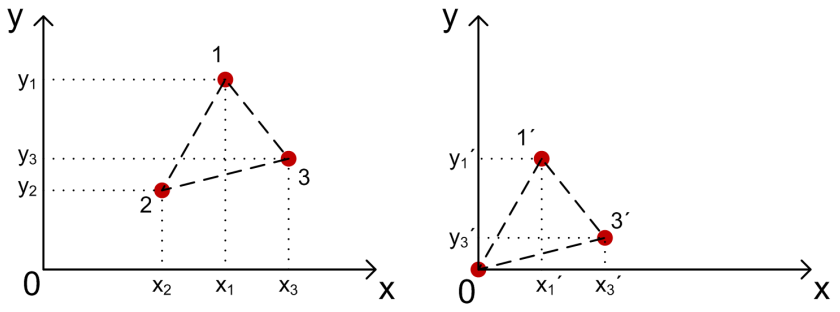


Fig. 6. Translation of reference points. Translated point 2 is the center of the axial system

Translated points 1 and 3 are then expressed using polar coordinates (Fig. 7). Polar coordinates $[r_1, \varphi_1]$ of point 1 (similarly for point 2) are calculated from Cartesian coordinates $[x_1, y_1]$ using the following equations:

$$r_1 = \sqrt{x_1^2 + y_1^2} \tag{5}$$

$$\varphi_1 = \begin{cases} 0, & \text{if } x = y = 0 \\ \arcsin\left(\frac{y}{r}\right), & \text{if } x \geq 0 \\ -\arcsin\left(\frac{y}{r}\right) + \pi, & \text{if } x < 0 \end{cases} \tag{6}$$

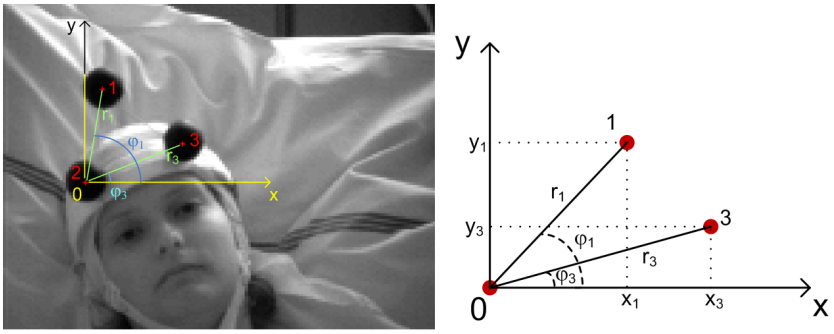


Fig. 7. Point location expressed using polar coordinates

The triangle rotation in space is estimated using a so-called combination table. The combination table contains rotation in space and the corresponding polar coordinates of points 1 and 3. It is not necessary to include all theoretically possible rotational values – only those physically possible when the patient is lying on a bed are included (Table 1).

Table 1. Combination table parameters

	Head movement	Rotation range [°]	Step
Rotation around X axis	Forward and backward	0...30	2^o
Rotation around Y axis	Side movement	-30...30	5^o
Rotation around Z axis	Rotation around vertical axis	-45...45	5^o

The final table contains 3952 possible combinations of rotations and the calculation time on a mid-range CPU (Intel Core2Duo T5250, 1.50GHz) took less than a second. Polar coordinates are determined for each frame of the recorded video (Fig. 7 left). It is evident that the obtained values will not correspond directly to the values in the combination table. This is because the centers of

the reference points are not always precisely detected and also due to the step size in the combination table. The combination table look-up process determines the value of rotation so that the calculated values of polar coordinates are matched with the least possible deviation. Figure 8 shows the graphical representation of the said table via polar coordinates for points 1 and 3 (Fig. 8). Each point is associated with specific rotational angle. Red points represent the polar coordinates obtained from the picture. Neighboring points from the combination table are then looked up and the head rotation is determined. While the determined rotation angle is not precise, it is sufficient to determine the approximate eye location and has no negative effect on the success rate of the blink detection algorithm.

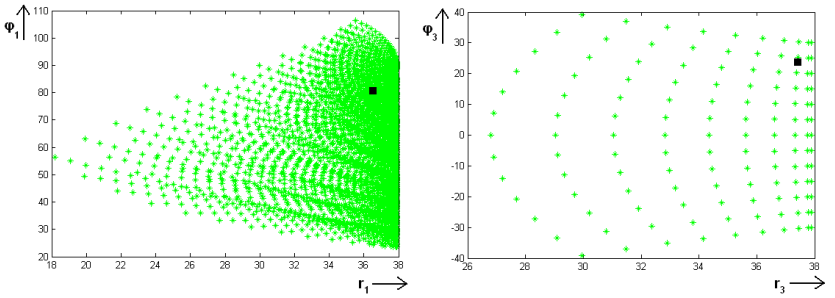


Fig. 8. Location of the searched point (black square) in polar coordinates of the combination table

2.4 Eye Coordinates

The eye distance from the reference triangle varies and must thus be determined from the reference frame. The reference frame is identical to the frame which was used to calculate the conversion ratio d. Eye coordinates $[x, y, z]$ are calculated with respect to reference point denoted as point 2 in previous text. After interactively setting eye position in the reference frame, the coordinates are added to the three-dimensional model (Fig. 3). The following steps are then performed to calculate eye position for any arbitrary frame in the recorded video:

1. Head rotation in all axes is determined,
2. The 3D model is rotated based on the detected rotation,
3. Model points are projected onto the defined plane using rectangular projection,
4. In-plane eye coordinates are determined from reference point 2,
5. Eye position is marked in the image using coordinates determined in previous point.

Figure 9 shows the eye detection algorithm steps.

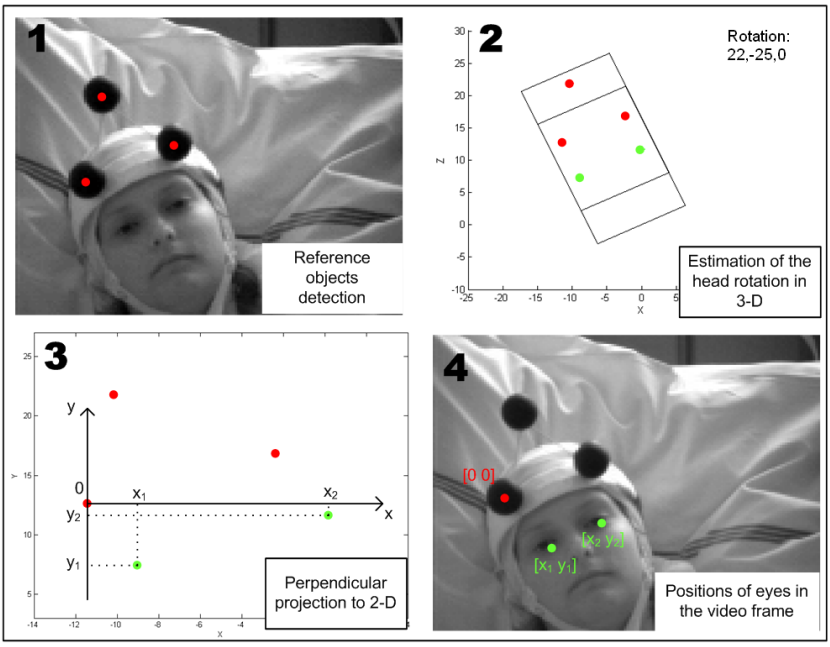


Fig. 9. Eye position determined using rotation of three-dimensional model

Fig. 10. Determined eye position (green points) in still reference frames

3 Results and Conclusion

Figure 10 shows the results of eye detection for various head rotations. The obtained results indicate the detection algorithm is not entirely precise, mainly

because the centers of the reference points are not always accurately detected due to local brightness variations. However, this drawback does not negatively influence the correct function of the subsequent eye blink detection algorithm and artifact marking in EEG recording. The proposed method for eye tracking in compressed video feeds is an important step towards the successful detection of eye artifacts. Detection of the said artifacts will enable the automatic removal of false delta activity during pre-processing of EEG records for automated analysis. The proposed algorithm can also be an alternative to other eye movement ranking algorithms based on EOG evaluation such as [6].

Acknowledgement. This paper is supported by the following project: University Science Park of the University of Zilina (ITMS: 26220220184) supported by the Research&Development Operational Program funded by the European Regional Development Fund.

References

1. Bridzik, R., Novák, V., Penhaker, M.: Simultaneous Recording of Physiological Parameters in Video-EEG Laboratory in Clinical and Research Settings. In: Lim, C.T., Goh, J.C.H. (eds.) ICBME 2008. IFMBE Proceedings, vol. 23, pp. 1042–1043. Springer, Heidelberg (2009)
2. Panning, A., Al-Hamadi, A., Michaelis, B.: A color based approach for eye blink detection in image sequences. In: 2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA). IEEE (2011)
3. Divjak, M., Bischof, H.: Real-time video-based eye blink analysis for detection of low blink-rate during computer use. In: First International Workshop on Tracking Humans for the Evaluation of their Motion in Image Sequences, THEMIS 2008 (2008)
4. Babušiak, B., Mohylová, J.: Eye-blink artifact detection in the EEG. In: Dössel, O., Schlegel, W.C. (eds.) WC 2009. IFMBE Proceedings, vol. 25/4, pp. 1166–1169. Springer, Heidelberg (2010)
5. Babusiak, B., Gála, M.: The eye-blinking artifacts detection and elimination in the EEG record. Programmable Devices and Embedded Systems 9(1) (2009)
6. Penhaker, M., Stula, T., Cerny, M.: Automatic Ranking of Eye Movement in Electrooculographic Records. In: 2010 Second International Conference on Computer Engineering and Applications (ICCEA), vol. 2. IEEE (2010)

Guidelines for the Eye Tracker Calibration Using Points of Regard

Pawel Kasprowski, Katarzyna Harężlak, and Mateusz Stasch

Silesian University of Technology, Institute of Informatics
{kasprowski,katarzyna.harezlak,mateusz.stasch}@polsl.pl

Abstract. Eye movement data may be used for many various purposes. In most cases it is utilized to estimate a gaze point - that is a place where a person is looking at. Most devices registering eye movements, called eye trackers, return information about relative position of an eye, without information about a gaze point. To obtain this information, it is necessary to build a function that maps output from an eye tracker to horizontal and vertical coordinates of a gaze point. Usually eye movement is recorded when a user tracks a group of stimuli being a set of points displayed on a screen. The paper analyzes possible scenarios of such stimulus presentation and discuses an influence of usage of five different regression functions and two different head mounted eye trackers on the results.

Keywords: eye movements, eye tracking, calibration.

1 Introduction

Eye movement data may be used for many various purposes. In most cases it is used to estimate a gaze point - that is a place where a person is looking at. Most devices registering eye movements, called eye trackers, return information about relative position of an eye, without information about the gaze point. To obtain this information it is necessary to build a function that maps an output from the eye tracker to horizontal and vertical coordinates of a gaze point. It is typically done using information about eyes position when an examined person is looking at a set of points (called Points of Regard or PoRs). There are several problems that must be addressed when preparing such a function:

- How many points to use
- How to locate the points - i.e. point's layout
- How long to present the stimulus in each point
- What type of mapping function to use
- Which measurements to use as mapping function input
- How to check the validity of the function

The paper discusses some of issues enlisted above using two different eye trackers and a considerable amount of data registered during a couple of sessions.

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_21, © Springer International Publishing Switzerland 2014

The main contribution of the paper is defining guidelines which may be used when preparing own calibration procedures.

While it seems obvious that more calibration points and a longer presentation of points gives more data to build mapping model, it should also be taken into account that too complicated calibration procedure is inconvenient for participants. Participants may be tired or annoyed with a long preparation phase and tend to loss their concentration during the calibration process itself and - what may be even worse - during the subsequent experiment. Therefore, the main objective of the calibration step is to gather and analyze sufficient amount of data during a procedure that is as short and simple as possible.

The pattern which is most widely (eg. [13,18]) used to present stimuli to participant is the square grid, which typically consist of 9 to 25 points. Ramanauskas et al. [13] confirm what was mentioned above, that higher number of points in this configuration usually results with better performance. Ohno et al. [19] propose an eye tracking system which uses only 2 calibration points, however this setup requires to maintain very strict relations between a camera, an IR illuminator and an eye.

2 Experiment Setup

There were two different eye trackers used during the experiment.

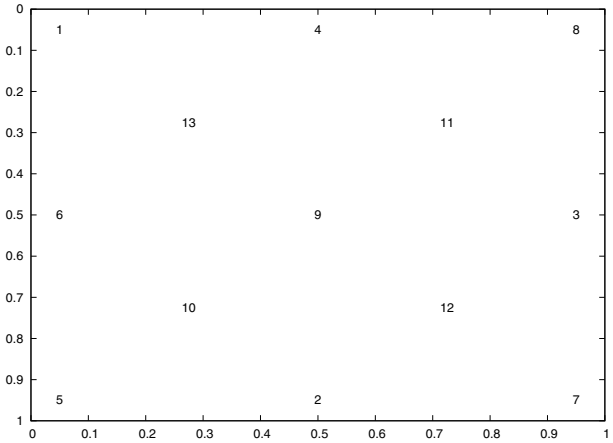
The first one was a head mounted Jazz-Novo eye tracker (product of Oberconsulting) that records eye positions with 1000Hz. It uses Direct Infra Red Oculography (IROG) and utilizes pairs of IR emitters and sensors. The optoelectronic transducers are located between the eyes, thus hiding the sensor assembly behind the *shadow* of the nose.

The second eye tracker was the VOG head-mounted eye tracker developed with a single CMOS camera with USB 2.0 interface (Logitech QuickCam Express) possessing 352 x 288 sensor and lens with IR-Pass filter. The camera was mounted on the arm attached to head and was pointing at the right eye. The eye was illuminated with single IR LED placed off the axis of the eye that caused *dark pupil* effect, which was useful during a pupil detection. The system generated 20 - 25 measurements of a center of a pupil per second.

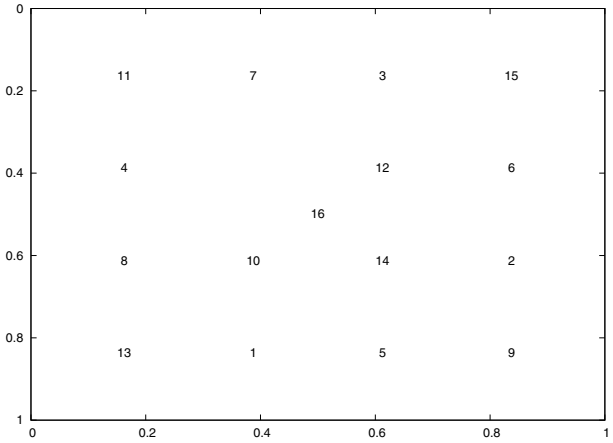
Both eye trackers were used in a usual points of regard calibration experiment. The participants of the experiment were looking at a stimulus presented on a screen. The stimulus was a circle pulsating on the screen to attract participant's attention. There were 29 different stimulus locations (Fig. 1). The stimulus was displayed for about 3 seconds in each location. The order of stimuli presentation was the same for each session.

In both cases the experiment was done on a 1280 x 1024 (370mm x 295mm) flat screen. The eye-screen distance was 500mm, vertical gaze angle was 40 deg and horizontal gaze angle was 32 deg.

There were overall 88 sessions with 39 sessions for Jazz-Novo eye tracker and 49 sessions for VOG eye tracker.



(a) Training points



(b) Testing points

Fig. 1. Points used during sessions

3 Calibration Algorithms

The objective of an eye tracker is to determine a gaze point - that is a place where the user is looking at - with the best possible accuracy. There are three main factors that influence the accuracy of an VOG eye tracker:

- Quality of the registered image
- Quality of the algorithm used for extracting image features
- Quality of the algorithm that is used to map the extracted features to point of regard (gaze point)

It is worth emphasizing that the aim of described study is to analyze how to improve the third of the aforementioned elements, that is mapping image features to point of regard (PoR).

The data was divided into training and testing part. The first 13 stimulus locations (points) were used for training (Fig. 1(a)). All tests and errors of the estimation were always calculated on the last 16 points (Fig. 1(b)). Various combinations of 13 points were used to create 61 sets differing in the number of points used and their locations. The sets can be divided into five groups differing in the number of points. There were 19 sets of 5 points, 20 sets of 7 points, 9 sets of 9 points and 12 sets of 11 points prepared. Additionally, there was a one *full* set of all 13 points used. Due to limited space, the detailed description of the sets is not presented here.

Each of defined sets was then used to build a model mapping an eye tracker output to gaze coordinates on the screen. Such a model consists of two functions:

$$\begin{aligned} x_s &= f(x_e, y_e) \\ y_s &= f(x_e, y_e) \end{aligned} \tag{1}$$

where x_e and y_e represent data obtained from an eye tracker and x_s and y_s are estimated gaze coordinates on a screen.

There are multiple possible regression functions to be used. In this study three types of such functions were used: the polynomial functions, the artificial neural network (ANN) and the support vector regression (SVR).

3.1 Polynomial Functions

The most common choice for a calibration function is usage of polynomial functions, which can differ in the degree and number of terms. Two comprehensive studies analyzed possible solutions [3,1]. There were three classic functions used in this work: a linear function, a quadratic function and a cubic function with all possible terms.

Linear equation

$$\begin{aligned} x_s &= A_x x_e + B x y_e + C_x \\ y_s &= A_y x_e + B_y y_e + C_y \end{aligned} \tag{2}$$

Quadratic equation

$$\begin{aligned} x_s &= A_x x_e^2 + B_x y_e^2 + C_x x_e + D_x y_e + E_x \\ y_s &= A_y x_e^2 + B_y y_e^2 + C_y x_e + D_y y_e + E_y \end{aligned} \tag{3}$$

Cubic equation

$$\begin{aligned} x_s = A_x x_e^3 + B_x y_e^3 + C_x x_e^2 y_e + D_x x_e y_e^2 + E_x x_e y_e \\ + F_x x_e^2 + G_x y_e^2 + H_x y_e + I_x y_e + J_x \\ y_s = A_y x_e^3 + B_y y_e^3 + C_y x_e^2 y_e + D_y x_e y_e^2 + E_y x_e y_e \\ + F_y x_e^2 + G_y y_e^2 + H_y y_e + I_y y_e + J_y \end{aligned} \tag{4}$$

For each function the coefficients were calculated based on training points values using Levenberg-Marquardt algorithm.

3.2 ANN

The second type of function was an artificial neural network (ANN). An activation network with sigmoid function as an activation function was used. The network was trained until the total train error was lower than 0.1, using the Back Propagation algorithm, with normalized samples recorded during a session. Configuration of the network consisted of two neurons in the input layer, 10 neurons in one hidden layer and two neurons as the output. ANN has been already used in several eye tracing applications [6,17].

3.3 SVR

The Support Vector Regression (SVR) [14] was of the last of analyzed types. The RBF kernel with parameters $C = 10$ and $\gamma = 8$ was used. The similar function has been utilized for an eye tracker calibration in [12] and [9] but in completely different setups.

4 Results

Experiments were conducted for all sets of calibration points and for all sessions what gave 11.925 (39 sessions x 61 sets x 5 functions) models for Jazz-novo and 14.935 (49 sessions x 61 sets x 5 functions) models for VOG eye tracker. Every model represented combination of a function, a session and a set of training points. All models were checked using 16 testing points from the same session (Fig. 1(b)). The error represented in degrees (E_{deg}) was calculated based on them. Additionally, it was determinant coefficient (R^2) calculated independently for both axes.

$$E_{deg} = \frac{1}{n} \sum_i \sqrt{(x_i - \widehat{x_i})^2 + (y_i - \widehat{y_i})^2} \tag{5}$$

$$R^2 = 1 - \frac{\sum_i (y_i - \widehat{y_i})^2}{\sum_i (y_i - \bar{y})^2} \tag{6}$$

where y_i, x_i represent an observed value, $\widehat{y_i}, \widehat{x_i}$ represent a value calculated by model and $\bar{y}, \bar{x}$ is the mean of observed values.

It must be emphasized, that it takes some time an eye to react to stimulus position change to fixate on another position. Such occurrence is called *saccadic latency* and lasts approximately 100-300 msec. In many cases this first fixation is not accurate and is corrected. During earlier experiments (not published yet) it was calculated that the safest range of measurements to include for further studies is from 700 msec. to 1800 msec. after the stimulus position changed. Therefore, only these measurements were taken into account in both training and validation phases. It gave about 20-30 samples per point for VOG eye tracker

and more than 1000 samples for the Jazz-Novo eye tracker. As so big number of features was computationally difficult for the ANN and the SVR methods, Jazz-Novo data was downsampled to 50Hz by calculating a median for every 20 subsequent measurements. Such downsampling process didn't affect results of the calibration, which was checked using the polynomial functions.

4.1 Model Validity

While checking the correctness and the accuracy of defined models, it occurred that some models gave completely incorrect results for testing samples. It resulted in R^2 coefficient values to become lower than zero, which means that the modeled curve had errors higher than the average of testing samples. It was decided to reject such models as completely not feasible. Additionally, the number of rejections was calculated for each function and group (i.e. sets with the same number of points). The model was rejected if any of R_x^2 or R_y^2 coefficients values for that model was lower than zero. Such strict condition resulted in rejection of 14.7% models for VOG eye tracker. The rejection percentage for different groups and functions is presented in Table 1. It can be observed that the rejection percentage was lower when more calibration points were taken into account. For 5 points almost 90% of models were not feasible for the cubic (x^3) function. It must be remembered that a cubic function requires calculation of 18 parameters so it needs more data than e.g. a linear function. The ANN and the SVR functions were stable for each class and the number of rejections was always lower than for the polynomial functions. Similar results - but with higher rejection rates - were achieved for the Jazz-novo eye-tracker. The only important difference was that the ANN and the SVR functions gave much higher rejection rates for this device.

Table 1. Rejections percentage for VOG eye tracker

Group	x^1	x^2	x^3	ANN	SVR
5	16.2%	32.4%	87.2%	10.4%	15.9%
7	5.7%	6.1%	35.1%	2.2%	3.8%
9	3.9%	2.7%	14.5%	1.4%	0.9%
11	2.6%	2.0%	5.1%	0.5%	0.3%
13	2.0%	2.0%	2.0%	0.0%	0.0%

4.2 Multidimensional Analysis of the Recorded Data

Filtered data was analyzed taking few aspects into account. It was interesting to check how the average error changed when various number of points constituted training set and what is an influence of a function used to build a regression model. And finally a question on the impact of a calibration points layout on defined model accuracy was asked. All issues were considered separately for both eye trackers used in the experiments.

Table 2. Errors by function

VOG			Jazz		
Func	E_{deg}	**SD**	**Func**	E_{deg}	**SD**
x^2	2.717	1.846	x^2	4.012	1.895
SVR	2.786	1.680	x^1	4.070	2.0188
x^1	3.135	1.965	x^3	4.671	1.762
ANN	3.268	1.650	ANN	5.388	1.672
x^3	3.485	1.770	SVR	7.019	1.705

Functions Comparisons. The first step of the analysis mentioned above was to compare accuracy of different functions using E_{deg} value - that is average angular error. As can be seen in Table 2 the best average results for the VOG system gave the quadratic function x^2. However the difference between x^2 and SVR is not significant ($p = 0.07$). Significant differences can be noticed in relation to other functions.

For Jazz the best function was x^2 although the difference between x^1 and x^2 was not significant. What was interesting that the SVR and the ANN functions were significantly worse for the Jazz-novo eye-tracker.

When errors for various number of points were analyzed, it occurred that for 5-points groups and VOG results the linear function (x^1) became better than SVR and the significant difference can only be noticed for the ANN and the cubic (x^3) function being the worst methods (Table 3). For Jazz x^1 function is significantly the best function for 5-points calibration. On the contrary for 13-points group the SVR method became the best method and the linear function x^1 became significantly the worst one. For the Jazz-novo device all polynomial functions were the best ones but with no significant differences (Table 5).

Table 3. Errors by function for sets with 5 points

VOG			Jazz-novo		
Func	E_{deg}	**SD**	**Func**	E_{deg}	**SD**
x^1	4.427	1.945	x^2	3.415	2.029
x^2	4.786	1.783	x^1	3.561	2.130
ANN	6.037	1.623	SVR	3.588	1.867
x^3	6.082	1.618	ANN	3.986	1.797
SVR	8.100	1.1486	x^3	5.421	1.638

Number of Points Comparison. When average errors were calculated for set with various numbers of points it occurred that, as it could be expected, 13-points set gave lowest errors (Table 5). However, the differences between 13 and 11 points groups were not significant for both eye trackers. Additionally, when

Table 4. Errors by function for 13 points set

VOG			Jazz-novo		
Func	E_{deg}	**SD**	**Func**	E_{deg}	**SD**
SVR	1.995	1.401	x^3	3.349	1.585
x^3	2.138	1.510	x^2	3.629	2.017
x^2	2.323	1.572	x^1	4.079	2.321
ANN	2.328	1.385	ANN	4.302	1.554
x^1	2.930	1.894	SVR	5.301	1.799

similar comparison was done only for models that used x^2 function (Table 6), it turned out that the differences between groups 13,11,9 and 7 were not significant. It shows that higher number of calibration points not necessarily causes lower error rates.

Table 5. Errors by number of points in set

VOG			Jazz-novo		
Points	E_{deg}	**SD**	**Points**	E_{deg}	**SD**
13	2.341	1.594	13	4.455	2.737
11	2.522	1.629	11	4.710	2.679
9	2.806	1.656	9	5.130	2.769
7	3.073	1.704	7	5.546	2.817
5	3.722	1.983	5	6.338	2.697

Table 6. Errors by number of points in set for x^2 function

VOG			Jazz-novo		
Points	E_{deg}	**SD**	**Points**	E_{deg}	**SD**
13	2.323	1.572	13	3.629	2.017
11	2.396	1.628	11	3.662	1.948
9	2.461	1.699	9	3.801	1.933
7	2.580	1.799	7	3.891	1.770
5	3.415	2.029	5	4.786	1.783

Finding the Best Models. Searching for the best models, the comparison of the E_{deg} errors was performed. All results were sorted with ascending order and ten sets with the lowest error values were analyzed. The analysis was done both for each function independently and for all functions together, however data from different eye trackers were treated as autonomous sets.

In Table 7 there were presented the best results in form of ten sets with the lowest E_{deg} average error values. Once again, due to limitation of the paper size, only results subset concerning polynomial functions were showed. It can be observed that for both devices the set 13_1 - that is the reference set using all training points - together with the most complicated polynomial function x^3

Table 7. Best ten sets with lowest E_{deg}

Jazz-novo				VOG			
Function	**Points**	**No**	E_{deg}	**Function**	**Points**	**No**	E_{deg}
x^3	13	1	3.349	x^3	11	9	2.058
x^2	7	16	3.357	x^3	13	1	2.138
x^1	7	13	3.377	x^2	11	9	2.256
x^2	11	12	3.450	x^2	9	2	2.261
x^2	9	7	3.476	x^2	7	1	2.265
x^2	7	13	3.480	x^3	11	1	2.267
x^2	9	2	3.518	x^2	7	7	2.270
x^2	11	1	3.545	x^3	11	12	2.287
x^2	11	7	3.547	x^2	7	10	2.289
x^2	11	10	3.573	x^2	11	8	2.291

provided one of the best results (the lowest E_{deg} error). But it turned out that the differences between this set and other sets with lower number of points are not significant. Especially for Jazz-novo there are two models (x2_7_16 and x1_7_13) for which the results are almost the same. It shows that the number of calibration points not necessarily results in lower error rates.

Interesting situation can be observed when analyzing the SVR and the ANN functions. It occurs that for the VOG system SVR outperformed all other functions (with 9 of 10 best models) and ANN gave results comparable to the polynomial functions (with the best result 2.166 for ANN_11_8). Quite differently in case of Jazz-novo eye-tracker the SVR methods provided very bad results (with 5.13 E_{deg} for set 13_1) and ANN was only a fraction better (with 4.30 for set 13_1). Such bad results require further studies.

In the next step sets differing in number and layouts of points were analyzed in terms of their usefulness for building a calibration model. This analysis, based on ten sets for which the lowest E_{deg} error values were obtained, was done for each of the functions used. It was expected that groups of the chosen sets would be dominated by sets with higher number of calibration points. The conducted studies confirmed this assumptions for cubic, ANN and SVR methods where sets with 11 points constituted majority of a group. However, in case of linear and quadratic functions, greater diversity of sets types was observed. There were sets with 5, 7 and 9 points especially for the first of mentioned methods.

These findings regard both types of eye-trackers although some differences were noticed. It concerned the quality of results achieved for the particular functions. The average error rates obtained in case of the Jazz-novo eye-tracker with regard to linear and quadratic functions were lower than for cubic, ANN and SVR ones, even though they were calculated using lower number of points. The opposite situation was determined in case of the second eye-tracker. The cubic, ANN and SVR methods provided the lower average error values than other polynomial functions.

Point Location Comparison. As it was shown in the previous section, groups with lower number of points may be as good as groups with higher number of points when a correct function is used. In this section different sets with the same number of points were compared to find out if the points location influences error rates. For this purpose the average error for all analyzed functions and for each set of points was calculated. The results for the VOG eye tracker are gathered in Table 8. Analysis of these results showed that in case of the VOG eye tracker there were significant differences for all five points group (19 sets). Among them the best set 5_7 (points: 3,6,9,12,13) giving average error 2.8 deg and the worst set 5_5 (points: 1,7,9,12,13) with an average error equal to 7.74 deg can be mentioned. After deeper layout examination it occurred that the lowest errors are calculated for sets containing point number 3 and 6 (north and south positions on the screen) and additionally of some points in the middle (10,11,12,13). Utilizing points in corners of the screen resulted in significantly higher errors. Similar situation was observed for the VOG system in case of seven points group (20 sets) although the differences were not so significant (from 2.78 deg for set 7_2 to 4.06 for set 7_11). There were no significant differences in points layout for sets with 9 and 11 points noticed.

Studying results for the second eye tracker Jazz-novo it was observed that there were significant differences in 5-points group as well. However, the results varied from that obtained in the VOG case. Although the same two sets (5_5 and 5_6) were the worst ones, there were sets for which Jazz-novo worked much

Table 8. Errors for sets ordered by group and error rate (VOG)

Set	E_{deg}	Set	E_{deg}	Set	E_{deg}
5_7	2.81	7_2	2.79	9_2	2.56
5_2	3.27	7_7	2.83	9_8	2.64
5_8	3.28	7_18	2.89	9_1	2.64
5_15	3.31	7_13	2.90	9_6	2.68
5_1	3.39	7_14	2.92	9_7	2.73
5_19	3.74	7_8	2.93	9_5	2.78
5_17	3.86	7_15	2.93	9_9	2.97
5_3	3.87	7_17	2.93	9_10	3.07
5_18	3.93	7_1	2.94	9_3	3.28
5_12	3.93	7_10	3.04	11_9	2.28
5_4	4.12	7_16	3.05	11_1	2.36
5_11	4.17	7_9	3.07	11_10	2.37
5_13	4.31	7_6	3.11	11_8	2.42
5_16	4.40	7_5	3.22	11_2	2.48
5_14	4.47	7_4	3.32	11_7	2.48
5_10	4.60	7_3	3.62	11_5	2.59
5_9	6.13	7_19	3.71	11_12	2.62
5_6	7.38	7_12	3.72	11_11	2.63
5_5	7.75	7_20	3.76	11_6	2.66
		7_11	4.06	11_4	2.69
				11_3	2.70

better than VOG eye tracker. The 5_9 set with points on the left side of the screen can be taken as an example. In general, points located in the corners of the screen did not influence significantly Jazz-novo outcome - 5_19 set with points 1,5,7,8,9 was the fourth of best models. As it was stated above, their usage in case of the VOG system entitled higher error rates.

Differences for all other groups (7,9 and 11) were not significant but similar trends to the VOG system findings were observed.

5 Conclusions

The main goal of the research was to check repeatability of results for both various regression functions being used and various numbers and layouts of calibration points. To achieve this goal two different eye trackers were used. Using them two environments were developed to test five various functions operating on 61 sets of calibration points. These sets differed with numbers of points and their position on the screen. The obtained results were compared in terms of type of device, type of function and type of calibration points sets. There were some significant differences found. For instance simpler regression functions like x^1 or x^2 operated better than more complicated (like x^3 or SVR) on sets with lower number of points. Some specific sets with 7 points gave results comparable to 13 points sets. In the same time it was showed that calibration results highly depend on the type of eye tracker: (1) Points locations in screen corners were not a good solution for VOG while gave very good results for Jazz-novo. (2) Regression functions SVR and ANN worked very good for VOG eye tracker and very bad for Jazz-novo. (3) x^2 function outperformed x^1 function for VOG but the results of both were similar for Jazz-novo. Therefore, pointing out the best calibration set, which worked well for all functions and for both devices, turned out to be difficult.

Because the presented studies focused on making the comparison described above, the further research are planned to be aimed at an improvement of the particular results separately for each eye tracker.

References

1. Blignaut, P., Wium, D.: The effect of mapping function on the accuracy of a video-based eye tracker. In: Proceedings of the 2013 Conference on Eye Tracking South Africa, ETSA 2013, pp. 39–46. ACM, New York (2013)
2. Brolly, X.L.C., Mulligan, J.B.: Implicit calibration of a remote gaze tracker. In: Conference on Computer Vision and Pattern Recognition Workshop, CVPRW 2004, pp. 134–134 (2004)
3. Cerrolaza, J.J., Villanueva, A., Cabeza, R.: Taxonomic study of polynomial regressions applied to the calibration of video-oculographic systems. In: Proceedings of the 2008 Symposium on Eye Tracking Research & Applications, ETRA 2008, pp. 259–266. ACM, New York (2008)
4. Darrien, J.H., Herd, K., Starling, L.-J., Rosenberg, J.R., Morrison, J.D.: An analysis of the dependence of saccadic latency on target position and target characteristics in human subjects. BMC Neuroscience 2(1), 13 (2001)

5. Duchowski, A.T.: A breadth-first survey of eye-tracking applications. Behavior Research Methods, Instruments, & Computers 34(4), 455–470 (2002)
6. Essig, K., Pomplun, M., Ritter, H.: A neural network for 3d gaze recording with binocular eye trackers. IJPEDS 21(2), 79–95 (2006)
7. Huey, E.B.: The Psychology & Pedagogy of Reading. The Macmillan Company (1908)
8. Jacob, R.J.K.: The use of eye movements in human-computer interaction techniques: What you look at is what you get. ACM Transactions on Information Systems 9, 152–169 (1991)
9. Martinez, F., Carbone, A., Pissaloux, E.: Gaze estimation using local features and non-linear regression. In: 2012 19th IEEE International Conference on Image Processing (ICIP), pp. 1961–1964. IEEE (2012)
10. Moré, J.J.: The levenberg-marquardt algorithm: implementation and theory. In: Watson, G.A. (ed.) Numerical Analysis. LNCS, vol. 630, pp. 105–116. Springer, Heidelberg (1978)
11. Morimoto, C.H., Mimica, M.R.M.: Eye gaze tracking techniques for interactive applications. Comput. Vis. Image Underst. 98(1), 4–24 (2005)
12. Noris, B., Keller, J.-B., Billard, A.: A wearable gaze tracking system for children in unconstrained environments. Computer Vision and Image Understanding 115(4), 476–486 (2011)
13. Ramanauskas, N.: Calibration of video-oculographical eye-tracking system. Electronics and Electrical Engineering 8(72), 65–68 (2006)
14. Smola, A.J., Schölkopf, B.: A tutorial on support vector regression. Statistics and Computing 14(3), 199–222 (2004)
15. Sugano, Y., Matsushita, Y., Sato, Y.: Calibration-free gaze sensing using saliency maps. In: 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2667–2674. IEEE (2010)
16. Villanueva, A., Cabeza, R.: Models for gaze tracking systems. Journal on Image and Video Processing 2007(3), 4 (2007)
17. Zhu, Z., Ji, Q.: Eye and gaze tracking for interactive graphic display. Machine Vision and Applications 15(3), 139–148 (2004)
18. Stampe, D.M.: Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems. Behavior Research Methods, Instruments, & Computers 25(2), 137–142 (1993)
19. Ohno, T., Mukawa, N., Yoshikawa, A.: Freegaze: A gaze tracking system for everyday gaze interaction. In: Proceedings of the 2002 Symposium on Eye Tracking Research & Applications, ETRA 2002, pp. 125–132. ACM, New York (2002)
20. Nyström, M., Andersson, R., Holmqvist, K., van de Weijer, J.: The influence of calibration method and eye physiology on eyetracking data quality. Behavior Research Methods 45(1), 272–288 (2013)

BlinkMouse-On-Screen Mouse Controlled by Eye Blinks

Joanna Marnik

Rzeszów University of Technology,
Al. Powstanców Warszawy 12, 35-959 Rzeszów, Poland
jmarnik@prz-rzeszow.pl
http://jmarnik.sd.prz.edu.pl/

Abstract. The paper concerns issues related to assistive technologies for disabled people. An interface which emulates a computer mouse, handled using eye blinks, is presented in the paper. A computer should be equipped with a low cost webcam to use the interface. Icons related to different directions of mouse cursor movements and real mouse buttons are sequentially presented to the user. An action corresponding to currently shown icon is performed when the user keeps his eyes closed for a specified period of time.

The interface is now used by a disabled girl with SMA (Spinal Muscular Atrophy). The part responsible for an analysis of an eyes behaviour was prepared as a dynamic linked library (DLL), so it can be used for other purposes.

Keywords: computer vision, assistive technologies.

1 Introduction

Nowadays almost everybody can have a computer equipped accordingly to his needs without much difficulty. It is particularly important for disabled people who cannot communicate with others in usual manner. A computer equipped with a special devices or software interfaces can allow such people to participate in social life, learn a profession and take up a job. This results in increased independence and life activity of disabled people. Disabled children can learn and play using such systems. Their teachers and therapists can use computer programs designed for educational and therapeutic purposes in their work.

People with serious motor impairments, such as spinal muscular atrophy (SMA), Amyotrophic Lateral Sclerosis (ALS) or spinal cord damages, are probably in the most difficult life situation. They often communicate with others using alternative communication systems, such as picture symbols (black and white pictures called pictograms or color pictures in the case of Picture Communication System – PCS). The symbol illustrating what the person wants to say is usually indicated by specified signal, which can be generated by disabled person. In the case of the above diseases face expressions or eye blinks are often used as these signals.

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_22, © Springer International Publishing Switzerland 2014

To communicate with the use of pictograms or PCS it is required that a person, which flips through the pages of the book with symbols with pictures, assists the disabled person. Computer program can substitute such person. Moreover, a speech synthesizer can read statements chosen by disabled user. It allows disabled person to feel more independent of others. Moreover, well tailored user interface can make communication more convenient.

Children using an alternative way of communication, such as pictograms, often do not learn reading and writing. It impairs their mental development and decreases their potential for participation in social life. It can be changed by providing tools for handling a computer equipped with special programs for communication and education. A virtual keyboard and a special interface which enables to emulate a computer mouse can give an access to the whole of a computer functionality. Virtual keyboard is standard component of Windows operating systems. There are many similar solutions freely available in the Internet (see e.g.`http://www.techsupportalert.com/best-free-onscreen-keyboard-osk.htm`.) Most of them are inconvenient in the use, but it is possible to freely download solutions which incorporates innovative technologies, such as word prediction, which makes it possible to reduce the number of clicks needed to write a text [1,2]. Thus, providing a disabled user with the interface for mouse handling, which is tailored to his capabilities, is very important for such person.

We present some solutions of this kind which can be used by people with considerable motor impairment (Sect. 2). Our aim was to give a 12 years old girl suffering from SMA the possibility to communicate with the others by means of a computer equipped with one of such interfaces. At first we experienced with b-Link system [21]. It allows to handle a computer using eye blinks and it is freely available from the Internet. Unfortunately, the system proved to be useless. We found out that it resulted from wrong face position relative to the camera. Namely, eyes line was not parallel to the bottom line of the face image. Fulfilling this condition in regard to the girl who lies on the bad, was impossible without an additional camera set-up equipment. Moreover, the graphical interface seemed to be too complicated and too little intuitive for the child. Inability to configure the application with a traditional computer mouse was also very inconvenient (it was possible to do it only using eye blinks, which was very arduous). The drawback of b-Link is also that it occupies a lot of screen area and there is no possibility to change it.

These remarks induced as to develop much simpler interface dedicated to children, without loss of functionality of b-Link. We decided to prepare an universal software interface for eye blinks detection, which could be used for other purposes. We decided to base on a blink detection method proposed in [22], because its high reliability in varying environment condition and high accuracy was reported by the authors. The interface is described in Sect. 3. Our main goal was to design an on-screen mouse application, which is intuitive for the children and will allow the girl mentioned above to handle the computer without help of any assistant person. We named it *BlinkMouse*. The solution is presented in Sect. 4.

Experiments we have done with the *BlinkMouse* and their results are described in Sect. 5. Conclusions and our plans for future work are included in Sect. 6.

2 Related Work

Several solutions, which enable for handling computer by disabled, are available on the market, but they are usually very expensive and they do not always meet user's expectations. Nowadays motor impaired users without possibility to use standard input devices, such as mouse and keyboard, can handle a computer by means of biomedical signals, face gestures or eye gaze tracking. These signals can be interpreted by special interfaces installed on user's computer. Interfaces based on biomedical signals, such as EEG, SSR – Sympathetic Skin Response [3] or muscular activity (EMG – electromyography) [4,5,6], require special input devices, which allow to record data related to intentional user's action.

To use interfaces mentioned above it is necessary to get physical device, which often is out of reach of disabled person. Moreover, the user may find that such device falls short of his expectations. That is why software interfaces can be interesting alternative to above solutions. Today, all common computers are equipped with a camera, so vision interfaces can be easily available for everybody. Software interfaces, in contrary to these which are based on biomedical signals, are non-invasive, so they are are comfortable in the use. Vision interfaces designed for severely motor disabled persons can be divided into head movement based interfaces, eye-gaze tracking interfaces, eye-blinks based interfaces and face gestures based interfaces.

2.1 Head Movement Based Interfaces

The *Camera Mouse* system tracks the computer user's movements with a video camera and translates them into the movements of the mouse pointer on the screen [7]. A person with severe motion impairments can use the *Camera Mouse* with on-screen text-entry software to communicate. The system is available as a free download at http://www.cameramouse.org.

The *Camera Mouse* software limits the user to left-click commands which the user executes by hovering over a certain location for a predetermined amount of time. For users who can control head movements and can wink with one eye while keeping their other eye visibly open, the "blink and wink interface" was designed [9,8]. It allows complete use of a typical mouse, including moving the pointer, left and right clicking, double clicking, and click-and-dragging.

Similar solution, named *HeadMouse* [10] is also freely available in the Internet at http://www.tecnologiasaccesibles.com/en/headmouse.htm.

2.2 Eye-Gaze Tracking Interfaces

Eye-gaze tracking interfaces are suitable for those disabled persons who can not move their heads and limbs, but retained the possibility to move their eyes. Such

interfaces are able to detect and follow user's gaze and are non-invasive, but additional light source, typically infrared and a special video camera are usually necessary. The light source is reflected from the eye and sensed by the camera. The information is then analysed to extract the eye rotation from changes in reflections [11]. To be useful for disabled person, the eye tracker must be closely related to the software, which should respond adequately to the user's intention [12]. In gaze-based communication systems for disabled freezing gaze for more than a predefined time interval (named dwelltime) is recognised as a user's command. This kind of interfaces are still developed, for example at the University of Pavia [13]. Similar solutions are developed by Kocejko, Bujnowski and Wtorek in Multimedia Systems Department of Gdańsk University of Technology [14]. The drawback of infrared camera-based interfaces is necessity of following complicated calibration procedure which can be difficult for small children.

2.3 Eye-Blinks Based Interfaces

Another kind of vision-based interfaces for hands-free computer handling can use an eye blinks. The blinks, when used together with eye-gaze tracking interfaces, can replace sending commands with the use of dwelltime. Many methods for blink detection were proposed in literature. One kind of such interfaces track eyelid movement with the use of optical flow [15] or normal flow [16]. Blink detection with the use of Gabor filter to extract arcs of the eye is proposed in [17]. The eye region is detected using template matching method. The method allows to measure the distance between detected top and bottom arc in eye region. The authors assert that the method has perfect success rate and it is able to work in real time.

The *Blink Link* system, which enables communication using sequences of long and short blinks, allows for blink detection in real time using correlation with an open eye template [18,19]. The system automatically detects a user's eye blinks and accurately measures their durations. Voluntary long blinks triggers mouse clicks, while involuntary short blinks are ignored. The location of the eyes is determined automatically through the motion of the user's initial blinks. Subsequently, the eye is tracked by correlation across time, and appearance changes are automatically analyzed in order to classify the eye as either open or closed at each frame. No manual initialization, special lighting, or prior face detection is required. The system works with inexpensive USB cameras and runs at a frame rate 30 fps. Experiments on eight test subjects yielded an overall detection accuracy of 95.3%.

We mentioned in the Sect. 1 about b-Link system [21]. The authors of the system proposed an algorithm of eye blink detection, estimation of eye blink duration and interpretation of sequence of blinks in real time. The main idea of the algorithm is to measure the area of visible part of an eye and to analyse the eye blink dynamics. The Haar-like face detection is performed first. Obtained face region is used to determine eyes location. Known geometrical dependencies between elements of human faces are taken into consideration there. Next step

is eye tracking. A template matching technique is used for his purpose. Eye blinks are detected and analysed using an active contour model. The authors classify detected eye blinks as short (< 200 ms) or long (≥ 200 ms). Short eye blinks are assumed to be spontaneous and they are discarded during analyse. The algorithm was used in two applications: BlinkWriter and BlinkBrowser. The first application has the form of on-screen keyboard designed to be handled by eye blinks. The scanning technique similar to that of Microsoft's build-in OSK was used there. The letters on the keyboard were arranged so that the most frequently used ones could be selected by shorter sequences of eye blinks. The second application allows to surf on the Internet by eye blinks. It involves an on-line mouse functionality. The average time needed for entering a single character was 10.7 s. for BlinkWriter and 9.5 s. for BlinkBrowser. Moving the mouse cursor diagonally of the screen using BlinkBrowser took 9.5 s. on average.

2.4 Face Gestures Based Interfaces

Another interface, named *LipMouse*, allows a user to use a computer using lip movements and gestures. It was proposed by Dalka and Czyżewski [23]. They used a cascade of boosted classifiers working with Haar-like features to detect face in a video stream from a standard web camera. Next, lip's region is extracted on the basis of a lip's shape approximation and a descriptor is calculated. The descriptor is constructed of a luminance histogram, statistical moments and co-occurrence matrices. It constitutes an input to an artificial neural network. The interface allows to recognize three lip gestures: mouth opening, sticking out the tongue and forming puckered lips.

The *Eyebrow Clicker* is a camera-based human computer interface system which automatically detects when a user raises his or her eyebrows and then triggers a mouse click [19]. It implements a form of simulated binary switch. Results demonstrated overall eyebrow raises detection accuracy of 89%. When the user raises his or her eyebrows, the binary switch is activated and a selection command is issued.

3 Interface

Our interface has the form of a Dynamic-Link Library (DLL) with `CBlink` class, which was designed for detecting eye blinks. The interface was prepared in C++ language, using Microsoft Visual Studio 2010. The open source OpenCV 2.4 library was used for image processing and image analysis tasks. The algorithm for blink detection is presented in Fig. 1.

It is implemented in the `WykrywajMrugn` method of `CBlink` class. The algorithm uses eyes' templates and two parameters `T` and `min_duration`. First parameter, `T`, is related to the similarity of eye appearance to their templates. The similarity is calculated using template matching. The eye template is a gray scale image of the eye. It is found using Haar-like classifier, which was previously trained to detect eyes. The template is compared to overlapped image regions of

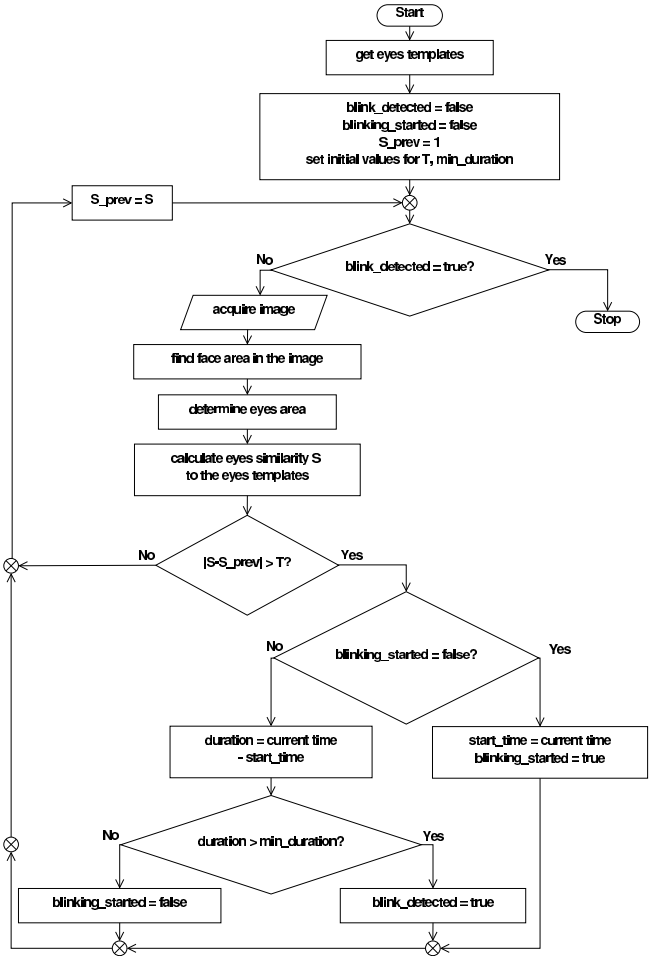


Fig. 1. Blink detection algorithm

the same size. To speed up calculations, only eyes' region is taken into account. The region is marked out by restricting the face area to its upper part (Fig. 2). The face area is found using another Haar-like classifier, which was previously trained to detect faces.

Separate calculations are performed for each eye. The left eye template is used for left part of the eyes' region and the right eye template for its right part. Assuming that R is a subregion of an eye area of image I, which has the same size as a template *templ* and is related to the pixel (x, y), the value of the similarity $S(x, y)$ of R and *temp* is calculated using following formula:

$$S(x,y) = \frac{\sum_{x',y'} \left(templ(x',y') \cdot R(x+x', y+y')\right)}{\sqrt{\sum_{x',y'} templ(x',y')^2 \cdot \sum_{x',y'} R(x+x', y+y')^2}} \tag{1}$$

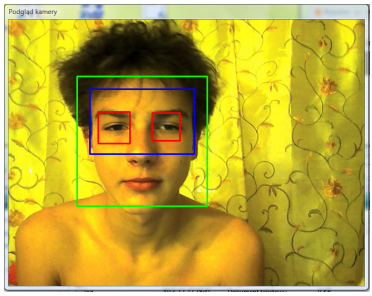


Fig. 2. Regions used during eye blink detection: face region (outermost rectangle), eye region determined on the basis of the face position (middle rectangle) and both eyes regions (two internal rectangles)

where (x, y) indicates the location of the pixel in eye image to which calculated value is assigned, x' and y' are coordinates of the template points.

As the result, an array of similarity values for each pixel contained in the eye region is obtained. Higher values indicate those locations which are more similar to the template. We assume that the eye is located in the area for which the highest similarity value (S^l_{max} for the left eye, S^r_{max} for the right eye) was achieved. Obtained values belong to the range $[0, 1]$. Total similarity S_{max} is calculated as an average value of similarities S^l_{max} and S^r_{max}.The threshold `T` is used to determine if the eyes changed their state from opened to closed or vice versa. A default value of `T` is 0.1. It is compared with the absolute value of difference $D(k)$ between eyes' similarity obtained for current frame k and minimal or maximal value of similarities obtained for $n = 10$ previous frames. If $|D| > T$ and similarity $S_{max}(k)$ related to current frame k is less than $\max\left(S_{max}(k-i)\right), i = 1, 2, ..., n-1)$, closing eyes action is stated. If S_{max} is grater than $\min\left(S_{max}(k-i)\right)$, opening eyes action is stated. If both of this events are detected, then the time that the eyes were closed is stored. It is accessible for the user of `CBlink` class through the method `CzasZamknieciaOczu`.

Second parameter of the algorithm, `min_duration`, determines a threshold of the eye blink duration. If the eyes are closed longer than the threshold value, a signal indicating that the user blinked can be generated. Default value for this threshold is set to 500 ms. It allows to discard unintentional blinks.

The parameters' values can be easily changed, because there are proper methods available for the user of our interface. Moreover, the interface offers additional functionalities, which can be useful during designing the applications for disabled users. Namely, it allows to rotate images obtained from the camera so, that the line connecting corners of user's eyes is perpendicular to the bottom edge of the image. The rotation angle is calculated after indicating outer corners of the users eyes. If can be set by invoking a proper method of the `CBlink` class of our interface. If the flag is switched on, images are rotated before eye blinks detection.

If the face detection process is not necessary, e. g. when the person which controls computer by eye blinks can not move his head, there is a possibility to

omit face detection stage. It makes processing time a little shorter. It can be done by setting proper flags of the interface. Remaining functionalities of the interface are:

- it provides a separate method for the generation of eye templates,
- it makes possible to get information about cameras available in the system,
- it makes possible to indicate which camera source should be used by the interface,
- it makes possible to set the resolution of the camera used during detection of eye blinks,
- it makes possible to incorporate image rotation before eye detection process, if it is necessary, i.e. when the eyes are not parallel to the bottom border of the face image,
- it makes possible to screen user's face images,
- it makes possible to determine which eye was used for blink,
- it makes possible to generate sound prompt indicating when the minimal blink time interval elapsed.

4 BlinkMouse

We have prepared *BlinkMouse* – an on-screen mouse destined for the 12 years old girl, who suffers from SMA. The mouse is handled by eye blinks. It uses the interface described in Sect. 3. After running the application a main window appears on the screen (see Fig. 3). Simultaneously, camera preview window is displayed. At this stage a help of an assistant if the disabled person is necessary to prepare the application to work by means of eye blinks. The assistant should set up the camera in proper position and click one of two buttons available at the main window, namely the "Point out eyes line" button or the "Take eyes templates" button. First button is intended for pointing out outer corners of eyes. It is necessary, if the line of eyes is not parallel to the bottom line of the image. If this way of starting the eye controlled mouse it chosen, the image rotation is performed as an element of blink detection process. If second button will be pressed the image rotation will be omitted. Pressing both buttons causes taking eyes' templates, which are used during blinks detection (see Sect. 3). Now, the disabled user can control his computer by eye blinks. He or she can move the mouse cursor up, down, left and right, perform mouse clicks (also double click) and drag and drop objects. Currently available action is shown on the screen by icons, one for each action. The icons appear sequentially near current mouse cursor position. Their appearance is related to the action, which is invoked by the icon. Namely, the directions correspond to arrows, the click action to icon with check symbol, the drag action to hand symbol.The duration time for the icon visibility can be set by the user. The default value is 2 sec. The action corresponding to the currently visible icon starts after the intentional eye blink is detected. The blink is treated as intentional if the eyes stay closed longer then a specified time duration. The default lower limit value is set to 500 ms, upper limit value is set to 1 500 ms. These values can be changed by the user.

Fig. 3. BlinkMouse: main window and preview window

A double-click action can be activated by closing eyes for longer time than half of the upper time duration limit of the intentional eye blink, while the mouse button click action is active. Activating the mouse cursor movement at the direction shown by the proper icon can be performed by the intentional eye blink while the icon is visible. The cursor stops after closing eyes action is detected. If the intentional blink will be affirmed next actions according to the sequence will be shown. If it turns out that it was the spontaneous blink, i.e. the eyes stayed closed shorter than lower time duration limit for the intentional eye blink, the cursor movement will be resumed. This allows the user to pinpoint the position on the screen. Speed of the mouse cursor movement can be set in `Options` menu (see Fig. 3).

The experience in the use of *BlinkMouse* shown that its functionality can be improved by afford the user an opportunity to enable to go to the next action in the sequence without necessity to wait until the time the individual icon is shown will elapse. It is possible in the *BlinkMouse* by closing eyes for the time shorter than lower limit of the intentional blink while the sequence of actions is presented to the user. At first the disabled girl for which the *BlinkMouse* was designed had difficulties in determining when she can open her eyes to accept the action. That is why we added additional functionality to the application allowing to generate sound prompts after the minimal time for accepting the action elapses. It can be easily switched off in the `Options` menu.

Main window of the *BlinkMouse* contains check box which enables to switch on and off the camera preview window. By default the preview window is closed after the eye templates generation. The demonstration of the *BlinkMouse* can be seen at our web site `http://kia-wizja.prz.edu.pl/main/projects/`.

5 Experiments

The *BlinkMouse* was tested on the laptop with Microsoft Windows Vista, Intel(R) Core(TM) 2 Duo CPU, 2.4 GHz, 4 GB RAM with HP Webcam working with the resolution 640×480. A group of 7 volunteers (4 adults aged 37 to 44 years and 3 children aged 12 to 15 years) were asked to handle the computer without using the keyboard and the mouse device. Instead they were to use the *BlinkMouse* and the virtual keyboard *FreeVK*. We used a test mode of the *BlinkMouse*, in which a log file is generated for each user. It contains an information about the detected eye blinks (i.e. blink duration, the time it occurred). An average processing time of the frame and its standard deviation during the blink detection process were also calculated. The same parameters were calculated for the eye matching. The blinks reported in the log file can be confronted with the real ones, because a video file with the camera view was also recorded.

Before the experiment each participant was informed how to use the *BlinkMouse*. None of them had a problem with understanding it, and everyone started to use the on-screen mouse without any questions.

The results are presented in Tab. 1. The results showed that the standard deviation of average processing time is significantly larger for some users. From the observation of the user's behaviour recorded in video file we conclude that it is depended on the user's mobility. If the user often changes his position in the relation to the monitor, the average processing time and its standard deviation increase. Hence average fps decreases. It is worth mentioning that one child participating in the experiments wore glasses. It did not disturb him to use the application.

Table 1. Results of the experiments

User	Age	Frame processing time		Eye matching time		fps
		avg. [ms]	std [ms]	avg. [ms]	std [ms]	
Adult 1	44	23	4.9	9	2.2	43
Adult 2	44	17	2.0	9	1.3	59
Adult 3	43	19	2.3	10	1.8	53
Adult 4	37	20	2.2	9	1.9	50
Child 1	15	23	5.1	12	3.3	43
Child 2	14	19	2.0	9	1.4	53
Child 3	12	18	3.9	9	2.3	56

All participants commented that using on-screen keyboard is inconvenient to write text. Most of their time during experiment they spend on playing computer games. They also suggested that the mouse cursor can move faster, because they lost their patience waiting the cursor to reach the destination point.

6 Conclusion

We prepared our own interface of the eye detection method based on the template matching technique in the form of DLL library. Such tool can also be used

in other applications, eg. to suggest people suffering from *Computer Vision Syndrome* that they should blink regularly [24] or to prevent drivers from falling asleep.

We used the interface in the on-screen mouse, which we named *BlinkMouse*. The mouse was intended for the girl which suffers from SMA. That is why it is intuitive and has facilities, such as sound prompt indicating that blink duration is sufficient to be recognised as user command, which allows to quickly start using it. The *BlinkMouse* in connection with an on-screen keyboard gives the girl, the possibility to handle a computer by the eye blinks. She uses it now during her educational process. She also uses it to play computer games (the memory game is her favourite one) and to view photos.

The experiments we conducted showed that the blink detection interface allows for a real-time applications. It works well with common laptop with build-in webcam. However, experiences showed that the algorithm which we implemented (see Sect. 3) has a little drawback. Namely, if the user looks beyond the screen, it can be recognised as the eye blink. The solution for this is to modify the algorithm, for example by analysing the visibility of the pupil. Adding the possibility to change cursor movement speed on-the-fly also may be useful.

References

1. Ashtiani, B., Scott MacKenzie, I.: BlinkWrite2: An Improved Text Entry Method Using Eye Blinks. In: Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications, ETRA 2010, pp. 339–345 (2012)
2. Teixido, M., Arroyo, I., Palleja, T., Tresanchez, M., Fernandez del Viso, A., Rebate, C., Palacin, J.: Implementacíon de VirtualKeyboard como Herramienta de Mejora de la Accesibilidad a las TIC. In: IV Congreso Internacional de Diseño, Redes de Investigación y Tecnología Para Todos, Madrid, España, pp. 443–450 (2011)
3. Wada, C., Masuda, F.: Basic Research of Input Support Device by Using Sympathetic Skin Response. In: Miesenberger, K., Klaus, J., Zagler, W.L., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 429–436. Springer, Heidelberg (2006)
4. Chung-Min Wu et al.: An Eye Input Device for Persons with the Motor Neuron Diseases. Biomed. Eng. Appl. Basis Commun. 25 (2013)
5. Chin, C., Barreto, A., Alonso Jr., M.: Electromyogram-Based Cursor Control System for users with Motor Disabilities. In: Miesenberger, K., Klaus, J., Zagler, W.L., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 905–912. Springer, Heidelberg (2006)
6. Felzer, T., Nordmann, R.: Evaluating the Hands-Free Mouse Control System: An Initial Case Study. In: Miesenberger, K., Klaus, J., Zagler, W.L., Karshmer, A.I. (eds.) ICCHP 2008. LNCS, vol. 5105, pp. 1188–1195. Springer, Heidelberg (2008)
7. Betke, M., Gips, J., Fleming, P.: The Camera Mouse: Visual Tracking of Body Features to Provide Computer Access For People with Severe Disabilities. IEEE Transactions on Neural Systems and Rehabilitation Engineering 10, 1–10 (2002)
8. Betke, M.: emPower, an Infrastructure for Remote Assessment of Interfaces for Individuals with Severe Motion Impairments. In: Assistive Technology: From Rresearch to Practice
9. Missimer, E., Betke, M.: Blink and wink detection for mouse pointer control. In: Proceedings of the 3rd ACM International Conference on Pervasive Technologies Related to Assistive Environments, PETRA 2010, Article No. 23, pp. 23:1–23:8 (2010)

10. Teixido, M., Guillamet, A., Palleja, T., Tresanchez, M., Fernandez del Viso, A., Rebate, C., Palacin, J.: Implementación del Ratón Virtual HeadMouse como Herramienta de Mejora de la Accesibilidad a las TIC. In: IV Congreso Internacional de Diseño, Redes de Investigación y Tecnología Para Todos, Madrid, España, pp. 433–442 (2011)
11. Crane, H.D., Steele, C.M.: Generation-V dual-Purkinje-image eyetracker. Applied Optics 24(4), 527–537 (1985)
12. Donegan, M., Oosthuizen, L., Bates, R., Gaunys, G., Hansen, J.P., Joos, M., Majaranta, P., Signorile, I.: User Requirements Report with Observations of Difficulties Users are Experiencing. Communication by Gaze Interaction, COGAIN (2005)
13. Porta, M., Ravarelli, A.: Some Eye Tracking Solutions for Severe Motor Disabilities. In: Hippe, Z.S., Kulikowski, J.L., Mroczek, T., et al. (eds.) Human – Computer Systems Interaction. AISC, vol. 98, pp. 417–431. Springer, Heidelberg (2012)
14. Kocejko, T., Bujnowski, A., Wtorek, J.: Eye-mouse for disabled. In: Hippe, Z.S., Kulikowski, J.L. (eds.) Human-Computer Systems Interaction. AISC, vol. 60, pp. 109–122. Springer, Heidelberg (2009)
15. Lalonde, M., Byrns, D., Gagnon, L., Teasdale, N., Laurendeau, D.: Real-time eye blink detection with GPU-based SIFT tracking. In: Proceedings of the Fourth Canadian Conference on Computer and Robot Vision, CRV 2007, pp. 481–487 (2007)
16. Heishman, R., Duric, Z.: Using Image Flow to detect Eye Blinks in Color Videos. In: Proceedings of the the Eight IEEE Workshop on Applications of Computer Vision, WACV 2007, p. 52 (2007)
17. Arai, K., Mardiyanto, R.: Comparative Study on Blink Detection and Gaze Estimation Methods for HCI, in Particular, Gabor Filter Utilized Blink Detection Method. In: Proceedings of the Eighth International Conference on Information Technology: New Generations, ITNG 2011, pp. 441–446. IEEE Computer Society (2011)
18. Chau, M., Betke, M.: Real Time Eye Tracking and Blink Detection with USB Cameras. Department of Computer Science Technical Report BUCS-2005-012, Boston University (2005)
19. Grauman, K., Betke, M., Lombardi, J., Gips, J., Bradski, G.R.: Communication via eye blinks and eyebrow raises: Video-based human-computer interfaces. Universal Access in the Information Society 2(4), 359–373 (2003)
20. Królak, A., Strumiłło, P.: Fatigue monitoring by means of eye blink analysis in image sequences. In: International Conference on Signals and Electronics Systems, Łódź, Poland, pp. 219–222 (2006)
21. Królak, A., Strumiłło, P.: Eye-Blink Controlled Human-Computer Interface for the Disabled. In: Hippe, Z.S., Kulikowski, J.L. (eds.) Human-Computer Systems Interaction. AISC, vol. 60, pp. 123–133. Springer, Heidelberg (2009)
22. Chau, M., Betke, M.: Real time eye tracking and blink detection with USB cameras. Boston University Computer Science Technical Report No. 2005-12 (2005)
23. Dalka, P., Czyżewski, A.: Controlling Computer by Lip Gestures Employing Neural Networks. In: Szczuka, M., Kryszkiewicz, M., Ramanna, S., Jensen, R., Hu, Q. (eds.) RSCTC 2010. LNCS, vol. 6086, pp. 80–89. Springer, Heidelberg (2010)
24. Yan, Z., Hu, L., Chen, H., Lu, F.: Computer Vision Syndrome: A widely spreading but largery unknown epidemic among computer users. Computers in Human Behavior 24(5), 2026–2042 (2008)

Bio-amplifier with Programmable Gain and Adjustable Leads for Basic Measurement of Bioelectric Signals

Branko Babusiak, Stefan Borik, and Michal Gala

University of Zilina, Dept. of Electromagnetic and Biomedical Engineering,Univerzitna 1, 01026 Zilina, Slovakia
branko.babusiak@fel.uniza.sk

Abstract. This article describes a four-channel amplifier for human neurophysiologic electrical activity measurement. The advantage of the device, in combination with the developed software, lies in the real-time display of various human electrical signals such as ECG, EEG and EOG. Signal samples are stored in a binary file on the hard drive for further analysis. Combination of leads is changeable depending on the measurement type and characteristics thereof. Gain of individual leads is adjustable by digital potentiometers. The various types of measurement are presented in the article in order to verify device function.

Keywords: bio-amplifier, programmable gain, changeable leads, ECG, EEG, EOG.

1 Introduction

In practice it is often necessary to measure different human electrical activity by a single device. There are many commercial multi-channel bio-amplifiers with excellent precision. However, commercial devices are very expensive for general purpose applications. In the next sections we describe the design of four-channeled bio-amplifier for electrical activity measurements. The resulting device can be used for measurements of ECG, EOG and EEG with maximum of four channels. The experimental measurements are shown in further section [1].

2 Bio-amplifier Features

The device is primarily targeted for biological electrical signal measurements in the frequency range up to 160 Hz. The device is named BioAmp based on its main function. Because of different amplitude of various bio-signals, e.g. ECG has amplitude around 1 mV, EEG around 100 μV; it is appropriate to change signal gain depending on the type of measured signal. Change of gain helps to maintain stronger signal in defined output voltage range and weaker signals such as EEG are stretched to whole voltage range. Normally the device operating voltage is supplied from USB. Output analog signal is in the 0 - 3.3V

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_23, © Springer International Publishing Switzerland 2014

range. Alternatively, the device may be supplied from external DC power source connected via 2.1 mm power jack connector. When external power source is connected, supply from USB is automatically switched off and output analog signal is in range from 0 V to 5 V. Wider voltage range allows measurement of signals with higher gain without amplitude cut-off. It is possible to connect a maximum of five electrodes to the device. One of electrodes has function of noise suppression - DRL electrode. The electrodes are connected to device via a pair of DB9 connectors which are commonly used for serial ports. Pinout of connectors is compatible with connectors of commercial Biopac electrode set. The Biopac systems are used across multiple disciplines including human physiology, pharmacology, exercise physiology and many other biological applications. All four channels are bipolar which means they measure the difference between electrode pairs. The said pairs are interchangeable by switching position of jumper on the board. Each channel is sampled at a rate of 320 times per second. The resolution of AD converter is 12-bits. Communication with PC is provided by USB virtual serial port.

3 Hardware Design

3.1 Analog Part

The analog part of device is schematically displayed in the Fig. 1. The input electrode signals first pass through voltage followers. This isolates the output circuit so that the input is not affected in any way by the output. There are four input electrodes and one Driven Right Leg (DRL) electrode. DRL circuit is providing a high common mode rejection at power line frequency (50 Hz) [1][2].

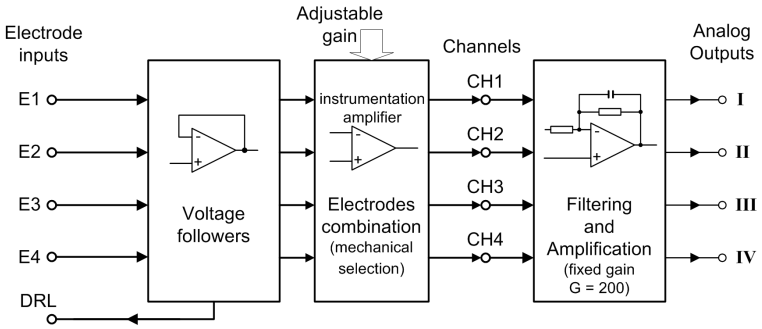


Fig. 1. Block diagram of the analog part

Each channel measures the difference between electrode pair. One electrode is permanently connected to inverting input of instrumentation amplifier. One of other electrodes is not connected to inverting input but arbitrary connected to

non-inverting input of amplifier (Fig. 2). This is realized by switching position of jumper on the board.

The difference between the chosen pair of electrodes is amplified by instrumentation amplifier INA333. The INA333 provides very low offset voltage (25μV, G $>$ 100) and high common-mode rejection (100 dB at G $>$ 10). The gain of INA333 is given by resistor RG connected between inputs:

$$G = 1 + \frac{100k\Omega}{R_G} \tag{1}$$

Programmable gain is achieved by replacing resistor RG with digital potentiometer DS1804-100 (Fig. 2). The DS1804-100 allows change of resistance in range from 0 to 100 kΩ, by step of 1 kΩ. The gain is adjustable via software running on the PC. The gain G1 of amplifier can be set within the range from 2 to 101.

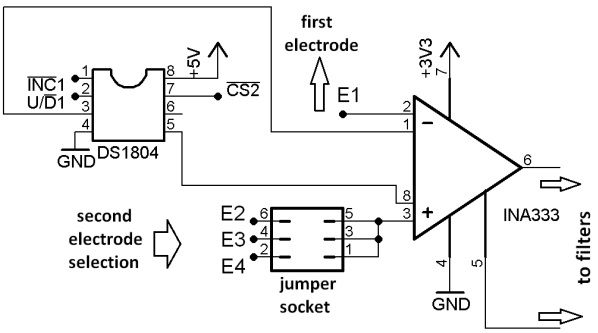


Fig. 2. Programmable gain of INA333 provided by digital potentiometer DS1804. The selection of second electrode is performed by changing jumper position in on-board socket

The last segment of analog part provides analog filtration of signal (Fig. 3). The passband filter consists of pair of operational amplifiers. First of them (IC7A in Fig. 3) acts as highpass filter with cut-off frequency at 0.5 Hz. The cut-off frequency in Hz, according to Fig. 3, is computed as

$$f = \frac{1}{2\pi R_{12} C_3} \tag{2}$$

In ECG measurements, the highpass filter with cut-off frequency 0.5 Hz is used to remove fluctuation of ECG isoline due to breathing. Second amplifier (IC7B in Fig. 3) acts as a lowpass filter with cut-off frequency at about 160 Hz. The cut-off frequency of lowpass filter is computed in the same way as in

the previous case. This filter is also known as anti-aliasing filter. The minimal sampling frequency according to sampling theorem is then

$$F_S = 2.F_{MAX} = 2.160 = 320Hz \tag{3}$$

Half of supply voltage is connected to the non-inverting pin of amplifiers in order to shift the zero line of the output signal.

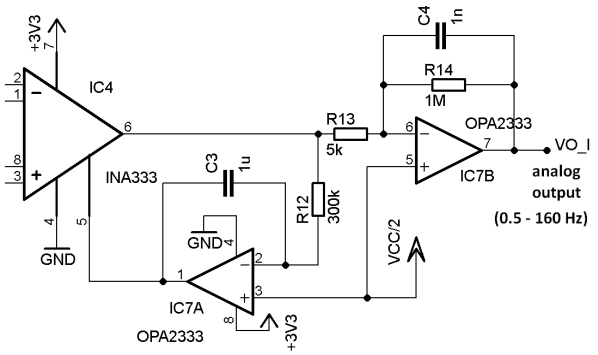


Fig. 3. Analog filtration of the signal

The gain of output amplifiers is set to fixed value G2 = 200. The total gain G of each channel is given by

$$G = G_1.G_2 \tag{4}$$

The total gain G in (3) can vary within the range from 400 to 20 200. The output stage of analog part provides analog filtration in frequency range from 0.5 to 160 Hz.

3.2 Digital Part

The outputs of analog part are connected to AD converter inputs. Low-power converter ADS7844 by Texas Instruments is used in this process. The ADS7844 is an 8-channel, 12-bit sampling AD converter with synchronous serial interface. Only four single ended channels are used. The AD converter is controlled by 8-bit microcontroller ATmega88A by Atmel. ATmega88A is the "heart" of the bio-amplifier device. It is operating at 18 MHz (external crystal oscillator). It sets resistance of digital potentiometers, thus setting the gain of each channel and communicates with PC by using UART interface. Current PCs lack a standard serial port (RS232); therefore UART interface is converted to USB interface by using FT232RL integrated circuit. A shift register is necessary because many microcontroller pins are needed to control the AD converter, set the resistance

of digital potentiometers and communicate with PC. In our case we used the 74HC595 shift register. The 74HC595 is 8-bit serial-in and parallel-out shift register. The shift register is used for controlling four digital potentiometers.

3.3 Power Supply

The device is supplied directly from USB bus or externally from DC voltage adaptor. In case of supply from external power source, the supply from USB is automatically switched off. This is realized by utilizing a relay circuit in Fig. 4. The advantage of external supply consists in wider voltage range in which a signal can be spread out. If device is supplied from USB, the reference of AD converter is connected to +3.3 V and all amplifiers are supplied by this voltage. The reference input of 12-bit ADC determines the analog input voltage range and the LSB size as

$$LSB = \frac{V_{REF}}{4096} \tag{5}$$

If V_{REF} = 3.3 V then LSB size is approximately 806 μV. In case of external power supply, the reference of AD converter is connected to +5 V and all amplifiers are supplied by this voltage. The LSB size is then approximately 1.22 mV.

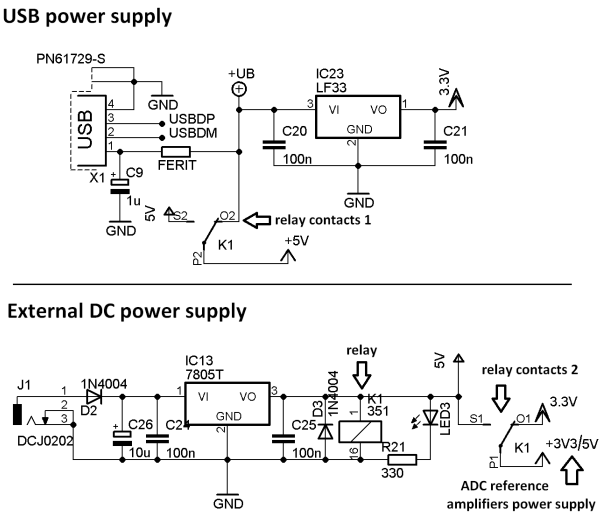


Fig. 4. Types of possible power supply. Automatic supply switching is provided by relay

3.4 Construction of BioAmp Device

The BioAmp device consists of two separate single-layered PCBs connected together through connector (Fig. 5). The bottom PCB includes the analog part

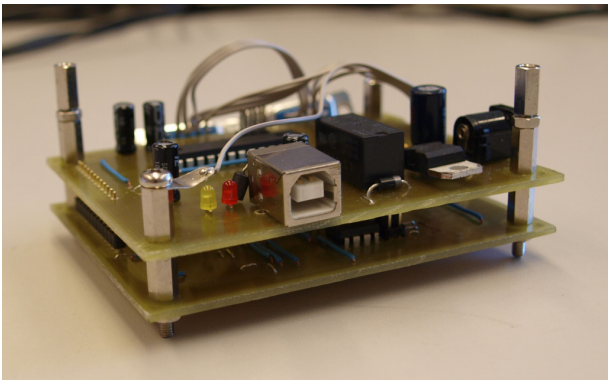


Fig. 5. Construction of BioAmp device

with digital potentiometers. The digital part and power supply circuits are placed on the top PCB.

The electronics are encapsulated in the shielded box made of aluminum (Fig. 6). The shielded box effectively reduces noise from surroundings.

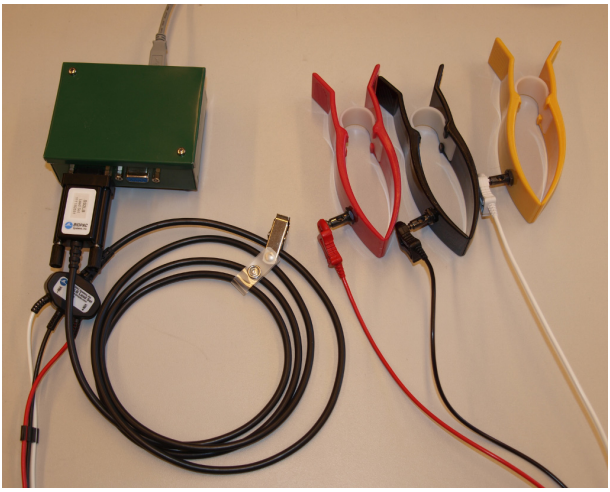


Fig. 6. BioAmp device with ECG clamp electrodes

4 Data Transfer and Software

The device communicates with PC via virtual serial port. Baud rate is set to 38 400 with one stop bit and no parity bit. Each sample consists of two bytes (type uint16). Bytes are transmitted according to following timing diagram (Fig. 7).

Data is transferred when RTS pin is set high. When RTS pin is set low, the data is transmitted until the last byte of the fourth channel is transmitted. The DSR pin indicates the busy state of the microcontroller. When DSR is high, the microcontroller is idle. When AD conversion or data transfer are in progress, the DSR pin is set to low.

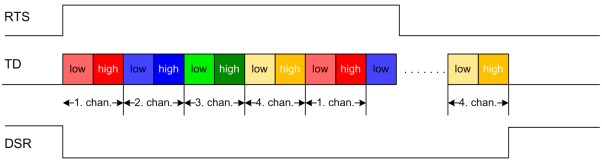


Fig. 7. Serial timing diagram of data transfer

The software in C# programming language was developed for the purpose of displaying the measured signals. The software was named BioAmpMonitor in accordance with its principal function. The BioAmpMonitor displays measured signals in real time (Fig. 8). Moreover, the BioAmpMonitor application allows setting the resistance of digital potentiometers on-line. By changing the resistance, as was mentioned previously, the gain of each channel is also modified. Increasing or decreasing the gain is visible on the signal progress immediately. Gain of each channel is set individually by trackbars (Fig. 9).

Besides online gain setting it is possible to invert particular channel when necessary. The BioAmpMonitor allow saving of multiple channels to binary file for further analysis in another software such as MATLAB.

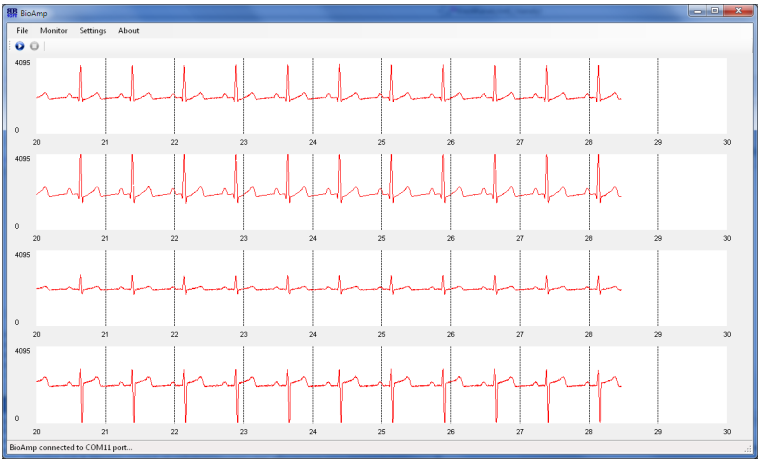


Fig. 8. The BioAmpMonitor application. Measurement of standard limb leads ECG (first three channels)

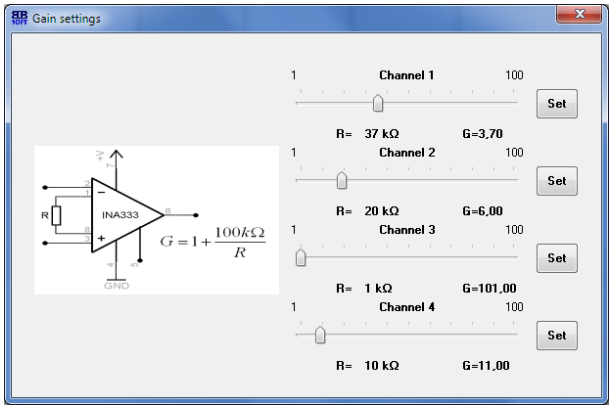


Fig. 9. Channel by channel gain settings

5 Measurements

5.1 Measurement of ECG

The BioAmp device can be used for measurement of ECG. The measurement of standard limb leads and placement thereof is shown in Fig. 10. The human ECG was replaced by calibrated patient simulator MPS450 from Fluke Biomedical. The last electrode was connected to arbitrary place on chest (not depicted in Fig. 10). The first three channels in Fig. 6 refer to standard limb leads; the fourth channel refers to the voltage between chest electrode (one of V1 - V6) and the left arm. The gain for all channels was set to G1 = 11. The signal in second channel is inverted according to standard procedure by software function. Because of low power dissipation the smaller version of BioAmp can be battery powered and used for long term ECG analysis of Holter's type examination as was described in [3].

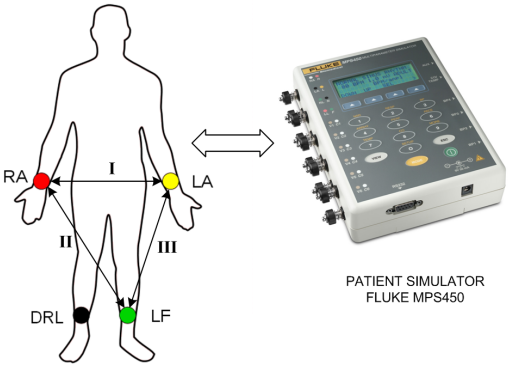


Fig. 10. The electrode placement for ECG measurement (limb leads). The patient is replaced by patient simulator

5.2 Measurement of EEG

In this subsection we demonstrate the measurement of alpha activity in the EEG. The measurement of EEG was performed according to electrode placement in Fig. 11. Particular channels are numbered by roman letters. It is appropriate to use this placement to measure alpha activity (8 - 12 Hz) because alpha activity is produced primarily in posterior parts of the brain. The alpha activity rises with closed eyes and attenuates with eyes opened. Cup electrodes with conductive electrode paste were used for the measurement. The examined subject has closed eyes for first 40 seconds; then the eyes were opened for next 40 seconds.

The EEG was saved as binary file by the BioAmpMonitor software and then analyzed in MATLAB. Fig. 12 shows the EEG signal from channel I (P3 - P4) in time of transition between eyes closed/opened. Fig. 12a depicts raw signal

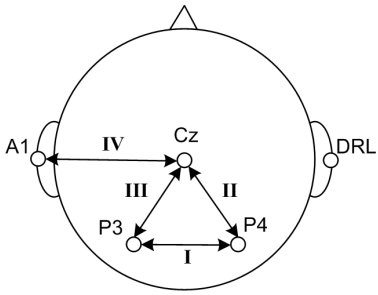


Fig. 11. Electrode placement and channel configuration for EEG measurement

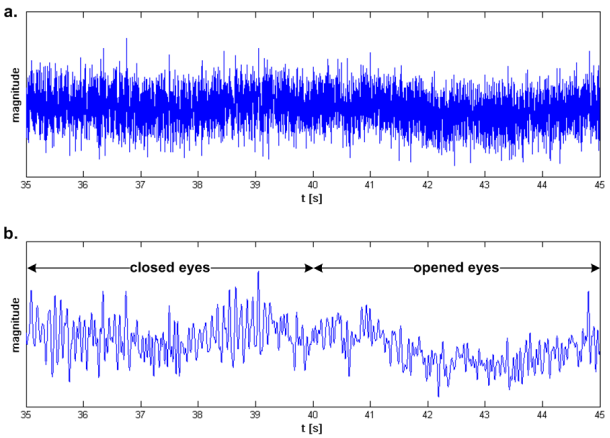


Fig. 12. The EEG signal from channel I (10 seconds). Raw noised EEG (a) and filtered EEG - low-pass filter 20 Hz (b)

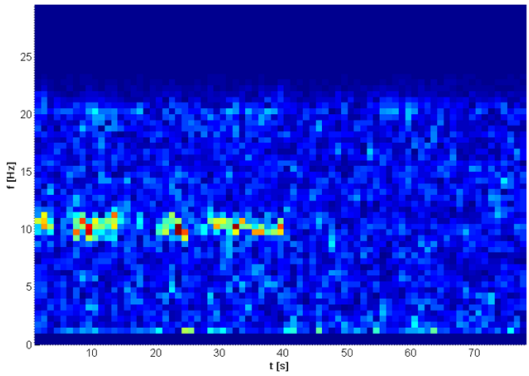
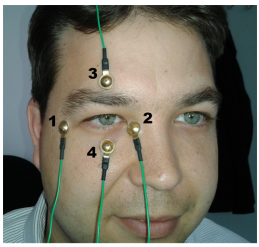


Fig. 13. Spectrogram (STFT) of entire EEG signal (channel I). Presence of alpha activity (8 - 12 Hz) in first 40 seconds (closed eyes)

from BioAmp. The measured EEG contains large amount of noise from electrical wiring. Fig. 12b shows the filtered signal. A low-pass filter with cut-off frequency of 20 Hz was used. The function of BioAmp in EEG measurement is verified by Short-Time Fourier Transform visualization (STFT) in Fig. 13. The higher magnitude of alpha activity is visible in the first forty seconds. In the second part of signal (after opening the eyes) there is low magnitude of alpha activity. Similar results were visible in all measured channels. Up to now, the BioAmp device has been successfully used for EEG measurement in application of EEG biofeedback and the brain-computer interface based on SSVEP.

5.3 Measurement of EOG

Electrode placement for measurement of EOG is shown in Fig. 14. Cup electrodes in combination with conductive electrode paste were used. The fifth electrode (DRL) is connected to the left ear (not visible in Fig. 14).



Channel	Electrode pairs
I	1 - 2
II	2 - 3
III	3 - 4
IV	4 - 1

Fig. 14. Electrode placement for EOG measurement; electrode configuration is shown on the right

Channel I and III refer to horizontal and vertical eye pupil movement, respectively. These two channels are displayed in Fig. 15. The aim of the measurement was preliminary detection of pupil movement in four different directions - up, down, left and right. The observed movement in particular direction is denoted in Fig. 15.

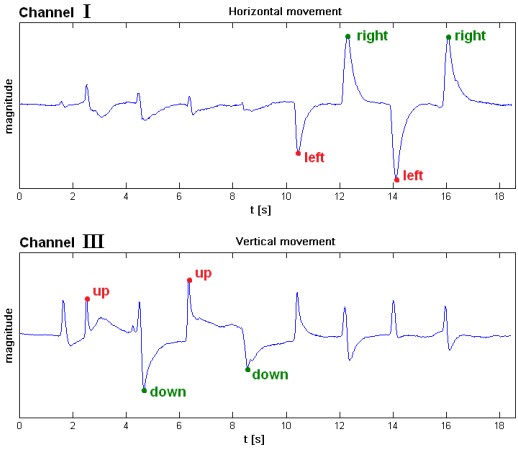


Fig. 15. Measured EOG signals in horizontal and vertical direction; signals were preprocessed by moving average window

The analysis and processing of EOG can be used for human-computer interface, e.g. controlling mouse cursor by eye movements or for automatic ranking of eye movement during sleep [4]. The BioAmp device is appropriate for all mentioned applications. For the future it will be great to develop embedded bio-amplification device for monitoring human electrical activity with basic frequency analysis as is described in [5].

6 Conclusion

This article describes the design of a 4-channel bio-amplifier with programmable gain and switchable electrode pairs per channel. The strength of the device, in combination with the developed software, lies in the real-time display of various human electrical signals such as ECG, EEG and EOG. Signal samples are stored in a binary file on the hard drive for further analysis. Various types of measurements were presented in the article in order to verify device function. The device can be used for general diagnostic purposes because of its relatively small frequency range (up to 160 Hz).

References

1. Wong, A., Pun, K.-P., Zhang, Y.-T., Choy, C.-S.: An ecg measurement ic using driven-right-leg circuit. In: Proceedings of the IEEE International Symposium on Circuits and Systems, ISCAS 2006, p. 4. IEEE (2006)
2. Winter, B.B., Webster, J.G.: Driven-right-leg circuit design. IEEE Transactions on Biomedical Engineering (1), 62–66 (1983)
3. Penhaker, M., Stula, T., Augustynek, M.: Long-term heart rate variability assessment. In: Osman, N.A.A., Abas, W.A.B.W., Wahab, A.K.A., Ting, H.-N. (eds.) BIOMED 2011. IFMBE Proceedings, vol. 35, pp. 532–535. Springer, Heidelberg (2011)
4. Penhaker, M., Stula, T., Cerny, M.: Automatic ranking of eye movement in electrooculographic records. In: 2010 Second International Conference on Computer Engineering and Applications (ICCEA), vol. 2, pp. 456–460. IEEE (2010)
5. Machacek, Z., Srovnal, V.: Automated system for data measuring and analyses from embedded systems. In: Proceeeding of the 7th WSEAS International Conference on Automatic Control, Modeling and Simulation, Prague, Czech Republic, pp. 1–6 (2005)

Development of Diagnostic Stroke Ontology – Preliminary Results

Teresa Podsiadly-Marczykowska[1], Bogdan Ciszek[2], and Artur Przelaskowski[3]

[1] Nalecz Institute of Biocybernetics and Bioengineering,
Trojdena 4, 02-109 Warsaw, Poland
tpodsiadly@ibib.waw.pl
[2] Center for Biostructure Research, Warsaw Medical University,
Chałubińskiego 5, 02-004 Warsaw, Poland
bciszek@ib.amwaw.edu.pl
[3] The Faculty of Mathematics and Information Science,
Warsaw University of Technology, Koszykowa 75, 00-662 Warsaw, Poland
arturp@mini.pw.edu.pl

Abstract. The paper presents the preliminary results of the research aiming at the development of the Diagnostic Stroke Ontology. The described machine processable knowledge representation was designed for the applications supporting stroke diagnosis and management of stroke patients. Applied methodology of ontology development was methontology with subsequent activities: specifications of requirements, domain knowledge acquisition, conceptualization and formalization of conceptual model followed by evaluation. Presented in the article structure and content of Stroke Diagnostic Ontology (DStrokeOnto ontology) include salient ontology modules, classes in modules, relations between modules, main classes definition and details of ontology metrics and expressivity. The DStrokeOnto ontology is syntactically correct, model completeness and logical consistency was verified by two experts in neurology and brain anatomy (excluding sub-module of Imaging Procedure For Stroke - in progress) as good quality interface to CT-based system of computer-aided stroke diagnosis.

Keywords: stroke, ontology, diagnosis.

1 Introduction

According to Gruber [1] ontology is a formal, explicit specification of a shared conceptualization. Formal means that ontology is machine readable and processable, explicit specification stipulates the necessity of explicit and unambiguous definitions of the concepts described by ontology, shared refers to the fact that ontological model is meant to contain consensual knowledge, and finally conceptualization denotes an abstract model of a phenomenon existing in the world created by identification of the relevant concepts and the relations existing between them. Representing domain knowledge in machine processable format ontologies

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_24, © Springer International Publishing Switzerland 2014

can be incorporated into computer-based applications thus, they can be applied in the development of clinical decision support systems [2].

The paper is organized as follows: second section provides the review of relevant ontologies in biomedicine, neurology and brain related research; third section gives basic notions of applied ontology development methodology; fourth section presents the resulting DStrokeOnto ontology describing: purpose, range and potential uses, ontology structure and content, salient ontology modules and classes, relations between modules, main classes definition and details of ontology metrics and expressivity. Section five concludes the paper.

2 Related Research – Review of Relevant Ontologies in Biomedicine, Neurology and Brain Related Research

Ontologies are now pervasive in biomedicine, where they serve as a primary means to standardize terminology, that's why at the beginning of the work on DStrokeOnto ontology we searched the web hoping to find ontologies at least partially related to our topic. The results of the search are forming part of the knowledge acquisition process for the DStrokeOnto ontology.

To find ontological knowledge resources related to stroke we performed web search using special purpose semantic search engine "Watson" and searching through bio ontologies gathered in BioPortal[1] The search revealed no ontology strictly dedicated to stroke diagnosis, although classical key word search using Google indicated three stroke related, potentially useful ontologies, developed in projects shortly presented below.

2.1 NeuroWeb Reference Ontology

NeuroWeb Reference Ontology has been developed in the cadre of NEUROWEB project [3], supporting research concerning cerebrovascular association studies, intended as the search for statistical correlations between a patient genotype and a phenotype. In this project the phenotype refers to the description of patients' pathological state, thus it is formulated on the basis of clinical data collected during the diagnostic process. According to the [3] authors, the knowledge about cerebrovascular fenotypes and genotypes is formulated in the form of the NeuroWeb Reference Ontology. Unfortunately, large containing 16 pages paper describing NeuroWeb Ontology does not provide detailed description of ontology structure and content, and the authors have not granted the possibility of browsing or downloading the ontology file. In that situation it is impossible to evaluate the usefulness, or the possibility of even partial ontology reuse[2].

[1] BioPortal - a web portal that provides access to a library of biomedical ontologies and terminologies (`http://bioportal.bioontology.org`) via the National Center for Biomedical Ontology (NCBO) Web.

[2] It should be noted that the prerequisite for even partial ontological model reuse is the availability of ontology file. Unavailability of positively evaluated owl files of ontological models is a major obstacle in realizing the idea of knowledge reuse.

2.2 OntoNeuroBase Ontology

OntoNeuroBase Ontology has been developed as part of NeuroBase [4] project, aiming to facilitate collaborative research in neuroimaging through a federated database system based on semantic web technologies. The cornerstone and focus of the project is the design of a common semantic model providing a metadata for data and software tools to be shared, a multi-components, formal OntoNeuroBase ontology. This ontology is meant to conceptualize a specific domain of neuroimaging applications in order to facilitate sharing of neuroimaging data. Currently targeted particular applications concern neuroimaging techniques used in epilepsy, Alzheimer's disease and the task of visual cortex exploration. This ontology is partially related to our topic, because we plan to include the description of neuroimaging stroke data content and meta-data as one of the modules of DStrokeOnto ontology, but this task is already in progress. Also in this case, ontology file is not provided.

2.3 NeuMORE (NeuroMOtor REcovery) Ontology

The basic aim of the NeuMORE Ontology [5] is to present in the formal way the knowledge related to the therapy and rehabilitation of patients after brain stroke. The final objective of NeuMORE Ontology, yet not obtained, is supporting the identification of factors determining the optimal, individual treatment and rehabilitation strategies for patients after brain stroke. Those factor should help researchers in better predicting stroke survivors' neural motor functional recovery level, and help physical therapists customize most effective, individual intervention therapy plans for stroke survivors.

The other purpose of NeuMORE Ontology is enrichment of existing data mining techniques using the mechanisms of reasoning in ontologies. Combination of these two techniques, called by authors the semantics-enhanced data mining is supposed to allow for more effective discovering of the knowledge in data.

NeuMORE Ontology file is provided for browsing in BioPortal[3]. At the moment the ontology contains 33 named classes[4]. Ontology classes represents general notions, like for example Common Bio Concepts (with three subclasses Material Entity, Processual Entity and Clinical Characteristics). In this case, browsing of the ontology file allowed to state that the model does not provide content of interest to us, it can be concluded that the NeuMORE Ontology, after completion of the work, will be complementary to DStrokeOnto ontology.

2.4 Foundational Model of Anatomy (FMA) Ontology and Other Top-Level Ontologies

In the diagnostic process, information about patient comordibities, the results of the laboratory tests and neurological findings play an important role. Because those notions represent general medical concepts, it was decided to define

[3] `http://bioportal.bioontology.org/ontologies/NEUMORE`
[4] Defined only by name, without detailed definitions.

them in DStrokeOnto using top-level reference ontologies related to those fields, namely The Foundational Model of Anatomy Ontology[5] [6] (FMA Ontology), clinical healthcare terminology ontology [7], (SNOMED CT ontology), Chemical Entities of Biological Interest Ontology [8] (ChEBI ontology), and Human Disease Ontology (HDO) [9].

All four, above mentioned top-level ontologies although not directly involved to either stroke diagnosis or with brain diseases, can serve as reference for general medical concepts defined in DstrokeOnto.

3 Stroke Diagnostic Ontology Development Methodology

In the field of ontology engineering, many ontology methodologies have been proposed to date. To develop our Stroke Ontology we have chosen METHONTOLOGY [10] methodology because of its grounding in software engineering methodologies, allowing for creation of ontologies at knowledge level, following a life cycle based on evolving prototypes. METHONTOLOGY divides ontology development activities in two general groups, development oriented and supporting. The first group includes such activities as specification, conceptualization, formalization, the second is composed of knowledge acquisition, evaluation and documentation.

METHONTOLOGY is suitable for building domain ontologies from scratch, although it underlines and support the importance of knowledge reuse. The Stroke Ontology has been constructed following main developmental and supporting activities proposed by METHONTOLOGY:

1. Specification of requirements — stating what the motivation for ontology building is, specifying the model purpose, restricting its domain and range and declaring its intended users;
2. Knowledge acquisition — the goal of this activity is to acquire knowledge in a given domain, from experts using some kind of manual or semi-automatic process or by using ontology learning i.e. automatic process;
3. Conceptualization of domain knowledge — identification of concepts in the domain of interest, their attributes and values;
4. Formalization — transformation of the conceptual model from previous stage into formal, computable model using formal ontology languages and ontology editing tool;

[5] The Foundational Model of Anatomy Ontology [6] is a reference ontology representing spatial and structural description of concepts and their relationships enabling symbolic modeling of the whole of human anatomy in a computable form, including anatomy of the brain, at high level of granularity (FMA ontology currently includes more than 75,000 distinct anatomical classes and more than 200 types of relationships). The FMA ontology, created by Structural Informatics Group at the University of Washington (`http://fma.biostr.washington.edu`), explicitly represents a coherent body of declarative knowledge about human anatomy. It was used as a reference biomedical domain ontology for anatomical concepts defined in DStrokeOnto ontology.

5. Evaluation — a syntactic (technical) and logical assessment of the ontology quality and specification fulfillment.

4 DStrokeOnto Ontology Presentation – Aim and Range, Main Modules Content, Salient Classes

The main goal of the ontology is to provide formalized medical knowledge necessary for reliable and sound stroke diagnosis. The intended end-users of the Stroke Ontology knowledge base are developers of the applications supporting stroke diagnosis and management of stroke patients, final users of those medical applications are neurologists clinicians with various levels of expertise, radiologists and paramedics.

Different techniques that can be applied for collecting information about the range of the ontology are: brainstorming sessions, scenarios and use cases, interviews with users and domain experts, finally competency questions[6] (CQs). The assessment of the Stroke Ontology range was guided by semi-structured interviews with two experts neurologists, collection of use cases and clinical algorithms, finally construction of ontology CQ.

Our medical experts are two professors from Medical University of Warsaw, active experts in clinical neurology, brain anatomy, specializing in the diagnosis and treatment of stroke and other brain related diseases, authors and co-authors of scientific publications in neuroscience, teaching medical students and residents.

Interviews with medical experts allowed to establish at first attempt broad categories of information necessary to obtain sound stroke diagnosis such as: clinical characteristics of stroke patient, stroke related neurological findings, laboratory tests, comorbidities and clinical conditions constituting stroke risk or mimicking stroke, stroke type assessment and imaging procedures for stroke diagnosis confirmation and treatment assessment.

4.1 Knowledge Acquisition

This stage of ontology development has a purpose to gather knowledge resources necessary to reliably and completely describe domain of the ontology. The knowledge can be generally divided into ontological (biomedical domain ontologies, top-level ontologies) and non-ontological (state-of-the-art medical hand books

[6] CQs method has been proposed for the first time by Gruninger [11], it consists in constructing a set of questions that ontology knowledge base should be able to answer. CQs and their responses also play the role of requirement specification against which the ontology can be evaluated. CQs should be defined in a stratified manner, with higher-level questions requiring the response given by questions from lower levels. CQs can be written in natural language, or they can be formalized and answered in ontology query languages. As a definition of the range of the Diagnostic Stroke Ontology, as much as serving as an evaluation aid we have formulated around 150 competency questions1 grouped in 12 broad categories.

on neurology and anatomy, relevant scientific papers on stroke diagnostic process).

The results of ontological knowledge resources search was presented in Section 2. To sum up, NeuroWeb Reference Ontology seems the most suited for reuse in our project, unfortunately the ontology file is inaccessible. OntoNeuroBase Ontology aim is only relatively related to our, the purposes of NeuMORE Ontology and DStrokeOnto ontology are complementary.

In the absence of reusable ontological knowledge resources we decided to build ontology from scratch using non-ontological knowledge and classical knowledge acquisition techniques[7] such as structured and non-structured interviews with our medical experts, informal and semi-formal analysis of medical texts, to name only a few [12,13,14,15] and top level bio-medical ontologies for concepts references. Four top-level ontologies, FMA, SNOMED-CT, ChEBI and HDO has been used to reference concepts in DStrokeOnto ontology.

4.2 Conceptualization of Domain Knowledge – General Presentation of DStrokeOnto Ontology Content, Main Modules, Salient Classes and Relations

The goal of this stage is to detect, define and organize ontology concepts into hierarchies, metadata and attributes are used to refine concepts definition. During conceptualization acquired knowledge is structured and organized to obtain independent domain knowledge representation.

Normalized ontology is organized as orthogonal modules formed by sets homogenous hierarchies of classes joined with class definition used for restrictions and complex descriptions. This technique of constructing general ontology structure has been proposed in [16] and has subsequently been used in constructing ontologies in biomedicine. Modular, normalized ontology enables to achieve explicit ontology structure transparent for human authors' and for machine inference, support model reuse and maintainability.

To achieve these goals, DStrokeOnto ontology has been structured at first level into three orthogonal modules: Patient Description Module, Characteristic Module and External Reference Module (all modules are described in more details below and graphically presented in Figure 1).

The module inter-relations are straightforward, Characteristic Module provides concepts for Stroke Patient Description class Module, External Reference Module is used to provide references for concepts in Characteristic Module. Patient DescriptionModule contain at present one [8] class Stroke Patient Description which serve as information frame defining body of information necessary to reliably assess stroke diagnosis.

[7] In general knowledge acquisition consists of following three phases: (1) informal to identify potential concepts in the domain of interest, (2) grouping to structure terms into provisional categories, (3) refining and identification of semantic cross-references between topics/areas.

[8] We plan to complement in the future the module with classes representing different stroke diagnostic patterns.

Characteristics Module is divided into ten sub-modules. In general, they are coherent orthogonal modules, mostly used to define classes used as range fillers in properties of class Stroke Patient Description. Exception to this rule are four classes, Abbreviated Human Anatomy, Stimuli Type and Modifier For Stroke and Stroke Classification. First three classes are used as fillers in class properties ranges but in sub-modules of Characteristics Module, namely in Stroke Related Neurological Findings , Laboratory Test for Stroke, Clinimetric Scale for Stroke and Imaging Procedures for Stroke. The subclasses of fourth exception class, Stroke Classification, are used to assess patient stroke type using subsumtion reasoning in ontology, more details in section DStrokeOnto application- use case.

General structure of DStrokeOnto ontology,its main modules and their salient mutual relations are schematically presented in Figure 1, definition of the class Stroke Patient Description is presented in Figure 2 and described in more details in Patient Description Module.

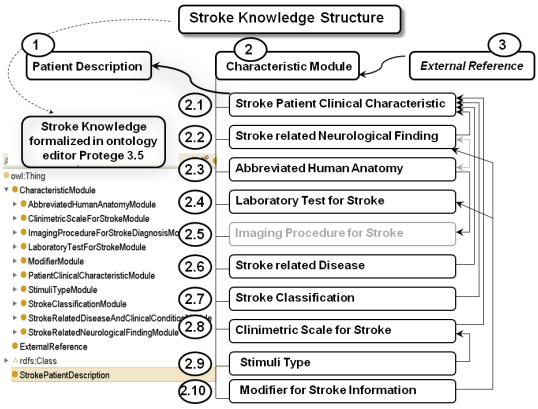


Fig. 1. Schematic representation of Stroke related Diagnostic Knowledge; the whole body of knowledge in divided into three modules, their relation are marked with arrows; characteristic Module provides concepts for Stroke Patient Description class Module, External Reference Module is used to provide references for concepts in Characteristic Module; marked in grey module Imaging Procedure for Stroke needs further elaboration

Patient Description Module. Medical data needed for reliable stroke diagnosis contain results from different medical knowledge areas such as: patient demographic and life style, neurological and clinical findings, laboratory tests, imaging data, stroke related diseases and clinical conditions, stroke type classifications and clinimetric scales proving means for quantitative assessment of patient state. Those information are represented as different sub-modules of Characteristic Module. The relations between Stroke Patient Description class and corresponding Sub-modules of Characteristic Module are presented in details in Fig. 2.

In DStrokeOnto Stroke Patient Description class definition is expressed using 13 object properties, their ranges are filled with classes from Characteristic

Module forming information frame needed for the reliable assessment of stroke diagnosis (see Figure 2). Exemplary detailed specification of filler classes for three selected properties of class Stroke Patient Description is provided below.

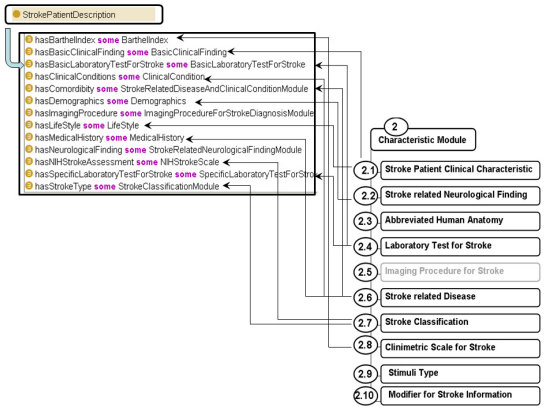


Fig. 2. Information frame defining Stroke Patient Description is constructed using 13 object properties linking stroke patient description with classes modeling adequate medical knowledge defined in Characteristics Module of the DStrokeOnto

Selected three important properties of the class Stroke Patient Description, along with their range filler classes are as follows:

1. hasBarthelIndex object property, has filler class BarthelIndex (containing 10 subclasses) from the sub-module Clinimetric Scale for Stroke (no. 2.8 in Fig. 1) of Characteristic Module;
2. hasBasicClinicalFinding object property, has filler class BasicClinicalFinding (containing 4 subclasses) from the sub-module Stroke Patient Clinical Finding (no. 2.1 in Fig. 1) of Characteristic Module;
3. hasComordibities object property, has as filler one of the subclasses of the class ClinicalCondition (containing 12 subclasses) from the sub-module StrokeRelatedDisease (no. 2.6 in Fig. 1) of Characteristic Module;
4. hasNeurologicalFinding object property, has as filler one of the subclasses of the class Stroke Related Neurological Finding (containing at the moment 45 subclasses) from the sub-module Stroke Related Neurological Finding (no. 2.2 in Fig. 1) of Characteristic Module.

External Reference Module. External reference Module contain instances of top-level reference medical ontologes, FMA, SNOMED CT CheBI and HDO. Classes representing general medical concepts in Characteristics Module have been annotated with the notions from top-level reference ontologies such as:

1. SNOMED CT — comprehensive, clinical healthcare terminology available in the form of the ontology [7]; concepts from SNOMED CT ontology are used

as references for concepts in Stroke Related Diseases and Stroke Related Neurological Finding Modules;
2. FMA — Foundational model of Anatomy Ontology, a reference ontology for the domain of human anatomy [6]; concepts from FMA Ontology are used as references for concepts in Abbreviated Human Anatomy Module;
3. ChEBI — Chemical Entities of Biological Interest Ontology [8]; concepts from ChEBI Ontology are used as references for concepts in Laboratory Test for Stroke Module;
4. Human Disease Ontology — provides biomedical community with consistent, reusable and sustainable descriptions of human disease terms [9]; concepts from SNOMED CT ontology are used as references for concepts in Stroke Related Diseases module.

DStrokeOnto Ontoloogy Metrics and Expressivity. The Stroke Diagnostic Ontology contains at the moment 456 classes (12 defined), 233 properties (103 objects, 130 datatype) and 77 restrictions. Ontology DL expressivity is ALCHOF(D). It means that OWL-DL constructors used to specify DStrokeOnto ontology allow for concept intersection, full universal quantification, atomic negation and limited existential quantification (AL) , complex concept negation (C), property hierarchy (H), use of singleton sets (O), functional properties (F) and datatypes (D).

4.3 DStrokeOnto Ontology Formalization

Formalization goal is to build formal i.e. computer processable representation of domain knowledge acquired in former ontology development stages. Stroke Ontology formalization has been performed using Protégé-2000 [17] ver. 3.5 ontology editor. Protégé-2000 is an open-source tool developed at Stanford Medical Informatics as a platform for ontology building, it enables users to build ontologies Web Ontology Language (OWL) [18]. To implement the Stroke Diagnostic Ontology the OWL-DL ontology language has been used, in order to take advantage of its expressivity and inference capabilities[9].

4.4 DStrokeOnto Ontology Evaluation

Full evaluation of the DStrokeOnto should confirm its formal, syntactic consistency and completeness of domain knowledge representation. The design and formalization of ontologies are complex tasks, during which ontology engineers need to be supported by methodologies and tools enabling detection and repairing of errors, i.e. ensuring correctness of the resulting ontology. Tools allowing for checking, among other, ontology syntactic correctness are reasoners, i.e. programs that infer logical consequences from a set of explicitly asserted facts or

[9] OWL DL — includes all OWL language constructs with restrictions such as type separation, supports maximum expressiveness without losing computational completeness and decidability of reasoning systems.

axioms. Typically reasoners are embedded in ontology editors providing standard reasoning services such as: concepts satisfiability and redundancy, subsumption reasoning, instance checking and retrieval and assessment of the knowledge base. Use of the reasoner throughout[10] ontology formalization is a good practice allowing to avoid modeling errors, concept and ontology inconsistency[11], class redundancy, thus ensuring syntactic correctness of the ontology.

DStrokeOnto ontology is consistent and does not include redundant classes, its syntactic correctness was confirmed by Pellet reasoner (ver. 1.5.2) embedded in Protégé-2000 ontology editor, used during model development.

One of the standard methods of assessing completeness of the domain knowledge representation is the comparison of declared ontology domain and range expressed using CQ in natural language with the results of queries representing CQ expressed in SPARQL [19]. To that end it is necessary to translate CQ from natural language to SPARQL, and to execute them against ontology knowledge base, finally to assess query results. This comparison will be performed after population of the DStrokeOnto with patient data.

4.5 DStrokeOnto Use Case

Stroke subtype can be assessed using one of the four systems: OCSP clinical classification [20], and CCS [21], TOAST [22], ASCO [23]. The OCSP defines four stroke subtypes assessing anatomical localization of the disease using only patient neurological finding. OCSP stroke subtypes are: Total Anterior Circulation Stroke, Partial Anterior Circulation Syndrome, Posterior Circulation Syndrome and Lacunar Syndrome. A fundamental goal of any disease classification is to generate reliable patients subgroups with discrete phenotypic, therapeutic, and prognostic features. The OCSP classification although relatively simple, achieved only moderate inter-rater reliability [24]. Factors leading to disagreement among raters in stroke type classification occur from two main sources: the ambiguities in patient data and the differences in knowledge and experience of raters [24]. The second source of inter-rater opinion variability can be addressed using ontology inference mechanism, our first application of the DStrokeOnto ontology concerns automatic assessment of stroke subtype using the OCSP classification. To that end we introduced in the DStrokeOnto ontology defined classes modeling OCSP subtypes. The definition of those classes contains restrictions on patient neurological finding fulfilling OCSP criteria. In assumed, manually constructed ontology hierarchy the descriptions of real patients data are represented as instances of the Patient Description ontology class, classes representing OCSP subcategories have no instances. In automatically computed ontology hierarchy instances of classes Patient Description are inferred, based on their definitions,

[10] Reasoners can also be used at ontology publish time providing already made inference for ontology users, or in ontology-based applications as a querying mechanism (especially for small ontologies).

[11] Unsatisfiable (inconsistent) concepts are fundamental modeling errors, because they are unable to have instance, i.e. they are unable to represent any individual. They are equivalent to the empty set, ontology containing inconsistent classes is inconsistent.

also as instances of classes which model OSCP types. This application has the potential to improve patients categorization in stroke related research. The definition of classes modeling OCSP subcategory Total Anterior Circulation Stroke is presented in Fig. 3.

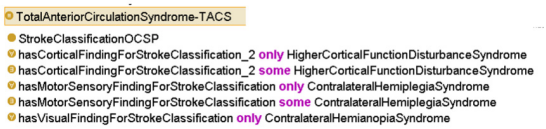


Fig. 3. Definition of the class modeling OCSP subcategory Total Anterior Circulation Stroke

5 Conclusions and Future Work

The article presents DStrokeOnto ontology, one of the first attempts to formalize the whole body of medical knowledge necessary for stroke diagnosis. DStrokeOnto has been developed using Protégé-2000 (ver. 3.5) ontology editor, it contains at the moment 456 classes (12 defined), 233 properties (103 objects, 130 datatype) and 77 restrictions. The model is consistent and does not include redundant classes, its syntactic correctness was confirmed by Pellet reasoner (ver. 1.5.2). Ontology logical evaluation has been performed by two neurologists, domain experts, to completely finish ontology development we need to elaborate in more details the sub-module Imaging Procedure For Stroke, perform once more full logical model evaluation and populate ontology knowledge base with instance i.e. representative, real stroke cases data.

Full development of ontology as central repository of knowledge will enable future works i.e. the design of a number new functionalities facilitating stroke emergency. Possible solutions include: (i) stroke case database oriented to clinical facts, integrated with ontology to extend its usefulness (case-oriented retrieval, diagnosis support); (ii) numerical complement (i.e. numerical descriptors) of conceptually and object-oriented knowledge structure for semantic data processing, computer aided diagnosis etc.; (iii) knowledge-oriented teleinformatic platform bringing the formalized models to clinical and pre-clinical paths and decision-making procedures resulting in computerized support of improved stroke emergency, tele-consultation activity and data mining.

Acknowledgement. This work was supported by the polish National Science Center grant no. 2011/03/B/ST7/03649,2012-2014, Computer Understanding of Medical Images by Integrating Signals Acquisition, and the Representation of Information and Knowledge Models.

References

1. Gruber, T.: translation approach to portable ontology specifications. Knowledge Acquisition 5, 199–220 (1993)
2. Bodenreider, O., Stevens, R.: Bio-ontologies: current trends and future directions. Brief Bioinform. 7(3), 256–274 (2006)

3. Colombo, G., Marico, D., et al.: An ontological modeling approach to cerebrovascular disease studies: The NEUROWEB case. J. Biomed. Inform. 43(4), 469–484 (2011)
4. Temal, L., Michel Dojat, M., et al.: Towards an ontology for sharing medical images and regions of interest in neuroimaging. J. Biomed. Inform. 41(5), 766–778 (2008)
5. Townsend, C., et al.: Ontology-Based Knowledge Acquisition for Neuromotor Functional Recovery in Stroke. In: Proc. IEEE Int. Conf, Bioinf. & Biomed., Hong Kong (2010)
6. Rosse, C., Mejino, J.L.V.: The Foundational Model of Anatomy Ontology. In: Burger, A., Davidson, D., Baldock, R. (eds.) Anatomy Ontologies for Bioinformatics: Principles and Practice, pp. 59–117. Springer, New York (2007)
7. http://bioportal.bioontology.org/ontologies/SNOMEDCT
8. http://www.ebi.ac.uk/chebi/
9. http://www.disease-ontology.org/
10. Fernández-López, et al.: Methontology: From ontological art towards ontological engineering. In: Proc. Spring Symp. Ontological Engin (AAAI), pp. 33–40 (1997)
11. Gruninger, M., et al.: Methodology for the design and evaluation of ontologies. In: Proc. Int. J. Conf. AI Wrkshp. Basic Ontological Iss. in Knowldge. Sharing, Montreal (1995)
12. Merrit's Neurology, 10th edn. Lippincott Williams & Wilkins (2000)
13. Kwiatkowski, J.R., Larry, B., et al.: Guidelines for the Early Management of Patients With Ischemic Stroke: A Scientific Statement From the Stroke Council of the American Stroke Association. Stroke 34, 1056–1083 (2003)
14. Markus, H.: Genes for Stroke. J. Neurol Neurosurg. Psychiatry 75, 1229–1231 (2004)
15. Shang, W.Y., Liu, J.Y.: Stroke subtype classification: A comparative study of ASCO and modified TOAST. J. Neurological Sciences 314, 66–70 (2012)
16. Rector, A.: Modularisation of domain ontologies implemented in description logics and related formalisms including OWL. In: Proc. K-CAP 2nd Conf Knwldg Capture, Florida, pp. 121–128. ACM Press (2003)
17. http://protege.stanford.edu
18. http://www.w3.org/TR/owl-features
19. Query Language for RDF, W3C Recommendation (January 15, 2008), http://www.w3.org/TR/rdf-sparql-query/SPARQL
20. Bamford, J., Sandercock, P., et al.: Classification and natural history of clinically identifiable subtypes of cerebral infarction. Lancet 337, 1521–1526 (1991)
21. Ay, H., Benner, T., et al.: A computerized algorithm for etiologic classification of ischaemic stroke: the Causative Classification of Stroke System. Stroke 38, 2979–2984 (2007)
22. Adams Jr., H.P., Bendixen, B.H., et al.: Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial TOAST Trial of Org 10172 in Acute Stroke Treatment. Stroke 24, 35–41 (1993)
23. Amarenco, P., Bogousslavsky, J., Caplan, L.R., Donnan, G.A., Hennerici, M.G.: New approach to stroke subtyping: the A-S-C-O (phenotypic) classification of stroke. Cerebrovascular Disease 27, 502–508 (2009)
24. Dewey, H., Macdonnel, R., Donnan, G., Freeman, E., Thrift, A., Sharples, C.: Inter-rater reliability of stroke sub-type classification by neurologists and nurses within a community-based stroke incidence study. J. Clin. Neurosci. 8(1), 14–17 (2001)

Ensemble Classifier Systems for Headache Diagnosis

Konrad Jackowski[1,2], Dariusz Jankowski[2], Paweł Ksieniewicz[2], Dragan Simić[3], Svetlana Simić[4], and Michał Woźniak[2]

[1] IT4Innovations, VŠB-Technical University of Ostrava, 17. listopadu 15/2172, 708 33 Ostrava - Poruba, Czech Republic

[2] Wrocław University of Technology, Department of Systems and Computer Networks, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland, dariusz.jankowski@pwr.wroc.pl

[3] University of Novi Sad, Faculty of Technical Sciences, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia

[4] University of Novi Sad, Faculty of Medicine, Hajduk Veljkova 1-9, 21000 Novi Sad, Serbia

Abstract. Headache, medically known as cephalalgia, may have a wide range of symptoms and its types may be related and mixed. Its proper diagnosis is difficult and automatic diagnosis is usually rather imprecise, therefore, the problem is still the focus of intensive research. In the paper we propose headache diagnosis method which makes the decision on the basis of questionnaire only. It distinguished among 11 headache classes, which taxonomy is provided. The paper presents results of experiments which aim at selecting the best classification algorithm including several classical machine learning methods as well as ensemble approach. Results of experiments carried on dataset collected in University of Novi Sad confirm that the automatic classification system can gain high accuracy of classification for the problem under consideration.

Keywords: headache diagnosis, ensemble classifier systems.

1 Introduction

One of the most popular application area of computer decision support systems is medicine. According to [1] about 11% of such systems are devoted to medical domain and about 21% of research articles on applications of computer software which assist in decision making, are illustrated by medical cases. There are many works which describe computer-aided decision support systems: some present reviews of working medical decision support software [2], whereas others are related to the problems involved in choosing the best classification method for the particular medical task. In this work we will focus on the second area.

The aim of the classification task is to assign a given object to one of the predefined category. This decision is made on the basis of the selected object's characteristics. Nowadays developing more efficient and accurate recognition algorithms still remains an important topics for the machine learning community.

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_25, © Springer International Publishing Switzerland 2014

In the article we will present the preliminary results of our work on computer aided decision support system dedicated to headache diagnosis. Nowadays headache is one of the common disease which has high negative impact on productivity and the life quality of the affected persons. Unfortunately, the precise diagnosis of the headache type is very complex and usually imprecise, thus a high quality classifier dedicated to this task is desirable diagnostic tool.

The main contribution of this paper is the extensive comparative analysis of classification methods trained on the basis of dataset, which includes medical cases of patients suffered from migraine. The dataset was gathered by the team from the University of Novi Sad [3]. Now, the work on rule acquisition is in progress.

The content of the work is as follows. Section 2 introduces the medical problem. In the next section, we describe a mathematical model of the headache diagnosis. Then we present the experimental evaluation of the selected classifiers for the problem under consideration. The last section concludes the paper.

2 Headache Diagnosis

Headache, medically known as cephalalgia, is a continuous pain in the head or neck region. As the brain has no pain receptors, headaches are not felt in the brain. The pain is caused by disturbances of the pain-sensitive structures surrounding the brain. Headache may have many possible causes, which means that it is a non-specific symptom.

Our knowledge and understanding of headache have changed significantly in the twenty-first century. Headache pathophysiology and genetics are still in progress and new diagnosis systems and treatments are available for headache disorders.

In 2004, World Health Organization (WHO) accepted International Classification of Headache Disorders (ICHD-2) published by the International Headache Society (IHS), which the first version of ICHD was proposed in 1988 [5]. Those classification system become a standard for headache diagnosis and clinical research. The IHS provides operational definitions for all headache types.

The ICHD-2 divides headaches into two broad categories: the primary headache disorders (without organic cause) and the secondary headache disorder (where etiological cause can be determined).

The most frequent epidemiologic studies reported that over 90% of world's population suffer from headache. Tension-type headache and migraine are the most common among the general population (20.8% and 15% respectively) [8]. Secondary headaches occurs in a very small percent of the population. While secondary headache diagnosis can be easily established, tension-type headache (TTH) and migraine recognition can be a problematic even for an experience physician.

TTH is the most common form of headache, which can affects both woman and man. Pain radiates bilaterally (rarely unilaterally) from forehead to the occiput. It may present as a band-like nonpulsatile ache or tightness in frontal,

temporal and occipital regions. TTH changes during time and then tends to remain constant without progressive symptomatology. Usually associated with this type of headache are sleep problems such as: difficulty in falling asleep, awaking at night or early awaking. The trigger for acute episodic tension-type headache often is stress, sleep deprivation, hunger, uncomfortable position, bad posture or eyestrain. Regrettably, there are no specific tests for TTH.

Migraine is a genetic neurological disease. Typically the pain is pulsating and placed on one side of the head. Potential symptoms are: nausea and vomiting, increased sensitivity to sound/light/odors and pain, which is worsened by any physical activity. There are several types of migraine, however we may distinguish two major categories: migraine without aura and migraine with aura. Typical aura is accompanied by visual and/or sensory and/or speech symptoms, gradual development, duration no longer than one hour. Migraine without aura involves headaches not accompanied by an aura. It is clinically clear that certain headache types, especially migraines, in women have their own specifics. Relevant to clinical characteristics and based on research conducted by MacGregor and al. [9] migraine concerns women may be divided to: menstrual and non-menstrual. Further on, menstrual migraine could be divided on: pure menstrual migraine and menstrually-related migraine.

3 Materials and Methods

Primary headaches may have a wide range of symptoms: pain of different intensity and occurring, irritability, feeling *hyper*, sensitivity to light/sounds/motion, blind spots, weakness, pain, pressing etc. Also triggers that bring on headaches can interact. Some researches suggest that migraine and tension headaches may be related. Therefore diagnosing the cause could be difficult. Patients with disabling headache may be misdiagnosed as tension-type headaches instead of migraines. Physicians should be sensitive to the needs of headache sufferers and be aware of the latest advances in treatment.

However, it is possible to automatize diagnosis when patients with primary headaches considering that diagnosis here is made on the basis of questionnaire. Automation of the process would help reduce situations in which doctors, due to everyday busy schedule and concentration on far more rare secondary headaches forget to ask questions crucial for primary headache diagnosis.

Having in mind questions that physician asks in everyday clinical practice, and in accordance with ICHD, we have put together questionnaire about headache characteristics based on which we have made the diagnosis of a headache type (Table 1). The questionnaire is so conceived that it can help diagnose migraine, tension-type headache, and other headache types. It should be stressed that phrase "other headache types" does not mean other primary headaches but all the other headaches. Furthermore, it is possible, when it comes to migraine, based on the questionnaire, to determine whether it is a migraine with an aura or a migraine without aura. According to frequency of attacks, tension-type headaches could be divided in: rare episodic tension-type headache, frequent episodic tension-type headache, chronic tension-type headache.

There are other questions beside those based on criteria from ICHD-2 and they are related to: age when the first headache occurred, the frequency of headache attacks, headache triggers, localization, intensity and quality of the pain, related symptoms etc. The last two questions are only for female respondents, and they are important in classification menstrual migraine as sub-type of migraine.

Answers to these questions provide wider picture of a certain headache type, considering that every headache type and sub-type has its specifics and also delivers useful information for the building the predictive model.

In total, the questionnaire contains 30 questions. Provided answers can be one of three types: true/false, select one of the answers or typing the appropriate value. All the answers were converted into numerical values for classification. The full version of the questionnaire is available on-line[1].

3.1 Experimental Data

Our research were conducted on 579 patients (in an age between 20 to 67) on the area of Novi Sad (Republic of Serbia).

Table 1 presents a distribution of the diagnosis in 11 previously described migraine classes. The training set is used to train the models, while the test one enables to assess the models.

Table 1. Data distribution

Class	Diagnostic group	Sets		
		Training	Test	Total
1	Migraine without aura in men	8	8	16
2	Migraine without aura and pure menstrual migraine	4	3	7
3	Migraine without aura and menstrual related migraine	25	24	49
4	Migraine without aura and non-menstrual migraine	16	15	31
5	Migraine with aura in men	3	2	5
6	Migraine with aura and menstrual related migraine	18	17	35
7	Migraine with aura and non-menstrual migraine	13	13	26
8	Rare episodic tension type headache	58	58	116
9	Frequent episodic tension type headache	50	49	99
10	Chronic tension type headache	5	4	9
11	Other headache type	93	93	186
	Total	**290**	**289**	**579**

4 Problem Statement and Classification Algorithms

One essential issue while designing decision support system is selecting the best classification algorithm, i.e. algorithm which is able to effectively create proper

[1] http://www.kssk.pwr.wroc.pl/wp-content/uploads/downloads/2014/02/QUESTIONARE_headache.pdf

model of the problem in hand. In the range of inductive learning algorithms the information are usually stored in a form of learning set.

$$LS = \{(x_1, j_1), (x_2, j_2), \cdots, (x_N, j_N)\} \tag{1}$$

where x_n denotes the attribute set of *n-th* sample in the set, where

$$x_n = \left\{x_n^{(1)}, \cdots, x_n^{(d)}\right\} \in R^d \tag{2}$$

and j_N is its corresponding class label. It belongs to predefine and finite set of possible classes $j_n \in \mathcal{M} = \{1, \cdots, M\}$.

In the case of medical problems a record in database can consist of number of features, which represents results of medical examinations, and the diagnosis given by a physician. Usually a relation between the features and diagnosis is hard to define explicitly in form of the rules, therefore, the model has to be built automatically by classification algorithm Ψ.

$$\Psi : X \to \mathcal{M} \tag{3}$$

The task is not trivial. Firstly, as it was mentioned, the model of the relation is not known, therefore, it is hard to recommend the most suitable classification algorithm. Secondly, the classifier has to be able to generalize knowledge, i.e. be able to build the model which will be correct not only for available learning set but also for further records which will become the subject of the decision. Therefore, it is recommended to test number of classification algorithms in attempt to find the best one. Naturally such an approach makes sense if a set of tested classifiers varies in term of classification function, on training algorithm.

For the medical task in hand, the authors decided to test the following set of elementary classifiers: Naïve Bayes, Multilayer Perceptron, Support vector machine, Decision Tree, K-nearest neighbours. Elementary description of each algorithm is given below.

4.1 Elementary Classification Algorithms

Naïve Bayesian Classifier (NB) [10] assumes that attributes are independent when the class label is known. For parameter estimation Naïve Bayes uses the method of maximum likelihood. New objects receive class labels which have the highest *posterior* probability. Despite their naive design and oversimplified assumptions, the Naïve Bayes classifier happens to outperform more sophisticated methods.

*Multilayer Perceptron (*NN*)* is a neural network which consists of a set of source nodes forming the input layer, one or more hidden layers of computation nodes, and an output layer of nodes. Each node is a neuron or processing element. The input signal layer-by-layer is propagated through the network. In the training process (technique called backpropagation) the network must learn the dependencies between input and output nodesset of input (which means adapting the

weights in all neurons) [11]. The power of an Multilayer Perceptron network with even only one hidden layer is surprisingly large.

*Support Vector Machine (*SVM*)* is non-probabilistic binary linear classifier. Training algorithm builds a model (set of hyperplanes) to separate data with as wide as possible clear gap. Good separation is achieved when distance to the nearest training data point of any class (functional margin) is the largest. When data are not linearly separable in original space, the problem may be mapped into a much higher-dimensional space (so-called kernel trick) [12].

*Decision tree (*J48*)* is the WEKA implementation of the C4.5 classifier [13]. Induction of decision tree has two phases: (i) growth or build, (ii) prune. In the first one, the decision tree selects the best test attribute for a root test node and then recursively whole process is repeated. The pruning phase cuts off insignificant nodes or even subtrees. Algorithms usually constructs decision tree in top-down manner, by choosing the *best test attribute* that split the learning set most homogeneous [14]. The most popular homogeneity measures are: entropy, information gain (e.g. ID3, C4.5) and Gini index (e.g. CART).

*K- Nearest Neighbours (k-*NN*)* is a method that predicts a value of class label based on the k closest training samples in the feature space (where k is a positive integer, typically small one, defined by user). A commonly used distance metric for continuous variables is Euclidean distance but another metric can be used (e.g., for text classification). This method is preferred because it is easy to implement computationally and has high classification accuracy.

4.2 Ensemble Classifier System

Having got diverse set of classifiers has two major advantages. Firstly, it allows to find the most appropriate one among the others. The quality of the pretenders can be measured simply by misclassification rate, i.e. factor of samples drawn from available database which has been misclassified by given algorithm $P_e(\Psi)$.

Nonetheless, in practice it might happened that the performance of the best classifier still does not meet our expectation. One can point out many reasons:

(a) insufficient, not representative learning set, which does not consists enough information for building proper model of the decision problem;
(b) presence of noise of different kinds which spoils the data (e.g., noise injected by medical equipment, presence of irrelevant attributes, wrong description given by physician, etc.);
(c) too high complexity of the model, which cannot be approximated by simple classifiers (e.g., large set of predefined classes, complicated density distributions of attributes in feature space, etc.);

(d) not efficient learning algorithm which cannot fully extract the information from learning set. Here, one can point out the second advantage of diverse classifier set, i.e. possibility of creating ensemble classifier system which makes a decision in collective manner. Usually the system consists of committee of elementary classifiers which classify given sample independently.

$$\Pi^{\Psi} = \{\Psi_1, \Psi_2, \cdots, \Psi_K\} \tag{4}$$

Next, the system fuses classifiers responses in order to make final decision.

$$\bar{\Psi}(x) = F(\Psi_1(x), \Psi_2(x), \cdots, \Psi_K(x)) \tag{5}$$

where $\bar{\Psi}$ is ensemble classifier system, and F stands for combination rule. Selecting the best combination model is hard to do, because usually it is not possible to check all the combination due to high computational effort required. On the other hand, there are some heuristic algorithm available, but they do not guarantee obtaining the best results.

Authors decided to test two popular ensemble algorithms: *Majority Voting*, and *Weighted Combination of Discriminating Functions*.

*Majority Voting (*MV*)* . In this model the sample is labeled with the class which is most frequently represented in the set of responses.

$$\bar{\Psi}(x) = i \Leftrightarrow \sum_{k=1}^{K} \delta(\Psi_k(x), i) = \max_{i \in \mathcal{M}} \sum_{k=1}^{K} \delta(\Psi_k(x), i) \tag{6}$$

where δ stands for Kronecker delta.

There are two main advantages of majority voting model. Firstly it is intuitive, and secondly, it is relatively easy for implementation and does not require any additional training procedure as it does not have any additional parameters which have to be adjusted.

*Weighted Voting (*WMV*)* can be given as an example of more sophisticated approach in which a contribution of each classifiers in the ensemble into final decision depends on the quality of the classifiers and should be counter proportional to misclassification rate of the classifier.

$$\bar{\Psi}(x) = i \Leftrightarrow \sum_{k=1}^{K} w_k \delta(\Psi_k(x), i) = \max_{i \in \mathcal{M}} \sum_{k=1}^{K} w_k \delta(\Psi_k(x), i) \tag{7}$$

In the tested model the following normalized weight was used.

$$w_k = \frac{1 - P_e(\Psi_k)}{\sum_{t=1}^{K} (1 - P_e(\Psi_t))} \tag{8}$$

4.3 Feature Selection

As it was stated in the previous section, the accuracy of the classification systems can be negatively affected by presence of irrelevant attributes. Therefore, while designing classification algorithm, one can apply feature selection procedure which aims at choosing the subset of most useful attributes.

Keeping in mind, that classification algorithms use different models and learning algorithms, it seems to be clear, that feature selection procedure shall be done for each algorithm separately as the importance and usefulness of the attributes can differ from one classification algorithm to other. Attributes subset which would be the most appropriate for one algorithm might diminish the quality of the other [15,16,17]. There are many potential benefits of variable and feature selection: simplifying data visualization and data understanding, reducing the measurement and storage requirements, shortening training and utilization time, enhanced generalisation. The last one is the most important from our point of view. Therefore, we decided to apply *Backward Feature Elimination* [18]. In the backward elimination procedure one starts with all the features. In subsequent iterations the less useful feature is left out. The model is retrain and the number of variables can be further reduced by backward elimination algorithm. The total number of iterations is $n * (n + 1)/2 - 1$, where n is number of input variables. It is often argued, that forward selection is computationally more efficient, but in backward elimination process the importance of variables is assessed in the context of other included variables.

5 Evaluation of Classification Algorithms

As it was stated in Section 2, headache diagnosis problem in hand is compound decision problem due to its complexity, large number of classes, and limited size of empirical data collected. Therefore designing decision support system requires some experimental investigations while selecting the best solution architecture.

5.1 Objectives

The following objectives were set while preparing the plan of experiments.

1. Comparative analysis of several classification algorithms for evaluation of their usefulness for headache diagnosis.
2. Assessing the possibility of elevating the accuracy of classification by incorporating feature selection.
3. Evaluation of advantage taken from application of ensemble classification algorithms.

5.2 Experimental Environment

All experiments were carried on in KNIME framework using WEKA implementation of elementary classifiers (i.e., NB, NN, SVM, J48, k-NN). Ensemble classification algorithms (i.e., MV, and WMV) were implemented by authors. In both tested ensemble systems the committee consisted of all elementary classifiers.

Table 2. Summary of feature selection process

Classifier	Feature subset length	Selected features
NB	15	Q4, Q5, Q5, Q6, Q11, Q14, Q15, Q16, Q18, Q20, Q22, Q25, Q26, Q27, Q30
NN	9	Q4, Q5, Q6, Q7, Q8, Q27, Q28, Q29, Q30
SVM	20	Q4, Q5, Q6, Q7, Q8, Q10, Q12, Q13, Q15, Q16, Q17, Q19, Q21, Q22, Q24, Q25, Q26, Q27, Q28, Q30
J48	20	Q3, Q4, Q5, Q7, Q8, Q9, Q13, Q14, Q15, Q16, Q19, Q20, Q21, Q22, Q24, Q25, Q26, Q27, Q28, Q30
k-NN	20	Q1, Q2, Q3, Q4, Q5, Q6, Q8, Q9, Q10, Q11, Q13, Q14, Q16, Q18, Q20, Q22, Q25, Q26, Q28, Q30

In case of WMV, values of the weights have been calculated according to formula (8).

The Backward Feature Selection algorithm implemented in KNIME was used for feature selection. The subsets of features selected for different classifiers differ. Summary of feature selection process are presented in Table 2.

In order to ensure the reliability of the tests all the experiments was carried on using 5x2 cross validation methods. Additionally, comparative analysis results obtained by ensemble systems vs elementary classifiers was confirmed by application of statistical 5x2 cross validation F test with *0.05* confidence level. The null hypothesis stated that have the same accuracy. Therefore, rejection the null hypothesis means that the difference in accuracy between two compared classifiers is statistically significantly.

Results of experiments are presented in Table 3. It presents misclassification rate of tested algorithms (mean value and respective variance of all repetition of 5x2 cross validation procedure.

5.3 Observations and Conclusions

1. The quality of elementary classifiers differ significantly and varied in range between 61% of misclassification for k-NN with full attribute set and 33% for J48 decision tree. That fact shows how difficult is prior selecting appropriate classification algorithm as one cannot know if what would be resulting quality. Therefore, procedure of experimental evaluation of large range of available classification algorithms seems to be good solution while designing decision support system.
2. Application of feature selection methods allowed for elevation accuracy of the classification. In all cases, the classifiers which made classification based on properly selected subset of features obtained the better results that respective algorithms working on the basis of full feature set. Almost in all cases the difference were statistically proven (bolded results). In the case of k-NN the difference is close to 20%. Even in case of decision tree (J48), which

Table 3. Evaluation of classification algorithms with full and selected set of features for headache diagnosis problem; bolded results indicate statistically better feature classification method (in row); underlined score indicate the best classification algorithm for given feature selection method (in column); below the best score, list of all outperformed classifiers are given, for which the superior position was confirmed with statistical test

	Full attribute set		**Selected attribute set**	
Classifier	Error mean	Error variance	Error mean	Error variance
NB	0.4014	0.0019	0.3416	0.0010
NN	0.4045	0.0008	0.3489	0.0006
SVM	0.5544	0.0005	0.3579	0.0004
J48	0.3634	0.0012	0.3392	0.0004
k-NN	0.6097	0.0004	0.4135	0.0005
MV	0.3365	0.0005	0.2743 statistically better then SVM, k-NN	0.0003
WMV	0.3230 statistically better then MV, NN, SVM, J48, k-NN	0.0004	0.2763	0.0003

inherently has strong ability for feature selection, application of backward feature selection allowed to improve the accuracy by 3%.

3. For both feature selection methods, the best results were obtained by ensemble systems (underline scores). Taking advantage from diverse knowledge collected in elementary classifiers improved the accuracy by several percent points.
4. Nonetheless, the statistical tests do not always confirm statistical difference between the best ensemble system and elementary classifiers. Looking for a reason, one has to notice relatively high variance of performance of some classifiers (ex. NB). It indicate that the algorithm showed poor stability of learning process which could be negatively affected by variation of learning set. Therefore, one of arising challenge is elevating ensemble stability.
5. The problems with stability of learning procedure could be caused by relatively small size of available learning set and rather poor representation of some classes, as it was discussed in Section 2.

6 Conclusions

Cephalalgia widely known as the headache is common disease which affects society and diminish our life quality. After a years of intensive researches precise taxonomy of headache have been developed, nonetheless, diagnosis is very hard as the disease have a wide range of symptoms and the borders between the types is very hard to establish. The paper presented preliminary comparative study of classification systems which can distinguished among 11 types of headaches on the basis of patient responses for questions from questionnaire. The presented

approach is the continuation of the previous work on the headache classification, but former work consider simplified headache taxonomy. Looking for the most suitable machine learning algorithm several classical classification algorithms were tested including classifier ensemble approaches. The last one proved their high performance showing the highest accuracy. That fact ensures the authors that ensemble approaches shall become basis for further works. It is planned to enrich the system with module aiming at classification made with means of rule based algorithm.

Acknowledgement. The work was supported by the statutory funds of the Department of Systems and Computer Networks, Wroclaw University of Technology.

This paper has been elaborated in the framework of the project Opportunity for young researchers, reg. no. cz.1.07/2.3.00/30.0016, supported by Operational Programme Education for Competitiveness and co-financed by the European Social Fund and the state budget of the Czech Republic.

References

1. Liebowitz, J. (ed.): The Handbook of Applied Expert Systems. CRC Press (1998)
2. Kaplan, B.: Evaluating informatics applications - clinical decision support systems literature review. International Journal of Medical Informatics 64, 15–37 (2001)
3. Simić, S., Simić, D., Cvijanović, M.: Clinical and socio-demographic characteristics of tension type headache in working population. HealthMED 6(4), 1341–1347 (2012)
4. Kuncheva, L.I.: Combining Pattern Classifiers. Methods and Algorithms. Wiley (2004)
5. Headache Classification Committee of the International Headache Society. Classification and diagnostic criteria for headache disorders, cranial neuralgias and facial pain. Cephalalgia 8(suppl. 7), 1–96 (1988)
6. Brown, M.R.: The classification and treatment of headache. Medical Clinics of North America 35(5), 1485–1493 (1951), PMID 14862569
7. Ad Hoc Committee on Classification of Headache. "Classification of Headache". JAMA 179(9), 717–718 (1962), doi:10.1001/jama.1962.03050090045008
8. Vos, T.: Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380(9859), 2163–2196 (2012), doi:10.1016/S0140-6736(12)61729-2, PMID 23245607
9. MacGregor, E.A., Hackshaw, A.: Prevalence of migraine on each day of the natural menstrual cycle. Neurology 63(2), 351–353 (2004)
10. John, G.H., Langley, P.: Estimating Continuous Distributions in Bayesian Classifiers. In: Eleventh Conference on Uncertainty in Artificial Intelligence, San Mateo, pp. 338–345 (1995)
11. Cybenko, G.: Approximation by superpositions of a sigmoidal function. Mathematics of Control, Signals, and Systems 2(4), 303–314 (1989)
12. Boser, B.E., Guyon, I.M., Vapnik, V.N.: A training algorithm for optimal margin classifiers. In: 5th Annual ACM Workshop on COLT, pp. 144–152. ACM Press, Pittsburgh (1992)

13. Quinlan, R.: C4.5: Programs for Machine Learning. Morgan Kaufmann (1993)
14. Quinlan, J.R.: Induction of Decision Trees. In: Machine Learning, vol. 1, pp. 81–106. Kluwer Academic Publishers (1986)
15. Hall, M.A.: Correlation-base feature selection for discrete and numeric class machine learning. In: Proceedings of the 17th International Conference on Machine Learning, pp. 359–366 (2000)
16. Kohavi, R., John, G.: Wrappers for feature subset selection. Artficial intelligence 97(1-2), 273–324 (1997)
17. Guyon, I., Weston, J., Barnhill, S., Vapnik, V.: Gene selection for cancer classification using support vector machines. Machine Learning 46(1-3), 389–422 (2002)
18. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. The Journal of Machine Learning Research (2003)

Part V

Signal Processing

Rate Adaptive Atrioventricular Delay in the Cardiac Pacing

Kazimierz Pęczalski[1], Dariusz Wojciechowski[2], and Tadeusz Palko[1]

[1] Warsaw University of Technology, Poland
K.Peczalski@mchtr.pw.edu.pl
[2] IBBE PAS, Poland

Abstract. Physiological shortening of atrioventricular (AV) delay with increasing of heart rate is well documented in healthy human beings. The specialized algorithms were developed that control the AV delay. The goal of our study is the assessment of the impact of varying heart rate on the optimal AV delay in paced patients. The SV values were assessed by the Impedance Cardiography (IC) and Penaz Plethyzmography (PP) at the same time for each patient. The statistical analysis has showed no statistically significant differences between results obtained simultaneously by both evaluation methods. Three types of AV optimal changes following pacing rate increase were found: 1. physiological (shorter), 2. neutral (no change), 3. atypical (longer).

Keywords: atrioventricular delay adaptation, cardiac pacing, cardiac output optimization.

1 Introduction

The hemodynamics of the pacemaker patients depends not only on the general determinants of cardiac functions, such as contractility of the myocardium, preload and afterload, but also on the factors resulting from the patient conduction defect and the applied pacing mode. The physiological pacing is characterized by the physiological sequences between atrial and ventricular contractions that are influenced by the exercise adequate sinus rate control, and the orthograde stimulus conduction and propagation.

The DDDR (dual channel prohibited and synchronized pacing according to pacemaker code) mode of pacing is usually applied for the patients with chronotropic incompetence. The hemodynamics of atrioventricular (AV) synchronous paced heart is related to the rate adaptation, AV synchrony, retrograde AV conduction and AV interval. The cardiac output during exercise depends most strongly on the variation of the heart rate . The stroke volume can be increased during exercise by about 50% versus the heart rate increase by approximately 250%. So, in chronotropic incompetence patients the adaptive control based on the heart rate seems to be the best and the most effective solution. Since their introduction, the automatic adaptation of the pacing rate to the momentary metabolic needs of the patients improved significantly their exercise tolerance.

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_26, © Springer International Publishing Switzerland 2014

The crucial part of all rate adaptive pacemakers is the algorithm which converts changes of the heart rate signal to changes of the pacing rates. The heart rate sensor signal indicates whether the pacing rate should be increased or decreased. The ratio of heart rate variation to the pacemaker rate change depends on patients general condition, age, disease etc. The wide range of the heart rate sensors have been tested. Presently the activity and respiration (minute ventilation) sensors are most widely applied. The activity sensors are based on a piezocrystal or an accelerometer sensors. The ventilation sensors are based on the impedance measurements. The impedance methods are applied also in other sensors such as stroke volume or preejection time interval sensors.

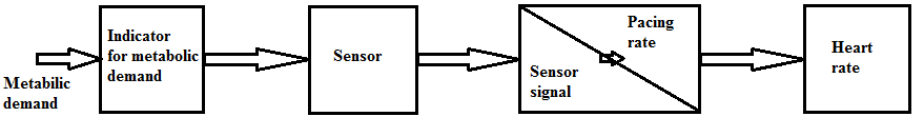


Fig. 1. An open loop algorithm

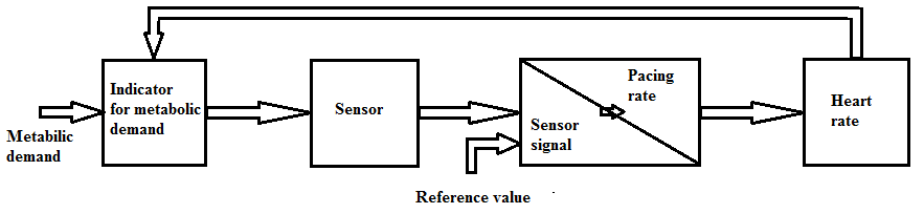


Fig. 2. A closed loop algorithm

Generally two concepts of rate adaptation are applied - open and closed loop [1]. The one way open loop algorithm is presented in Figure 1. The sensor delivers the electric signal to the algorithm and the algorithm converts the changes of the signal to the changes of pacing rate. The most important factor in this process is the step response of pacemaker rate change to the change of unit sensor signal. In an open loop algorithms such factor is known as a rate response curve or slope. The closed loop system sensor (Figure 2) has more direct relationship with the pacing rate. The sensor signal is influenced by metabolic demands and by the momentary pace rate in opposite way. The metabolic demands increase causes increase of sensor signal and heart rate increase causes decrease of sensor signal. The pacing rate is increased or decreased until the sensor signal has reached its reference value within the predetermined error value. The pacing rate is stable when the sensor signal is equal to the reference value.

The conduction of electrical impulses by the heart is influenced by the autonomic nerve activity. This autonomic control is most apparent at the AV node. The sympathetic activation increases conduction velocity in the AV node by increasing the rate of depolarization. This leads to more rapid depolarization of

adjacent cells, which leads to a more rapid conduction of action potentials. The sympathetic activation of the AV node reduces the normal delay of conduction through the AV node, thereby reducing the time interval between atrial and ventricular contractions. The parasympathetic activation increases conduction delay at the AV node and increases the time interval. The sympathetic activation increases and parasympathetic activation decreases the heart rate. Therefore the heart rate increase is followed by the AV delay decrease and the heart rate decrease is followed by the AV delay increase. Physiological shortening of AV with increasing heart rate is well documented in healthy human beings. The special algorithms were developed that control the AV delay. The principle of AV control algorithm are presented in Figure 3.

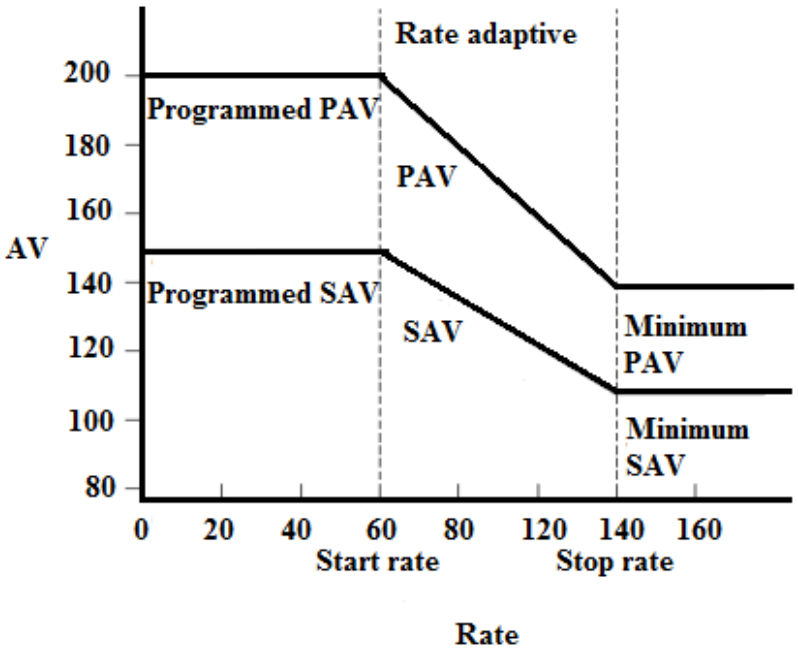


Fig. 3. The AV delay control algorithm

The AV delay is adapted to heart rate in range from 60 to 120 beats per minute (bps) with step 10 ms/10 bps for paced cycles (PAV) and 5 ms/10 bps for sensed cycles (SAV). The range and step values are adjustable. The Rate Response Cardiac Pacing with adaptive AV delay function is applied in dual channel pacing (DDDR), Cardiac Resynchronization Therapy (CRT) and DDDR Implanted Cardioverter Device (ICD) treatment. Above listed methods are dedicated to critical heart failure patients when the optimization of heart rate (HR) and the AV delay gaining of cardiac output is crucial.

2 Aim of the Study

The goal of our study is the assessment of the impact of varying HR on the optimal AV delay in paced patients.

3 Methodology

Group of 20 patients (mean age 77.4 years +/- 4,6 years) consists of 8 females (mean age 77.6+/- 5.5) and 12 males (mean age 77.3+/-4.1) treated by DDD were enrolled in comparative study.

The criteria of enrolment were as follow:

1. More than 80% of paced heart beats in holter recording of pacemaker system (stored in internal memory of pacemaker circuit),
2. 100% of paced heart beats during examination.

The criterion of maximal stroke volume (SV) were applied for two utilized methods of AV delay optimization. The SV values were assessed by the Impedance Cardiography (IC) and Penaz Plethyzmography (PP) at the same time for each patient. The following settings of pacing were applied: AV coupling from 100 to 200 msec. with step of 20 msec. for pacing rates of 70 and 90 beats per minute.

4 Results

The statistical analysis, done with the Wilcoxon test ($p<0.05$), has showed no statistically significant differences between the AV optimal values obtained simultaneously by both evaluation methods for the two applied pacing rates.

Table 1. Distribution of types of reaction

	Type of reaction		
Measurement method	Typical	Neutral	Atypical
IC	7	6	7
PP	6	8	6
The same for both methods	6	6	6

Three types of AV optimal changes following pacing rate increase were found:

- physiological (shorter),
- neutral (no change),
- atypical (longer).

The numbers of patients presenting particular reaction for each applied method are presented in Table 1.

5 Discussion

Use of rate adaptive AV delay in patients treated by cardiac pacing is controversial [2,3]. The common used optimization algorithm is based on changing of AV delay until best E and A wave separation is seen on mitral inflow of pulsed wave Doppler of echocardiography [4]. In some patients the shortening of AV delay, with increasing of heart rate, improves hemodynamic parameters of the heart activity (such as SV or Cardiac Output). Currently, optimization of AV delay is mainly performed during resting conditions however some centers report optimization during exercise [5]. We postulate to introduce Impedance Cardiography (one of commercially produced systems was recently certified by FDA) or Penaz Plethysmography as a standard methods for optimization of AV delay in cardiac pacing. Recommended methods compared to ultrasound methods (echocardiography) are simpler, can be performed by less trained operators and the measure procedure takes less time. The cases of increase of optimal AV delay with heart rate increase should be confirmed on wider group of patients. The new algorithm consist of typical and atypical response of AV delay value to heart rate variations should be taken under consideration.

6 Conclusions

1. The Impedance Cardiography or Penaz Plethyzmography can be applied for estimation of maximal value of cardiac output and optimal atrioventricular delay,
2. The AV adaptation algorithms should be applied after determination of patient reaction to the rate increase.
3. Individualized pacemaker programming should be performed for each patient treated by cardiac pacing.

References

1. Alt, E., Barold, S.S., Stangl, K.: Rate adaptive Cardiac Pacing. part 1. Springer, Heidelberg (1993)
2. Melzer, C., Bondke, T., Koerber, T., Nienaber, A., Baumann, G., Ismer, B., Chen, W.K.: Should we use the rate-adaptive AV delay in cardiac resynchronization therapy-pacing? Europace 10(1), 53–58 (2008)
3. Bogart, M., Meine, M., Tuinenburg, A., Maskara, A., Loh, P., Doevendans, P.: Cardiac resynchronization therapy beyond nominal settings: who needs individual programming of the atrioventricular and interventricular delay? Europace 14(12), 1746–1753 (2012)
4. Singh, J., Abraham, W., Chung, E., Rogers, T., Sambelashvili, A., Coles, J., Martin, D., Belmont, C.A.: Wadsworth: Clinical response with adaptive CRT algorithm compared with echocardiography optimized atrioventricular delay: a retrospective analysis of multicentre trials. Europace, 1622–1628 (November 2013)
5. Bogard, M., Kirkels, J., Hauer, R., Loh, P., Doevendans, P., Meine, M.: Should we optimize cardiac resynchronization therapy during exercises? J. Cardiovascular Electrophysiology 21(11), 1307–1316 (2010)

Decision Support System for Real-Time Monitoring of Patients with Cardiovascular Risks

Wojciech Tylman[1], Tomasz Waszyrowski[2], Andrzej Napieralski[1], Zbigniew Kulesza[1], Rafał Kotas[1], Paweł Marciniak[1], Radosław Tomala[1], Maciej Wenerski[1], and Marek Kamiński[1]

[1] Łódź University of Technology, Department of Microelectronics and Computer Science, ul. Wólczańska 221/223, 90-924 Łódź, Poland
tyl@dmcs.p.lodz.pl
http://www.dmcs.p.lodz.pl/web/tyl/
[2] Norbert Barlicki University Clinical Hospital No. 1 in Łódź, ul. Kopcińskiego 22, 90-153 Łódź, Poland

Abstract. This paper presents decision support system for estimation of patient's general state with particular stress laid on cardiovascular and pulmonary diseases. It works in conjunction with a standard medical monitor, which provides input signals such as temperature, blood pressure, pulseoxymetry, ECG, ICG, etc. The signals are preprocessed and analysed by a set of algorithms, the core of which is based on Bayesian networks. As an output the system should give information about detected problems, possible future threats, possible causes and suggestions for further treatment. From the hardware point of view the system is implemented in DaVinci DM3730 embedded microprocessor. Use of the presented system might be especially important in places such as Accident Emergency Departments and Admission Rooms, where small medical team has to take care of many patients in various general conditions.

Keywords: decision support, sudden cardiac arrest, Bayesian networks.

1 Introduction

Cardiovascular diseases are one of the most common causes of death in the world [8]. They are also characterised by often sudden onsets of the severe stage of illness, leading to a rapid deterioration of patient's state. These include Sudden Cardiac Arrest (SCA), the most common cause of hospital mortality. If such deterioration takes place, immediate preventive steps have to be taken in order to sustain patient's life. It would be even better to *predict* the onset of the severe stage in order to take preventive steps before actual deterioration takes place.

The experiences with Modified Early Warning Score (see Section 2) show that even such a simple approach allowed to reduce the risk of SCA and hospital mortality [2]. It is therefore tempting to expand this approach to make it automatic, based on more data pertaining to the condition of the patient and utilising advanced reasoning algorithms in order to predict problems with patient's health. This paper presents research aimed at developing such a system.

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_27, © Springer International Publishing Switzerland 2014

The usefulness of the proposed concept is particularly evident in situations when the doctors may have difficulties with devoting enough attention to the particular patients. These includes admission rooms and Accident and Emergency departments, where limited staff may have to deal with a suddenly increased number of cases. Moreover, in some circumstances the doctor may not be present at all: ambulances are a typical example of such situation.

The remainder of this paper is organised as follows: Section 2 introduces Modified Early Warning Score, a commonly used indicator of patient's condition and its change. Section 3 discusses recent developments in the field of decision support systems for cardiovascular diseases. Section 4 outlines the proposed concept, including the hardware issues and the role of a typical medical monitor. Section 5 presents the issues relating to the Bayesian network – the central reasoning element of the system, including an example of the network structure and problems related to the Bayesian network simulation software. Section 6 concerns the proposed user interface, while Section 7 offers closing remarks and plan of future work.

2 Modified Early Warning Score

Modified Early Warning Score (MEWS) is a simple scale used in hospitals and by paramedic services to quickly determine the patient's condition and changes of this conditions. Although there are various MEWS scales, the most common inputs required to compute the score are systolic blood pressure, heart beat rate, respiratory rate, body temperature and level of consciousness (using the Alert, Voice, Pain, Unresponsive scale). The simplicity of the scale allows the score to be computed without the need of any computing device, simply by referring to a compact chart summarising the possible reading values, as presented in Table 1. The rise of the score by a value of 2 is treated as an indication for consultation with a doctor, while the rise by a value of 4 as an indication for immediate transfer to the Intensive Care Unit.

Even though MEWS has proven effectiveness, it is a very simple scale. It does not make use of data currently often available even in the simplest paramedic scenarios, such as for example the Electrocardiography (ECG) signal, which carries enormous amount of data important for assessing the cardiovascular problems. Moreover, the concept of combining the input data using a table is possible only in very simple cases of few inputs.

It should be noted that there are in use much more elaborate scales, such as e.g. the Acute Physiology, Age, and Chronic Health Evaluation (APACHE III) scale. They allow to include much more input data, but this makes the use of such scale difficult in the paramedic or Emergency Department environment: all parameters have to be entered into a computer program manually. As APACHE III uses 20 input parameters, and determining some of them requires laboratory tests, it cannot be used in the same way as MEWS.

Table 1. A sample MEWS chart

Readings, observations	Score 3	2	1	0	1	2	3
Systolic blood pressure	<45%	30%	15% down	Normal for patient	15% up	30%	>45%
Heart rate [beats/min]	-	<40	41-50	51-100	101-110	111-129	>130
Respiratory rate [breaths/min]	-	<9	-	9-14	15-20	21-29	>30
Temperature [°C]	-	<35	-	35.0-38.4	-	>38.5	-
AVPU	-	-	-	A	V	P	U

3 Cardiovascular Computerized Decision Support Systems

In recent years, a number of computerised decision support systems for use in diagnosis of cardiovascular diseases have been designed. Some of them are based on advanced concepts of soft computing. For example, in [6] the authors propose a system for diagnosis of chest pain utilising rule-based inference. The system is based on the knowledge learned from real patient data. In [4], the authors describe another system for chest pain diagnosis, this time built on the ontology approach and Bayesian Networks. Yet another method is outlined in [13], where data mining algorithms and fuzzy models are used as a basis for the decision support system.

Although such computerised systems turn out to improve quality of care [5], [1], [11], none of them automatically acquires the data about the patient in real-time. Instead, they are based on patient records or additional questionnaires, which, similarly to APACHE III discussed in the previous section, make them difficult or impossible to use in cases when the patient record is not available or there is no time to enter all the required data. They also cannot respond to the quickly changing patient's state. The system proposed in this paper aims to overcome these difficulties.

4 Proposed System Overview

4.1 Role of a Medical Monitor

In today hospital and paramedic practice medical monitors are commonly used. These are devices that acquire constantly a number of signals that are indications

of a patients health. The signals typically include ECG, body temperature, respiration rate, blood oxygen saturation and blood pressure. More advanced models also have possibilities of performing some analyses on the received signals and virtually all support some means of exporting the acquired data: either by saving them on removable flash memory, or by continuous transmission using popular network standards (typically IEEE 802.3/Ethernet).

The proposed solution relies on the medical monitor as a source of data about patient's condition. Such approach has a number of advantages:

- reduced cost, as it utilises the infrastructure already present in the hospitals/ambulances,
- simpler operation, as there is no need to attach additional wires to the patient,
- simpler development, as there is no need to develop the part responsible for acquiring signals.

On the other hand, this approach requires a medical monitor that exports the acquired data in a continuous manner, it is also necessary to understand the format of data employed by the monitor. For development purposes an Emtel FX2000-MD (Figure 1) has been selected: it exports data over Ethernet connection and the manufacturer made available – for the purpose of research – the data format.

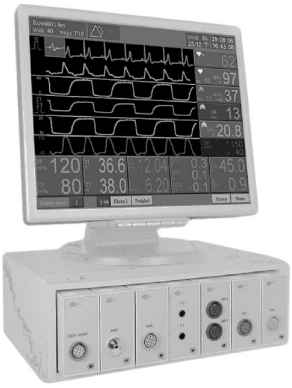


Fig. 1. Emtel FX2000-MD medical monitor

4.2 Hardware Considerations

Because the proposed solution does not have to acquire signals, its role will be to process data and interact with the medical personnel. Consequently, from the hardware point of view, the proposed solution is nothing else than a computer attached to the medical monitor: it must have significant computing power and keyboard and screen for interaction with the medical personnel.

However, employing of-the-shelf PC computer, notebook or tablet is not advisable. Such devices are not designed to operate in difficult conditions that can be encountered in ambulances or Accident and Emergency departments. Besides, they are equipped with a number of components that are unnecessary in the developed solution and only result in increased battery consumption, increased size or reduced reliability (additional communication interfaces or CD-ROM drives are examples of such components).

For this reason it has been decided to build the system around an embedded processor. Due to the previous experience of the team, Texas Instruments DaVinci DM3730 Digital Media Processor [12] has been chosen, which integrates in one integrated circuit an ARM Cortex-A8 processor, a TMS320C64x+ DSP and a POWERVR SGX graphics accelerator. The presence of an DSP may be advantageous for implementation of the processing algorithms. As a prototyping platform Embest DevKit8500 (Figure 2) utilising this processor has been selected. In order to provide interaction with medical personnel, an LCD touchscreen has been attached to the board.

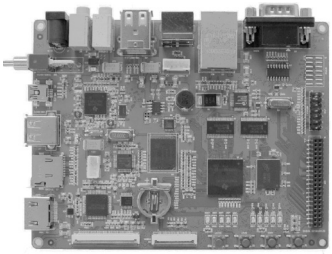


Fig. 2. Embest DevKit8500 prototyping board

4.3 Structure of Software

The key to obtaining the desired functionality is the software running on the embedded processor. The software consists of a number of interconnected blocks. These can roughly be divided into input/preprocessing blocks, AI blocks and output/presentation blocks. The task of the first is to gather data from the medical monitor and perform preliminary analyses on these data. The analyses can vary from very simple ones (averages, differences over time) to advanced (detecting predefined anomalies in the ECG data). The task of the AI blocks is to combine all input data in order to reach conclusion about patient's state and detect indications of state deterioration; an example of such block will be discussed in Section 5. The output/presentation blocks are responsible for presenting the results of analyses to the medical staff and also storing these for further reference.

5 Bayesian Network AI Module

Bayesian network has been chosen as the main AI approach in the proposed solution. It is planned that AI modules based on other approaches will also be constructed in order to allow comparison of performance; however, for the time being efforts are focused on the Bayesian network. Consequently, only this approach will be described in the paper.

5.1 Bayesian Networks

Bayesian networks employ directed acyclic graphs to encode qualitative relationships between variables. The nodes of the graph correspond to the variables, and the edges correspond to the dependences between them. In this way it is possible to express the independence between variables (by not connecting some of the nodes) and to encode the joint probability distribution of the variables in a very compact way. To encode the joint probability it is also necessary to specify the quantitative relationships between variables. These take form of conditional probabilities for the variables that depend on other variables and unconditional probabilities for the rest of them. For the most common case of discrete variables these are expressed using probability tables.

Inference with Bayesian networks involves setting the state of the observed variables (entering evidence) and applying one of numerous algorithms that allow to compute probabilities of the other variables. It is possible to freely choose the nodes for which the evidence is entered, so the Bayesian network is an ideal tool for reasoning in cases when some of the input data might be missing.

Another important feature of Bayesian networks is their ability to learn. It is possible to learn both the structure (graph) and the probabilistic tables. Again, a number of algorithms to perform these procedures are available.

For in-depth discussion of Bayesian networks refer to [10], [9], [3].

5.2 Proposed Bayesian Network Structure

As already mentioned, the task of the Bayesian network in the system is to fuse various pieces of data coming from the medical monitor, after the preprocessing procedures have been performed. The Bayesian network should be in this way capable of detecting a number of conditions that pose immediate danger to the patient.

A preliminary structure of such network is presented in Figure 3. The network is organized in such a way that the nodes (ovals) in the upper part represent the diseases that the network is designed to detect, while the other nodes represent the possible observations and symptoms. Among the latter, the nodes with no outgoing edges are strictly associated with specific observations, while those with outgoing edges may be observations (e.g., *Raised ST segments in two leads*), or are only supporting elements, grouping observations from other nodes (e.g., *ECG abnormalities*). In general, the direction of the arrow is from the diseases to symptoms.

It should be noted that the connection (edge) between a disease and a symptom does not necessarily implies that the symptom is present for the particular disease. In fact it indicates that the presence or *absence* of a symptom is typical for the disease. Consequently, in some cases the edge indicates that the symptom is not present in the disease. This is the case for example with the *Amplitude of ST vs. T in V6> 0.25* symptom and the *Early repolarization* disease.

The structure of the network will be further expanded to include more diseases or conditions requiring prompt actions of the medical staff. Moreover, to make the network operable, the numerical specification will have to be provided. Currently, such specification is being constructed based on a questionnaire utilising Likert scale [7]. Further work includes possibility of automatic adjusting of the numerical specification using patient database and doctor's diagnoses and decisions.

5.3 Bayesian Network Simulation Software

As the designed system runs on non-PC hardware, an important issue that had to be solved was to provide software implementing algorithms for Bayesian network computation. In general, a number of software packages exists, but they are targeted at the Intel x86/x64 architecture and cannot be used in the designed system, which is based on the ARM architecture. In order to solve this problem the team was working with the Decision Systems Laboratory at the University of Pittsburgh in order to produce a version of the SMILE library, developed Decision Systems Laboratory, compiled for the ARM Cortex-A8 processor. Because this work initially proved abortive, an alternative path has been pursued, aiming at providing an implementation of the computing algorithms written from scratch.

As a result of these parallel efforts, the team now is able to use two implementations of the algorithm: the SMILE library, which is a comprehensive approach containing a number of different algorithms and support for learning of the probability tables, and the team's own implementation, which is lightweight and currently lacks support for learning. Which of these will be finally used is up to the results of tests that are currently being conducted.

6 User Interface

In order to communicate with the medical personnel, the system is equipped with an LCD touch-screen. The designed software should therefore provide an appropriate Graphical User Interface (GUI). The obvious aim of the GUI is to provide the results of the analyses, the warnings concerning patient's condition and the indications suggesting possible treatment procedures. However, the GUI should also be able of collecting some data from the medical personnel, such data as are not available from the medical monitor, but are helpful in the operation of the algorithm.

The whole GUI is organised as a number of tabs that can be selected one at a time. Figure 4 presents two of such tabs (the names of some of the others can

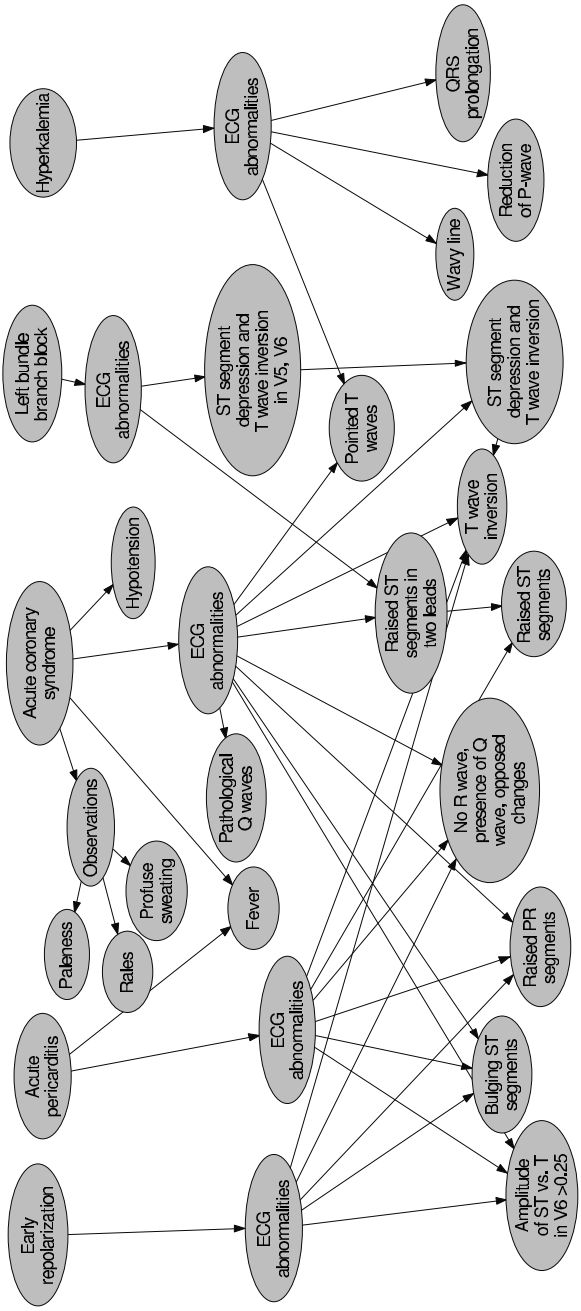
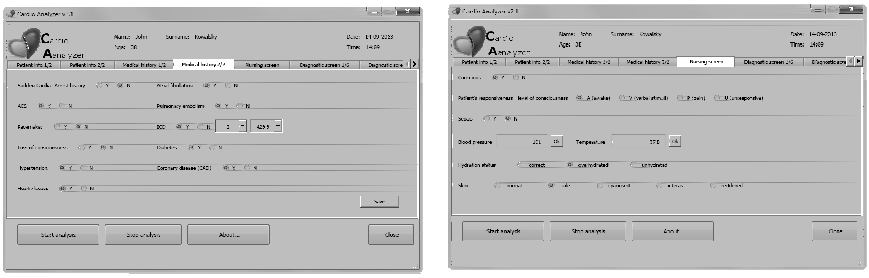


Fig. 3. Preliminary structure of the Bayesian network for detection of cardiovascular problems

be seen near the top of the screen). The former presents the data that can be entered based on the patient's medical history and contain mainly information about the most serious diseases that may immediately affect patient's condition and consequently should be taken into account by the decision algorithm. These data should be entered upon connecting the patient to the medical monitor. The latter is the screen that can be used by nurses in order to supply additional information not available through the medical monitor. Some of these (e.g., level of consciousness) are never measured by the medical monitor, while others (e.g., systolic blood pressure, temperature) may be available through the monitor, but this is not always the case.



(a) Medical history tab (b) Nurse's tab

Fig. 4. User interface

7 Conclusions

This paper presented work concerning development of a real-time monitoring system for patients, with emphasis on detecting situations related to cardiovascular diseases that may pose immediate risk to the patients' health or life. The aim of the system is to provide support for (para)medical staff, especially in situations where the doctors may not be immediately available or have to attend a large number of patients. The system utilises a standard medical monitor as the source of data.

The system in the current stage of development allows to gather data from the medical monitor in real time, perform preprocessing of such data and use them to compute the value of MEWS. Intuitive GUI is provided for interaction with medical staff. Two computation libraries for Bayesian networks are set-up for use on the embedded processor.

At this moment work focuses on the development of the Bayesian network, the central reasoning element of the system. The initial structure of the network is designed and the numerical specification is being constructed. Once this step is completed, the system will be subjected to clinical tests.

Acknowledgement. The graph of the Bayesian network has been designed using the GeNIe modelling environment. The system may use SMILE reasoning engine for graphical probabilistic model. GeNIe and SMILE are developed by the Decision Systems Laboratory of the University of Pittsburgh and available at http://genie.sis.pitt.edu/.

The project described in this paper has been financed through National Science Centre funding granted by decision number DEC-2011/01/B/ST6/04726.

References

1. Bright, T.J., Wong, A., Dhurjati, R., Bristow, E., Bastian, L., et al.: Effect of clinical decision-support systems: a systematic review. Ann. Intern. Med. 157(1), 29–43 (2012)
2. Burch, V.C., Tarr, G., Morroni, C.: Modified early warning score predicts the need for hospital admission and inhospital mortality. Emerg. Med. J. 25, 674–678 (2008)
3. Cowell, R., Dawid, P., Lauritzen, S., Spiegelhalter, D.: Probabilistic Networks and Expert Systems. Springer, New York (1999)
4. Farooq, K., Hussain, A., Leslie, S., Eckl, C., Slack, W.: Ontology-driven cardiovascular decision support system. In: Proceedings of the 5th International Conference on Pervasive Computing Technologies for Healthcare, pp. 283–286 (2011)
5. Guide to Community Preventive Services. Cardiovascular disease prevention and control: clinical decision-support systems (CDSS), `http://www.thecommunityguide.org/cvd/CDSS.html` (accessed February 23, 2014)
6. Kong, G., Xu, D.-L., Body, R., Yang, J.-B., Mackway-Jones, K., Carley, S.: A belief rule-based decision support system for clinical risk assessment of cardiac chest pain. European Journal of Operational Research 219(3), 564–573 (2012)
7. Likert, R.: A Technique for the Measurement of Attitudes. Archives of Psychology 140, 1–55 (1932)
8. Mendis, S., Puska, P., Norrving, B. (eds.): Global Atlas on cardiovascular disease prevention and control. WHO (2011)
9. Neapolitan, R.: Learning Bayesian Networks. Pearson Prentice Hall, Upper Saddle River (2004)
10. Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers, San Mateo (1988)
11. Piazza, G., Goldhaber, S.Z.: Computerized Decision Support for the Cardiovascular Clinician: Applications for Venous Thromboembolism Prevention and Beyond. Circulation 120(12), 1133–1137 (2009)
12. Texas Instruments DM3730 Digital Media Processor description, `http://www.ti.com/product/dm3730` (accessed December 20, 2013)
13. Tsipouras, M.G., Exarchos, T.P., Fotiadis, D.I., Bechlioulis, A.: Decision Support Systems for Cardiovascular Diseases Based on Data Mining and Fuzzy Modelling. In: Strategic Information Systems: Concepts, Methodologies, Tools, and Applications, pp. 1520–1530. IGI Global, Hershey (2010)

A Multivariate Sample Entropy of Differentiated Electrohysterographical Signals for an Identification of an Uterine Labor Activities

Dariusz S. Radomski

Nuclear and Medical Electronics Division of the Radiolelectronic Institute,
Warsaw University of Technology, Warsaw, Poland
d.radomski@ire.pw.edu.pl

Abstract. Electrohysterography measures a bioelectrical activity of a uterus. The methods which were published up to date still give inaccurate labour prediction/prediction of labour. This paper presents an application of multivariate sample entropy to differentiated 4-channels electrohysterographical signals. This nonlinear measure uses joint information from a multivariated signals. The obtained results confirm that a differentiation may improve sensitivity of a multivariate sample entropy on a uterine bioelectrical activity. It enables to recognize uterine contractions in labor with the sensitivity equaled too 87% and the specificity about 85%. These results should be verified in a prospective study.

Keywords: electrohysterography, multivariate sample entropy, nonlinear signal analysis.

1 Introduction

Monitoring of uterine contractions and a fetal heart rate during a 3^{rd} trimmest of a pregnancy is a crucial obstetrical procedure beside an interview, a clinical and ultrasonographical examination. From a clinical point of view, an identification of uterine contractions as well as changes in the cervix is very important. It allows to recognize an approaching labor. Electrohysterography may be useful for detecting a preterm labor risk. Preterm labors are still the main obstetrical problems which imply many neonatal complications. Some of them cause physical and mental disabilities e.g. infantile cerebral palsy [1]. Therefore, preterm labors generate many costs associated with neonatal intensive care and further medical and social rehabilitation. However, uterine contractions monitoring may be clinically useful to select the best way of delivery in post-term pregnancies and labors as an additional information. Post-term birth also produces a fetal ischemia causing newborns' complications [2]. Localization of a uterus in pelvis and an abdominal cave strongly limits measuring methods of uterine contractions. Routinely, these contractions are monitored by tocography, i.e. a dynamometer measures mechanical pressure of a contracted wall of a pregnant uterus. Despite visualizing a local uterine contractions, this method has some disadvantages Firstly, tocography is

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_28, © Springer International Publishing Switzerland 2014

local, i.e. it disables to visualize propagation of a myometrium contraction. Secondly, a correlation between measured mechanical pressure and an intrauterine pressure moving a fetus into a cerivix and a vagina is low. Thirdly, tocography is ineffective for obese women because a fatty tissue forcibly damps a mechanical pressure decreasing a sensitivity of this method [3]. Therefore, an alternative method called *electrohysterography* (EHG) is developing It measures bioelectrical activity of myometrium which proceeds its mechanical contractions. However, bioelectrical activity of a myometrium is not characterized by any constant pattern so EHG signals are rather similar to EMG or EEG signals than to ECG signals. Thus, it is very difficult to find such parameterization of EHG signals which highly correlate with a physiological state of an uterus. There are many methods published beginning with linear methods basing on a spectral analysis and ending with nonlinear methods [4,5]. The results obtained by Hassan *et al.* as well as our yet unpublished data strongly support hypothesis that a dynamic of EHG signals is nonlinear [6]. It is also conformable with the multiscale model of a uterine electrical activity developed by Marque *et al.*. This model contains nonlinear differential equations which describe changes of biopotentials measured by electrodes in EHG devices [7]. Among different methods of nonlinear signal analysis the *approximate entropy* (AppEn) as well as its unbiased version called *sample entropy* (SamEn) is most frequently used for biological signals. They are particularly useful when a regularity of the analyzed signal is associated with the interesting clinical state of a patient. Many results confirmed that values of AppEn and SamEn computed for EHG signals can predict contractile activities of an uterus. But this prediction was done on the basis of the values calculated for the singular *a prori* selected EHG signal [8]. There have been numerous results published which suggest that multichannel measuring of a bioelectrical uterine activity may be more useful. For example, this manner enables to reconstruct a spatial distribution of uterine biopotentials or to visualize how bioelectrical waves propagate during uterine contraction [9]. However, there are a few results presenting application of nonlinear analysis to multivariate EHG signals which use joint information contained in the EHG vector. The example of this method is a nonlinear correlation coefficient proposed by Hassan *et al.* [6]. Therefore, the goal of the presented work is to investigate a prediction accuracy of labor uterine activities based on values of multivariate sample entropy computed for a four elements vector of EHG signals. Moreover, a proposition to use a multivariate sample entropy for the differentiated EHG signal was also studied This idea stems from theoretical models and consideration noting that a mechanical contraction of a myometrium fiber is stimulated by ions flow but not by a constant ions concentration [10].

2 Measuring of a Uterine Electrical Activity

In contrast to other electrophysiological measurements neither standard measuring technique nor study design have been elaborated for EHG signals. It makes results comparing difficult.

2.1 Multichannel Measuring System for EHG Signals

The applied EHG measuring techniques produce a large variety of electrodes number and their localization on an abdominal skin starting from 2 channels and ending in 16 channels arranged as a 4×4matrix [11].

In the presented study 4 channels measuring system, specially designed for EHG signals, was used [12]. Each channel consisted of a bipolar electrode. Localization of electrodes on an abdominal skin was in accordance with methodology presented by Euliano *et al.* in [13], *i.e.* the *up* channel was placed over an uterine fundus, the *right* and the *left* channels were localized over adequate sides of a uterus and the *down* channel was put over the cervix. The reference electrode was laid on a right hip of a woman.

The pass band of the measuring system was 0.01 Hz- 3Hz. The sample frequency was equaled to 20 Hz. This system has also individually self tuning amplifiers which gains allowing for optimal using of A/D measuring range [12].

2.2 The Study Groups of Patients

Almost all performed studies evaluating labor prediction based on EHG signal had the case – control design. However, definitions of cases and controls varied across studies.

Usually, a case group contains patients being during the 2^{nd} period of a labor when a contractile activity of a uterus is maximal. The same criteria was assumed in the reported study. The case group contains 15 primigravidas being in 2^{nd} period of a unifetal labor.

Control groups of EHG signals generally contained signals registered with pregnant women who were not conscious of uterine contractions. This inclusive criteria may produce underestimation of a sensitivity of the studied EHG parameters. From biological considerations published by Garfield and Maner stemmed that already few months before labor increasing number of gap junctions and myometrial ion channel changes were observed [14]. Thus, they may produce bioelectrical activity of myometrial fibers seen in EHG signal despite long time to a labor.

Therefore, in contrast to other studies, in the presented study the control group consisted of 14 unpregnant women being in a folicular phase of a menstrual cycle ensuring possible low contractile activity of their uteruses. Localization of electrodes reflected proper parts of a uterus according to the individual size of this organ previously measured by USG. No woman belonging to the control group had pathology in her uterine cavity.

3 Multivariate Sample Entropy of Differentiated EHG Signals

A multivariate sample entropy is a generalization of a sample entropy for multivariate time series introduced by Ahmed and Mandic [15]. The main advantage

of this generalization is utilization of join information contained in a multivariate vector of an analyzed signal.

According to my best knowledge hitherto published manners of EHG nonlinear parameterization used raw samples of the signals. To ensure better prediction of uterine labor activities a multivariate sample entropy was calculated for differentiated four-variate EHG signals. Then, a multivariate sample entropy becomes sensitive to the dynamic of registered biopotentials instead of their instantaneous values.

Let us introduce a four elements vector of the differentiated signals $\mathbf{X}(t) = \left[x_1(t)\ x_2(t)\ x_3(t)\ x_4(t) \right]$, where $x_1 = e\dot{h}g_{up}, x_2 = e\dot{h}g_{left}, x_3 = e\dot{h}g_{right}, x_4 = e\dot{h}g_{down}$ denote respective components of a differentiated EHG signal after a standarization. The multivariate time series $\mathbf{X}(t)$ has a length equaled to T.

According to classical nonlinear methods the composite delay vector must be created in the following form::

$$\mathbf{X}_m(t) = \begin{bmatrix} x_1(t)\ x_1(t+\tau_1) \cdots x_1(t+(m_1-1)\tau_1) & \cdots \\ \cdots \quad x_2(t+\tau_2) \cdots x_2\left(t+(m_2-1)\,\tau_2\right) \cdots \\ \cdots \quad x_3(t+\tau_{3)} \cdots x_3(t+(m_3-1)\tau_3) & \cdots \\ \cdots \quad x_4(t+\tau_4) \cdots x_4(t+(m_4-1)\tau_4) \end{bmatrix} \tag{1}$$

$\mathbf{M} = \left[m_1\ m_2\ m_3\ m_4 \right]$ is the respective embedding vector and $\tau = \left[\tau_1\ \tau_2\ \tau_3\ \tau_4 \right]$ is the time delay vector. The lower index of $\mathbf{X}_m$ is $m = \sum_{i=1}^{4} m_i$.

Analogously to the sample entropy the multivariate sample entropy was estimated in the following procedure:

1. Create $T-n$composite delay vectors $\mathbf{X}_m(i) \in \mathrm{R}^m$, where $i = 1, 2, \ldots T-n$ and $n = \max\{\mathbf{M}\} \max\{\tau\}$
2. Compute the distance between any two composite delay vectors as the maximum of the following norm:

$$d\left[\mathbf{X}_m(i), \mathbf{X}_m(j)\right] = \max_{l=1,\ldots m} \left\{ |x(i+l-1) - x(j+l-1)| \right\} \tag{2}$$

3. For a given composite delay vector and an assumed similarity threshold r count the number of cases fulfilled the following condition:

$$d\left[\mathbf{X}_m(i), \mathbf{X}_m(j)\right] \leq r, i \neq j \tag{3}$$

   denoted by S_i^m and calculate a probability of similarity occurrence according to the formula:

$$B^m(r) = \frac{1}{T-n} \sum\nolimits_{i=1}^{T-n} B_i^m(r) \tag{4}$$

   where $B_i^m(r) = \frac{1}{T-n-1} S_i^m$
4. Extend the dimensionality of the multivariate delay composite vector from m to $m+1$ and form the $\mathbf{X}_{m+1}(t)$ according to the structure of (1).
5. For a given $\mathbf{X}_{m+1}(i)$ vector calculate the probabilities of similar elements analogously to the 4^{th} step, *i.e.* $B_i^{m+1}(r) = \frac{1}{4(T-n)-1} S_i^{m+1}$, where S_i^{m+1}

is the number of cases fulfilled the (3) inequality for the extended composite delay vector. Then, calculate the following quantity:

$$B^{m+1}(r) = \frac{1}{4(T-n)} \sum_{i=1}^{4(T-n)} B_i^{m+1}(r) \tag{5}$$

6. The multivariate sample entropy is calculated as:

$$\mathrm{MSamEn}(\mathbf{M}, , \tau, r) = -\ln \left[\frac{B^{m+1}(r)}{B^m(r)} \right] \tag{6}$$

4 Results

The multivariate sample entropy was applied to the differentiated 4-channels EHG signals measured in the control and case group. Fig. 1 presents an example of a raw signal registered during the 2^{nd} period of a labor.

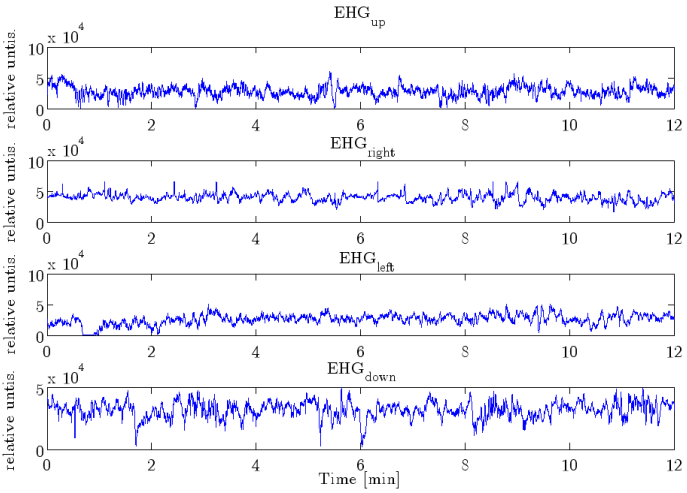


Fig. 1. The 4 channels, raw EHG signal registered during the 2^{nd} period of the labor for a randomly chosen patient

Fig. 2 shows the same but numerical differentiated EHG signal. We can note that a differentiation enhances bioelectrical bursts. It is particularly important when join multichannel information is used for parameterization of EHG signals. The multivariate sample entropy was calculated for the differentiated as well as raw EHG signals registered in the control and the case group to investigate an influence of a differentiation on a contraction prediction. The accuracy of this prediction was estimated as the area under ROC.

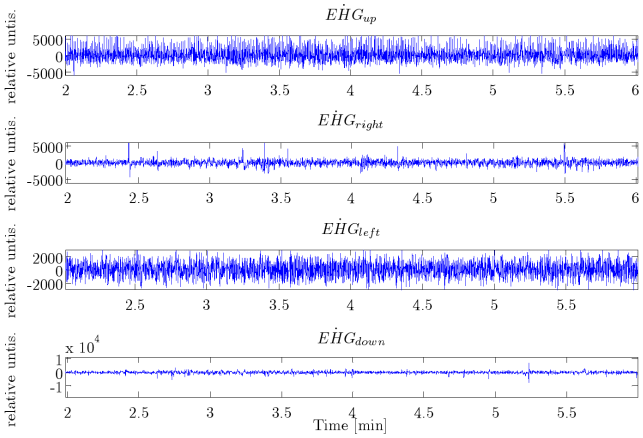


Fig. 2. The 4 channels, differentiated EHG signal registered during the 2^{nd} period of the labor for the same patient as in Fig. 1

Calculation of the multivariate sample entropy was conducted assuming that the embedding vector is $\mathbf{M} = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$, the delay vector is $\tau = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$ and the $r = 0.17TV$, where TV denotes the total variance of 4 elements vector computed as the trace of the covariance matrix. Increasing of the $\mathbf{M}$ and τ values did not significantly change the MSamEn values.

The Table 1 presents the mean and standard errors of the multivariate sample entropy obtained for the both groups of EHG signals

Table 1. The mean and standard errors of MSamEn for raw and differentiated EHG signals in the non-labor and labor group. * - the statistically significant results confirmed by Mann-Whitney test (p <0.0008)

Parameter	Non-labor 4 chan. EHG signal		Labor 4 chan. EHG signal	
	Mean	Std. Err.	Mean	Std. Err.
MSamEn for raw signals	0.389	0.142	0.418	0.217
MSamEn * for differentiated signals	1.126	0.346	0.450	0.467

The applied Mann-Whitney test shows that significant differences between the groups were observed only for the differentiated signals (p<0.0008).

Fig. 3 features the ROC obtained for the differentiated signals. The AUC = 0.873±0.074 and the cut-off point is equaled to 0.571. The sensitivity was equaled to 0.87 and the specificity was 0.86. These values seem to be promising.

Moreover, the mean value of MSamEn computed for differentiated signals was higher than the respective value obtained for raw EHG signals. These differences were statistical significant only in the non-labor group based on the sign test (p <0.001). This observation suggests that a signal differentiation enhances a nonlinear or chaotic character of the parameterized signal. A synchronization of a uterine bioelectrical activity associated with labor may cause more regular, i.e. more "linear" character of an EHG signal. This hypothesis seems to be confirmed by lack of a differentiation influence on the multivariate sample entropies in the labor group.

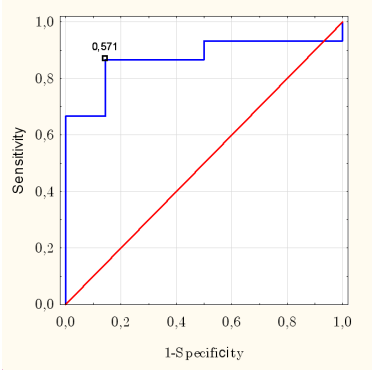


Fig. 3. The ROC obtained for MSamEn computed for differentiated EHG signals

5 Conclusion

Physiological mechanisms which control labor are still unknown. Thus, it is difficult to find an optimal parameterization method of electrohysterographical signals which corresponds to labor uterine activity. An application of multivariate sample entropy for this purpose was proposed. The original idea was a calculation of the multivariate sample entropy for differentiated EHG signals which significantly increased a relationship between estimated values and uterine activities. The obtained value of the area under ROC may be overestimated comparing to other methods because of using unpregnant women as the control group. Nevertheless, the presented results indicate that a multivariate sample entropy may clinically efficient parameterizes a multi - channel EHG signal when it is applied to a differentiated signal. Probably, a differentiation might increase a nonlinearity in EHG signals.

Acknowledgement. I thank to engineers Wojciech Zabolotny, Wojciech Zaworski and Agnieszka Podbielska for an elaboration of the EHG measuring system as well as two obstetricians Katarzyna Pankiewicz and Mikolaj Laniewski for an acquisition of the EHG signals.

References

1. Wolke, D., Chernova, J., Eryigit-Madzwamuse, S., Samara, M., et al.: Self and Parent Perspectives on Health-Related Quality of Life of Adolescents Born Very Preterm. The J. Pediatr. 162, 35–342 (2013)
2. Oyarzo, C., Bertoglia, P., Avenda, R., Bacigalupo, F., Escudero, A., Acuri, J., Escudero, C.: Adverse perinatal outcomes after the February 27th 2010 Chilean earthquake. J. Matern. Fetal Neonatal Med. 25, 1868–1873 (2012)
3. Euliano, T.Y., Nguyen, M.T., Marossero, D., Edwards, R.K.: Monitoring contractions in obese parturients: electrohysterography compared with traditional monitoring. Obstet. Gynecol. 109, 1136–1140 (2007)
4. Skowronski, M.D., Harris, J.G., Marossero, D.E., Edwards, R.K., Euliano, T.Y.: Prediction of intrauterine pressure from electrohysterography using optimal linear filtering. IEEE Trans. Biomed. Eng. 53, 1983–1989 (2006)
5. Diab, A., Hassan, M., Marque, C., Karlsson, B.: Quantitative performance analysis of four methods of evaluating signal nonlinearity: Application to uterine EMG signals. Presented at 34th Ann. Int. Conf. IEEE EMBS (2012)
6. Hassan, M.M., Terrien, J., Muszynski, C., et al.: Better Pregnancy Monitoring Using Nonlinear Correlation Analysis of External Uterine Electromyography. IEEE Trans. Biomed. Eng. 60, 1160–1166 (2013)
7. Laforet, J., Rabotti, C., Terrien, J., Mischi, M., Marque, C.: Toward a Multiscale Model of the Uterine Electrical Activity. IEEE Trans. Biomed. Eng. 58, 3487–3490 (2011)
8. Moslem, B., Khalil, M., Marque, C., Diab, M.O.: Complexity analysis of the uterine electromyography. Presented at 32nd Ann. Int. Conf. IEEE EMBS (2010)
9. Euliano, T.Y., Marossero, D., Nguyen, M.T., Euliano, N.R., Principe, J., Edwards, R.K.: Spatiotemporal electrohysterography patterns in normal and arrested labor. Am. J. Obs. Gyn. e1-7 (2009)
10. Bursztyn, L., Eytan, O., Jaffa, A.J., Elad, D.: Mathematical model of excitation-contraction in a uterine smooth muscle cell. Am. J. Physiol. Cell. Physiol. 292, C1816–C1829 (2007)
11. Jiang, D., Li, G., Lin, L.: Uterine electromyogram topography to represent synchronization of uterine contractions. Int. J. Gynaecol. Obstet. 97, 120–124 (2007)
12. Zabolotny, W.M., Podbielska, A., Grzanka, A., Radomski, D.: A four channels electrohysterograph with individually self tuning amplifier gains. In: Proc. IA STED Int. Conf. Biomed. Eng. (BioMed 2013), pp. 125–132 (2013)
13. Euliano, T.Y., Nguyen, M.T., Darmanjian, S., McGorray, S.P., Euliano, S.P., Onkala, S.P., Gregg, A.R.: Monitoring uterine activity during labor: a comparison of 3 methods. Am. J. Obstet. Gynecol. e1-6 (2013)
14. Garfield, R.E., Maner, W.L.: Physiology and electrical activity of uterine contractions. Semin. Cell Dev. Biol. 18, 289–295 (2007)
15. Ahmed, M.U., Mandic, D.P.: Multivariate Multiscale Entropy Analysis. IEEE Signal Proc. Let. 19, 91–94 (2012)

Correlation Matrices as a Tool to Analyze the Variability of EEG Maps

Hanna Goszczynska, Leszek Kowalczyk, and Bozenna Kuraszkiewicz

Nałęcz Institute of Biocybernetics and Biomedical Engineering
Polish Academy of Sciences, 4 Trojdena, str., 02-109 Warsaw, Poland
hgoszczynska@ibib.waw.pl

Abstract. The aim of this paper is to present the selected examples of possible applications of image of correlation coefficients matrix of EEG map series in the analysis of variation of the topography of the isopotential areas in EEG maps, and thus in the assessment of stationarity, spatio-temporal variability and trends of changes of bioelectric activity of the brain. The image of correlation coefficients matrix shows similiarity of all pairs of maps in a series. The choice of segmentation threshold of characteristic areas in images of the correlation coefficients matrix of EEG map series corresponding to the sequence similarity relationships in a series of maps was based on the results of research conducted on test series. Determining of the segmentation threshold value allowed for the presentation of research results of similar sequences in the examplary series of real EEG recordings. The obtained results consistent with the visual assessment of the similarity of the topographic distributions of isopotential areas in EEG maps form the basis for further research of usefulness of the analysis of the correlation coefficients matrix in the assessment of the dynamics of EEG mapping topography.

Keywords: EEG mapping, similarity measurements, correlation matrices.

1 Introduction

Analysis of EEG maps in relation to EEG signal analysis has two fundamental advantages: isopotential areas configuration analysis and topographical differences allow for the detection of changes in the bioelectric activity of the brain in pathologies or during the experiments and the analysis of EEG mapping topography benefits from the independence of isopotential areas contours distribution from the reference electrode. The change of the reference electrode moves only the level of the base-line of level of potentials [1].

Procedures of the mapping are available widely in commercial apparatuses for EEG routine examinations and are being used in the clinical practice. Maps automatically are being compared with referential maps for control groups. Available methods in electroencephalographic systems concern single maps and are not sufficient for some special applications. The available analysis methods of EEG

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_29, © Springer International Publishing Switzerland 2014

mapping both at commercial apparatuses for EEG routine examinations, as well as in packages for the EEG signal analysis developed in many academic centers [2] do not provide methods for similar sequences searching in long map series.

One of the most popular methods for assessing temporal dynamics of the distribution of scalp potentials developed by Lehmann and co-workers, and extended by statistical analysis compared distributions enables to separate time segments of the EEG signal characterized by a certain distribution of potentials – functional microstates [3,4,5]. Analysis of functional microstates is used, among others, in studies of spontaneous brain activity in psychology and psychiatry [6], as well as in the evaluation of the relationship between bioelectric and hemodynamic activity of the brain [7]. These methods concerned estimation of similarity of EEG maps sequences with some defined patterns of maps. Despite application of various techniques either using method described above or content-based and text-based image retrieval approach there are no relatively simple method for similar sequences detection which could be helpful in clinical practice for estimation of the dynamic of changes visible in EEG maps series.

The present work was performed under the research and development of method for the analysis of the EEG maps sequence similarity in long series which would allow separation repteded EEG time segments characterized not only by constant but also variable (eg, with a specific trend) isopotential areas distribution. The method involves calculating the matrix of correlation coefficients (MCC) for each pair of maps in the series and the generation of corresponding to this matrix grayscale image [8,9]. Below following [8] main assumptions of the method were quoted.

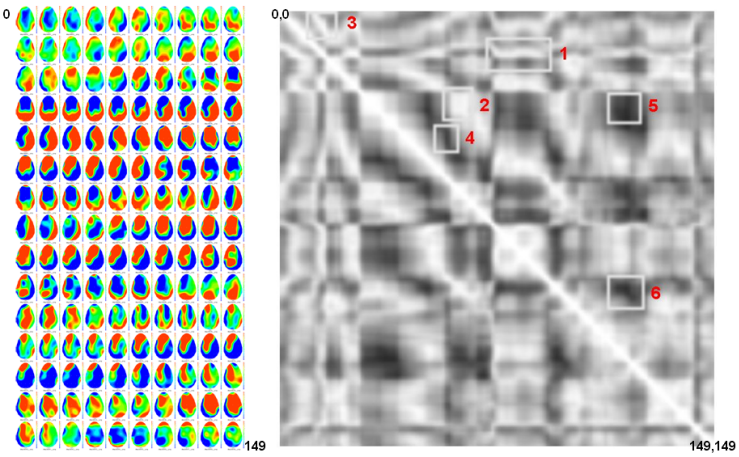


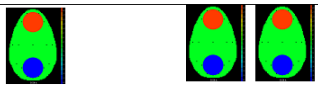
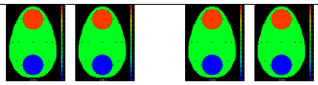
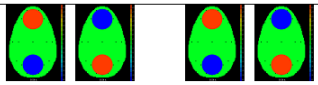
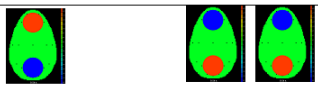
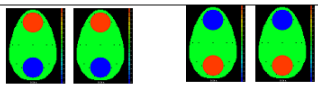
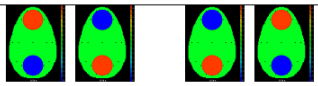
Fig. 1. A series of 150 maps from the period of seisure episode and the MCC image of the series; modified from [8] with permission

Figure 1 shows a series of 150 EEG maps generated every 10 ms, and MCC image for each pair of maps in this series. Image MCC shows all combinations of the level of similarity (in terms of correlation coefficients value) of individual sequences in a series of maps. The next step of analysis is location of specific areas in MCC image corresponding to a certain similarity relationship of EEG maps sequence. Examples of such areas are marked on the MCC image in Figure 1.

Six types of relationships of map sequence similarity in the series were defined: three relationships of type positive – positive and three relationship of type positive – negative (Table 1):

1. one map is similar to all maps in a given sequence *(Pos1-N)*,
2. one sequence is similar to another one, all maps in sequences are similar *(PosN-N)*,
3. respective consecutive maps in both sequences are similar, maps in a given sequence are not similar *(Pos1-1)*,
4. one map is a negative of all maps in a given sequence *(Neg1-N)*,
5. one sequence is a negative of another one, all maps in a given sequence are similar *(NegN-N)*,
6. respective consecutive maps in both sequences are negative, the maps in a given sequence are not similar *(Neg1-N)*.

Table 1. Areas of the MCC image indicated in Figure 1b (column II), patterns (column III) and sample test sequences satisfying the given relationships (column IV); modified from [8] with permission

I	II	III	IV
	Specific areas of MCC image	Patterns	Examples of two sequences
1 Pos1-N			
2 PosN-N			
3 Pos1-1			
4 Neg1-N			
5 NegN-N			
6 Neg1-1			

For each area pattern corresponding to the relationship was suggested (Table 1, col III).

For two maps similar one to each other determines the maps that the correlation coefficient meets the criterion, for example, is greater than the thresholds. Positive and negative of EEG maps are defined the maps in which the pixels with the same coordinates take the same values (according to a scale units), but with opposite signs. In interpretation of MCC image two maps are described as positive and negative when their correlation coefficient meets certain criteria, for example, is less than the assumed threshold. Proposed to use three methods specific areas location in the MCC image areas and thereby determining the sequence indexes meeting the above-mentioned similarity relations: visual (manual), segmentation method and matching the reference mask [8].

The proposed methods for locating specific areas rely mainly on choosing the appropriate threshold, in this case the value of the correlation coefficient, which allows for the segmentation of the respective areas and the indexation of sequences that meet the criteria of similarity. The aim of this study was to perform tests in order to select the segmentation thresholds for the relationship type positive - positive. Correlation coefficient sensitivity to changes in the topography of objects in a series of test images was investigated, taking into account the characteristic color scale used in the visualization of EEG potential maps and for the two definitions of correlation coefficient. The paper presents the results of tests and examples of search similar sequences in a sample series of EEG maps. The obtained results consistent with the visual assessment of the similarity of distributions of topographic EEG maps form the basis for further research of suitability of MCC images in the analysis of dynamics of changes in the EEG mapping topography.

2 Materials and Method

2.1 Test Maps

In order to choose the initial value of the segmentation threshold of MCC images of EEG map series MCC images of the test series containing the elementary objects imitating isopotential areas with a specific parameters changes were analyzed (Fig. 2):

- Linear change in dimension of the area – *test1*,
- Linear change in surface of the area – *test3*,
- Linear change of location of the area with constant size – *test2*,
- Linear change of the surface of the area in the vicinity of the area of constant size and value of the potential of opposite sign – *test4*.

Assumed initially segmentation threshold corresponds to 30% change in dimension of the area in the sequence *test1* (sample maps marked with an ellipse in Figure 2).

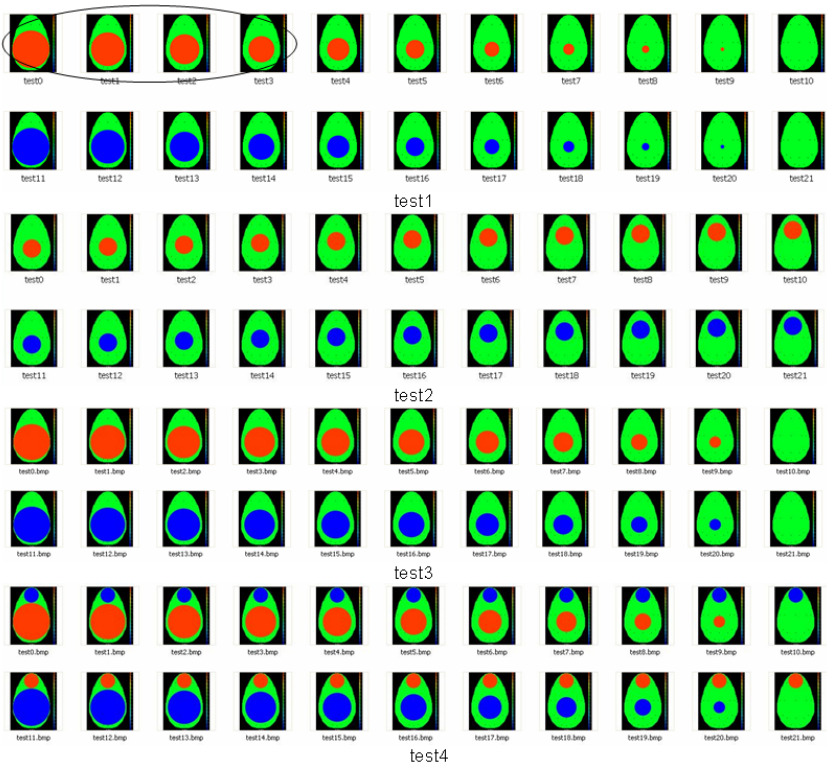


Fig. 2. Test sequences of maps: *test1* – the radius of the object linearly decreases by 25 pixels (in the range of 250-0), *test2* – center of the object moves linearly every 25 pixels, *test3* – the surface of the object decreases linearly (in the range 250-0), *test4* – surface of the object decreases linearly (in the range of 250-0) and an object of the constant surface in the negative color

2.2 Real Maps

Analysis of MCC image of selected EEG maps series included a series of 50 EEG maps (Fig. 3) from before seizure episode periods with different characteristics for two subjects [10]. Selection of map series (subject A from patient group and subject B from control group) was made based on visual analysis of the diagrams of sequence of two-dimensional histogram EEG map images [11]. A series of maps, for which histograms sequences show certain periodicity were chosen. Figure 4 depicts the histograms for the two series of EEG map from Figure 3. In histograms in planar form the values of normalized area are presented in color scale according to legend in right side of diagram. Maps with similar isopotential areas distributions have similar histograms, and the similarity of histograms does not mean topographical similarity.

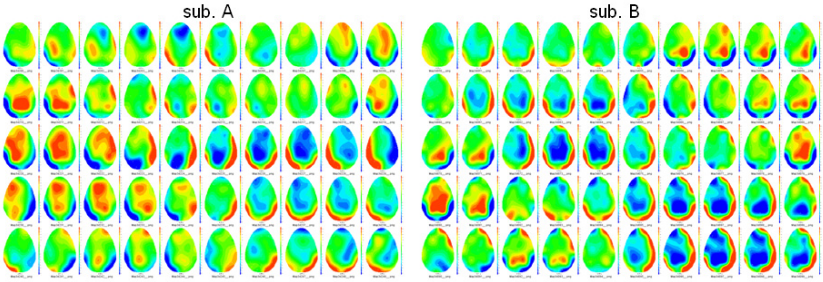


Fig. 3. Series of EEG maps generated every 10 ms in 17th ranges color scale presenting range potentials including ±20 mV from the period before seizure episode periods for two subjects

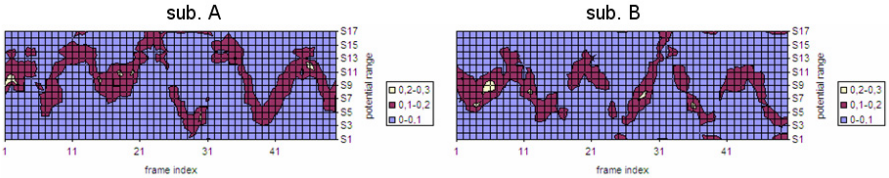


Fig. 4. Sequences of histograms for two series of EEG maps from Figure 3; normalized area are presented in color scale according to legend in right side of diagram

2.3 Similarity Measures

Analyzed matrices of correlation coefficients based on Euclidean distance measure without (Equation 1) and with normalization by the mean value (Equation 2) [8].

$$C_1(k,l) = \frac{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} I_k(i,j) * I_l(i.j)}{\sqrt{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} I_{k^2}(i,j) * \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} I_{l^2}(i,j)}} \tag{1}$$

$$C_2(k,l) = \frac{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} \left(I_k(i,j) - \bar{I}_k\right) * \left(I_l(i,j) - \bar{I}_l\right)}{\sqrt{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} \left(I_k(i,j) - \bar{I}_k\right)^2 * \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} \left(I_l(i,j) - \bar{I}_l\right) 2}} \tag{2}$$

where $C_1(k,l)$ and $C_2(k,l)$ are the correlation coefficients of the map I_k with the map I_l; k, l are the coordinates of the element of correlation matrix; m, n are the sizes of examined maps, i, j are the coordinates of compared maps and $\bar{I}_k$ and $\bar{I}_l$ are the mean values of gray levels of I_k and I_l maps, respectively.

Correlation analysis was carried out for a series of maps after converting maps images to grayscale images using two look up table (LUT) prefer the positive or negative values of potential (Fig. 5).

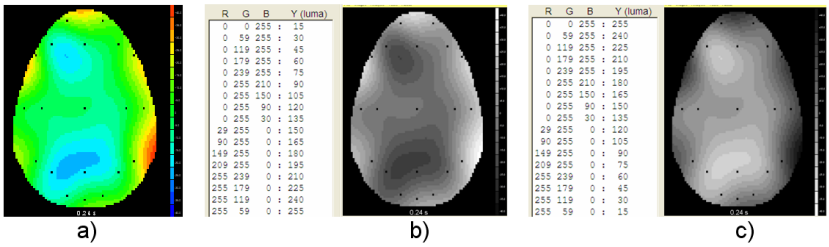


Fig. 5. Conversion of a color map image (a) on the grayscale image by two conversion tables LUT: prefers positive values LUT_pos (b) and negative values LUT_neg (c)

3 Results

3.1 Test Maps

Assumed segmentation threshold corresponding to 30% of change of object dimension (Fig. 2, the maps contained in the black ellipse) has been tested for other changes of parameters of the test objects. This threshold is respectively 0.96 (in the range 0, 1) and 0.87 (in the range -1 to 1) for MCC image calculated according to Equations 1 and 2. Figure 6 shows MCC images calculated according with Equations 1 and 2 and table LUT_pos after segmentation using specified above thresholds. MCC images for maps 12-21 of test sequences are equivalent to CMM images calculated according to table LUT_neg for maps 1-11.

3.2 Real Maps

For EEG map series segmentation thresholds equal to 0.95 and 0.9 for the MCC images calculated according to Equations 1 and 2 were assumed. The values of the thresholds were chosen experimentally taking into account the results of segmentation of MCC images for *test4* sequence, as the most similar to the real EEG maps. Figure 7 shows MCC images after segmentation calculated according to Equations 1 and 2 and table LUT_neg and LUT_pos and set of correlation coefficients (>0.95) calculated according to Equation 1 and table LUT_pos for maps 19-25 and 30-36 (pat. A) and for maps 5-11 and 17-23 (pat. B). In Figure 6 and 7 in red color values above the thresholds of segmentation were denoted.

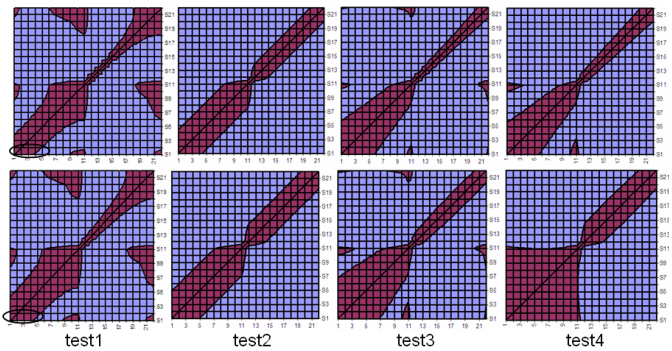


Fig. 6. MCC images after segmentation calculated according to Equation 1 (upper row) and the Equation 2 (bottom row) for four test sequences

4 Discussion and Conclusion

The results of tests confirmed that the threshold corresponding to 30% of the change in dimension of the area in the test sequence *test1* allows detection of the changes in other tested parameters of objects in the range corresponding to the visual similarity estimation, however, there is a strong dependence on the size of the investigated objects (Fig. 6). The exception is the result for the sequence *test4* using Equation 2, where assumed threshold equal to 0.87 is definitely too low, while the threshold equal to 0.96 (for Equation 1) is a bit too high for objects with smaller area. Thus, for the analysis of real EEG maps thresholds equal to 0.95 (for Equation 1), and 0.9 (for Equation 2) were proposed. The differences seen in MCC images after segmentation using two tables LUT indicate that both images must be analyzed in order to identify similar sequences with mainly positive (red areas in EEG maps) or a negative potentials (blue areas in EEG maps).

Assumed thresholds allow for the preliminary automatic indexation of similar EEG maps sequences (Fig. 7, set of correlation coefficients for selected areas of MCC images). Segmentation of MCC images for two maps series generated for EEG recordings with different characteristics show difference in trends of changes of isopotential areas topography. In series for sub. A dominant cyclic relationship is the relationship type *Pos1-1* (Fig. 7, yellow segments), while in series for sub. B dominant type of relationship is the relationship *PosN-N* (Fig. 7, yellow squares) with also occurring relationships type *Pos1-1*. The obtained preliminary results for MCC images segmentation of sample EEG map series suggest the usefulness of the method for the quantitative description of trends of in topography in EEG mapping.

Quantitative analysis of MCC image for EEG maps series may allow estimation of the dynamics of isopotential areas changes by indicating the sequences (EEG record segments) similar to given sequence (indicated two types of sequence similarity) and the ability to assess the relationship of time parameters

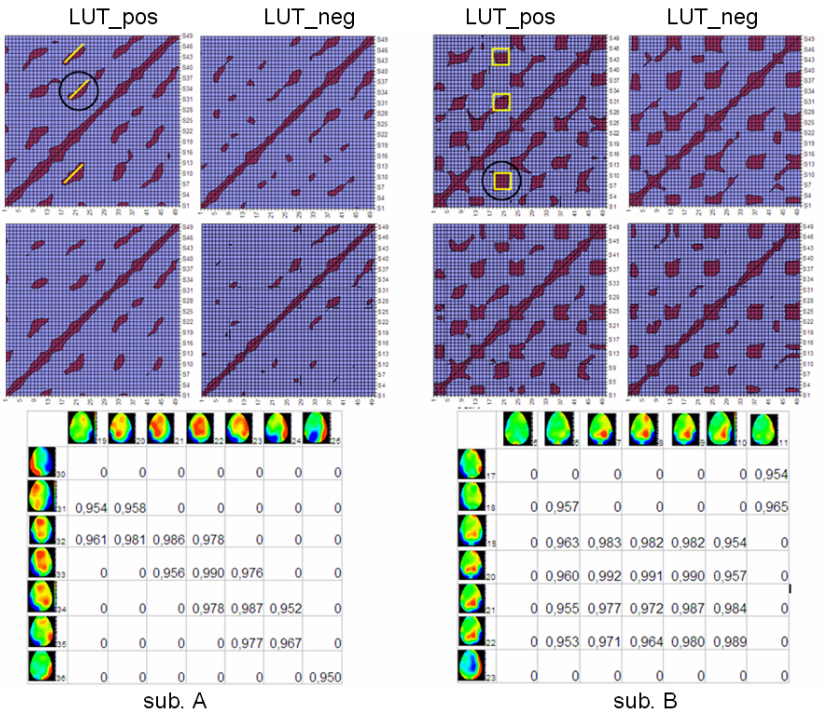


Fig. 7. MCC images after segmentation calculated according to Equation 1 (upper row) and Equation 2 (bottom row) and for two tables LUT; in yellow color selected areas corresponding relations *Pos1-1* (sub. A) and *PosN-N* (sub. B) and shows the values of correlation coefficients for specific areas marked with black circles in MCC images

of repeated sequences, such as duration, frequency of appearance and the total duration of the sequence. This analysis can be a tool to support such topographic measures proposed by Lehmann and co-authors as Global Field Power, Global Map Dissimilarity and segmentation of functional microstates [3,4,5].

Thanks to using the correlation matrix an aggregation of extensive one-dimensional data (EEG signal) or two-dimensional (EEG maps) to one picture containing information about the dynamic of the entire registered signal during the examination is possible. It seems to be even better approach than clustering of microstates. Certainly comparing two correlation matrices for two records of EEG signal is simpler than sets of microstates for these records.

The presented results depend to a large extent on the choice of the threshold value of the correlation coefficient and size of compared isopotential areas. In further work on the method of MCC image segmentation there will be analyzed the suitability of the use of adaptive threshold taking into account the size of compared objects – isopotential areas. Simple test image sequences with only a few values of the potential will be enriched through the introduction of areas

of potential values continuously changing from the maximum to the minimum value for a better representation of real EEG map. In order to select the optimal similarity measure there will be carried out detailed analysis of sets of correlation coefficients for MCC (Fig. 7) calculated according to Equations 1 and 2, or other measures of similarity. Also semantic interpretation of the maps similarity and analysis and processing of image descriptors [12] would enrich information that can be obtained from MCC image analysis.

References

1. Brunet, D., Micah, M., Murray, M.M., Michel, C.M.: Spatiotemporal Analysis of Multichannel EEG: CARTOOL. Computational Intelligence and Neuroscience, Article ID 813870, 15 pages. Hindawi Publishing Corporation (2011), doi:10.1155/2011/813870
2. Sylvain Baillet, S., Karl Friston, K., Oostenveld, R.: Applications for Electromagnetic Brain Mapping Using MEG and EEG. Computational Intelligence and Neuroscience, Article ID 972050, 4 pages. Hindawi Publishing Corporation (2011), doi:10.1155/2011/972050
3. Wackermann, J., Lehmann, D., Michel, C.M., Strik, W.K.: Adaptive segmentation of spontaneous EEG map series into spatially defined microstates. International Journal of Psychophysiology 14(3), 269–283 (1993)
4. Lehmann, D., et al.: EEG_microstates. Scholarpedia 4(3), 7632 (2009), `http://www.scholarpedia.org/article/EEG_microstates`, doi:10.4249/scholarpedia.7632
5. Koenig, T., Studer, D., Hubl, D., Melie, L., Strik, W.K.: Brain connectivity at different time-scales measured with EEG. Philos. Trans. R Soc. Lond. B Biol. Sci. 360(1457), 1015–1024 (2005)
6. Schlegel, F., Lehmann, D., Faber, P.L., Milz, P., Gianotti, L.: EEG microstates during resting represent personality differences. Brain Topogr. 25(1), 20–26 (2012)
7. Yuan, H., Zotev, V., Phillips, R., Drevets, W.C., Bodurka, J.: Spatiotemporal dynamics of the brain at rest – Exploring EEG microstates as electrophysiological signatures of BOLD resting state networks. NeuroImage 60, 2062–2072 (2012)
8. Goszczynska, H., Doros, M., Kowalczyk, L., Zalewska, E., Kolebska, K.: Detection of similar sequences in EEG maps series using correlation coefficients matrix. Machine Graphics & Vision 20, 73–92 (2011)
9. Goszczyńska, H., Kowalczyk, L.: Analiza czułości współczynnika korelacji na zmiany topografii map EEG - wyniki wstępne. In: Proc. of XVIII Krajowa Konferencja Biocybernetyki i Inżynierii Biomedycznej, Gdańsk, p. 1 (2013) (in Polish)
10. Goszczynska, H., Doros, M., Kowalczyk, L., Hoser, P., Kolebska, K.: Analysis of dispersion of isopotential areas values in EEG maps sequences. In: Burduk, R., Kurzyński, M., Woźniak, M., Żołnierek, A. (eds.) Computer Recognition Systems 4. AISC, vol. 95, pp. 549–557. Springer, Heidelberg (2011)
11. Goszczyńska, H., Podsiadły-Marczykowska, T.: Histograms of selected EEG maps sequences. Journal of Medical Informatics & Technologies 18, 19–27 (2011)
12. Podsiadly-Marczykowska, T., Goszczynska, H.: Ontology of EEG mapping - preliminary research. In: Piętka, E., Kawa, J. (eds.) ITIB 2012. LNCS, vol. 7339, pp. 183–198. Springer, Heidelberg (2012)

A Novel Approach to Robust Weighted Averaging of Auditory Evoked Potentials

Tomasz Pander, Stanisław Pietraszek, and Tomasz Przybyła

Silesian University of Technology, Institute of Electronics,
16 Akademicka St., 44-100 Gliwice, Poland
{tpander,spietraszek,tprzybyla}@polsl.pl

Abstract. The paper describes the robust weighted averaging method applied to averaging of auditory brainstem responses. This type of signals is characterized with extremely low signal-to-noise ratio. Suppression of noise that contaminates this type of signals can be achieved with the use of the averaging technique. The auditory evoked potentials are time-aligned and then the average template is determined. The weighted averaging operation can be regarded as special case of clustering. In this work the averaging process is formulated as the problem of certain criterion function minimization. The maximum likelihood estimator of location based on the generalized Cauchy distribution is used as the measure of dissimilarity function. The proposed methods performance is experimentally evaluated and compared to the reference methods in the presence of the artificial noise and in the case of real signals. The experiments show usefulness of the proposed method for robust weighted averaging of periodic signals, for instance the evoked potentials.

Keywords: auditory brainstem response, robust weighted averaging, General Cauchy distribution.

1 Introduction

Sensory evoked potentials (EPs) are time-aligned changes of the electrical activity of the brain recorded from the human scalp. These potentials represent aggregated electrical activity from a large number of temporally and spatially aligned neurons of the brain and arise as the response to a variety of controlled external stimuli [3,9]. The clinical utility of EPs is based on their ability to show abnormal sensory system conduction, to discover asymptomatic engagement of a sensory system, to help define the anatomic distribution of a disease process and monitor changes in a patient's neurological status [19]. In clinical practice the auditory (AEPs), visual (VEPs) and somatosensory (SEPs) are tested most frequently and they can be recorded in patients who are anaesthetised or comatose [3,19]. In general, the electroencephalogram (EEG) is extensively used within neurophysiology, cognitive neuroscience, cognitive psychology, and brain-computer interfacing [9].

The peak amplitudes and the latencies are commonly used for quantitative evaluation of EPs which last for a few hundreds of milliseconds, with it various

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_30, © Springer International Publishing Switzerland 2014

features categorized into early, middle and late components [3,19]. Early components of AEPs are auditory brainstem responses (ABRs) and auditory steady state response arising within 0-10 ms [3]. AEPs are much smaller than the EEG and early and middle components typically have a signal-to-noise ratio (SNR) in the range from -30 to -20 dB [3]. The most common method to isolate the AEP from background EEG is averaging the EEG responses to multiple identical auditory stimulations. The ABR can be regarded as the deterministic component for all stimuli (because it is time-locked) while the background noise (including spontaneous EEG) will vary and thus be reduced by averaging [17]. The assumptions for realize the averaging are the following: (i) the signal remains constant from trial to trial, (ii) the noise on any trial is uncorrelated with the noise on other trials, and (iii) the noise statistics remain stationary from trial to trial [5,7]. But if the noise varies from one trial to the next, averaging is less effective. The random phenomena, sudden transients as well as non-stationarities in the signals (caused both by a physiological adaptation process to the series of stimuli or by possible pathological evidences), do cause that the obtained average, in most of the cases, is only a very general information about a mean behaviour of the system under study.

From another point of view, the averaging methods are similar to the fuzzy robust clustering methods [2,6]. The averaging process can be regarded as a special case of clustering with only one prototype. The clustering method should be robust for data corrupted by outliers or heavy-tailed distributed noise. The influence of outliers on the averaged signal can be reduced by choosing a median as the aggregation operation. However the median averaging does not only remove the outliers but also the rest of data [11,13]. The alternative approach is based on the trimmed mean method [13]. The possibility of using the myriad cost function to develop a procedure for robust weighted averaging is presented in [15]. The approach based on L_p-norm is presented in [16]. An application of Vapnik ε-insensitive function allows to increase the robustness of the weighted averaging is presented in [11].

The goal of this paper is to show the novel approach to robust weighted averaging applied to processing of ABRs. This paper presents a robust cost function based on Generalized Cauchy probability distribution which plays role of the dissimilarity function in the weighted averaging method based on the minimizing of certain criterion function. The paper is divided into four sections. Section 2 presents the idea of the weighted averaging method based on the minimization of the scalar criterion function and introduces the proposed method. Section 3 describes the proposed experiments. Section 4 presents the obtained results and discussion. Finally, the conclusions are given in Section 5.

2 Methods

2.1 Weighted Averaging

Let us consider N sweeps of ABR potentials where $\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{iM}]^T$ is the ith sweep which consists of M samples and $1 \leq i \leq N$. Let $\mathbf{y} = [y_1, y_2, \ldots, y_M]^T$

is the averaged signal and $\mathbf{w} = [w_1, w_2, \ldots, w_N]^T$ is the weight vector that satisfies the following conditions:

$$\underset{1 \le i \le N}{\forall} \quad w_i \in [0,1], \quad \sum_{i=1}^{N} w_i = 1. \tag{1}$$

Each of the $\mathbf{x}_i$ sweep is the sum of a deterministic component and the noise $\mathbf{x}_i = \mathbf{s} + \nu_i$. The noise ν_i is a random component with zero mean and variance σ_i^2. The weighted average of N sweep waveforms is given by:

$$\mathbf{y} = \sum_{i=1}^{N} w_i \mathbf{x}_i. \tag{2}$$

The general idea of weighted averaging rests on the assumption that each signal $\mathbf{x}_i$ affects the resulting averaged signal $\mathbf{y}$ in the manner specified by the value of weight w_i. The simplest case of weighted averaging is the arithmetic averaging, where all weights are the same, equal to N^{-1}. Estimation of the weights values is a crucial for the process of averaging. The easy mathematical consideration shows that, given the variance σ_i^2 of each individual sweep, the weights are the following [10]:

$$w_i = \sigma_i^{-2} \left(\sum_{j=1}^{N} \sigma_j^{-2} \right)^{-1}. \tag{3}$$

This method is one of the reference methods and is denoted as the WA (the weighted averaging) method. There are different methods to estimate the noise variances or to estimate the optimal weights without direct knowledge of the noise variance. One of such approach is presented below.

2.2 Criterion Function Minimization

The WACFM (weighted averaging method based on criterion function minimization) is based on minimization the following scalar function [2,11]:

$$I_m(\mathbf{w}, \mathbf{y}) = \sum_{i=1}^{N} \sum_{j=1}^{M} (w_i)^m \rho(z_{ij}), \tag{4}$$

where $z_{ij} = x_{ij} - y_j$, $m \in (1, \infty)$ is the assumed weighting exponent and $\rho(\cdot)$ is a measure of dissimilarity for the vector argument. The eq. (4) can be rewritten as:

$$I_m(\mathbf{w}, \mathbf{y}) = \sum_{i=1}^{N} (w_i)^m \left(\sum_{j=1}^{M} \rho(z_{ij}) \right), \tag{5}$$

where $\sum_{i=1}^{M} \rho(z_{ij}) = \rho(\mathbf{z}_i)$ and then $\mathbf{z}_i = \mathbf{x}_i - \mathbf{y}$. The $\rho(\cdot)$ function is a measure of dissimilarity for a vector argument [11]. It plays a similar role to the cost

function in the robust M-estimator [2]. The task of searching for an optimal an optimal weight vector $\mathbf{w}^*$ and optimal averaged signal $\mathbf{y}^*$, can be formulated as follows:

$$I_m(\mathbf{w}^*, \mathbf{y}^*) = \min_{\mathbf{w},\mathbf{y}} I_m(\mathbf{w}, \mathbf{y}). \tag{6}$$

The problem of I_m minimization is considered as the constrained optimization problem and the method of Lagrange multipliers is applied [11] to find the optimal weights vector $\mathbf{w}$ which are given in the following way:

$$\underset{1\leq i\leq N}{\forall} \quad w_i = \frac{\rho\left(\mathbf{z}_i\right)^{1/(1-m)}}{\sum_{j=1}^{N}\left[\rho(\mathbf{z}_j)\right]^{1/(1-m)}}. \tag{7}$$

The robustness of the weighted averaging depends strictly on the function $\rho(\mathbf{z})$ which should be symmetric positive-definite function [2,11]. The weights w_i corresponds to membership (belonging) of $\mathbf{x}_i$ to the prototype (the averaged signal) [2].

If we assume that $\mathbf{y}$ is fixed, the next step of the algorithm consists in estimation of the $\mathbf{y}$ averaged signal. From the theory of M-estimators, the derivative of $\rho(z)$ is called the influence function $\psi(z) = \frac{\mathrm{d}\rho(z)}{\mathrm{d}z}$, while the weighted function is defined as $\varpi(z) = \frac{\psi(z)}{z}$ [12]. Then, for given data set $\mathbf{x} = [x_1, x_2, ..., x_N]$, using the weighted function of the dissimilarity function, the averaged signal $\mathbf{y}$ can be found by applying the fixed-point search algorithm which can be written as [16]:

$$\mathbf{y}^{(k+1)} = \frac{\sum_{i=1}^{N}(w_i)^m \cdot \varpi(\mathbf{x}_i - \mathbf{y}^{(k)}) \cdot \mathbf{x}_i}{\sum_{i=1}^{N}(w_i)^m \cdot \varpi(\mathbf{x}_i - \mathbf{y}^{(k)})}, \tag{8}$$

where the superscript (k) denotes the iteration number. The algorithm is regarded as convergent when $||\mathbf{w}^{(k+1)} - \mathbf{w}^{(k)}|| < \epsilon$, ϵ is a small positive value (e.g. $\epsilon = 10^{-6}$).

2.3 Generalized Cauchy Distribution-Based Cost Function

The family of Generalized Cauchy Distribution (GCD) has the feature that its pdf has the closed form for the whole family and also has algebraic tails which makes it suitable to model many impulsive process in real world of signals (the Cauchy distribution is the special case of the α-stable distribution for $\alpha = 1$ [8]). The pdf of the GCD has the following form [1]:

$$f(z) = a\sigma(\sigma^p + |z|^p)^{(-2/p)} \tag{9}$$

with $a = p\Gamma(2/p)/2(\Gamma(1/p))^2$; σ is the scale parameter and p is the tail constant. The family of GCD contains the Meridian [1] ($p = 1$) and Cauchy ($p = 2$) distributions as special cases. The adjustable and closed-form feature of the GCD makes it useful for robust estimation and filtering techniques [1]. Equation (9)

is the basis for determination of the maximum likelihood estimators of location; the cost function has the following form:

$$\rho(z) = \log\left(\sigma^p + |z|^p\right), \tag{10}$$

where $\sigma > 0$ and $0 < p \leq 2$. Using (5) and (10), the scalar criterion function can be rewritten as:

$$I_m(\mathbf{w}, \mathbf{y}) = \sum_{i=1}^{N} (w_i)^m \sum_{l=1}^{M} \log\left(1 + \left(\frac{|x_{il} - y_l|}{\sigma}\right)^p\right). \tag{11}$$

According to (7) the weights w_i are given by:

$$\underset{1 \leq i \leq N}{\forall} \; w_i = \frac{\left[\sum_{l=1}^{M} \log\left(1 + \left(\frac{|x_{il} - y_l|}{\sigma}\right)^p\right)\right]^{1/(1-m)}}{\sum_{j=1}^{N}\left[\sum_{l=1}^{M} \log\left(1 + \left(\frac{|x_{jl} - y_l|}{\sigma}\right)^p\right)\right]^{1/(1-m)}}. \tag{12}$$

The influence function of (10) has the form $\psi(z) = \frac{p|z|^{p-1}\mathrm{sgn}(z)}{\sigma^p + |z|^p}$ and the corresponding weighted function is given by:

$$\varpi(z) = \left(\frac{1}{p}|z|^{2-p}\left(\sigma^p + |z|^p\right)\right)^{-1}. \tag{13}$$

The averaged signal $\mathbf{y}$ is obtained with using formulas (8) and (13) in the following form:

$$\mathbf{y}^{(k+1)} = \frac{\sum_{i=1}^{N} (w_i)^m \mathbf{x}_i \left[\sum_{l=1}^{M} \frac{|x_{il} - y_l^{(k)}|^{2-p}}{p}\left(\sigma^p + |x_{il} - y_l^{(k)}|^p\right)\right]^{-1}}{\sum_{i=1}^{N} (w_i)^m \left[\sum_{l=1}^{M} \frac{|x_{il} - y_l^{(k)}|^{2-p}}{p}\left(\sigma^p + |x_{il} - y_l^{(k)}|^p\right)\right]^{-1}}. \tag{14}$$

Finally, using (12) and (14) the algorithm of the weighted averaging with criterion function minimization with the GCD cost function (WACFMGC) can be described as follows, where ϵ is a preset parameter:

1^o fix $m = 2$, σ and p, initialize $\mathbf{y}^{(0)} = \mathbf{0}$, set the iteration index $k = 1$,
2^o calculate $w_i^{(k)}$ for the kth iteration using the formula (12),
3^o update the averaged signal for the kth iteration $\mathbf{y}^{(k)}$ using the formula (14), and $\mathbf{w}^{(k)}$
4^o if $||\mathbf{w}^{(k+1)} - \mathbf{w}^{(k)}|| > \epsilon$ then $k \leftarrow k + 1$ and go to 2^o.

3 Experiments

The real ABR data was recorded with the participation of a young healthy man (24 years old) with a properly functioning sense of hearing. He sat in a

chair in a acoustically shielded room. Recordings were obtained using standard acquisition and stimulus parameters setup. The auditory stimulus was delivered as a rarefaction 'click stimulus' with a 0.1 ms pulsewidth from the standard TDH-49p supra-aural Telephonics headphones. The evoked ABRs were recorded with a bioamplifier at sampling frequency 48 kHz and bandpass filtered within 100-3000 Hz. Amplifier gain was 1000, A/D converter had 12 bits (providing the resolution of 0.084 μV/bit), amplitude range 5.5 mV. The example of the noisy ABR sweep and the averaged ABR are presented in Figure 1.

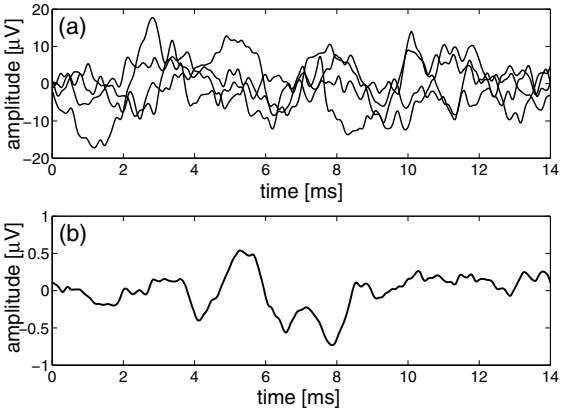


Fig. 1. (a) Examples of four raw ABRs waves, (b) the ABR wave after arithmetic averaging of 1800 sweeps at 40 dB normal hearing level. Stimulus is presented at zero time point

The simulated reference signal used in this study is based on ABR and consists of typical waves. The three sinc functions represent prominent ABR waves I, III and V [3]:

$$s(t) = 0.4\mathrm{sinc}\left[0.13\pi(4t-6+l)\right] + 0.4\mathrm{sinc}\left[0.13\pi(4t-14+l)\right] \\ -\mathrm{sinc}\left[0.13\pi(4t-24+l)\right], \tag{15}$$

where l determines the latency variation and in this work $l = 0$. The artificial reference signal is shown in Figure 2.

In order to simulate the real conditions of data recordings the symmetric α-stable (SαS) distribution [8] is applied to model the noise. The level of impulsiveness is controlled with the characteristic exponent α. The case of $\alpha = 2.0$ corresponds to the Gaussian probability distribution and $\alpha = 1.0$ corresponds to the Cauchy distribution. Figure 3 presents histogram of the characteristic

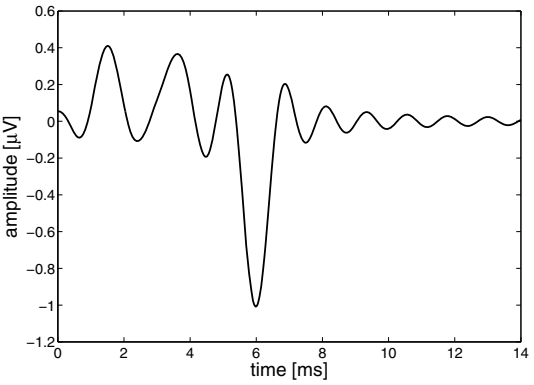


Fig. 2. Ideal artificial reference ABR wave

exponent α over 1000 sweeps of ABR. In this study α is calculated applying method presented in [14]. As can be seen, a certain part of sweeps are disturbed noise that has an impulsive nature. For that reason in this work α varies from 1.8 to 2.0 with the step 0.05.

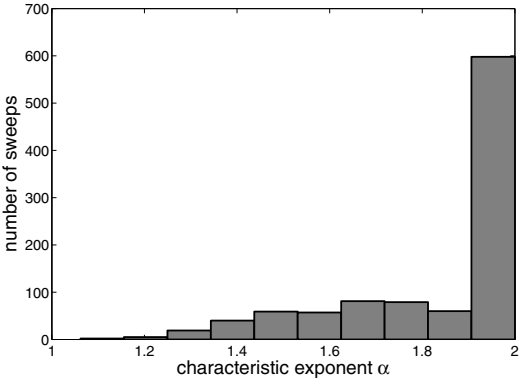


Fig. 3. Histogram of the characteristic exponent α calculated for 1000 ABRs sweeps

The class of the symmetric α-stable distributions does not have finite moments of the second or higher order. It means that the use of variance in the standard definition of the signal-to-noise ratio is meaningless [18]. For that reason the Geometric-SNR (GSNR) is applied [4]:

$$\mathrm{GSNR} = \frac{1}{2C_g}\left(\frac{A}{S_0}\right)^2, \tag{16}$$

where $C_g = e^{C_e} \approx 1.78$ is the exponential of the Euler constant, A is the amplitude of a modulated signal in an additive-noise channel with noise geometric

power S_0. The normalization constant $2C_g$ is used to ensure that the definition of the GSNR corresponds to that of the standard SNR if the channel noise is Gaussian. The geometric power definition is the following $S_0 = \exp\left(\frac{1}{N}\sum_{i=1}^{N}|x_i|\right)$ [4]. A series of 1000 ABRs sweeps are generated with the same deterministic component (given with eq. (15)) disturbed by the artificial noise, according to the two predefined simulation patterns — for the 1st, the 2nd, the 3rd, the 4th and the 5th part of ABR sweeps (each part containing 200 cycles), the GSNR values are equal to:

P1: -30, -20, -10, -15, -25 dB, respectively.
P2: -30, -30, -15, -30, -30 dB, respectively.

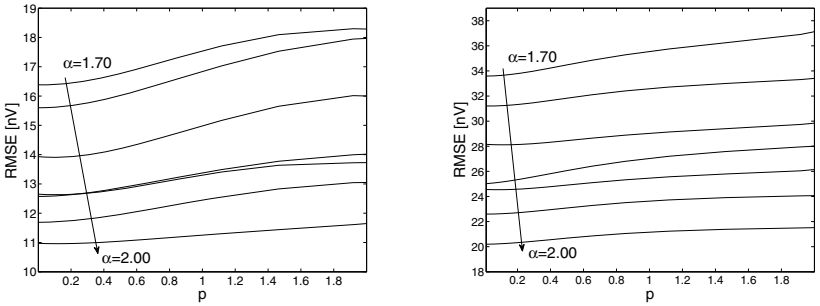
These two simulation patterns allow to test the efficiency and quality of the proposed methods of weighted averaging in the presence of artificial impulsive noise as well as real noise. The P1 and P2 patterns represent the case of lower and greater power noise variations.

4 Results and Discussion

The averaging process should not deform the signal. For that reason, the presented methods are evaluated using the root mean-square error (RMSE) between the deterministic component and the averaged signal. Subtraction of the deterministic component from the averaged signal gives the residual noise. The second index used to evaluate the quality of the tested methods, is the maximal absolute difference between the deterministic component and the averaged signal (MAX). The methods based on minimization of scalar criterion function are initialized with zero vector $\mathbf{y}$ and $m = 2$. The proposed methods performance is compared to that of the arithmetic mean (AM), the trimmed mean (TM) and the classical weighted averaging (WA) method. All experiments were done in MATLAB environment.

4.1 Selection of p

Selection of the optimal values for the p and σ parameters of the WACFMGC method is important but there is a concern that it will be attuned to the analysed signal only. In [1] the multiparameter estimation algorithm is presented that tries to solve this problem. This solution is suitable in the case of robust filtering [1], but in the case of cycles averaging it would require additional optimization of p and σ, leading to an increase of the computational complexity without the warranty to achieve the parameters optimal values. The influence of p on RMSE value obtained with the WACFMGC method is presented in Figure 4. The smallest values of RMSE are obtained for small p for both experiments and for further research will be used $p = 0.01$ as well as $\sigma = 1.0$.



(a) P1 experiment (b) P2 experiment

Fig. 4. Comparison of RMSE of the averaged 1000 ABRs using different p value

4.2 Efficacy of Averaging

The results of averaging are presented in Tables 1 and 2. The purpose of these experiments are to investigate the proposed method in the presence of the Gaussian noise ($\alpha = 2$) and nearly-Gaussian noise ($1.8 \leq \alpha < 2$). The smallest value of RMSE allows to achieve the highest signal-to-noise improvement.

Table 1. RMSE for averaged signals with artificial noise (P1 experiment)

α	AM	TM 25%	WA	WACFMGC $\sigma = 1$, $p = 0.01$
		RMSE [nV]		
1.80	47.87	20.57	14.60	**13.92**
1.85	56.95	22.54	13.57	**12.64**
1.90	37.38	21.79	13.08	**12.57**
1.95	51.44	21.33	12.42	**11.68**
2.00	31.53	21.20	11.69	**10.96**
		MAX [nV]		
1.80	241.54	74.27	**41.72**	41.81
1.85	462.99	75.30	40.34	**34.22**
1.90	139.02	63.53	**35.81**	36.38
1.95	1034.48	70.96	46.48	**40.13**
2.00	92.53	78.26	33.03	**30.14**

In the case of the P1 experiment, the RMSE value of the residual noise for the investigated methods are presented in first part of Table 1. The worst results (the highest value of RMSE) are achieved by a method of arithmetic averaging. The

performance of AM improves a little for higher value of the characteristic exponent α. To the slightly better results leads the trimmed mean method; RMSE values for all α remain more or less stable. The WA method allows to achieve significantly better performance then the methods of equal weights (AM and TM methods). But the best results are obtained with the proposed WACFMGC method. Both methods WACFMGC and WA reach lower values of MAX in comparison to reference methods. However, the WACFMGC method achieves the smallest value of MAX for $\alpha \geq 1.9$.

The P2 experiment is more difficult to the weighted averaging methods for two reasons. The first is the extremely low GSNR level and the second reason is the nearly the same level of GSNR. In this case, the weighted averaging may have a problem estimating the proper weight values. However, the results are similar to those of the P1 experiment to the benefit of the WACFMGC method. Only for $\alpha = 1.8$ the WA and WACFMGC methods have the same RMSE but the WACFMGC has the smallest MAX error. If $\alpha \geq 1.85$ the performance of WACFMGC is better than the WA method. The advantage from the use of this method is the ability to work with other types of noise than the Gaussian but it requires extra computational burden (at given ϵ, 3-4 iterations are performed).

Table 2. RMSE for averaged signals with artificial noise (P2 experiment)

α	AM	TM 25%	WA	WACFMGC $\sigma = 1$, $p = 0.01$
		RMSE [nV]		
1.80	160.03	48.75	**28.50**	**28.50**
1.85	77.11	48.92	26.24	**26.16**
1.90	63.82	51.01	25.08	**24.70**
1.95	59.71	49.49	23.44	**23.15**
2.00	49.64	48.26	21.12	**21.00**
		MAX [nV]		
1.80	3501.36	171.18	115.52	**111.38**
1.85	466.88	139.75	85.37	**82.76**
1.90	217.85	151.73	**91.22**	91.75
1.95	463.62	130.81	83.11	**76.75**
2.00	149.89	132.65	76.94	**74.17**

The results of ABR averaging of real signals are shown in Figure 5. Signal (a) was obtained using the arithmetic mean (equally weights), (b) using the weighted averaging method, (c) using the weighted averaging with criterion function minimization with the dissimilarity function based on the General Cauchy distribution with $p = 0.01$ and $\sigma = 1$, and (d) using the trimmed mean method.

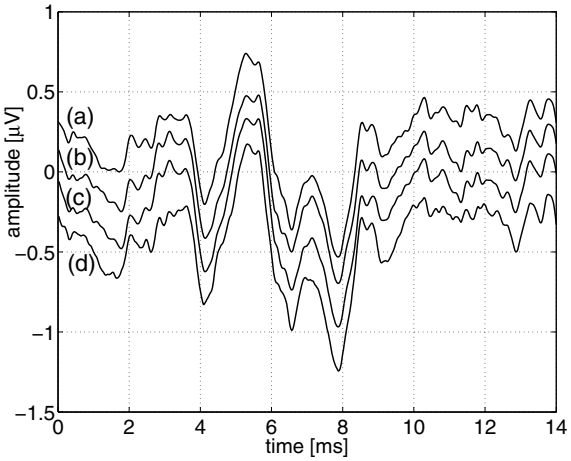


Fig. 5. Results of ABRs averaging for 1800 real sweeps: (a) AM method, (b) WA method, (c) WACFMGC method, (d) TM method. For better presentation signals (b), (c) and (d) are shifted by -0.2, -0.4 and -0.6 μV respectively

5 Conclusion

A new method of robust weighted averaging of periodic signals is presented in this work. It is applied to auditory brainstem responses averaging. The proposed method operation is based on minimization of scalar criterion function. The robust dissimilarity functions derived from the generalized Cauchy are applied. The special cases of the proposed method are obtained with use of the myriad and the meridian cost functions, respectively. The robustness of the proposed methods can be controlled either with the two parameters. Two patterns of simulation experiments were proposed to evaluate the respective methods. The methods were tested with the artificial signals and the real ABRs. The best capability of the noise suppression (the smallest value of RMSE which means the highest signal-to-noise improvement) in majority of cases had the method based on the generalized Cauchy distribution cost function. The obtained results show the usefulness of the presented WACFMGC weighted averaging method for ABRs processing. The presented methods can help to improve averaging of evoked potentials like ABRs or other biomedical signal cycles when the data are highly non-stationary and the signals are disturbed by the impulsive noise.

Acknowledgement. This work is supported by SUT grant BK-216/RAu3/2013 The authors would like to thank J.Zając and W.Kowalski (from EMAG, Katowice) for providing them with their ABRs data and the anonymous reviewers for their valuable suggestions.

References

1. Carrillo, R., Aysal, T., Barner, K.: A generalized Cauchy distribution framework for problems requiring robust behavior. EURASIP JASP (2010), doi:10.1155/2010/312989
2. Dave, R.N., Krishnapuram, R.: Robust clustering methods: a unified view. IEEE Trans. Fuzzy Syst. 5(2), 270–293 (1997)
3. De Silva, A., Sinclair, N., Liley, D.: Limitations in rapid extraction of evoked potentials using parametric modeling. IEEE Trans. Biomed. Eng. 59, 1462–1471 (2012)
4. Gonzalez, J., Paredes, J., Arce, G.: Zero-order statistics: A mathematical framework for the processing and characterization of very impulsive signals. IEEE Trans. Signal Process. 54, 3839–3851 (2006)
5. Hassan, U., Anwar, M.: Reducing noise by repetition: introduction to signal averaging. Eur. J. Phys. 30, 453–465 (2010)
6. Hathaway, R.J., Bezdek, J.C., Hu, Y.: Generalized fuzzy c-means clustering strategies using lp norm distances. IEEE Trans. Fuzzy Syst. 8, 576–582 (2000)
7. John, M., Dimitrijevic, D., Picton, T.: Weighted averaging of steady-state responses. Clin. Neurophysiol. 112, 555–562 (2001)
8. Kalluri, S., Arce, G.: Adaptive weighted myriad filter algorithms for robust signal processing in α-stable noise environments. IEEE Trans. Signal Process. 46, 322–334 (1998)
9. Kidmose, P., Looney, D., Ungstrup, M., Rank, M., Mandic, D.: A study of evoked potentials from ear-eeg. IEEE Trans. Biomed. Eng. 60, 2824–2830 (2013)
10. Laciar, E., Jané, R.: An improved weighted signal averaging method for high-resolution ecg signals. Comput. Cardiol. 28, 69–72 (2001)
11. Ł ęski, J.: Robust weighted averaging. IEEE Trans. Biomed. Eng. 49, 796–804 (2002)
12. Lee, Y., Kassam, S.: Generalized median filtering and related nonlinear filtering techniques. IEEE Trans. Acoust., Speech, Signal Process. 33, 672–683 (1985)
13. Leonowicz, Z., Karvanen, J., Shishkin, S.: Trimmed estimators for robust averaging of event-related potentials. J. Neurosci. Meth. 142, 17–26 (2005)
14. McCulloch, J.H.: Simple consistent estimators of stable distribution parameters. Comm. Stat. Simulat. Comput. 15, 1109–1136 (1986)
15. Pander, T.: An application of myriad m-estimator for robust weighted averaging. In: Gruca, A., Czachórski, T., Kozielski, S. (eds.) Man-Machine Interactions 3. AISC, vol. 242, pp. 269–276. Springer, Heidelberg (2014)
16. Pander, T., Przybyła, T., Czabański, R.: An application of the l_p-norm in robust weighted averaging of biomedical signals. JMIT 22, 71–78 (2013)
17. Plourde, G.: Auditory evoked potentials. Best Pract. Res. Clin. Anaesthesiol. 20, 129–139 (2006)
18. Shao, M., Nikias, C.L.: Signal processing with fractional lower order moments: α-stable processes and their applications. Proc. IEEE 81, 986–1009 (1993)
19. Walsh, P., Kane, N., Butler, S.: The clinical role of evoked potentials. J. Neurol. Neurosurg. Psychiatr. 76, 16–22 (2005)

Reproduction of Missing Parts of a Quasi Periodic Signals

Tomasz Przybyła, Tomasz Pander, and Robert Czabański

Silesian University of Technology, Institute of Electronics, Gliwice, Poland
Tomasz.Przybyla@polsl.pl

Abstract. In this paper we propose a reproduction method of missing parts of an analysed signal. The analysed signal must meet one essential requirement: the signal should contain repetitive parts. The reproduction process is carried out in a phase–space. The phase space is created by applying the Taken's theory to the analysed signal. In the proposed approach parts of the signal are reconstructed from a projective sub–space. The dimension of the projection sub–space is much lower than the phase–space dimension. An artificial generated signal is used in our numerical experiments. The median of a vicinity method is used as the reference method. The obtained results show advantages of the proposed approach.

Keywords: phase–space, projection sub–space, reproduction, Taken's theory.

1 Introduction

Most of a bio–electric signals manifest small (single mV) or very small (tens of μV) amplitudes. During such a signal recording an additional noise component is also recorded. The noise amplitude can exceed the amplitude of a usable signal [1]. A filtering process is applied for the noise suppression. Linear filters are successfully applied when a noise frequency band does not overlap a frequency band of the processed signal. Significant progress in biomedical signal processing was achieved by applications of digital filters [2]. Unfortunately, the wide frequency band of the noise that overlaps the frequency band of the processed signal makes the filters practically useless. The synchronized averaging technique was introduced to cope with overlapping frequency bands [3]. Due to increasing computational power of computers, methods from the field of nonlinear dynamics are taken into account. One of these methods is the nonlinear state-space projection method (NSSP). The nonlinear state-space projection filtering was successfully applied in the ECG noise reduction [4,5], the fetal ECG extraction [6] or the noise reduction in hydrologic time series [7]. For the projection subspace estimation, the NSSP methods involve principal component analysis (PCA) [15].

Another group of biomedical signals are indirectly recorded signals. An example of such signals is a fetal hear rate (FHR) signal. The most popular method

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_31, © Springer International Publishing Switzerland 2014

for the FHR monitoring is based on monitoring of a mechanical activity of a fetal heart using the Doppler ultrasound technique. It is particularly susceptible to motion artifacts coming from fetal movements or uterine contractions. These disturbances together with a transducer displacement are common causes of signal loss episodes. In such a case, the values of the fetal heart rate are unknown. The episodes are most often being replaced with the mean value of the fetal heart rate which is estimated from the neighbouring samples, or using a linear interpolation of the samples [8,9].

Paper [10] presents a recovery methods for FHR signals. The missing parts are computed as a median or a weighted mean of a vicinity. The weights are estimated based on a distance distribution of the vicinity elements.

This paper is organized as follows: the second section contains a signal representation in a phase–space as well as the description of reference method and the proposed approach. Results from the conducted numerical experiments are presented in Section 3. Conclusions complete this paper.

2 Methods

In this section we present a phase–space reconstruction method. Afterwards, a vicinity problem estimation is presented as well as the reproduction method of missing parts of the signal. The presented median method is taken as the reference reproduction method. Next, we present a conception of an estimation of the projective sub–space and a description of the proposed method as well. Finally, reconstruction in the phase–space of a projected point is shown.

2.1 Phase-Space Reconstruction Method

The applied technique is an outcome of the theory of nonlinear dynamical systems. In the proposed method it is assumed that variables describing a system depend on each other in an undetermined way. Observing one variable for a long time makes it possible to reproduce values of the other descripting variables of the system. At the t moment, the value of only one variable is known, but other variables affect the observed variable value at the $t + \tau$ moment. In deterministic dynamical systems, the post–transient trajectory of the system is frequently confined to a set of points in the phase–space, called an attractor [11]. The phase–space (or state–space, the both names are used interchangeable in this paper) can be reconstructed by the Takens embedding operation [12]. For a given signal x_N, the point in the reconstructed state-space is given by

$$\mathbf{x}_n = [x(n), x(n+\tau), \ldots, x(n+(m-1)\tau)], \tag{1}$$

where $x(n)$ is the processed signal, N is the length of the signal, τ is the time lag and m is the embedding dimension. The product $(m-1)\tau$ is the embedding window. In many applications, the time lag $\tau = 1$ is advantageous [4,13]. So, henceforth this time lag value will be used in this study. The whole processed

signal x_N which is represented by a vector in the time domain, in the phase-space is represented by the matrix $\mathbf{X}$ given by

$$\mathbf{X} = \begin{bmatrix} \vdots & & & \vdots \\ x_{n-m+1} & x_{n-m+2} & \dots & x_{n-1} & x_n \\ \vdots & & & \vdots \\ x_{n-1} & x_n & \dots & x_{n+m-3} & x_{n+m-2} \\ x_n & x_{n+1} & \dots & x_{n+m-2} & x_{n+m-1} \\ \vdots & & & \vdots \end{bmatrix} \tag{2}$$

where: x_n denotes the amplitude of a processed signal at the time stamp n. It is essential for the time- domain reconstruction, that the element x_n occurs m times in the phase-space representation.

2.2 Vicinity Determination

Let $\mathbf{x}_{n_0}^{(l)}$ be a point of the embedded space that represents a missing value at the l position (feature number). Let $\mathcal{X}$ denotes a set of points that lie in the hypersphere of a given radius ε and the origin at $\mathbf{x}_{n_0}^{(l)}$, i.e

$$\mathcal{X} = \left\{ \mathbf{x}_n \Big|_{n,\ n \neq n_0}^{\forall} \|\mathbf{w}_l \otimes \mathbf{x}_n - \mathbf{x}_{n_0}^{(l)}\| \leq \varepsilon \right\}, \tag{3}$$

where $\mathbf{w}_l$ is a weight vector which is defined as follows

$$\mathbf{w}_l(k) = \begin{cases} 1 & k \neq l, \\ 0 & k = l \end{cases}. \tag{4}$$

The $\otimes$ operation denotes multiplication of corresponding components of two vectors, i.e.

$$\mathbf{a} \otimes \mathbf{b} = \begin{bmatrix} a_1\ b_1 \\ a_2\ b_2 \\ \vdots \\ a_m\ b_m \end{bmatrix}$$

Hence, the missing sample (feature) can be computed as

$$\mathbf{x}_{n_0}^{(l)} = \text{median}\,(\mathcal{X}) \otimes \tilde{\mathbf{w}}_l, \tag{5}$$

where $\tilde{\mathbf{w}}_l$ is the complement to the $\mathbf{w}_l$ and is defined as follows

$$\tilde{\mathbf{w}}_l(k) = \begin{cases} 0 & k \neq l \\ 1 & k = l \end{cases}. \tag{6}$$

In many cases it is hard to choose the correct value of the vicinity radius ε. For small values of ε, the set $\mathcal{X}$ might be an empty set. In such a case, the recovery

of the missing sample is not possible. Thus, a value of ε should be increased according to an assumed method (e.g. $\varepsilon \leftarrow \varepsilon \cdot \beta$, where $\beta > 1$) and the vicinity has to be estimated again. This strategy increases the computation time. For large values of ε, the $\mathcal{X}$ might include distant points that can deteriorate the estimation result. Therefore the k–nearest neighbours (kNN) method is proposed as the way for the vicinity determination.

The k–nearest neighbors method is a method for classifying objects based on the closest training examples [14]. Unlike the previous method, it does not require a radius specification. Instead, the number of vicinity members has to be specified. So, the cardinal number of $\mathcal{X}$ is known and it is equal to the number of specified neighbours.

2.3 PCA and Projective Reproduction

Principal component analysis (PCA) is a technique for extracting a structure from high-dimensional data sets [15]. PCA is an orthogonal transformation of the coordinate systems in which the data are described. The new coordinate system (known as the principal coordinates) is obtained by the projection onto the so-called principal axes of the data. Let $\mathcal{X}^{(n)} = \left\{\mathbf{x}_1^{(n)}, \mathbf{x}_2^{(n)}, \ldots, \mathbf{x}_N^{(n)}, \right\}$ be a vicinity of $\mathbf{x}^{(n)}$, where $\mathbf{x}_i^{(n)} \in \mathbb{R}^m$. Each element from the dataset is described by m features associated with the time stamp (i). Determination of the principal axes begins with centering the data samples and then computing the sample covariance matrix, i.e.

$$\mathbf{C}_X = \frac{1}{N} \sum_{i=1}^{N} \left(\mathbf{x}_i^{(n)} - \bar{\mathbf{x}}^{(n)}\right) \left(\mathbf{x}_i^{(n)} - \bar{\mathbf{x}}^{(n)}\right)^T, \tag{7}$$

where $\bar{\mathbf{x}}^{(n)}$ is the sample mean,

$$\bar{\mathbf{x}}^{(n)} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_i^{(n)},$$

and $N = |\mathcal{X}^{(n)}|$ is the cardinal number of the dataset. The time stamp indicator is marked as the superscript and the subscript represents the vicinity item number for avoiding double subscripts.

The principal axes $\mathbf{w}_i$ $(1 \leq i \leq d)$ are equal to the eigenvectors that correspond to the largest eigenvalues of the covariance matrix $\mathbf{C}_X$. The projection onto d-dimensional principal space is a linear transformation of $\mathbf{x}_{(n)}$ which is performed according to the following equation

$$\mathbf{y}_{(n)} = \mathbf{W}^T \left(\mathbf{x}_{(n)} - \bar{\mathbf{x}}_{(n)}\right), \tag{8}$$

where $\mathbf{y}_{(n)} \in \mathbb{R}^d$ is the d-dimensional representation of the $\mathbf{x}_{(n)}$ and $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_d]$ is the projection matrix.

A reconstruction of the projected point $\mathbf{y}_{(n)}$ in the m-dimensional space is given by

$$\hat{\mathbf{x}}_{(n)} = \mathbf{W}\mathbf{y}_{(n)} + \bar{\mathbf{x}}_{(n)} = \mathbf{W}\mathbf{W}^T \left(\mathbf{x}_{(n)} - \bar{\mathbf{x}}_{(n)}\right) + \bar{\mathbf{x}}_{(n)}. \tag{9}$$

Let $\tilde{\mathbf{x}}_{n_0}$ be the nearest point to $\mathbf{x}_{n_0}^{(l)}$ taken from the vicinity $\mathcal{X}$, i.e.

$$\tilde{\mathbf{x}}_{n_0} = \min_{1 \le i \le K} \| \mathbf{w}_l \otimes \mathbf{x}_{(i)} - \mathbf{x}_{n_0}^{(l)} \|,$$

and the weights are defined by (4).

Thus, the missing sample (feature) can be calculated in the following way:

$$\mathbf{x}_{(n_0)}^{(l)} = \tilde{\mathbf{w}}_l \otimes \left[\mathbf{W}\mathbf{W}^\top \left(\tilde{\mathbf{x}}_{(n_0)} - \bar{\mathbf{x}}_{(n)}\right) + \bar{\mathbf{x}}_{(n)}\right], \tag{10}$$

where the weights $\tilde{\mathbf{w}}_l$ are defined by (6).

3 Numerical Experiment

The aim of the numerical experiment is to investigate the influence of the method parameters on the quality of the reproduced signal. For our experiments, a reproduction error is introduced. The reproduction error is defined as:

$$\xi_1 = \frac{1}{\mathcal{T}} \sum_{t=1}^{\mathcal{T}} |X_i(t) - X_o(t)|,$$

where $X_i(t)$ is the original (uncorrupted) signal, $X_o(t)$ is the reproduced signal and $\mathcal{T}$ is the gap duration.

During the reproduction process, the following parameters are fixed: the embedded dimension:

$$m = \{21, 31, 41, 51, 61, 71\},$$

and the number of neighbours:

$$k = \{21, 31, 41, 51\}.$$

In the first experiment a fully deterministic signal — Gaussian-modulated sinusoidal pulse – is used. The frequency of the modulated sinusoidal pulse is $f_p = 20$Hz, repetition frequency is $f_r = 2$Hz and the signal is sampled with the frequency $f_s = 100$Hz. This signal is subjected to a corruption process. The corruption lasts for $\mathcal{T} = \{5, 10, 15, 20, 50, 75, 100\}$ samples. For this time an amplitude of the signal is undefined and unknown. The purpose of this experiment is to find out the method that yields exact results for exact data (i.e. analytical). The signal is corrupted in the following way: a corruption (signal gap) begins in a flat (or nearly flat) part of the signal. Such location of the gap allows to investigate the reproduction process of a constant or a slow variable signal parts. Afterwards, the gap duration is increased until the gap encompass the fast changing parts of the signal.

Figure 1 depicts a part of the original signal (uncorrupted) as well as the corrupted signal and the reproduced signal. For the both reproduction method (i.e. the projective method and the reference method), the embedding dimension is $m = 51$ and the vicinity size is $K = 31$. For the projection method, the dimension of the projective sub–space is $d = 2$. On the top of the Figure 1 a part of the signal is presented. In the middle of the figure the corrupted signal is depicted, and at the bottom the reproduced signal obtained from the projective method is presented. For the reference method the obtained signal is exactly the same as signal from the projective method. The reproduction error computed for the both methods is proportional to $\xi_1 \propto 10^{-13}$ and it is caused by round–off computation errors. Hence, for the exact data (i.e. the noise–free signal), the proposed method results in exact results.

The purpose of the second experiment is to investigate the influence of a noise included in the reproduced signal on the reproduction error. The signal to noise ratio is defined as

$$SNR = 10 \log \frac{\sigma_s^2}{a^2 \sigma_v^2},$$

where σ_s^2 is the signal variance, σ_v^2 is the noise variance and a is an amplitude of the added noise. In our experiments, an additive noise model is used, i.e.

$$x(n) = s(n) + av(n)$$

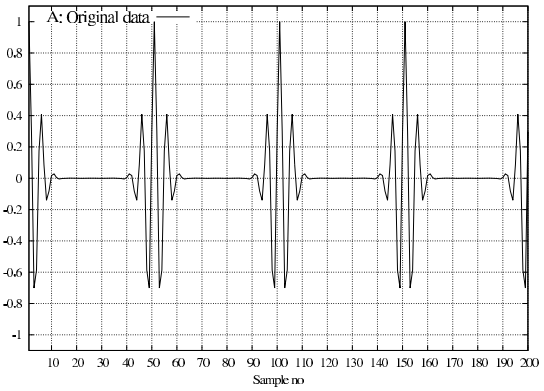
where $x(n)$ is a contaminated signal, a is an amplitude of the noise component and $v(n)$ is the noise. The Gaussian distributed noise with variance $\sigma_v^2 = 4$ is used as the noise component.
The corruption length (the gap size) changes in the same way as in the previous experiment. The SNR factor is set either 10 or 5 dB. An example of a reproduction of the noisy signal is presented in Figure 2. Figure 2(a) depicts the contaminated signal. The gap length is $\mathcal{T} = 100$ samples and the signal to noise ratio is $SNR = 5$dB. In Figure 2(b) the signal reproduced by the reference method is presented and Figure 2(c) depicts the reproduced signal by the projective method. The dimension of a projection sub–space is $d = 2$. The reproduction error for the projective method is $\xi_1 = 13.23$ and for the reference method is $\xi_1 = 8.17$.

The reproduced signal obtained by the reference method is not in line with expectation despite the lower reproduction error value. Therefore, a new reproduction quality index is introduced:

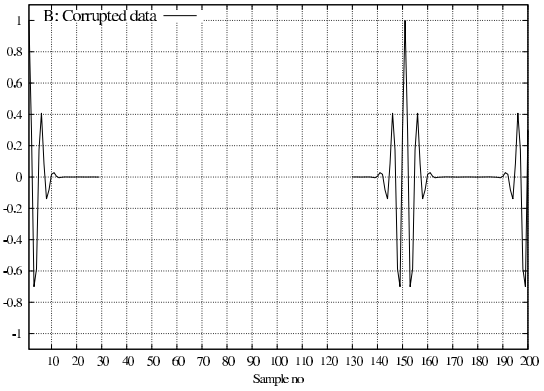
$$\xi_2 = |var(x_i) - var(x_o)|,$$

where $var(x_i)$ is the variance of the uncorrupted signal and $var(x_o)$ is the variance of the reproduced signal.

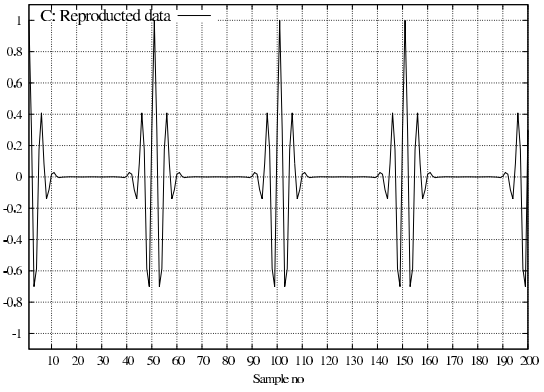
Table 1 and Table 2 contain the reproduction error values for SNR=5dB and different dimensions of a projection sub–space for the dimension of projective sub–space $d = 2$ and $d = 5$, respectively. Values of the embedding dimension m are in the leftmost column.



(a) original signal

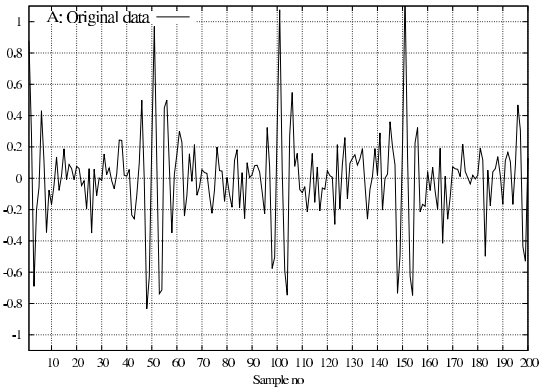


(b) corrupted signal

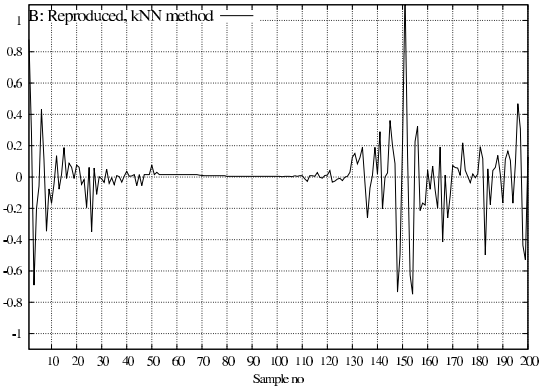


(c) reproduced signal

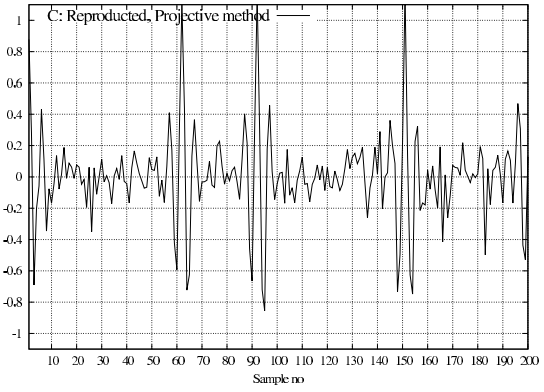
Fig. 1. Example of the reproduction of a deterministic signal. The signal corruption lasts for $\mathcal{T} = 100$ samples. The embedding dimension is $m = 51$, the vicinity size is $K = 31$ and the dimension of projection sub–space is $d = 2$.



(a) original signal



(b) reproduced signal by the reference method



(c) reproduced signal by the projective method

Fig. 2. Example of the reproduction of a contaminated signal; the signal gap is $\mathcal{T} = 100$ samples; the embedding dimension is $m = 51$, the vicinity size is $K = 31$ and the dimension of projection sub–space is $d = 2$.

Table 1. The reproduction error computed for the projection method. The signal to noise ratio is SNR=5dB. The leftmost column indicates the embedding dimension. The vicinity size is presented in the selected rows. The dimension of the projective sub–space is $d = 2$.

Projection method, d=2								
m	$\mathcal{T} = 10$				$\mathcal{T} = 20$			
	21	31	41	51	21	31	41	51
21	0.0183	0.0163	0.0236	0.0202	0.0211	0.0217	0.0218	0.0357
31	0.0166	0.0080	0.0001	0.0000	0.1942	0.1949	0.1951	0.1940
41	0.0154	0.0146	0.0301	0.0269	0.0240	0.0406	0.0460	0.0445
51	0.0382	0.0301	0.0314	0.0265	0.0446	0.0499	0.0517	0.0424
61	0.0365	0.0167	0.0218	0.0179	0.0392	0.0308	0.0371	0.0449
71	0.0199	0.0089	0.0138	0.0246	0.0472	0.0539	0.0465	0.0468
m	$\mathcal{T} = 50$				$\mathcal{T} = 100$			
	21	31	41	51	21	31	41	51
21	0.0249	0.0264	0.0313	0.0281	0.0168	0.0139	0.0001	0.0170
31	0.0232	0.0284	0.0297	0.0324	0.0123	0.0151	0.0158	0.0556
41	0.0204	0.0940	0.0915	0.0341	0.0200	0.0228	0.0190	0.0227
51	0.0296	0.0207	0.0187	0.0195	0.0204	0.0194	0.0199	0.0191
61	0.0341	0.0260	0.0234	0.0219	0.0137	0.0176	0.0211	0.0204
71	0.0259	0.0290	0.0303	0.0223	0.0186	0.0223	0.0191	0.0186

Table 2. The reproduction error computed for the projection method. The signal to noise ratio is SNR=5dB. The leftmost column indicates the embedding dimension. The vicinity size is presented in the selected rows. The dimension of the projective sub–space is $d = 5$.

Projection method, d=5								
m	$\mathcal{T} = 10$				$\mathcal{T} = 20$			
	21	31	41	51	21	31	41	51
21	0.0209	0.0171	0.0244	0.0217	0.0120	0.0177	0.0070	0.0312
31	0.0289	0.0187	0.0185	0.0179	0.1905	0.1915	0.1922	0.1920
41	0.0195	0.0061	0.0098	0.0242	0.0228	0.0177	0.0437	0.0427
51	0.0174	0.0275	0.0319	0.0185	0.0276	0.0258	0.0399	0.0428
61	0.0244	0.0372	0.0267	0.0168	0.0394	0.0497	0.0441	0.0520
71	0.0137	0.0123	0.0022	0.0023	0.0355	0.0483	0.0532	0.0525
m	$\mathcal{T} = 50$				$\mathcal{T} = 100$			
	21	31	41	51	21	31	41	51
21	0.0224	0.0214	0.0229	0.0258	0.0117	0.0073	0.0120	0.0099
31	0.0231	0.0245	0.0255	0.0258	0.0050	0.0117	0.0120	0.0099
41	0.0319	0.0287	0.0315	0.0293	0.0130	0.0161	0.0113	0.0204
51	0.0318	0.0270	0.0220	0.0205	0.0143	0.0096	0.0140	0.0141
61	0.0308	0.0258	0.0138	0.0205	0.0110	0.0092	0.0142	0.0138
71	0.0327	0.0328	0.0312	0.0282	0.0133	0.0113	0.0140	0.0147

Table 3. The reproduction error computed for the reference method. The signal to noise ratio is SNR=5dB. The leftmost column indicates the embedding dimension. The vicinity size is presented in the selected rows.

Reference method								
m	$\mathcal{T}=10$				$\mathcal{T}=20$			
	21	31	41	51	21	31	41	51
21	0.0287	0.0377	0.0363	0.0307	0.2132	0.2135	0.2138	0.2141
31	0.0424	0.0369	0.0330	0.0319	0.2123	0.2119	0.2122	0.2123
41	0.0285	0.0318	0.0328	0.0336	0.0373	0.0449	0.0431	0.0471
51	0.0266	0.0294	0.0297	0.0305	0.0283	0.0350	0.0397	0.0391
61	0.0277	0.0377	0.0330	0.0322	0.0483	0.0368	0.0366	0.0401
71	0.0324	0.0374	0.0307	0.0309	0.0399	0.0398	0.0380	0.0380
m	$\mathcal{T}=50$				$\mathcal{T}=100$			
	21	31	41	51	21	31	41	51
21	0.0962	0.0970	0.0971	0.0971	0.0879	0.0884	0.0885	0.0886
31	0.0963	0.0964	0.0967	0.0970	0.0881	0.0882	0.0883	0.0884
41	0.0330	0.0326	0.0332	0.0316	0.0228	0.0234	0.0246	0.0243
51	0.0269	0.0258	0.0279	0.0280	0.0208	0.0208	0.0212	0.0216
61	0.0304	0.0306	0.0301	0.0292	0.0188	0.0208	0.0206	0.0201
71	0.0289	0.0292	0.0290	0.0267	0.0197	0.0193	0.0209	0.0210

The Table 3 shows the results obtained from the reference method. The signal to noise ratio is $SNR = 5$dB. The table is organized analogously to Table 1.

For the most cases, the projection method results in comparable or lower values of the reproduction error ξ_2 to the reference method. The cases, when the projection method yields bigger (c.a. 1%) values of the reproduction error ξ_2 take place for a high ($m > 51$) embedding dimension and a low ($K = 21$) vicinity size. For these cases, the covariance matrix, during the PCA process, is rough estimated. This results in an improper projective sub–space estimation. When the vicinity size is equal or greater than the embedding dimension, the projective method yields lower values of the reproduction error.

4 Conclusions

In this paper we presented a reproduction method of missing parts of a signal. The reproduced signal should meet one requirement: it has to consists of similar or repetitive parts. The reproduction process is conducted in a phase–space reconstructed for the signal. The phase–space reconstruction is done by applying the Taken's theory. A missing sample in the time domain corresponds to a missing feature (coordinate) in the phase–space. The following steps are carried out to perform reproduction of a missing sample. In the first step, a vicinity in the feature space is estimated for each point with a missing feature. Afterwards, a projective sub–space is computed for the estimated vicinity. The dimension of

the projective sub–space is much lower than the dimension of the reconstructed phase–space. Next, for the point with a missing coordinate we seek for the nearest point and then the point is projected into sub–space. Finally, the missing feature is computed based on the reconstructed point in the phase–space taken from projective sub–space. In the conducted numerical experiments, two cases were taken into consideration. In the first experiment, a noise–free signal was corrupted and then the missing samples were reproduced. In the second case, the reproduced signal was contaminated with a Gaussian distributed noise. The signal to noise ratio was fixed. In the both cases, the results obtained for the proposed method proof the correctness of the presented approach.

Acknowledgement. Calculations were carried out using the computer cluster Ziemowit (http://www.ziemowit.hpc.polsl.pl) funded by the Silesian BIOFARMA project No. POIG.02.01.00-00-166/08 in the Computational Biology and Bioinformatics Laboratory of the Biotechnology Centre in the Silesian University of Technology.

References

1. Tompkins, W.J.: Biomedical digital signal processing: C–language examples and laboratory experiments for the IBM PC. Prentice Hall (1993)
2. Pahlm, O., Sornmo, L.: Data processing of exercisse ECG's. IEEE Trans. Biomed. Eng. BME-34, 158–165 (1987)
3. Jane, R., Rix, H., et al.: Aligment methods for averaging of high resolution cardiac signals: a comparative study of performance. IEEE Trans. Biomed. Eng. 38, 571–579 (1991)
4. Schreiber, T., Kaplan, D.: Nonlinear noise reduction for electrocardiograms. Chaos 6, 87–92 (1996)
5. Kotas, M.: Projective filtering of time warped ECG beats. Comp. in Biology and Medicine (38), 127–137 (2008)
6. Richter, M., Schreiber, T.: Fetal ECG extraction with nonlinear state space projections. IEEE Trans. Biomed. Eng. 45, 133–137 (1998)
7. Elshorbagy, A., Simonovic, S.P., Panu, U.S.: Noise reduction in chaotic hydrologic time series: facts and doubts. J. Hydrology (256), 147–165 (2002)
8. Kupka, T., Jeżewski, J., Matonia, A., et al.: Timing events in Doppler ultrasound signal of fetal heart activity. In: 26th IEEE/EMBS, San Francisco, USA, pp. 337–340 (2004)
9. Jeżewski, J., Wróbel, J., Labaj, P., et al.: Some Practical Remarks on Neural Networks Approach to Fetal Cardiotocograms Classification. In: 29th IEEE/EMBS, Lyon, France, pp. 5170–5173 (2007)
10. Przybyła, T., Pander, T., Wróbel, J., et al.: A recovery of FHR signal in the Embedded Space. In: XIII Mediterranean Conference on Medical and Biological Engineering and Computing, IFMBE Proc., vol. 41 (2013)
11. Kantz, H., Schreiber, T.: Nonlinear time series analysis. Cambridge Univ. Press (2004)

12. Takens, E.: Detecting strange attractors in turbulence. Lecture Notes in Math, vol. 898, pp. 366–381 (1981)
13. Kotas, M.: Robust projective filtering of time-warped ECG beats. Comp. Methods and Programs in Biomedicne (92), 161–172 (2008)
14. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification. Wiley-Interscience, New Jersey (2000)
15. Jollife, I.T.: Principal component analysis. Springer, New York (2002)

Pronunciation Error Detection Using Dynamic Time Warping Algorithm

Marcin Bugdol, Zuzanna Segiet, and Michał Kręcichwost

Silesian University of Technology, Faculty of Biomedical Engineering, Zabrze, Poland
marcin.bugdol@polsl.pl

Abstract. In the paper a pronunciation error detection method has been presented, wchich is based on word structural features. A low-complexity classifier has been proposed, that is not concentrated on a limited base of error patterns, but is flexible enough to find unspecified mispronunciations. Two classification variants using Dynamic Time Warping (DTW) algorithm has been tested on speech corpus containing recordings of 30 people.

Keywords: pronunciation error detection, structural features, Dynamic Time Warping (DTW), Polish language.

1 Introduction

Automatic speech recognition and automatic speaker recognition have been the objects of researchers' interest for the last few decades. Detection of pronunciation errors is a derivative of these problems, as it provides new applications to well-known solutions. The detection of incorrectly pronounced phonemes or their joints is studied mostly in the context of its usage in computer assisted language learning systems (CALL systems) [1].

The vast majority of systems which are currently being developed focus on pronunciation errors in Asian tonal languages. The studies concern mostly Mandarin Chinese [1], but also Taiwanese [2] and other ones. There are few studies that focus on the European languages (e.g. [3]). Based on authors' knowledge, the only work concerning Polish language (and Czech, Slovak, Russian and German) is the EURONOUNCE project [4,5,6,7]. The software designed within this project is meant to detect the pronunciation and prosody errors of the non-native speakers.

Several basic concepts of digital analysis of the pronunciation correctness can be distinguished. Approaches can be divided based on the choice of structural units that can be classified as incorrect or correct. Basic units of speech analysis are predominantly diphones (adjacent pairs of phones) [8] or triphones [9,10,11]. In [12] the detection of errors is based on individual phonemes, while in [13] the use of syllables as the basic units of analysis is suggested.

Pronunciation error detection methods can be also divided due to the detection of only specific errors or mispronunciation in general. The approaches

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_32, © Springer International Publishing Switzerland 2014

suggested to detect the latter are mostly based on Hidden Markov Models (HMM) [4,14,2]. Goodness of Pronunciation (GOP) [15,1,16,17] or Log-Posterior Probability (LPP) [12] are the examples of measures which determine the correctness of the phone by thresholding the values calculated upon the probabilities provided by HMM. This does not require specifying the troublesome phonemes in advance.

Most of the studies concentrate only on a small number of selected phones in order to simplify the task. Tests results for the 8 most common mistakes that the Japanese do in pronunciation of the Chinese language, were described in [1]. Similarly, the method presented in [12] is based on several error patterns. The models were built with the use of linguistic knowledge of frequently made mispronunciations. Combining the speech corpus with appropriate language rules increased the efficiency of detection of systematic errors. However, any other distortions could not be detected.

Many CALL systems use LPP as a measure of speech quality because of its robustness. Predefined errors detection requires different classifiers and measures of pronunciation quality. The Revised Log-Posterior Probability (RLPP) [12] is a modification of this methods. The LPP calculates the probability of correctness with regard to all phonemes in the network, while the RLPP considers only certain phonemes which were annotated as the most problematic. The aforementioned process narrows down the search field and increases the probability of correct classification at the same time.

Different methods of systematic errors detection are Likelihood Ratio, which explicitly specifies whether a phoneme is more similar to the correct or incorrect pattern included in the speech base [1], or Weigelt's algorithm which is based on the signal time parameters [15].

The method described in [1] suggests a different approach to the detection of errors. The idea is not based on the analysis of the absolute features of a signal segment (like in GOP or LR), but on the use of structural features which depend on other segments of a word. The distances between distributions of specific segments are calculated and placed in distance matrices of correctly and incorrectly pronounced word parts. Subsequently, the structural deviations between the distance matrices are calculated. On that basis, hypotheses about incorrectness of outliers may be posed.

The purpose of the use of structural features is to determine which phoneme has been pronounced incorrectly, in accordance to the comparison with other phonemes in the same word. As noted in [1], this approach shall not work when more than one phoneme in the word is incorrect, because the distances between mispronounced segments cannot be regarded as valid references. This problem may be solved by the use of the additional SVM classifier for each phoneme in the distance matrix. Each of the phonemes is pre-classified as incorrect or correct, which gives it the appropriate weight (0 or 1) for calculating the actual structural deviations.

The detection of pronunciation errors with SVM classifier has been also proposed in [12].

2 Dynamic Time Warping Method for Pronunciation Error Detection

The aim of this paper is to propose a pronunciation error detection method, which will not require an extensive speech corpus, will have relatively low computational complexity and will still maintain its effectiveness while working on the limited dictionary. It has been decided not to concentrate on systematic errors, but conduct a study that will test the algorithm robustness to finding diversified mispronunciations. Another objective of this work is to propose a system dedicated to Polish language and optimized for its pronunciation.

Phones have been used instead of syllables or diphones. The use of smaller structural units makes it possible to detect mistakes more precisely. According to [13], phonetic errors do not occur at the transitions between syllables, but rather around their center. This could help to optimize the operation of the system as searching could be conducted by taking only cores of syllables instead of all the phonemes. Unfortunately, the specifics of Polish pronunciation do not support this thesis, since phonemes at the end or beginning of a syllable may also be distorted (especially in case of such troublesome sounds as /r/). Because of this, the syllables were decided not to be used as units of analysis.

Similarly as in [1], the proposed method is based on the structural characteristics of the words pronounced correctly and incorrectly. Dynamic Time Warping (DTW) algorithm has been used as a source of metrics. The algorithm result is a matrix which contains the distances between various elements of compared vectors (Fig. 1). Subsequently, the algorithm calculates the optimal path of alignment in the matrix. The cost of this path is treated as a measure of similarity of patterns.

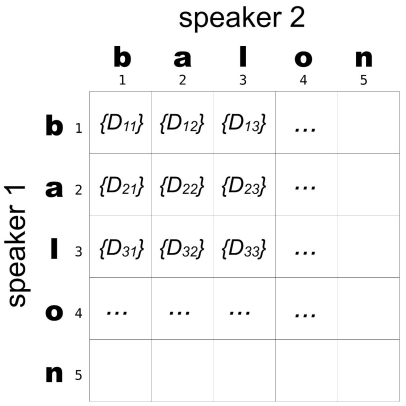


Fig. 1. The DTW matrix for the word "balon" recorded by two speakers. D_{ij} contains a distance between the i-th segment of speaker 1 recording and the j-th segment of speaker 2 recording

The great advantage of the DTW algorithm is the simplicity of its implementation. The possibility of using it in a pronunciation error detector will be shown in the subsequent parts of the paper.

3 Experiments

The study has been conducted on a set of over 400 recordings of words spoken correctly and incorrectly. The words came from a fixed vocabulary of 24 words of phonetically diversified structure (Table 1).

Table 1. Analyzed words with transcription according to International Phonetic Alphabet (IPA)

Word	Pronunciation	Word	Pronunciation
aparat	[aˈparat]	kubek	[ˈkubɛk]
balon	[ˈbalɔ̃n]	lekarz	[ˈlɛkaʃ]
cebula	[t͡sɛˈbula]	łazienka	[waˈʑɛ̃nka]
ćwiczenie	[t͡ɕfʲiˈt͡ʃɛɲɛ]	następny	[naˈstɛ̃mpnɨ]
dalej	[ˈdalɛj]	okno	[ˈɔknɔ]
ekran	[ˈɛkrãn]	pokaż	[ˈpɔkaʃ]
firanka	[fʲiˈrãnka]	poprzedni	[pɔˈpʃɛdʲɲi]
gałąź	[ˈgawɔ̃w̃ɕ]	rower	[ˈrɔvɛr]
haczyk	[ˈxat͡ʃɨk]	wstecz	[fstɛt͡ʃ]
jajko	[ˈjäjkɔ]	zamknij	[ˈzãmʲcɲij]
koniec	[ˈkɔ̃ɲɛt͡s]	zupa	[ˈzupa]
książka	[ˈcɕõw̃ʃka]	żona	[ˈʒɔ̃na]

Speech database for error detection testing should include recordings of people actually committing errors, such as foreign language learners. However, the databases for new methods testing are often constructed from recordings prepared by native-speakers, but containing artificial errors (e.g. swapped vowels) [1,18]. For this study, 30 people with correct pronunciation have been asked to deliberately introduce errors in part of the words they were supposed to say during recordings.

In order to simplify the research task, there was only one error in every recorded word - only one phoneme was substituted, inserted or deleted. The solution of the problem of greater number of errors in methods based on structural features of the signal is, as mentioned beforehand, described in [1]. This factor has been considered possible to solve by implementing additional classifiers in the future and therefore eliminated from the tests.

The recordings have been manually divided into phonemes. Mispronounced segments have been annotated in order to verify the detection results. An example of segmentation is shown in Fig. 2.

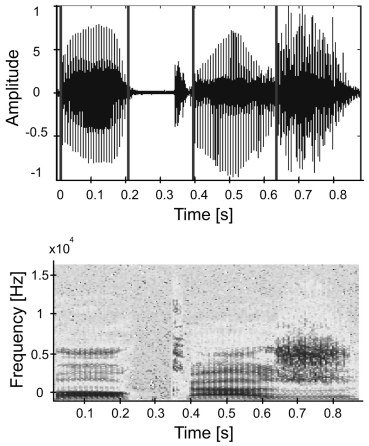


Fig. 2. Time plot and spectrogram of the isolated word. Vertical lines separate individual segments

Mel-frequency cepstral coefficient matrix was calculated for each separated segment. Subsequently, the length of the DTW optimal alignment path to each segment of a reference recording was determined, as shown in Fig. 3.

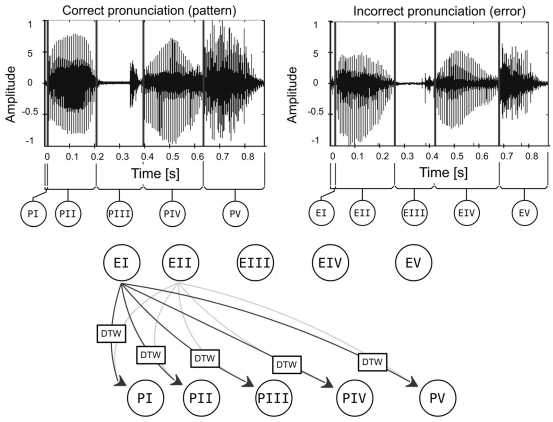


Fig. 3. Application of DTW in calculating inter-segment distances

The obtained values have been afterwards used to perform classification based on the location of the most similar segments of analyzed and reference words. Two different versions of algorithm have been applied.

The first method (DTW error detection) was based on finding the parts of the reference recording which were the most similar to the analyzed segment. If the minimum distance part position didn't correspond to the analyzed segment

location, this fragment was designated as incorrect. This method is schematically shown in Fig. 4. The decision process is performed for each segment separately.

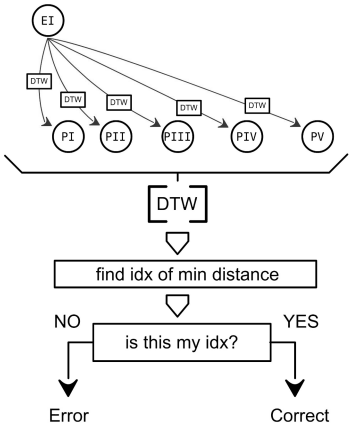


Fig. 4. DTW error detection method

The second method (Thresholding DTW error detection) is based on the first one and develops the same idea. The algorithm finds the minimum distance for each segment and normalizes them. Subsequently, it uses thresholding to determine whether the value is small enough to classify the phoneme as correct. If not, it is annotated as an error. The idea is presented in Fig. 5.

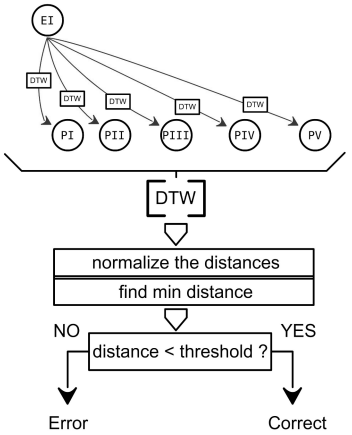


Fig. 5. Thresholding DTW error detection method

4 Results

Pronunciation error detection, as a classification problem, can return four types of outcomes [1,18] for each analyzed phoneme:

1. Correctly Accepted (CA), correct pronunciation classified as correct,
2. Correctly Rejected (CR), mispronunciation classified as incorrect,
3. Falsely Accepted (FA), mispronunciation classified as correct,
4. Falsely Rejected (FR), correct pronunciation classified as incorrect.

The optimal performance of the algorithm should demonstrate as large CA and CR counts as possible, as they sum all phone occurences that have been properly classified. The basic parameters that can be calculated given the foregoing outcome types are False Acceptance Rate (FAR), False Rejection Rate (FRR) and Average Error Rate (AER) [1]:

$$FAR = \frac{FA}{CR + FA} \tag{1}$$

$$FRR = \frac{FR}{CA + FR} \tag{2}$$

$$AER = \frac{FAR + FRR}{2} \tag{3}$$

AER, as the mean of the other two, gives the clearest view of the algorithm overall performance. Theoretically, the optimization of the system should consist of minimazing both of AER components (FAR and FRR) and keep balance between them. However, in CALL systems it is more important to keep FRR low. That is due to the fact that rejecting correct pronunciation could be more harmful to learners than accepting their pronunciation errors [18].

The Table 2 summarizes the values of all three parametres for the presented classification variants. The threshold method shows slightly better results, as AER is more satisfying and FRR is significantly lower than FAR. In this experiment, the threshold was set to keep FRR low. Nevertheless, choosing a different threshold will change the FRR to FAR ratio, so it is possible to rerun the test to minimize AER or balance its components. Future tests are supposed to include this option as well.

Table 2. The results of basic and thresholding classifier

Method	FAR	FRR	AER
DTW error detection	0.3824	0.4433	0.4128
Thresholding DTW error detection	0.5588	0.2062	0.3825

The Table 3 shows the comparison of our outcomes and the results of error detection methods described in other papers.

Table 3. Comparison of DTW error detection and other methods

Method	Source	FAR	FRR	AER
DTW error detection	**-**	**0.38**	**0.44**	**0.41**
Thresholding DTW error detection	**-**	**0.56**	**0.21**	**0.38**
Structural features + SVM	[1]	0.21	0.04	0.13
GOP	[18]	0.17	0.33	0.25
Error Patterns	[19]	0.47	0.13	0.30

It is worth noting that this comparison cannot be considered truly meaningful due to different experimental conditions in the analyzed studies. All of these methods (including ours) are based on the detection of mispronounced phonemes, not diphones or syllables. Unfortunately, this is probably the only factor that is common for them. The tests were conducted using speech databases in various languages ([1,19] - Chinese, [18] - Dutch, here - Polish) and the impact of language selection for the algorithm's performance has not been investigated yet. What is more, all methods, besides ours, concentrate on detection of predefined mispronunciations. Such approach improves the effectiveness of detection, yet introduces the risk of omission of some cases. The study described in [1] demonstrated poorer results for tests on actual recordings of non-natives than for data with pronunciation errors introduced artificially. As the authors reported, one of the reasons was that the range of non- natives errors appeared wider than expected. That means that the presumed pronunciation models cover most of the problems, but not all of them.

In order to show the problem of comparing the results of various studies, the Table 4, which contains scores of testing the same method (Goodness of Pronunciation) under different experimental conditions, is placed hereunder.

Table 4. GOP-based error detection results

Source	FAR	FRR	AER
[1]	0.75	0.12	0.43
[18]	0.17	0.33	0.25
[19]	0.42	0.24	0.33

The specifics of the study and the qualitative and quantitative diversity of speech databases have a very large impact on results. The main differences between the listed studies were language of interest (Chinese or Dutch) and the method of determining the threshold, which is necessary in GOP method. These variations proved to be enough to cause significant (from 10 to 18 percentage points) differences in the effectiveness of detection

Similarly, comparing the DTW thresholding method with other error detection systems connot be considered as truly fair. This is mostly because conducted

tests involved relatively small speech database. It is planned to collect larger number of recordings and carry out tests on a larger scale in the near future. However, the results obtained for the first implementation of the proposed classifier are promising and will probably increase through the next steps of system development.

5 Conclusion

In this paper, the use of DTW in structural feature-based pronunciation error detection method has been presented. Two classification variants have been tested on a Polish speech corpus and false rejection rates of 0.44 and 0.21 have been achieved.

This study has been designed to, inter alia, check the possibilities of finding errors which have not been specified in advance. The outcomes suggest that this approach can be effective, although, for now, the performance of the method is from 8 to 25 percentage points weaker than the other systems. Considering that this is only a prototype system, it can be assumed that it is possible to make the mispronunciation detection independent from the linguistic base and extrapolate it onto wider range of errors, and still maintain simple implementation status.

In the future, the speech corpus will be extended in order to conduct more complex experiments. Further steps of system improvement would contain automatic recording segmentation. It is also planned to run tests on diphones instead of phonemes, which could increase algorithm sensitivity to incorrect phone junctions.

References

1. Zhao, T., Hoshino, A., Suzuki, M., Minematsu, N., Hirose, K.: Automatic chinese pronunciation error detection using svm trained with structural features. In: SLT. IEEE (2012)
2. Liang, M.S., Hung, J.Y., Lyu, R.Y., Chin Chiang, Y.: Pronunciation error detection for computer assisted pronunciation teaching in mandarin. In: 6th International Symposium on Chinese Spoken Language Processing, ISCSLP 2008, pp. 1–4 (2008)
3. Russell, M., Series, R.W., Wallace, J.L., Brown, C., Skilling, A.: The star system: an interactive pronunciation tutor for young children, pp. 161–175 (2000)
4. Demenko, G., Wagner, A., Cylwik, N.: The use of speech technology in foreign language pronunciation training. Archives of Acoustics 35(5), 309–330 (2010)
5. Demenko, G., Cylwik, N., Agnieszka, W.: Applying speech and language technology to foreign language education (2009)
6. Demenko, G., Wagner, A., Cylwik, N., Jokisch, O.: An audiovisual feedback system for acquiring l2 pronunciation and l2 prosody. In: SLaTE 2009 (2009)
7. Cylwik, N., Wagner, A., Demenko, G.: The EURONOUNCE corpus of non-native polish for ASR-based pronunciation tutoring system. In: SLaTE 2009 (2009)
8. Rosenberg, A., Colla, A.: A connected speech recognition system based on spotting diphone-like segments–preliminary results. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 1987, vol. 12, pp. 85–88 (1987)

9. Saraclar, M., Nock, H.J., Khudanpur, S.: Pronunciation modeling by sharing gaussian densities across phonetic models, pp. 137–160 (2000)
10. Saraclar, M., Khudanpur, S.: Pronunciation change in conversational speech and its implications for automatic speech recognition, pp. 375–395 (2004)
11. Jokisch, O., Wagner, A., Sabo, R., Jaeckel, R., Cylwik, N., Rusko, M., Ronzhin, A., Hoffman, R.: Multilingual speech data collection for the assessment of pronunciation and prosody training in a language learning system. In: Proc. of Speech and Computer (SPECOM) (2009)
12. Xu, S., Jiang, J., Chen, Z., Xu, B.: Automatic pronunciation error detection based on linguistic knowledge and pronunciation space. In: ICASSP, pp. 4841–4844. IEEE (2009)
13. Ng, R.W.M., Hirose, K.: Syllable: A self-contained unit to model pronunciation variation. In: ICASSP, pp. 4457–4460. IEEE (2012)
14. Liang, M.S., Hong, Z.Y., Lyu, R.Y., Chiang, Y.C.: Data-driven approach to pronunciation error detection for computer assisted language teaching. In: Spector, J.M., Sampson, D.G., Okamoto, T., Kinshuk, C.S.A., Ueno, M., Kashihara, A. (eds.) ICALT, pp. 359–361. IEEE Computer Society (2007)
15. Strik, H., Truong, K.P., de Wet, F., Cucchiarini, C.: Comparing classifiers for pronunciation error detection. In: INTERSPEECH, pp. 1837–1840. ISCA (2007)
16. Witt, S.M.: Use of speech recognition in computer-assisted language learning (1999)
17. Witt, S.M., Young, S.J.: Phone-level pronunciation scoring and assessment for interactive language learning. Speech Communication 30(2-3), 95–108 (2000)
18. Kanters, S., Cucchiarini, C., Strik, H.: The goodness of pronunciation algorithm: a detailed performance study. In: SLaTE 2009 (2009)
19. Wang, Y.B., Shan Lee, L.: Improved approaches of modeling and detecting error patterns with empirical analysis for computer-aided pronunciation training. In: 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5049–5052 (2012)

Quantification of Linear and Non-linear Acoustic Analysis Applied to Voice Pathology Detection

Daria Panek, Andrzej Skalski, and Janusz Gajda

AGH University of Science and Technology,
Department of Measurement and Electronics,
Al. Mickiewicza 30, 30-059 Krakow
{dpanek,skalski,jgajda}@agh.edu.pl

Abstract. Present development of digital registration and methods of recorded voice processing are useful in detection of most pathologies and diseases of a human vocal tract. The recognition of the voice condition requires the creation of a model which is comprised of different acoustic parameters of speech signal. In this study a vector consisting of 31 parameters for analysing the speech signal was created. The speech parameters were extracted from time, frequency and cepstral domains. Using Principal Components Analysis the number of the parameters was reduced to 17. In order to validate the detection of the pathological voice signal, a tenfold cross-validation and confusion matrix were used. The goal and novelty of this work was the analysis of applicability of the parameters selectively used to assess the pathology.

Keywords: acoustics analysis, cepstral analysis, pathology detection, dysphonia, principal component analysis, cross validation.

1 Introduction

Currently, European standards emphasize the need for a comprehensive assessment of voice disorders with regard to objective methods. The value of acoustic analysis is increasingly appreciated as a diagnostic test for non-invasive and objective examination. In the broad sense of laryngeal and phoniatric diagnosis of voice and speech disorders acoustic analysis provides supporting and complementary studies. The value of these studies increased significantly throughout the last years with the introduction of high-speed digital voice analyzers. The detection accuracy of voice and speech disorders located in the larynx and in voice channel increases.

Acoustical analysis allows us to make physical description of the waveforms generated and emitted by the organ of the human voice and correlates well with the phoniatric state of proper and pathological voice. There are two methods to conduct voice analysis: classic, based on subjective assessment of voice examination and modern, based on objective acoustic analysis, spectrographic images, sonographic or time recording speech signal. The physical characteristics of the voice are determined to use the latest digital technology in the acoustic

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_33, © Springer International Publishing Switzerland 2014

analysis and a detailed statistical analysis of the results. Fidelity recording and processing the digital audio signal, which is a stochastic process, promotes the development of a growing number of measurable characteristics (characteristic for the different subjective characteristics) of human voices and allows good accuracy, objective assessment of discrete, unobtrusive overflow method, voice and speech disorders [1].

Currently, the best measurement methodology models, algorithms and different approaches for classification that could discriminate between normal and pathological voices are still being sought [2]. The performance of these systems is not perfect, but they are useful as an additional source of information for other laryngoscopial examinations [3,5].

In this work, the articulation in speech and its pathological deformation was examined. This includes tools and techniques used to detect deformations in the voice signal vocalised by an ill person and a healthy one. The aim was to select and describe those signal parameters that contain the most valuable information and show the highest sensitivity to speech deformations.

2 Material

The Saarbruecken Voice Database has been published online by Institute of Phonetics of the University of the Saarland [4]. It is a collection of voice recordings collected from more than 2000 people. Each of the recording sessions contains recordings of the vowels /a/, /i/, /u/ produced at normal, high, low, low-high-low pitch.

The length of the recordings with sustained vowels amounts to 1 – 4 seconds. All the recordings are sampled at 50 000 Hz with a resolution of 16-bit. The database contains 71 different well defined pathologies. In our work we used the recordings of the vowel /a/ of 850 women, of which 425 were healthy and 425 suffered from different voice disorders (167 suffered from hyperfunctional dysphonia,139 had vocal cord paresis, 119 suffered from other pathologies listed in the database) and 510 men, of which 255 were healthy and 255 were diagnosed with different pathologies, of which 46 men suffered from hyperfunctional dysphonia, 74 suffered from vocal cord paresis, 83 experienced laryngitis of which some also had leukoplakia. The rest of the men who underwent this examination suffered from other pathologies listed in the database. Recordings that were missing or damaged were excluded from the dataset. Only for women the analysis of the vowel /a/ was extended to include all intonations – low and high pitch. Because of the intrinsic differences in voice behaviour between men and women (and because the number of male and female speakers was not equal in all groups), parameters were statistically analysed for males and females separately.

3 Method

This study was carried out to assess the suitability of several methods for mapping speech signals in diagnostics of pathological speech. Each method was used

with reference to both correct and pathological speech samples. Firstly, the focus of examination was put on preliminary transformation of speech waveforms into a set of parameters whose values represented a basis for a diagnosis of the patient's disease.

The attention was drawn to the fact that the acoustic signal processing for a set of features whose values are so called parameters, is the basis for a description of the object's state in terms of diagnosis. Registration itself and its preliminary signal processing does not make it fully useful to the process of identifying and assessing changes in deformation and pathology. Therefore, it becomes necessary to develop and describe recorded phonetic tests using a set of parameters, which then, sorted out in the corresponding structure - a feature vector will be used to develop models of speech deformation. Such models can be the foundation of the recognition process, assessment of pathological changes or rehabilitation process. Analysis of the speech signals (Fig. 1) was performed with 31 parameters: root

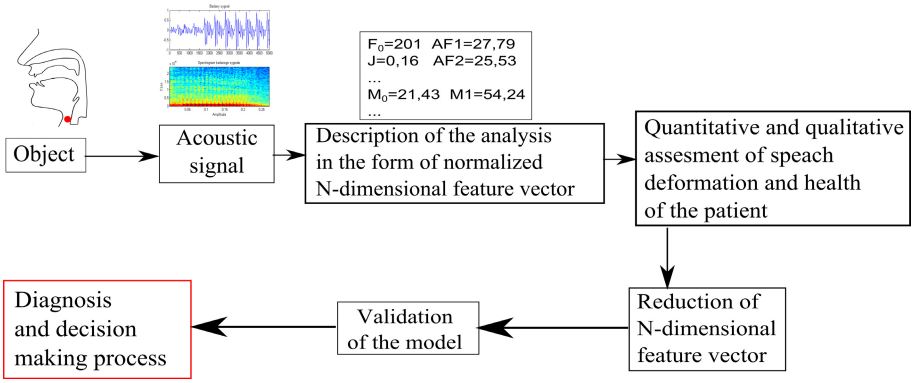


Fig. 1. Block diagram of conducted acoustic analysis

mean square value (RMS), fundamental frequency, jitter (J) and shimmer (S) coefficients, mean value of the signal, energy, average power, zero-order, first-order, second-order and third order moments, kurtosis, power factor, 1st, 2nd, 3rd formant's amplitude, 1st, 2nd, 3rd formant's frequency, maximum and minimum value of the signal and 10 mel-frequency cepstrum coefficients (MFCC). MFCC coefficients are widely used in speech recognition because they reflect well the auditory sensations by enhancing the audible frequency and are less sensitive to noise [6,7,8,9]. MFCC are designed to reflect the natural response of the auditory system to stimulation of the auditory speech sound.

A root-mean-square (RMS) value of the signal is used to estimate its loudness. In the first step of the analysis we proceeded with the normalisation where the aim was to bring the average amplitude to a target level. The normalisation was done for all signals so that their RMS value of each signal was the same.

Cepstral analysis was performed using the Fourier transformation of the sound to calculate the fundamental frequency. In order to facilitate the extraction of

spectral components before converting from time domain to the frequency domain, the signal was subjected to a windowing operation using a Hamming window. In order to improve the performance of this calculation and further purification of the spectrum the non-maximum elements were removed. The location of the maximum of the spectrum corresponds to the fundamental frequency [10]. Therefore, knowing the position of the maximum, we could determine the fundamental frequency of the analysed sample using the formula:

$$f_0 = \frac{l_s z}{wm} \tag{1}$$

where l_s – is the number of elements of the spectrum, z – is the audio sample rate,w – is the width of analysed window (number of samples), m – is the position of maximum cepstrum. The jitter coefficients (J) was calculated as an average deviation of the fundamental frequency from frequency f_0 in consecutive cycles. The jitter coefficient is presented in equation 2:

$$J = \frac{\sqrt{\frac{1}{2N-1}\sum_{i=1}^{2N-1}(f_i - f_{i-1})^2}}{\frac{1}{2N}\sum_{i=1}^{2N-1} f_i} 100\% \tag{2}$$

where f_i - the frequency in the i-th cycle. The shimmer coefficient (S) defines the variations of the fundamental tone amplitude from the average amplitude A_i in consecutive cycles, and is described by:

$$S = \frac{\sqrt{\frac{1}{2N-1}\sum_{i=1}^{2N-1}(A_i - A_{i-1})^2}}{\frac{1}{2N}\sum_{i=1}^{2N-1} A_i} 100\% \tag{3}$$

where the average amplitude is defined as:

$$\overline{A_0} = \frac{1}{N}\sum_{i=1}^{N} A_i \tag{4}$$

By integrating the square of the signal the variable E_x is formed, which is a measure of the energy carried by the signal x:

$$E_x = \int_{t_1}^{t_2} x^2(t)dt \tag{5}$$

where t_1 and t_2 are the beginning and the end of the sustained signal. The average power of the signal is defined as the average time derivative of the energy.

$$P_x = \frac{\int_{t_1}^{t_2}(x^2(t))^2 dt}{t_2 - t_1} \tag{6}$$

A large number of details, which are carried by a spectral analysis make it difficult to interpret and recognise the relevant information contained in the signal. Therefore, we determined from the frequency spectrum those features

that are useful in the analysis. Having defined signal time-frequency domain $G(t, f)$ parameters describing its shape can be determined. The first parameter describing the shape of the spectrum is the zero-order spectral moment which takes the form of:

$$M_0(t) = \sum_{i=0}^{\infty} G(t, f_i) \tag{7}$$

The zero-order moment is used to normalise the higher-order moments. The first-order moment can be interpreted as the centre of gravity of the spectrum (frequency-weighted average). This moment is used in the formulas for the calculation of higher-order central moments.

$$M_1(t) = \frac{\sum_{i=0}^{\infty} G(t, f_i) f_i}{M_0(t)} \tag{8}$$

Second-order moment is interpreted as the square of the spectrum width.

$$M_2(t) = \frac{\sum_{i=0}^{\infty} G(t, f_i)[f_i - M_1(t)]^2}{M_0(t)} \tag{9}$$

The third-order moment can be interpreted as the asymmetry of the spectrum – skewness. Standardised higher-order spectral moments are less suitable because they are correlated with each other.

$$M_3(t) = \frac{\sum_{i=0}^{\infty} G(t, f_i)[f_i - M_1(t)]^3}{M_0(t)} \tag{10}$$

Another parameter in the analysis was kurtosis, defined as flattening the spectrum measurement:

$$kurtosis = \frac{M_4(t)}{M_2(t)^2}. \tag{11}$$

Further calculation was done to get power factor characterised as the ratio of the relative power of the signal in the desired frequency f_0 wide-band $<f_d, f_g>$ to signal power across the bandwidth $<f_0, f_\infty>$, and is shown in eq. 12.

$$W_m(t) = \frac{\sum_{t=t_b}^{t_g} \sum_{f=f_{d2}}^{f_{g2}} G(t, f_i)}{\sum_{t=t_b}^{t_g} \sum_{f=f_{d1}}^{f_{g1}} G(t, f_i)} \tag{12}$$

where f_{g1}, f_{d1}- are the lower and upper frequency of the power wide-band, f_{g2}, f_{d2} – are the upper and lower frequency range of the selected frequency wide-band, t_b, t_g – are the beginning and the end of the recorded voice sample. The band selection followed the formants structure of the vowels calculated using the set of vowels recorded from the patients involved in this examination. The power coefficient was calculated only for the 1st power signal. The next computed parameters were the formants. Formant level can be defined from the envelope which can be drawn to enclose smoothly the harmonics within the

spectral maximum, i.e. the sound pressure level in dB of the envelope peak [5,9]. The frequency of the formant is measured as the frequency position of the envelope maximum (sequentially designated as $F1$, $F2$, $F3$... etc). Currently, a classical spectral analysis of voice signals methods is often supplemented with methods such as linear predictive analysis, wavelet analysis or homomorphic in the field of so-called cepstrum [6,9,11,12]. Cepstrum is determined as the inverse Fourier Transform of the logarithm of the signal spectrum giving a better picture of the structure of harmonic signal and allowing for the separation of existing noise in the transformed signal of the harmonic components in the same signal [7,13]. Many authors emphasize the importance of the cepstral factors in the diagnostic evaluation of pathological changes in the glottis [14,15]. In the present work the process of determination of the cepstral coefficients was extended to a so-called Melow filtration, which consists of an additional non-linear frequency scale signal spectrum transformation, yielding MFCC coefficients (Mel Frequency Cepstral Coefficients). Obtaining the MFCC coefficients required taking a few steps. Firstly the signal was divided into frames, the amplitude was obtained from each frame and the log was taken of these spectrums. Afterwards the results were converted into the Mel scale and the Discrete Cosine Transform (DCT) was applied. Mel filtering was applied using triangular band pass filters corresponding to the Mel scale. The number of filters was set to 10.

4 Feature Selection

All the parameters discussed above had different results. Statistically, much of this data is redundant and it is therefore useful to identify the method that can extract most significant information from the collected data. While sorting features according to their discriminant capacity it is necessary to get a stable and consistent result, which is reflected in the overall performance of the system [18]. For feature selection we used a method called Principal Component Analysis (PCA). Essentially, PCA transforms data orthonormally so that the variance of the data remains constant, but is concentrated in the lower dimensions [8]. The matrix of data being transformed consisted of all calculated parameters for every voice sample in this examination. Thus, there was a single matrix of data with all parameters. The covariance matrix of the data was created. PCA for the data set was calculated to determine the eigenvectors, which are necessary for PCA [16]. As a result, a set of principal components was obtained, with the variance order from the highest to the lowest, which means that the most important data was extracted with minimum disruption to the original data collection. Once the data had been reduced and the principal components values extracted, we took 17 out of 31 parameters that covered up to 90% variance of the initial parameters.

5 Validation

The next step in the implementation of this study was the evaluation of the operator classification quality using cross-validation [19,20]. The operator

classification represents the result of the analysis of individual principal components taken into further analysis. For validation, we used a tenfold cross validation in which we randomly selected roughly 90% of the data, which represented the set of learners and then we used the remaining 10% of cases to test for classification. We iterated this procedure 10 times. To classify the data a K-means function was used. Cross validation was done separately using recordings of vowel /a/ at normal pitch for female and male recordings with healthy and pathology state and for female /a/ vowel at high and low pitch. As a result of the cross-validation a confusion matrix was constructed. A confusion matrix is a tool used to analyse the operation of the classifier, here parameters (Table 1).

Table 1. An example of confusion matrix used in the analysis

		Results from classification		
		healthy	**pathology**	
Diagnosed	**healthy**	True Positive (TP)	False Positive (FP)	**Precision Positive predictive value** = TP/(TP+FP)
	pathology	False Negative (FN)	True Negative (TN)	**Precision Negative predictive value** = TN/(FN+TN)
		Sensitivity =TP/(TP+FN)	**Specificity** =TN/(TN+FP)	

In order to make a qualitative assessment, the accuracy, precision, sensitivity and specificity were calculated (Table 2). The accuracy of the classifier determines what percentage of parameters from the test set were correctly assigned to their respective classes and obtained at the stage of testing. The precision meant the ratio of the number of cases actually correct, which have been classified by the system as correct to all classified by the system as correct. The sensitivity relates to the test's ability to indentify positive results, whereas the specificity relates to identifying negative results.

Due to unsatisfactory results for female data set, the test was repeated for the vowel /a/ at a high and low pitch. The results are shown in Table 3.

Comparing the calculated characteristics in the confusion matrix for the vowel /a/ using PCA with different intonations, we came to the conclusion that the analysis of the vowel /a/ at a high pitch gave the most accurate indications of the patient's healthy condition and the pathological one – the accuracy of 77.5%,

Table 2. Results obtained from the confusion matrix using 31 primary parameters and PCA for healthy(H) and pathological(P) voice samples

vowel /a/ *normal pitch*	**31 param**		**PCA**		**31 param**		**PCA**	
	female				***male***			
	H	***P***	***H***	***P***	***H***	***P***	***H***	***P***
accuracy [%]	81.4		74.8		100.0		100.0	
precision [%]	74.6	88.0	80.4	69.1	100.0	100.0	100.0	100.0
sensitivity [%]	86.0	–	72.2	–	100.0	–	100.0	–
specificity [%]	–	78.0	–	78.0	–	100.0	–	100.0

Table 3. Results obtained from the confusion matrix for healthy and pathological female voices with different intonations

female ***vowel /a/***	*high pitch*		*low pitch*	
	healthy	***pathology***	***healthy***	***pathology***
accuracy [%]	77.5		73.3	
precision [%]	83.7	71.3	79.0	67.5
sensitivity [%]	73.4	–	72.1	–
specificity [%]	–	79.9	–	78.6

precision of almost 84% for the healthy females and 71 pointing the pathology ones. The same dependence showed the sensitivity and specificity that were around 1% higher for a high pitch than normal one. Results from the vowel /a/ with a low pitch showed slightly lower results than a normal pitch, expect for the specificity that was slightly higher, but at the same time it was still lower than the one associated with a high pitch. There was no need to repeat the test for male recordings due to 100% correct classification Results based on 31 original parameters for female recordings show higher accuracy and sensitivity than PCA. While analysing PCA the precision for healthy women was higher than the one for all original parameters, pathology precision at the same time was still lower. The specificity for both methods for normal pitch was the same.

6 Conclusion

The open and free database available online has been used in the context of pathology detection. The substantial amount and various types of recordings included in this database made it possible to conduct different and interesting tests. Based on the conducted studies it was shown that the use of the calculated parameters and their subsequent reduction significantly differentiated between acoustic characteristics of the pathological speech and those of the healthy one. An integrated acoustical analysis of deformed pathological speech was discussed in this paper. The analysis employed among groups of patients showed that

speech pathology caused by various laryngeal diseases can be computed using acoustical methods. In order to expedite and simplify the calculations the Principal Components Analysis was conducted and led to obtaining 17 out of 31 parameters. The final analysis included just these factors. Analysis of the 17 parameters analysed for women showed lower accuracy and sensitivity than 31 original parameters at normal pitch when determining whether the patient was healthy or ill. The precision of pathology detection was higher when PCA was used. The specificity did not change.

According to the statistical accuracy of the pathological voice diagnosis obtained satisfactory results showing 100% compatibility classification obtained for the male voices analysing original vector of parameters and PCA, whereas for the female ones it proved more complicated.

The best results were achieved with high intonation for the female recordings giving an approximately 3% higher result of accuracy than normal intonation and almost 4% higher than low intonation. Other calculations like precision, specificity and sensitivity showed an upward trend. The reason might be that for the same vowel spoken at different pitches the relationship between the second harmonic and the first formant can change, causing the amplitude of this harmonic to be artificially amplified or attenuated [21]. The results show that the techniques discussed here could be used for detecting pathological voices.

Acknowledgement. This work was funded by the Ministry of Science and Higher Education in Poland under the Diamond Grant program, decision number 0136/DIA/2013/42 (AGH 68.68.120.364).

References

1. Deliyski, D.D.: Multi-dimensional acoustic analysis of spasmodic dysphonia. In: Proc. in the ASHA Convention, Atlanta (1991)
2. Gogh, C.D.L., Festen, J.M., Verdonck-de Leeuw, I.M., Parker, A.J.: Acoustical analysis of tracheoesophageal voice. Speech Communication 47, 160–168 (2005) ISSN 0167-6393
3. Martinez, D., Lleida, E., Ortega, A., Miguel, A., Villalba, J.: Voice Pathology Detection on the Saarbruecken Voice Database with Calibration and Fusion of Scores Using MultiFocal Toolkit. Advances in Speech and Language Technologies for Iberian Languages Commu-nications in Computer and Information Science 328, 99–109 (2012)
4. Barry, W.J., Putzer, M.: Saarbruecken Voice Database. Institute of Phonetics, University of Saarland, `http://www.stimmdatenbank.coli.uni-saarland.de/`
5. Fant, G.: Acoustic Theory of Speech Production With Calculations based on X-Ray Studies of Russian Articulations, Mouton, The Hague (1970) ISBN: 9027916004
6. Maciel, C.D., Pereira, J.: Identifying healthy and pathologically affected voice signals. IEEE Signal Processing Magazine 27(1), 120–123 (2010)
7. Arroyave, J.R.O.A., Bonilla, J.F.V., Trejos, E.D.: Acoustic Analysis and Non Linear Dynamics Applied to Voice Pathology Detection: A Review. Recent Patents on Signal Processing (2012)

8. Loughran, R., Walker, J., O'Neill, M., O'Farrell, M.: The Use of Mel-frequency Cepstral Coefficients in Musical Instrument Identification. Routes/Roots, Michigan (2008)
9. Engel, Z.W., Klaczynski, M., Wszolek, W.: A Vibroacoustic Model of Selected Human Larynx Diseases. International Journal of Occupational Safety and Ergonomics (JOSE) 13(4), 367–379 (2007)
10. Wuyts, F.L., De Bodt, M.S., Molenberghs, G., Remacle, M., Heylen, L., Millet, B., Van Lierde, K., Jan, R., Van de Heyning, P.H.: The Dysphonia Severity Index: An Objective Measure of Vocal Quality Based on a Multiparameter Approach. Journal of Speech, Language and Hearing Research 43, 796–809 (2000) ISSN 1092-4388
11. Osowski, S.: Sieci neuronowe w ujęciu algorytmicznym, WNT, Warszawa (1996)
12. Tadeusiewicz, R., Izworski, A., Wszolek, W.: Pathological speech evaluation using the artificial intelligence methods. Med. Biol. Eng. Comput. (2007)
13. Noll, A.: Short-term spectrum and 'cepstrum' techniques for vocal pitch detection. J. Acoust. Soc. Am. 41, 293–300 (1964)
14. Awan, S.N., Giovinco, A., Owens, J.: Effects of vocal intensity and vowel type on cepstral analysis of voice. In: Presented at the 39th Annual Symposium: Care of the Professional Voice, Philadelphia (2010)
15. Mehta, D.D., Deliyski, D.D., Zeitels, S.M., Quatieri, T.F., Hillman, E.R.: Voice Production Mechanisms Following Phonosurgical Treatment of Early Glottic Cancer. Ann. Otol. Rhinol. Laryngol. 119(1), 1–9 (2010)
16. Bishop, C.M.: Pattern Recognition and Machine Learning, pp. 559–599. Springer Science, Singapore (2006)
17. Methods based on Principal Components Analysis and the concept of Eigenface, Metody oparte na Analizie Glownych Skladowych i koncepcji Eigenface, http://icsolutions.pl/
18. Orozco-Arroyave, J.R., Murillo-Rendon, S., Alvares-Meza, A.M., Arias-Londono, J.D., Delgado-Trejos, E., Vargas-Bonilla, J.F., Castellanos-Domingues, C.G.: Automatic Selection of Acoustic and Non-linear Dynamic Features in Voice Signals for Hypernasality Detection. In: Interspeech, pp. 529–532 (2011)
19. Refaeilzadeh, P., Tang, L., Liu, H.: Cross Validation, Encyclopedia of Database Systems (EDBS), p. 6. Arizona State University, Springer (2009)
20. Delgado-Trejos, E., Sepulveda-Sepulveda, F.A., Castellanos-Domnguez, G.: Robustness Improvement of Hypernasal Speech Detection by Acoustic Analysis and the Rademacher Complexity Model. In: Advances in Biomed. Research, pp. 159–162
21. Epstein, M.A., Payri, B.G.: The effects of vowel quality and pitch on spectral and glottal flow measurements of the voice source, Lecture, University of California, Los Angeles.

Bronchopulmonary Dysplasia Prediction Using Support Vector Machine and Logit Regression

Marcin Ochab and Wiesław Wajs

AGH University of Science and Technology,
30 Mickiewicza 30-059 Kraków, Poland
marcin.ochab@labor.it.pl, wwa@agh.edu.pl

Abstract. The paper presents BPD (Bronchopulmonary Dysplasia) prediction for extremely premature infants after their first week of life. SVM (Support Vector Machine) and LR (Logit Regression) are used as classifiers. Data was collected thanks to the Neonatal Intensive Care Unit of The Department of Pediatrics at Jagiellonian University Medical College and includes 109 patients with birth weight less than or equal to 1500g. Fourteen different risk factor parameters were considered and all 2^{14} combinations were analyzed. Classifier based on six feature LR model provides accuracy up to 82%, while SVM one turns out to be generally much worse, providing in best case scenario 80% of accuracy. In addition, the article discusses the influence of the model parameters selection on prediction quality.

Keywords: bronchopulmonary dysplasia, support vector machine, logit regression, prediction, prematurity, low-birth-weight infant.

1 Introduction

Bronchopulmonary dysplasia (BPD) is a chronic pulmonary mordibity affecting premature infants [1,2]. It is most common among children with low birth weights and the ones who received prolonged mechanical ventilation to treat respiratory distress syndrome [3,4]. It affects nearly a third of infants with birth weight lower than 1000g [5]. Due to the fact that the disease is poorly understood, many projects are focused on identifying its factors of risk. Since it is diagnosed in a 28th day of life [6], it would be highly preferred to predict such a result after the end of the first week, which would make an early prevention of the disease possible [7]. Therefore an intensive work has been done to define a classifier, which based on static parameters gathered after birth and dynamic ones collected during the first week of life would be able to predict the diagnosis. Although several prediction models of BPD [8,9,10,11,12,13,14,15,16,17] used in research have been reported, none of them could be used in common clinical practice due to the variety of reasons.

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_34, © Springer International Publishing Switzerland 2014

2 Related Works

As mentioned before there are numerous works related to BPD, its risk factors and prediction [18,19,20,21]. The most popular one is the analysis of static data whose main features are gestational age and birth weight. The other factors considered are admission of surfactant, presence of patent ductus arteriosis (PDA) or respiratory support. Apart from that dynamical data (which is much harder to obtain) is analyzed in more sophisticated models. Most of such parameters are: arterial blood gas variables like fraction of inspired oxygen (FiO_2) or alveolar-arterial ratio (AA) [22] (which is respiratory distress degree measure); blood gas levels like oxygen saturation of arterial hemoglobin (SpO_2) and its deviation, mean value etc. [23] or even time series analysis [24]; heart beat and its derivatives.

$$AA = \frac{pO_2}{p_{ATM} \cdot FiO_2 - pCO_2}, \tag{1}$$

where pO_2 – oxygen partial pressure, p_{ATM} – atmospheric pressure, pCO_2 – carbon dioxide partial pressure, FiO_2 – fraction of inspired oxygen.

Some of the papers introduced race and ethnicity or sex as factors which seem to be promising but require a very big set of data. It should be indicated that the vast majority of studies uses logit regression (LR) for prediction. Best LR models gain about 73% to 82% of accuracy. Many of authors mention use of support vector machine(SVM) [25] in future works however it is difficult to find them. For this reason it is very interesting to compare LR BPD prediction with SVM one.

3 General Ideas of Used Methods

3.1 Logistic Regression

Probability of the dependent variable equalling a BPD positive diagnosis ($Y = 1$), on condition that explanatory variables (features of specific case) equals $X = (x_1, x_2, .., x_n)$ we define as:

$$P(Y = 1|x_1, x_2, .., x_n) = \frac{e^{a_0 + \sum_{i=1}^{n} a_i x_i}}{1} + e^{a_0 + \sum_{i=1}^{n} a_i x_i}, \tag{2}$$

where x_i – explanatory variables (feature values), a_i – regression coefficients, n – number of features.

Unlike linear regression when we assume normal distribution of the independent variables and because explanatory variables variance are not equal we can not use the method of least squares to obtain regression coefficients. Thus they are usually calculated using maximum likelihood estimation, maximizing likelihood function (L) or minimizing its negative logarithm using learning data:

$$L = \prod_{Y_k=1} p_k \prod_{Y_k=0} (1 - p_k), \tag{3}$$

$$ln(L) = \sum_{k=1}^{m} [Y_k \cdot ln(p_k) + (1 - Y_k) \cdot ln(1 - p_k)], \tag{4}$$

where k – observation number (learning case), Y_k – diagnosis for case k, p_k – expected probability of observation k, m – number of observations.

Having a_i regression coefficients, we can easily predict BPD positive diagnosis probability of case X using Eq. 2.

3.2 Support Vector Machine

We define the learning data D divided in two classes c as:

$$D = \{(X_k, c_k) | X_k \in R^n, c_k \in \{1, -1\}\}_{k=1}^{m} \tag{5}$$

We are looking for hyperplane

$$W \bullet X + b = 0 \tag{6}$$

which separates classes and provides maximum margin as on Fig. 1, which is the same as the problem of minimizing L:

$$L(W) = \frac{||W||^2}{2} + c \cdot \sum_{k=1}^{m} \varepsilon_k, \tag{7}$$

with conditions:

$$f(X_k) = \begin{cases} 1, & \text{if} \quad W \bullet X_k + b \geq 1 - \varepsilon, \\ -1, & \text{if} \quad W \bullet X_k + b \leq -1 + \varepsilon, \end{cases} \tag{8}$$

where ε – slack variable, c – penalty parameter for each point wrongly classified.

Thanks to the kernel functions for non linear separable problems, we can transform original data from n dimensional space to p dimensional ($p > n$, as on Fig. 2) in which there is much higher likelihood that they will be linear separable.

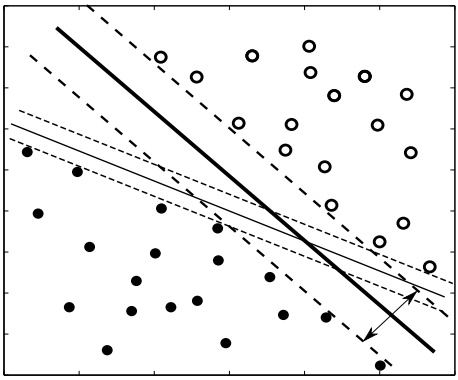
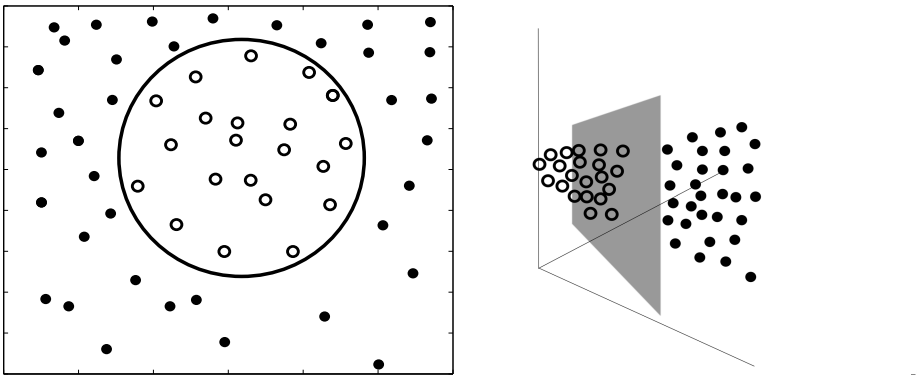


Fig. 1. Maximizing margin in SVM method



(a) data in R^2 space (b) data transformed to R^3 space

Fig. 2. Making data linear separable with dimensional transformation

4 Data and Methods

Data was collected thanks to the Neonatal Intensive Care Unit of The Department of Pediatrics at Jagiellonian University Medical College using our own software. It includes 109 patients born prematurely with birth weight less than or equal to 1500g admitted no later than on the second day of life. For 46 of them BPD have been diagnosed after fourth week of life. To build a suitable model 14 different features mentioned in literature were considered:

- Binary such as: presence of patent ductus arteriosus (PDA) [26] , use of a respirator ($RESPIMV$) during the first week of life, administration of surfactant ($SURFACT$) [27] in the same period.
- Real–Valued such as: birth weight ($BWEIGHT$) (550–1500g), gestational age ($GAGE$) (22–34 weeks), alveolar-arterial ratio (AA) (0.05–1) measured during patient admission, a percentage of the time during first week for which the oxygen saturation of hemoglobin was less than 85% ($LOW85$) (0.03%–12.45%) or higher than 94% ($HIGH94$) (14.56%–99.02%), average number of heartbeats per minute ($BPMMEAN$) [3](124.69–161.42), mean and standard deviation of oxygen saturation ($SPO2MEAN$, $SPO2DEV$) (accordingly 89.89%–98.99% and 1.19–7.98) and their trends (first day to first week ratio:$BPMMEAN\_TR$, $SPO2MEAN\_TR$, $SPO2DEV\_TR$) (accordingly 0.8–1.18, 0.96–1.07 and 0.51–2.36).

We performed calculations using LR and SVM algorithm in Matlab environment. It should be mentioned that all of the below results were exactly the same with or without parameter values normalization — we assume Matlab libraries do it automatically. Based on a few arbitrary chosen and tested models it has been found that quite good results for SVM prediction are obtained using Radial Basis Function (RBF) as a kernel function with $\sigma = 20$, and quadratic programming library from Optimization Toolbox as method of finding the separating

hyperplane (use of sequential minimal optimization and least-squares method gives less accurate results). We have done an initial search for models with these parameters. It is noteworthy that results obtained with simple polynomial functions are, unexpectedly, almost equally satisfactory; however, no more research has been done on that matter. To compare results the cross-validation method was used (each patient was treated as a test sample while all other data was learning set). Accuracy (ACC) defined as below was considered as preliminary result measure. In the last experiment stage for a number of the best models their sensitivity(TPR) and specificity(SPC) was also obtained:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}, \tag{9}$$

$$TPR = \frac{TP}{TP + FN}, \tag{10}$$

$$SPC = \frac{TN}{TN + FP}, \tag{11}$$

where TP – True Positives, FP – False Positives, FN – False Negatives, TN – True Negatives.

5 Prediction

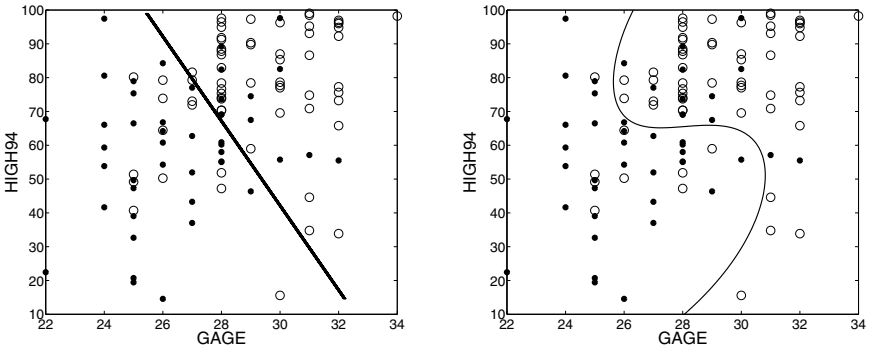
We reviewed all of the 2^{14} possible combinations of models containing 2 to 14 parameters. For the SVM, calculations were performed with mentioned $\sigma = 20$. The accuracy of that results stays within the range 40.37% to 78.90% for SVM and 57.63%–82.79% for LR . For further research we have chosen about 120 best SVM models whose accuracy was accordingly 77.06%–78.90% (for the LR case there was only a selection of the best results required since there was no algorithm parameters to optimize). We preformed an optimization of σ parameter based on review values 10,15,20,25,30. To gain a better comparison between models for each of the best 120 SVM ones and σ values a method similar to Jacknife [28] was used. Calculations were repeated 30 times, each time randomly excluding 30 samples of data. This way deviation and mean value of accuracy, sensitivity and specificity were obtained, which gives an estimate on the models 'sensitivity' to data structure (it might be important when calculating on such a little data set as 109 patients). The test was repeated once again excluding the data of only 10 random patients. For LR algorithm Jackknife was also used same way. As a sample two parameters model result was presented on Fig. 3.

6 Results

The most essential results are presented in Table 1 . To compare, in each presented model ACC, TPR, SPC were obtained using both methods: LR and

Table 1. Models and methods comparison

Params count	BWEIGHT	GAGE	RESPIMV	AA	PDA	SURFACT	SPO2MEAN	SPO2DEV	LOW85	HIGH94	BPMMEAN	BPMMEAN_TR	SPO2DEV_TR	SPO2MEAN_TR	Items excluded with Jacknife	ACC	TPR	SPC	Method	Model comment
																mean value				
															10	82.79%	84.20%	81.73%	LR	best ACC
															30	80.85%	82.72%	79.36%	LR	
															10	70.01%	64.56%	74.01%	SVM σ=30	
6		•		•	•			•			•	•			30	69,00%	61.91%	74.00%	SVM σ=30	
															10	82.74%	85.55%	80.82%	LR	
															30	82.52%	84.87%	80.81%	LR	
															10	67.11%	62.06%	70.73%	SVM σ=30	
6		•			•			•			•	•	•		30	66.71%	58.43%	72.41%	SVM σ=30	
															10	82.67%	87.50%	79.09%	LR	
															30	81.12%	85.78%	77.62%	LR	
															10	70.48%	65.77%	73.83%	SVM σ=30	
6	•	•			•	•			•		•				30	69.90%	64.58%	73.69%	SVM σ=30	
															10	82.59%	84.65%	81.07%	LR	
															30	81.46%	81.53%	81.30%	LR	
															10	72.10%	64.12%	77,88%	SVM σ=20	
5		•			•				•		•	•			30	71.40%	63.94%	76.69%	SVM σ=30	
															10	82.01%	84.38%	80.34%	LR	
															30	80.64%	80.08%	81.02%	LR	
															10	75.38%	71.32%	78.31%	SVM σ=15	best LR ACC
4		•			•				•		•				30	74.68%	69.58%	78.16%	SVM σ=25	4 features
															10	81.29%	88.70%	75.90%	LR	
															30	81.05%	88.52%	75.34%	LR	best LR&SVM
															10	80.15%	78.73%	81.17%	SVM σ=30	ACC 3 params,
3		•				•			•						30	79.19%	79.84%	78.71%	SVM σ=15	best SVM ACC
															30	81.06%	84.39%	78.62%	LR	
															10	80.30%	83.48%	78.01%	LR	
															30	71.38%	52.67%	84.98%	SVM σ=20	best LR ACC
2		•				•									10	71.02%	51.68%	84.90%	SVM σ=30	2 features
															30	79.69%	78.67%	80.47%	LR	
															10	79.63%	78.11%	80.76%	LR	
															10	79.39%	76.24%	81.67%	SVM σ=25	best SVM ACC
4		•			•			•						•	30	77.50%	72.36%	81.15%	SVM σ=30	4 features



(a) Logistic regression (b) Support Vector Machine

Fig. 3. Prediction result sample for two parameters ($GAGE$, $HIGH94$)

SVM, with different Jacknife parameters. First three 6-parameter models and one 5-feature have the best accuracy which is more than 82% - the highest was 82.79% with 84.20% sensitivity and 81.73% specificity. Next we present models with the highest ACC for each method which have 4 , 3 and 2 parameters. It might be surprising that 4-parameter LR model gained as much as 82.01%. Most TPR and SPC results compared to ACC fall within the five percent deviation, which is quite satisfactory. In general we observed that the SVM models are less accurate than the LR ones. Moreover the higher number of parameters were considered the worse results we received. The best SVM model gained 80.15% accuracy with $TPR = 78.73\%$ and $SPC = 81.17\%$ for only three-parameter, 79.39% for four-feature and only 77.43% for five ones. What might also be interesting is that even two-parameter LR model gained 81.06% of accuracy with $TPR = 84.39\%$ and $SPC = 78.62\%$. We also observed that the more data we exclude from the test results the worse accuracy we achieve (generally results with Jacknife equal 10 are better) which is promising - it shows that overfitting does not occur and classifier is well generalizing to cases not known during learning.

7 Discussion

Both SVM and LR prediction models results can provide more than 80% of accuracy, but SVM is significantly worse especially since it gets confused in case of more than four features. Taking into account lower accuracy, much bigger complexity and time needed for calculations SVM seems to be an inadequate method for the considered problem. On the other hand it might be a good idea to recalculate results with LibSVM library instead of the Matlab's which has fewer algorithm tuning options.

We confirmed that one of the most important risk factors mentioned in literature [3,4,6,7,26,27,29] is the $GAGE$ and exists in almost all of the

models with acceptable accuracy. Most classifiers presented in the literature consist $BWEIGHT$ and $RESPIMV$ parameters. Unexpectedly the highest ACC $RESPIMV$-containing model has six features and gained 81.69% which is just on the 50th place in our ranking. However $BWEIGHT$-containing one was indeed on the third place with 82.67% ACC, but it was the only one in the first twenty, which is a group with ACC higher than 82%. Among the most frequently mentioned parameters there is FiO_2 which depends on the AA (Eq. 1) feature used in our work that is indeed present in the best model, but also only in three others of the best twenty. On the other hand PDA feature appeared in 15 out of the best 20 models which is consistent with published results [6,7,23,26]. The same situation occurs with $BPMMEAN$ parameter, which is the second most important factor after GAGE in our results - it is in all of the first 20 models. The only one parameter related to SpO_2 which is in more than half best 20 classifiers is LOW85.The average importance features are $SPO2DEV$, $SPO2DEV\_TR$ and $SURFACT$, while $SPO2MEAN\_TR$, $HIGH94$ parameters seem to have even less effect on the occurrence of diseases.

As a final conclusion we confirmed [30] that prediction of BPD after 7th day of life is possible with the accuracy higher than 82%, and not only using one particular prediction model but we found more than twenty different of such (unfortunately none of them using SVM).

References

1. Horbar, J.D., Badger, G.J., Carpenter, J.H., Fanaroff, A.A., Kilpatrick, S., LaCorte, M., Phibbs, R., Soll, R.F.: Members of the Vermont Oxford Network. Trends in mortality and morbidity for very low birth weight infants, 1991-1999. Pediatrics 110, 143–151 (2002)
2. Stoll, B.J., Hansen, N.I., Bell, E.F., Shankaran, S., Laptook, A.R., Walsh, M.C., Hale, E.C., Newman, N.S., Schibler, K., Carlo, W.A., et al.: Eunice Kennedy Shriver National Institute of Child Health and Human Development Neonatal Research Network. Neonatal outcomes of extremely pre- term infants from the NICHD Neonatal Research Network. Pediatrics 126, 443–456 (2010)
3. Jobe, A.H.: The new bronchopulmonary dysplasia. Current Opinion in Pediatrics 23(2), 167 (2011)
4. Groothuis, J.R., Makari, D.: Definition and outpatient management of the very low-birth-weight infant with bronchopulmonary dysplasia. Advances in Therapy 29(4), 297–311 (2012)
5. Walsh, M., et al.: Summary proceedings from the bronchopulmonary dysplasia group. Pediatrics 117(3), S52–S56 (2006)
6. Tapia, J.L., Agost, D., Alegria, A., Standen, J., Escobar, M., Grandi, C., et al.: Bronchopulmonary dysplasia: incidence, risk factors and resource utilization in a population of South American very low birth weight infants. Journal de Pediatria (Rio J) 82(1), 15–20 (2006)
7. Farstad, T., Bratlid, D., Medbø, S., Markestad, T.: Bronchopulmonary dysplasia–prevalence, severity and predictive factors in a national cohort of extremely premature infants. Acta Paediatrica 100(1), 53–58 (2011)
8. Ryan, S.W., Nycyk, J., Shaw, B.N.: Prediction of chronic neonatal lung disease on day 4 of life. Eur. J. Pediatr. 155, 668–671 (1996)

9. Subhedar, N.V., Hamdan, A.H., Ryan, S.W., Shaw, N.J.: Pulmonary artery pressure: early predictor of chronic lung disease in preterm infants. Arch. Dis. Child. Fetal. Neonatal Ed. 78, F20–F24 (1998)
10. Romagnoli, C., Zecca, E., Tortorolo, L., Vento, G., Tortorolo, G.: A scoring system to predict the evolution of respiratory distress syndrome into chronic lung disease in preterm infants. Intensive Care Med. 24, 476–480 (1998)
11. Toce, S.S., Farrell, P.M., Leavitt, L.A., Samuels, D.P., Edwards, D.K.: Clinical and roentgenographic scoring systems for assessing bronchopulmonary dysplasia. Am. J. Dis. Child. 138, 581–585 (1984)
12. Corcoran, J.D., Patterson, C.C., Thomas, P.S., Halliday, H.L.: Reduction in the risk of bronchopulmonary dysplasia from 1980-1990: results of a multivariate logistic regression analysis. Eur. J. Pediatr. 152, 677–681 (1993)
13. Noack, G., Mortensson, W., Robertson, B., Nilsson, R.: Correlations between radiological and cytological findings in early development of bronchopulmonary dysplasia. Eur. J. Pediatr. 152, 1024–1029 (1993)
14. Yuksel, B., Greenough, A., Karani, J.: Prediction of chronic lung disease from the chest radiograph appearance at seven days of age. Acta Paediatr. 82, 944–947 (1993)
15. Bhutani, V.K., Abbasi, S.: Relative likelihood of bronchopulmonary dysplasia based on pulmonary mechanics measured in preterm neonates during the first week of life. J. Pediatr. 120, 605–613 (1992)
16. Kim, Y.D., Kim, E.A., Kim, K.S., Pi, S.Y., Kang, W.: Scoring method for early prediction of neonatal chronic lung disease using modified respiratory parameters. J. Korean. Med. Sci. 20, 397–401 (2005)
17. Bhering, C.A., Mochdece, C.C., Moreira, M.E., Rocco, J.R., Sant'Anna, G.M.: Bronchopulmonary dysplasia prediction model for 7-day-old infants. J. Pediatr. (Rio J) 83, 163–170 (2007)
18. Rojas, M.A., Gonzalez, A., Bancalari, E., Claure, N., Poole, C., Silva-Neto, G.: Changing trends in the epidemiology and pathogenesis of neonatal chronic lung disease. J. Pediatr. 126, 605–610 (1995)
19. Marshall, D.D., Kotelchuck, M., Young, T.E., Bose, C.L., Kruyer, L., O'Shea, T.M.: Risk factors for chronic lung disease in the surfactant era: a North Carolina population-based study of very low birth weight infants. North Carolina Neonatologists Association. Pediatrics 104, 1345–1350 (1999)
20. Oh, W., Poindexter, B.B., Perritt, R., Lemons, J.A., Bauer, C.R., Ehrenkranz, R.A., Stoll, B.J., Poole, K., Wright, L.L.: Neonatal Research Network. Association between fluid intake and weight loss during the first ten days of life and risk of bronchopulmonary dysplasia in extremely low birth weight infants. J. Pediatr. 147, 786–790 (2005)
21. Ambalavanan, N., Van Meurs, K.P., Perritt, R., Carlo, W.A., Ehrenkranz, R.A., Stevenson, D.K., Lemons, J.A., Poole, W.K., Higgins, R.D.: NICHD Neo- natal Research Network, Bethesda, MD. Predictors of death or bronchopulmonary dysplasia in preterm infants with respiratory failure. J. Perinatol. 28, 420–426 (2008)
22. Gilbert, R., Keighley, J.: The arterial/alveolar oxygen tension ratio. An index of gas exchange applicable to varying inspired oxygen concentrations. Am. Rev. Respir. Dis. 109, 142–145 (1974)
23. Stoch, P.: Prediction of BronchoPulmonary Dysplasia in preterm neonates using statistical and artificial neural network tools (Thesis or Dissertation style) Ph.D. dissertation, AGH University of Science and Technology, Kraków, pp. 60–72 (2007) (in Polish)

24. Kuenzel, L.: Predicting and undestanding bronchopulmonary dysplasia in permature infants. Stanford Undergraduate Research Journal
25. Burges, C.J.C.: A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery, vol. 2, pp. 121–167. Kluwer Academic Publishers, Boston (1998)
26. Sosenko, I.R., Bancalari, E.: New Developments in the Pathogenesis and Prevention of Bronchopulmonary Dysplasia. The Newborn Lung: Neonatology Questions and Controversies: Expert Consult-Online and Print 217 (2012)
27. Cunha, G.S., Mezzacappa-Filho, F., Ribeiro, J.D.: Risk Factors for Bronchopulmonary Dysplasia in very Low Birth Weight Newborns Treated with Mechanical Ventilation in the First Week of Life. Journal of Tropical Pediatrics 51(6), 334–340 (2005)
28. Jones, H.L.: Jacknife estimation of functions of stratum means. Biometrika 61(2), 343–348 (1974)
29. Ali, Z., Schmidt, P., Dodd, J., Jeppesen, D.L.: Bronchopulmonary dysplasia: a review. Archives of Gynecology and Obstetrics, 1–9 (2013)
30. Laughon, M.M., et al.: Prediction of bronchopulmonary dysplasia by postnatal age in extremely premature infants. American Journal of Respiratory and Critical Care Medicine 183(12), 1715 (2011)

Part VI

Biomechanics

Overview of the Ocular Biomechanical Properties Measured by the Ocular Response Analyzer and the Corvis ST

Magdalena Jedzierowska, Robert Koprowski, and Zygmunt Wróbel

University of Silesia, Institute of Computer Science,
41-200 Sosnowiec, Będzinska 39, Poland
{mjedzierowska,koprow,wrobel}@us.edu.pl

Abstract. Two commercially used noncontact tonometers: the Ocular Response Analyzer (ORA) and the Corvis ST (CST) are presented. The devices measures biomechanical properties of the cornea that are essential for better understanding of the optical and the geometrical properties of the cornea. Characteristic parameters for both ORA and Corvis ST are described. One of the main ORA parameter is corneal hysteresis (CH) which is supposed to represent the viscoelastic properties of the cornea. The main advantage of Corvis ST is a high-speed Scheimpflug camera capable of taking two-dimensional images of a cross-section of the cornea during its deformation, which can possibly give additional information about the biomechanical status of the cornea.

Keywords: biomechanical properties, tonometry, intraocular pressure, cornea.

1 Introduction

The biomechanical properties of the cornea are essential for better understanding of the optical and the geometrical properties of the cornea [1]. Moreover, the prospective uses of such measurements may improve the knowledge of the behavior of the cornea in certain diseases [2] and in the result of surgeries [3,4,5]. Recently investigators have focused on the impact of corneal biomechanics on intraocular pressure (IOP) measurements [6,7]. The influence of central corneal thickness (CCT), corneal curvature (K) and age on IOP measurements seems to be already well known. The "true" IOP is considered to be overestimated in eyes with thick corneas and underestimated in those with thin cornea [8]. According to Kotecha et al. [9] IOP is also overestimated with increasing age. However, many authors pointed out that besides the CCT, K and age, IOP is strongly affected by the biomechanical properties which may induces measurement error [3,6,7,9,10,11,12].

Currently the Goldmann applanation tonometer (GTA) is still the "gold standard" of tonometres. Nevertheless it is commonly known that the IOP measurements are affected by cornea properties. To date only two commercially used noncontact tonometers can measure the IOP as well as biomechanical properties.

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_35, © Springer International Publishing Switzerland 2014

In 2005 The Ocular Response Analyzer (ORA; Reichert Ophthalmic Instruments, Depew, NY) was introduced as the first device claiming to provide the biomechanical properties of the cornea and their correlation with intraocular pressure (IOP) [11]. Lately in 2010 Oculus launched a new noncontact tonometer Corvis ST (Corneal Visualization Scheimpflug Technology, Corvis ST; Oculus Optikgerate, Inc., Wetzlar, Germany) integrated with an ultra-high speed Scheimpflung camera capable of imagining dynamic corneal deformation upon applanation with an air-puff.

The purpose of this article is to revive the current literature regarding the biomechanical parameters measured by the two commercially used noncontact tonometers: the Ocular Response Analyzer (ORA) and the Corvis ST (CST).

2 Corneal Structure and Material Properties

The main structural component of the cornea is collagen. It has a high tensile strength and provides a resilient, protective coat to the globe [10]. The whole structure of the human cornea is highly organized. In [10] authors pointed out that the result of this organization is a variable thickness of the cornea. The cornea is the thinnest in the center and on it's periphery it is thicker. What is more it is anisotropic, that is, it displays different physical properties when stress is applied in different directions. It needs to be mentioned that these properties are not constant, they vary with age, eye pathologies and changes in cornea structure, which results in modified corneal biomechanics [10].

One of the biomechanical properties is the elasticity, which refers to how a material deforms in response to applied stress. Graphically elasticity is a slope of the linear region of stress-strain relationship. The elastic constant is called Young's Modulus (E). High elastic modulus means the specimen stretches less (indicates a stiffer material). In [13] in *exvivo* studies Hjortdal showed that the cornea performes a non-linear stress-strain relationship. What is more the region and direction also significantly influenced the regional strains in the cornea. The Young's Modulus is the highest at the center and para-center in the meridional direction and at the limbus in the circumferential directions.

It is also essential that, like the most biological tissues, cornea is viscoelastic and therefore the loading and unloading process is different. That means that cornea exhibits hysteresis.

3 Overview of the Ocular Response Analyzer

3.1 Principle of ORA Operation

The Reichert ORA is a noncontact tonometer, which uses a precisely metered collimated air pulse to move cornea inward, past a first applanation (flattening), and into a slight contractility. When the pressure applied to the eye decreases cornea passes through a second applanation and returns to its normal curvature. The whole process of cornea deformation is recorded by an electo-optical infrared

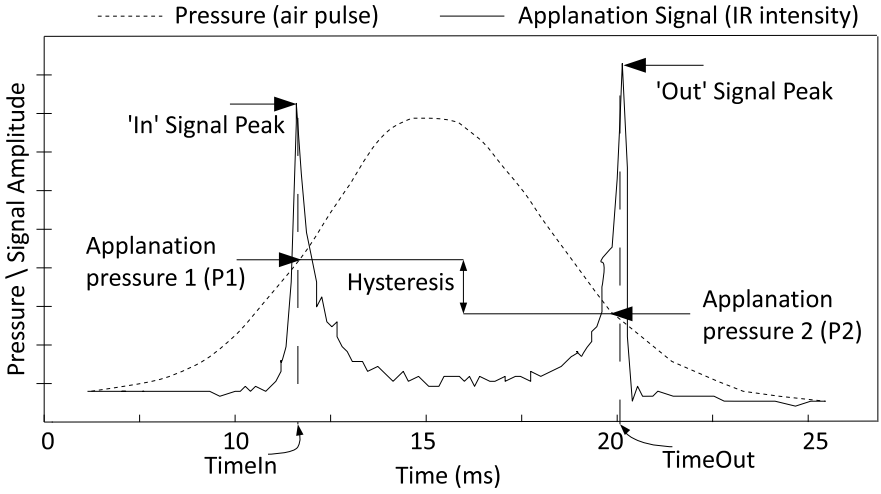


Fig. 1. A typical applanation - pressure plot determined by the ORA (Ocular Response Analyzer)

(IR) detection system (Fig. 1). On one hand the air-pulse pressure is a symmetrical signal with Gaussian distribution. On the other hand the IR intensity signal is asymmetric and defines two precise applanation times, which correspond with two peaks produced by inward and outward applanation events. These two peaks correspond to two pressures of both applanations (P_1 and P_2). As it was mentioned previously cornea is viscoelastic and primarily because of this the values of pressure at points P_1 and P_2 (applanation pressure) are different (Fig. 1). Another factors is the energy adsorption during rapid viscous damping which causes the delays in the inward and outward applanation signal peaks. The second applanation pressure is always lower than the first applanation pressure [11,14].

Based on the ORA processed signals a series of unique features can be described.

3.2 Cornea Hysteresis, Corneal Resistance Factor and Further Parameters

One of the main ORA parameter is corneal hysteresis (CH), which is the difference between the two pressures at the inward and outward applanation ($CH = P_1 - P_2$). CH is a numerical value and is supposed to represent the viscoelastic properties of the cornea. The ORA displays two different values of preassure: IOPg and IOPcc. IOPg - the Goldman-correlated IOP is calculated as an average of the two applanation pressures P_1 and P_2. IOPcc - corneal compensated IOP is the compensate value of pressure calculated from the equation:

Table 1. Parameters derived from Ocular Response Analyzer (ORA)

Parameter	Description
IOPg ($mmHg$)	Goldman correlated intraocular pressure
IOPcc ($mmHg$)	Corneal compensated intraocular pressure
CH ($mmHg$)	Corneal hysteresis
CRF ($mmHg$)	Corneal resistance factor
TimeIn (ms)	Time from starting until the first applanation
TimeOut (ms)	Time from starting until the second applanation

$P_2 - 0.43 \times P_1$ [15]. In addition ORA provides the parameter called: corneal resistance factor (CRF). It is derived from the formula: $P_1 - k \times P_2$, where k is a constant value that was empirically determined in the order to associated CRF with CCT more [10]. Some authors indicate also TimeIn and TimeOut (time at points P_1 and P_2) as significant parameters measured by the ORA (Table 1) [12].

Most researchers in their studies presented that the estimation of cornea biomechanical parameters was useful to evaluate the influence of corneal properties on IOP measurements. Felipe et al. [16] demonstrated that the IOPcc values were less influenced by cornel properties than IOP measures by the Goldman applanation tonometer. In his study GAT measurements were highly influenced by CCT, whereas the IOPcc were not significantly correlated with CCT, indicating that the IOPcc is a value which is independent of corneal thickness. The same independency was presented in [12] where authors showed that all parameters measured by ORA had a positive correlation with CCT, besides the IOPcc.

Moreover Franco and Lira [12] pointed out the relationship between the cornea hysteresis (CH) and corneal resistance factor (CRF) and corneal thickness (CCT). In eyes with thinner cornea both the CH and CRF were lower, though the cornea resistance factor had stronger correlation. The same findings were showed in [17,18]. Authors suggested that CRF had a better correlation because it is a measure of the cornea's material properties and may reflect the cornea's elastic properties better. CRF is originally calculated from a function of pressures P_1 and P_2 in order to maximize the correlation with central corneal thickness and obtained independency of IOP [18]. Corneal hysteresis has been also found to be negatively correlated with IOP [17].

Shah et al. [2] in his studies demonstrated that the hysteresis (CH) in normal eyes were higher than in keratoconic eyes. The mean value of hysteresis in normal eyes was 10.7 $mmHg$ and in keraoconic eyes 9.6 $mmHg$. These values were in accordance with another studies [4,14], however Luce [11] observed lower values: the mean hysteresis in the normal population was 9.6 $mmHg$ and 8.1 $mmHg$ in keratoconic eyes. It needs to be mentioned that it is also a significant weakness in CH parameter concerned with keratoconic patients. The hysteresis alone cannot differentiate normal corneas from the mild keratoconus [2].

ORA applications in eyes after refractive surgeries are well known. In [11] author finds that in patients after LASIK (laser assisted in-situ keratomileusis)

cornea hysteresis was lower than in normal eyes. All corneas after LASIK had a radically reduced applanation, which means that the optical reflectance was lower, as well as the signal width and amplitude. Luce [11] indicated that such results presented CH as a parameter that characterizes the cornea resistance during tests, which he claims to represent the combined effects of corneal thickness, rigidity, hydration and other factors. The subsequent work [12] confirmed the lower hysteresis for the post-LASIK eyes.

Another surgical procedure where the biomechanical properties are expected to be modified is corneal cross-linking (CXL) - treatment for keratoconus. Recent studies [19] reported that the CH and CRF remained unchanged 1 year after CXL. The absence of relevant changes confirmed previous ORA results [20]. So far no studies using ORA managed to show the changes in CH and CRF that would correlate with the increased stiffness after CXL procedure.

Recently a few groups have been working on developing a new, more complicated features of ORA, based on the waveform signal, with the objective of improving sensitivity and specificity in ocular diseases diagnosis [21].

4 Overview of the CORVIS ST

4.1 Principle of CORVIS ST Operation

The Corvis ST instrument is a noncontact tonometer that is equipped with a pachymetry function. CST measured IOP by the analysis of cornea displacement after air impulse. The ultrahigh-speed Scheimpflug camera takes 4330 $frames/sec$ covering up to 8.5 mm horizontally to monitor cornea response to a metered collimated air pulse with symmetrical configuration and fixed maximal pump pressure of 25 kPa. The light source is an ultraviolet free blue LED light with a wavelength of 455 nm. The measurement process takes 30 ms and is recorded on a video. Images are of excellent resolution (640×480 pixels) and 576 measuring points [22,23]. Firstly UHS Scheimpflug camera recorded the image of cornea before the air puff in the natural convex shape. Then the air jet moves cornea inwards, trough the first applanation, the phase of the highest contractility (HC), and next the cornea moves outwards though the second applanation achives its natural shape. Complex algorithms for edge detection of the corneal contours are applied for every image. The IOP is determined based on the first applanation time [23]. According to Imbert-Fick's law the IOP is calculated by deviding the air pressure by the applanation area.

4.2 Specific Biomechanical Parameters Measured by Corvis ST

During the cornea's displacement several parameters are measured (Table 2). The deformation amplitude (DA) is the maximum amplitude at the moment of the highest concavity. When the cornea reaches the HC moment also a radius of curvature of a circle that fits to corneal concavity is measured. Length of the flattened cornea at the first (A_1 length) and second (A_2 length) applanation and

corneal velocities - speed of the corneal apex at the first (A_1 velocity) and second (A_2 velocity) applanation is measured as well. Another parameter measured is the distance (peak distance) between the two apexes of the cornea at the time of the highest concavity. The timing of characteristic moments is also detected: time from starting until the firs applanation (A_1 time), time from starting until the second applanation (A_2 time) and time from starting until the HC of the cornea is reached (HC time). The pachymetric values (Pachy) and the cornea curvature are detected in the natural shape of cornea, when there is no air pulse [23,24,25].

Table 2. Parameters derived from Corvis ST

Parameter	Description
DA (mm)	Deformation amplitude
Radius (mm)	Radius of curvature at the time of highest concavity
A_1 length (mm)	Length of flattened cornea at the first applanation
A_2 length (mm)	Length of flattened cornea at the second applanation
A_1 velocity (m/s)	Speed of the corneal apex at the first applanation
A_2 velocity (m/s)	Speed of the corneal apex at the second applanation
Peak distance (mm)	Distance between the two apex of the cornea at the time of highest concavity
A_1 time (ms)	Time from starting until the first applanation
A_2 time (ms)	Time from starting until the second applanation
HC time (ms)	Time from starting until the highest concavity of cornea is reached
Pachy (μm)	Central corneal thickness

What is more, Corvis ST displays the real-time information for a patient just after an air impulse. Another features of Corvis ST: Scheimpflug images of the 1st and 2nd applanation moments and of the highest concavity of the cornea and a slow-motion video of the corneal deformation after air pulse.

Due to the UHS Scheimpflug camera Corvis ST displays three graphs: the displacement of the cornea apex (deformation amplitude) during the corneal flatting/ relaxation process, the change in applanation length during the acquisition and the corneal velocity. The deformation of cornea (Fig. 2) is expected to depend on IOP as well as on structural stiffness of the cornea and other nonlinear parameters.

In [26] a high-speed swept source optical coherence tomography was used for acquisition of anterior segment displacement during the cornea deformation. In this work authors indicate that in all examined cases, lower IOP produced higher displacement of the cornea, though, the variation of cornea displacement is linked to the IOP value. What is more they emphasize the fact that corneal biomechanics is also one of the major factors that influenced the dynamics of cornea apex. Another studies have shown that electrical activity of the heart could also influenced the corneal apex displacement [27].

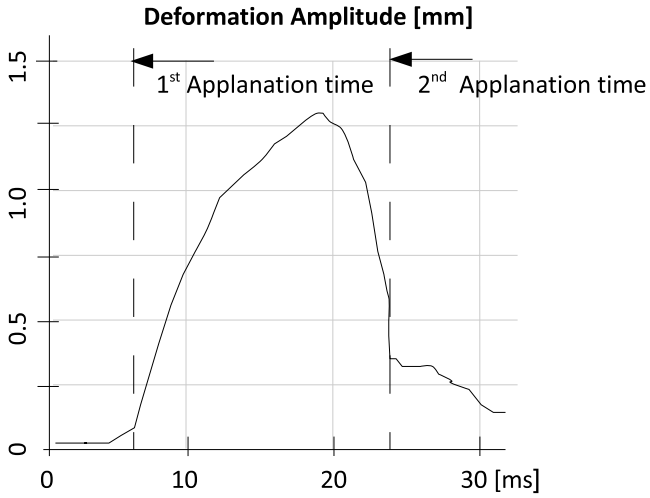


Fig. 2. Deformation amplitude plot determined by Corvis ST

In [28] authors showed a relationship between applied force (air) and dynamic response of the cornea. Like in [26] the deformation was higher with the rise of the applied air pressure. By building a simple model of material properties of the cornea Kempf et al. [28] suggested that the steep rise in displacement of the cornea can be only explained by non-linear material properties of the cornea. The observed time shift between the applied pressure and the deformation couldn't be explained by the proposed model.

The outward phase of cornea movement is also interesting, when the cornea returns to its normal position. Like it's seen in Fig. 2 after second applanation we can observe a slower decline in the deformation amplitude. Authors suppose it's caused by the dynamic response of the internal structure of human eye (not the cornea only).

Nemeth et al. [24] conducted a series of measurements in order to analyse the repeatability of parameters measured by Corvis ST and verify their correlation in healthy eyes. Scientist reported that data recorded by Corvis ST showed excellent repeatability for IOP and pachymetric values. Highest concavity as well as first applanation time showed only good repeatability, however other parameters had a weak repeatability. On the other hand a significant correlation of Corvis ST parameters was obtained for keratometric values what suggested that corneal curvature influenced the measured data.

Comparing the mean IOP values in healthy eyes in [24] was 15.02 $mmHg$ (range: 10.83 to 26.17) in [22] IOP was quite similar: 14.6 $mmHg$ (range: 11.0 to 17.8) and in [25] the IOP value was the highest and amounted 17.73 $mmHg$ (range: 7.00 to 32.20). However, in [25] statistical group of healthy patients was the biggest.

In [22] authors compared measurements obtained using Goldman applanation tonometer (GAT), Topocon noncontact tonometer (NCT) and Corvis ST (CST) in normal and glaucoma eyes. The results presented showed no statistically significant difference in IOP measurements between all three devices. IOP values measured with CST showed correlation with both: GAT and NCT IOP values, however CST correlated better with NCT. Based on statistical parameters in this study authors registered better intraobserver variability for CST than that for GAT and NCT. Authors speculated that this might be explained by the fact that CTS IOP measurements technique is more objective, and external factors like biomechanical properties affected CST IOP values less. However, the scientist concluded that IOPs measurements are not interchangeable between devices and pointed out that CST may give lower IOP than the other instruments.

Valbon et al. [25] examined the age relation with parameters measured with Corvis ST. Only the HC time was significantly associated with age. Other parameters, like DA and the curvature radius at the time of highest concavity, showed positive linear correlation only with young patients.

5 Summary

A review of current literature suggested that the cornea's biomechanical properties have an essential influence on the IOP measurements, affect surgical outcomes and the development and progression of corneal diseases. In this study we presented two noncontact tonometers capable of measuring such parameters: the Ocular Response Analyzer (ORA) and the Oculus Corvis ST. Currently the clinical application as well as limitation of ORA is widely known, therefore there is a need for finding more selective ORA parameters. The newest device, Corvis ST equipped with UHS Scheimpflug camera, has an enormous potential as a tool to assess *invivo* the biomechanical properties of the cornea. However, further studies are needed to expand the application and knowledge of the corneal biomechanical properties.

References

1. Dupps, W.J., Wilson, S.E.: Biomechanics and wound healing in the cornea. Exp. Eye Res. 83, 709–720 (2006)
2. Shah, S., Laiquzzaman, M., Bhojwani, R., Mantry, S., Cunliffe, I.: Assessment of the biomechanical properties of the cornea with the ocular response analyzer in normal and keratoconic eyes. Invest. Ophthalmol. Vis. Sci. 48, 3026–3031 (2007)
3. Ambrosio, R., Nogueira, L.P., Caldas, D.L., Fontes, B.M., Luz, A., Cazal, J.O., Alves, M.R., Belin, M.W.: Evaluation of Corneal Shape and Biomechanics Before LASIK. Int. Ophthalmol. Clin. 51, 11–39 (2011)
4. Ortiz, D., Pinero, D., Shabayek, M.H., Arnalich-Montiel, F., Alio, J.L.: Corneal biomechanical properties in normal, post-laser in situ keratomileusis, and keratoconic eyes. J. Cataract Refract. Surg. 33, 1371–1375 (2007)

5. Pepose, J.S., Feigenbaum, S.K., Oazi, M.A., Sanderson, J.P., Roberts, C.J.: Changes in corneal biomechanics and intraocular pressure following LASIK using static, dynamic, and noncontact tonometry. Am. J. Ophthalmol. 143, 39–47 (2007)
6. Ogbuehi, K.C., Osuagwu, U.L.: Corneal biomechanical properties: Precision and influence on tonometry. Cont. Lens Anterior Eye (2013)
7. Liu, J., Roberts, C.J.: Influence of corneal biomechanical properties on intraocular pressure measurement: quantitative analysis. J. Cataract Refract. Surg. 31, 146–155 (2005)
8. Doughty, M.J., Zaman, M.L.: MAJOR REVIEW Human Corneal Thickness and Its Impact on Intraocular Pressure Measures? A Review and Meta-analysis Approach 44 (2000)
9. Kotecha, A., White, E.T., Shewry, J.M., Garway-Heath, D.F.: The relative effects of corneal thickness and age on Goldmann applanation tonometry and dynamic contour tonometry. Br. J. Ophthalmol. 89, 1572–1575 (2005)
10. Kotecha, A.: What biomechanical properties of the cornea are relevant for the clinician? Surv. Ophthalmol. 52(suppl. 2), S109–S114 (2007)
11. Luce, D.A.: Determining in vivo biomechanical properties of the cornea with an ocular response analyzer. J. Cataract Refract. Surg. 31, 156–162 (2005)
12. Franco, S., Lira, M.: Biomechanical properties of the cornea measured by the Ocular Response Analyzer and their association with intraocular pressure and the central corneal curvature. Clin. Exp. Optom. 92, 469–475 (2009)
13. Hjortdal, J.O.: Regional elastic performance of the human cornea. J. Biomech. 29, 931–942 (1996)
14. Touboul, D., Roberts, C., Kerautret, J., Garra, C., Maurice-Tison, S., Saubusse, E., Colin, J.: Correlations between corneal hysteresis, intraocular pressure, and corneal central pachymetry. J. Cataract Refract. Surg. 34, 616–622 (2008)
15. Nessim, M., Mollan, S.P., Wolffsohn, J.S., Laiquzzaman, M., Sivakumar, S., Hartley, S., Shah, S.: The relationship between measurement method and corneal structure on apparent intraocular pressure in glaucoma and ocular hypertension. Cont. Lens Anterior Eye 36, 57–61 (2013)
16. Medeiros, F.A., Weinreb, R.N.: Evaluation of the influence of corneal biomechanical properties on intraocular pressure measurements using the ocular response analyzer. J. Glaucoma 15, 364–370 (2006)
17. Kotecha, A., Elsheikh, A., Roberts, C.R., Zhu, H., Garway-Heath, D.F.: Corneal thickness- and age-related biomechanical properties of the cornea measured with the ocular response analyzer. Invest. Ophthalmol. Vis. Sci. 47, 5337–5347 (2006)
18. Kamiya, K., Hagishima, M., Fujimura, F., Shimizu, K.: Factors affecting corneal hysteresis in normal eyes. Graefes Arch. Clin. Exp. Ophthalmol. 246, 1491–1494 (2008)
19. Greenstein, S.A., Fry, K.L., Hersh, P.S.: In vivo biomechanical changes after corneal collagen cross-linking for keratoconus and corneal ectasia: 1-year analysis of a randomized, controlled, clinical trial. Cornea 31, 21–25 (2012)
20. Vinciguerra, P., Albe, E., Mahmoud, A.M., Trazza, S., Hafezi, F., Roberts, C.J.: Intra- and postoperative variation in ocular response analyzer parameters in keratoconic eyes after corneal cross-linking. J. Refract. Surg. 26, 669–676 (2010)
21. Lam, A.K.C., Chen, D., Tse, J.: The Usefulness of Waveform Score from the Ocular Response Analyzer. Optom. Vis. Sci. 87, 195–199 (2010)
22. Hong, J., Xu, J., Wei, A., Deng, S.X., Cui, X., Yu, X., Sun, X.: A new tonometer-the Corvis ST tonometer: clinical comparison with noncontact and Goldmann applanation tonometers. Invest. Ophthalmol. Vis. Sci. 54, 659–665 (2013)

23. Ambrosio Jr, R., Ramos, I., Luz, A., Faria, F.C., Steinmueller, A., Krug, M., Belin, M.W., Roberts, C.J.: Dynamic ultra high speed Scheimpflug imaging for assessing corneal biomechanical properties. Rev. Bras. Oftalmol. 72, 99–102 (2013)
24. Nemeth, G., Hassan, Z., Csutak, A., Szalai, E., Berta, A., Modis, L.: Repeatability of ocular biomechanical data measurements with a Scheimpflug-based noncontact device on normal corneas. J. Refract. Surg. 29, 558–563 (2013)
25. Valbon, B.F., Ambrosio Jr, R., Fontes, B.M., Alves, M.R.: Effects of age on corneal deformation by non-contact tonometry integrated with an ultra-high-speed (UHS) Scheimpflug camera. Arq. Bras. Oftalmol. 76, 229–232 (2013)
26. Alonso-Caneiro, D., Karnowski, K., Kaluzny, B.J., Kowalczyk, A., Wojtkowski, M.: Assessment of corneal dynamics with high-speed swept source optical coherence tomography combined with an air puff system. Opt. Express. 19, 14188–14199 (2011)
27. Danielewska, M.E., Iskander, D.R., Kowalska, M., Kasprzak, H.T.: Phase dependencies between longitudinal corneal apex displacement and cardiovascular signals: is the ocular pulse influenced by the electrical activity of the heart? Clin. Exp. Optom. 95, 631–637 (2012)
28. Kempf, R., Kurita, Y., Iida, Y., Kaneko, M., Mishima, H.K., Tsukamoto, H., Sugimoto, E.: Understanding eye deformation in non-contact tonometry. In: Conf. Proc. IEEE Eng. Med. Biol. Soc., vol. 1, pp. 5428–5431 (2006)

Numerical Analysis of Taylor-Type External Fixator by Means of FEM

Mateusz Pawlik[1] and Marcin Basiaga[2]

[1] Scientific Group "Synergia" Faculty of Biomedical Engineering, Silesian University of Technology, Zabrze, Poland
mateusz.pawlik@windowslive.com
[2] Department of Biomaterials and Medical Devices Engineering, Faculty of Biomedical Engineering, Silesian University of Technology, Zabrze, Poland
marcin.basiaga@polsl.pl

Abstract. The main aim of this work was evaluate the displacements and strains and stresses in components of bone-fixator system using Finite Element Method. The Taylor-type fixator was chosen for the analysis as a one of the most technically advanced and the biggest therapy opportunities giving external fixator. Through the possibility of full and free manipulation of bone fractures (possible displacements and rotations in three independent planes) it is used in complex fractures therapy, complicated bones deformation therapy, distraction osteogenesis and much more. The first step was preparation of Taylor-type fixator geometrical model which was developed on a basis of a real model. To carry out an analysis it was necessary to develop a tibia bone model in which the fracture crack of 1mm breadth was simulated. Next a grid for finite-element method calculations was generated for the geometrical models. Subsequently it was necessary to specify and set the edge conditions to reflect appropriately the phenomena taking place in the real system. The system was loaded with axial force in range 100-1500N. Calculations were realized for fixator made of 316L steel. On a basis of obtained result, the highest values of reduced stress were observed in the fixator frame, in elements connecting telescopic strut with holder constrained in fixator full ring. For load $F > 1400$N the yield stress of frame material was exceeding, what could cause its damage. In turn, stress in bone didn't exceed its compression strength. The biomechanical analysis may form the basis for improving the geometry of analyzed fixator and optimising a selection of the mechanical properties of the material used to manufacture them.

Keywords: external fixation, Taylor-type fixator, numerical analysis, FEM.

1 Introduction

Human body is complicated mechanism exposed to various type of injuries as well as congenital or acquired defects in construction. The injuries can include, among others fractures and construction defects are most often associated with

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_36, © Springer International Publishing Switzerland 2014

bone deformities. In the case of the most complex defects and injuries of long bones is currently the most effective treatment is the use of an external fixator. The most important tasks for that type of devices are:

- Carrying load from one bone part to the other by means of fixator frame.
- Stabilizing and securing the bone fragments from moving relative to each other.
- Enabling stimulating bone micromovement incidence.

Currently, one of the most technologically advanced and offering the greatest potential for therapy is the Taylor-type external fixator. Six independent telescopic struts applied in it provide full freedom of bone parts manipulation - displacements and rotations in three independent planes are possible. Above attributes qualify Taylor-type apparatus to use it in complex fracture therapy, healing complicated bone deformities, distraction osteogenesis and many others [1].

During the continuous growth of injuries and defects of the bone system as a result of such traffic accidents and loss of physical activity increases the need for effective methods of treatment of the above mentioned diseases. The literature on biomechanical aspects of these implants, however, scant and mainly concerns analyses of dislocations, strains and stresses and their functional purposes [2,3,4]. Such an analysis forms the basis for the optimisation of the geometric features of implants and the selection of mechanical properties of the material used to make it. This work concentrates on the strength analysis of selected external fixator by means of a finite-element method.

2 Materials and Methods

In this thesis the subject of analysis was Taylor-type external fixator attached to the fractured tibia. On the basis of the real model - Fig.1a. - realistic form of fixator was developed using Autodesk Inventor Professional 2013 software - Fig.1b. Next, model was simplified to execute FEM analysis by means of Ansys Mechanical APDL 14.0 software - Fig.1c. For executing numerical analysis it was necessary to develop geometrical model of tibia bone, in which the 1mm wide fracture was simulated.

Analyzed stabilizer system consists of two complete rings connected together by means of six telescopic double-swivel struts - geometry was developed on a basis of available documentation [5]. Connection of bone and fixator was realized by four 2 mm diameter Kirschner wires and four 6mm diameter half pins. A mesh for finite-element method calculations was generated for the geometrical models - Fig. 2. For discretization of analyzed object Solid187-type elements was selected. The element is defined by 10 nodes having three degrees of freedom at each node: translations in the nodal x, y, and z directions. In order to make the calculations, it was necessary to specify and set the initial and boundary conditions to reflect appropriately the phenomena taking place in the real system. The following conditions were assumed [6,7,8,9,10,11,12,13]:

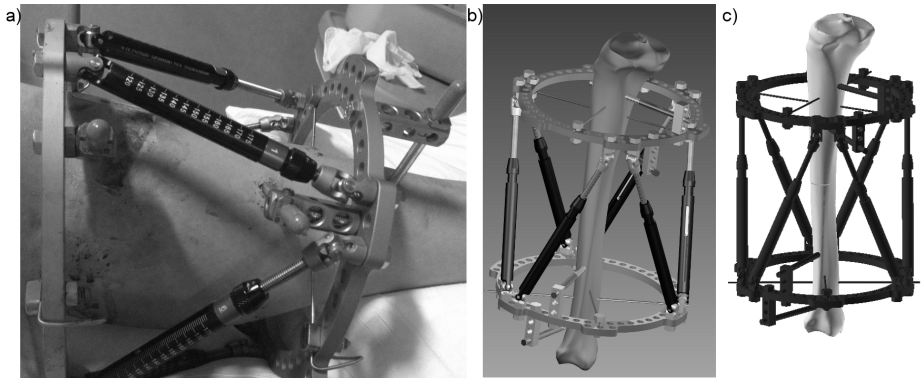


Fig. 1. Constructional form of Taylor-type external fixator: a) real model, b) geometrical model, c) simplified model

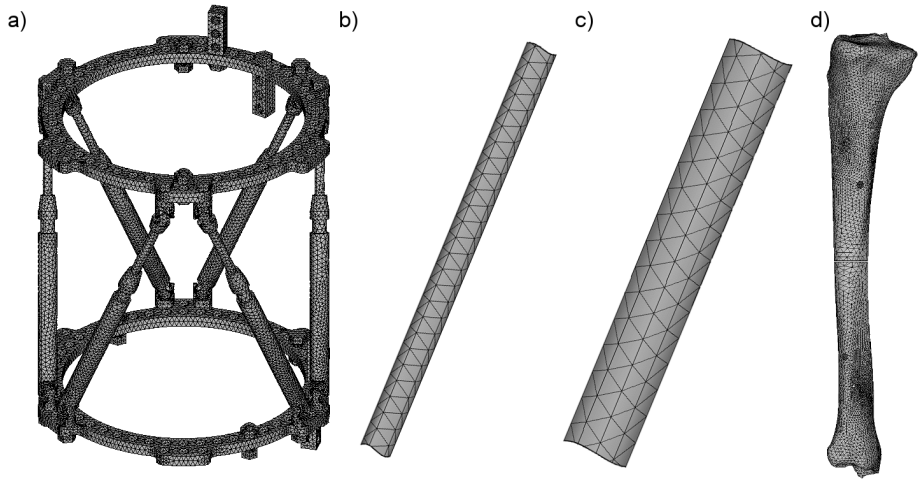


Fig. 2. Discretized models: a) fixator frame, b) Kirschner wire, c) half pin, d) tibia bone

- System clamped on the distal joint surface of tibia bone, what constrained bone displacements along X, Y, Z axes.
- System loaded with an axial force in range 100-1500N on the proximal joint surface of the tibia.

The analysis included designation of displacements, deformations and stresses at the individual elements of bone-fixator system according to set load. The following material properties were assumed [6,7,8,9,10,11,12,13]:

- Kirschner wires, half pins, fixator frame - stainless steel 316L $E = 200\ 000$ MPa, $\upsilon = 0.3$,
- Tibia - $E = 18\ 600$ MPa, $\upsilon = 0.3$.

3 Results and Discussions

Results of strength analysis of bone-fixator system for the different variants of the load F are presented in Table 1. The analysis of these results shows that the distribution of deformations, strains and stresses reduced both the individual elements of the stabilizer, as well as bone fragments. Based on the obtained

Table 1. Results of numerical analysis of bone-fixator system

	Bone displacement in fracture, mm			Stress, MPa				Strain, %			
Load, N	x	y	z	Wire	Half pin	Bone	Frame	Wire	Half pin	Bone	Frame
100	0.27	0.87	0.55	177	70	31	198	0.10	0.04	0.19	0.11
200	0.27	0.93	0.57	185	78	33	246	0.10	0.04	0.20	0.14
300	0.28	0.96	0.54	193	82	35	271	0.11	0.05	0.21	0.15
400	0.29	0.96	0.53	207	88	39	309	0.12	0.05	0.24	0.17
750	0.29	0.97	0.59	249	128	50	420	0.14	0.06	0.30	0.24
1000	0.29	0.97	0.76	274	145	57	493	0.15	0.07	0.35	0.28
1500	0.29	0.99	0.93	416	200	91	838	0.21	0.10	0.49	0.42

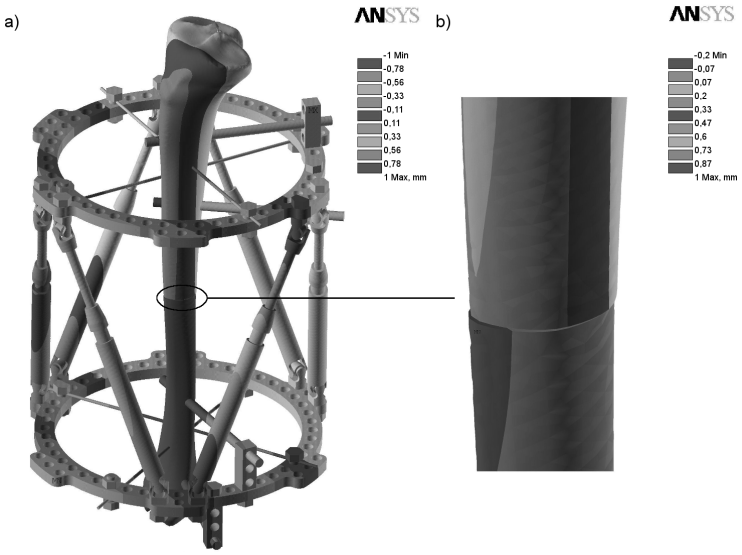


Fig. 3. Example distribution of displacement in bone-fixator system: a) Taylor-type fixator, b) tibia bone

results it was found that the biggest displacement values were observed in the bone fracture along Y axis and averaged 0.99 mm - Fig.3. These values didn‘t exceed the permissible value of the fracture gap about 1 mm, which determines the correct bone union.

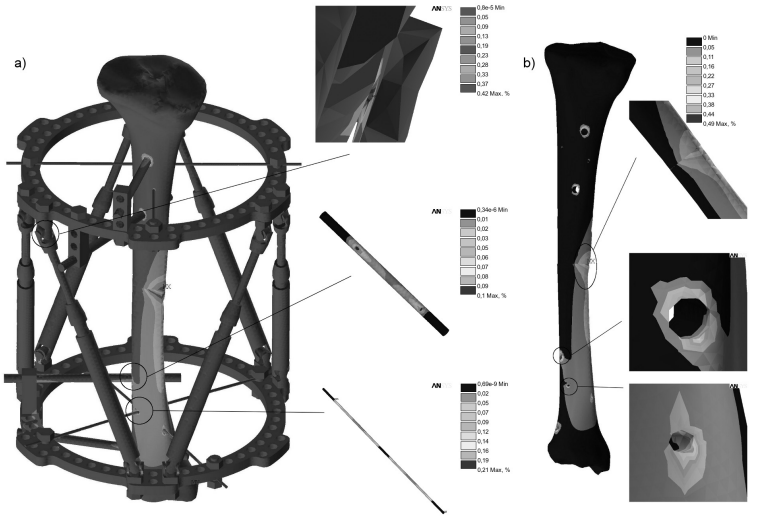


Fig. 4. Example distribution of strain in bone-fixator system: a) Taylor-type fixator, b) tibia bone

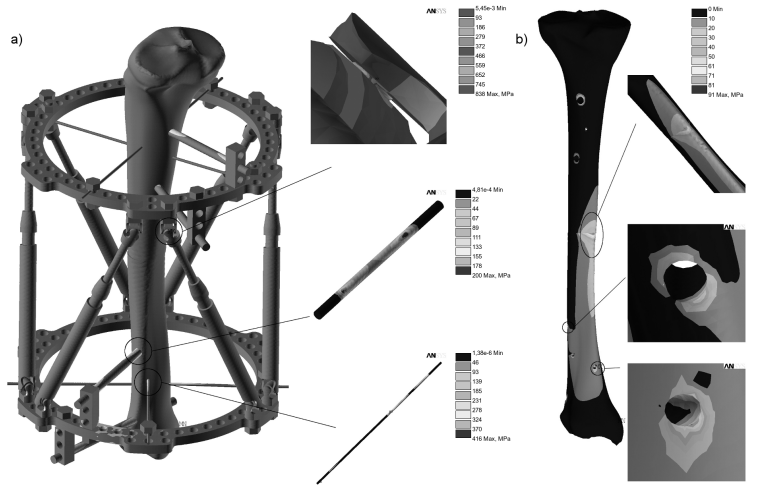


Fig. 5. Example distribution of stress in bone-fixator system: a) Taylor-type fixator, b) tibia bone

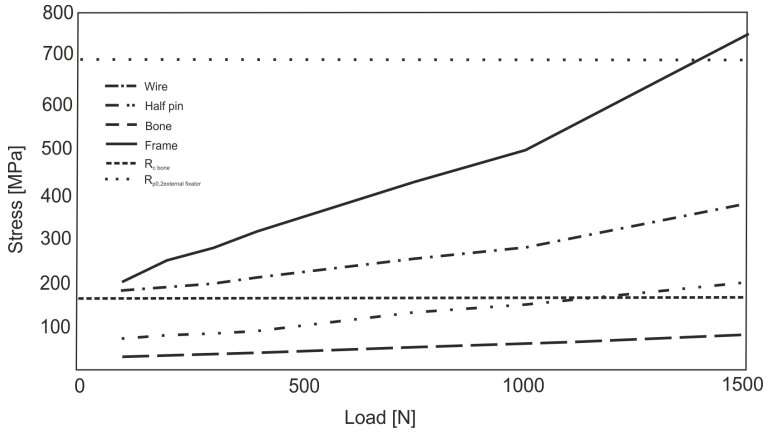


Fig. 6. Relation of reduced stresses as a function of the applied loading

It was found additionally that maximum strain and stress for the maximum force F=1500N were observed for connecting elements telescopic strut with holder constrained in fixator full ring (frame) and equals $\epsilon_{max} = 0.42\%$, $\sigma_{max} =$ 838MPa respectively - Table 1, Fig.4, 5. The values of stress and strain in other parts were much smaller and equals $\epsilon_{max} = 0.21\%$, $\sigma_{max} = 416$MPa for wire, $\epsilon_{max} = 0.10\%$, $\sigma_{max} = 200$MPa for half pin and $\epsilon_{max} = 0.49\%$, $\sigma_{max} = 91$MPa for bone respectively - Table 1. Obtained in the analysis of stresses don't exceed the yield strength of the metal biomaterial ($Rp_{0,2} = 690$MPa) for loads of the range F = 100-1400 N - Fig.6. It assures fixator elements strain in elastic range. However, the load force $F > 1400$N following border crossing point of $Rp_{0,2}$ frame material which in turn can cause damage. Additionally, it was also found that obtained stress values in bone didn't exceed its compression strength ($Rc = 160$MPa).

4 Conclusions

This paper is a preliminary attempt to analyze a complex system, which is the Taylor stabilizer. The main aim of this work was evaluate the displacements, strains and stresses in components of bone-fixator system using Finite Element Method. The analysis was selected type stabilizer Taylor, who is one of the most technically advanced and the biggest therapy opportunities giving external fixator. The results of the presented analysis allows a better understanding of the mechanism of action of the analyzed system. The analysis is the first of its kind realized numerical analysis using the finite element method. Research include only experimental work of external fixators such as: Iliazarov [14,15,16,17]. The results of numerical analysis by means of finite element method provides a valuable methodology for the proper design of the form of geometric features, the selection of the mechanical properties of the stabilizer and the degree of

strengthening metallic biomaterial. On the basis of the conducted study it was concluded that the most dangerous place to be exposed to damage is connecting elements telescopic strut with holder constrained in fixator full ring. Affirmed, that the values of stresses don't exceed the yield strength of the metal biomaterial ($Rp_{0,2} = 690$MPa) for loads of the range F = 100-1400 N. It assures fixator elements strain in elastic range. However, the load force $F > 1400$N following border crossing point of $Rp_{0,2}$ frame material which in turn can cause damage. Additionally, it was also found that obtained stress values in bone didn't exceed its compression strength ($Rc = 160$MPa).

The next stage of the study will include consideration of a stabilizer for use in the treatment of bone deformity, and also in the treatment of bone lengthening by distraction osteogenesis and experimental tests in order to compare them with the results obtained using the finite element method (correctness of the numerical model). Such a comprehensive study should be an important source of information for orthopaedic surgeons.

References

1. Choudari, M.: Taylor Spatial Frame. In: Kulkarni (ed.) Textbook of Orthopaedics and Trauma, 2nd edn. Jaypee Brothers Publishers (2009)
2. Seide, K., Weinrich, N., Wenzl, M.E., Wolter, D., Jürgens, C.: Three-dimesnional load measurements in an external fixator. Journal of Biomechanics 37, 9:1361–9:1369 (2004)
3. Philip Drijber, F.L.I., Bryan Finlay, J., Dempsey, A.J.: Evaluation of linear finite-element analysis models assumptions for external fixation devices. Journal of Biomechanics 25, 8:849–8:851 (1992)
4. Prat, J., Juan, J.A., Vera, P., Hoyos, J.V., Dejoz, R., Peris, J.L., Sánchez-Lacuesta, J., Comín, M.: Load transmission through the callus site with external fixation systems: Theoretical and experimental analysis. Journal of Biomechanics 27, 4:469–4:478 (1994)
5. Taylor, S., John Taylor, C.: Six axis external fixator strut. EP1239784 B1 (April 16, 2006)
6. Tejszerska, D., Świtoński, E., Gzik, M.: Biomechanika narządu ruchu człowieka. Wyd. Politechniki Śląskiej, Gliwice (2011)
7. Ramatowski, W.: Stabilizatory płytkowe: Zespol i Polfix. Agencja Wydawnicza Zebra, Kraków (1998)
8. Basiaga, M., Paszenda, Z., Szewczenko, J., Kaczmarek, M.: Numerical and experimental analysis of drills used in osteosynthesis. Acta of Bioengineering and Biomechanics 13, 4:29–4:36 (2011)
9. Basiaga, M., Paszenda, Z., Szewczenko, J.: Biomechanical behavior of surgical drills in simulated conditions of drilling in a bone. In: Piętka, E., Kawa, J. (eds.) Information Technologies in Biomedicine. AISC, vol. 69, pp. 473–481. Springer, Heidelberg (2010)
10. Kiel, M., Marciniak, J., Basiaga, M., Szewczenko, J.: Numerical Analysis of Spine Stabilizers on Lumbar Part of Spine. In: Piętka, E., Kawa, J. (eds.) Information Technologies in Biomedicine. AISC, vol. 69, pp. 447–456. Springer, Heidelberg (2010)

11. Marciniak, J., Szewczenko, J., Walke, W., Basiaga, M., Kiel, M., Mańka, I.: Biomechanical analysis of lumbar spine stabilization by means of transpedicular stabilizer. In: Information Technologies in Biomedicine. ASC, vol. 47, pp. 529–536. Springer (2008)
12. Ziebowicz, A., Kajzer, A., Kajzer, W., Marciniak, J.: Metatarsal osteotomy using double-threaded screws - biomechanical analysis. In: Piętka, E., Kawa, J. (eds.) Information Technologies in Biomedicine. AISC, vol. 69, pp. 465–472. Springer, Heidelberg (2010)
13. Kiel, M., Marciniak, J., Szewczenko, J., Basiaga, M., Wolanski, W.: Biomechanical analysis of plate stabilization on cervical part of spine. Archives of Materials Science and Engineering 38(1), 41–47 (2009)
14. Walke, W., Marciniak, J., Paszenda, Z.: Biomechanical Behaviour of Double Threaded Screw in Tibia Fixation. In: Information Technologies in Biomedicine. ASC, vol. 47, pp. 521–528 (2008)
15. Nishimura, N., Yamano, Y.: Hallus valgus operation using Iliazarov external fixator. Foot and Ankle Surgery 3, 183–188 (1997)
16. Courvoisier, A., Sailhan, F., Thevenin-Lemoine, C., Vialle, R., Damsin, J.P.: Congenital tibial deficiencies: Treatment using Ilizarov's external fixator. Orthopaedics & Traumatology: Surgery & Research 95, 431–436 (2009)
17. Gessmann, J., Jettkant, B., Seybold, D.: Mechanical stress on tensioned wires at direct and indirect loading: A biomechanical study on the Ilizarov external fixator. Injury 42, 1107–1111 (2011)

Influence of Radiation on Mechanical Properties of Ultra High Molecular Weight Polyethylene (UHMWPE)

Magdalena Grygiel and Marcin Kaczmarek

Silesian University of Technology, Faculty of Biomedical Engineering
Department of Biomaterials and Medical Devices Engineering, Zabrze, Poland
magdalena.grygiel@polsl.pl

Abstract. Ultra high molecular weight polyethylene (UHMWPE) is used in medicine, mainly in the acetabulum and the sliding elements of endoprosthesis. Benefits, which have determined its use is primarily a very good tribological properties, toughness, durability, resistance to biological fluids and good mechanical properties. Unfortunately UHMWPE is not free from drawbacks, which mainly include the low resistance to frictional wear a low yield strength, the tendency to creep and aging. An important issue exposed in the literature, is the change of the polyethylene physical properties by ionizing radiation, which influences the degree of crystallinity, and consequently, a decrease in mechanical properties. Therefore, the aim of the study was to determine the influence of X-rays on mechanical properties of polyethylene used for acetabula of hip endoprosthesis. The originality of the paper consists in the fact that the tested material was irradiated with the true value of radiation doses to which the patients are exposed during clinical treatment after hip replacement, in order to check their safety. Test specimens were irradiated with different X-ray doses. In the course of the study the visual inspection of the polyethylene surface, density measurement (immersion test), static tensile and bending tests as well as hardness test were carried out. As the result of this study, no influence of ionizing radiation on the mechanical properties (hardness, Young's modulus, tensile strength) of the tested material was observed.

Keywords: UHMWPE, ultra high molecular weight polyethylene, ionizing radiation, mechanical properties, acetabulum, joint endoprosthesis.

1 Introduction

About two million implantations of an artificial joint, or spine implant is performed every year in the world [1]. Some components of implants are made of a polymeric material, mainly from UHMWPE. The ultra-high molecular weight of polyethylene is its unique property, which resulted in a polymer with good physical and useable features. Suitable structure, mechanical properties and wear resistance determined the use of UHMWPE in orthopedics, mainly for sliding elements (bearings and acetabulum) of joint endoprosthesis [2,3,4].

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_37, © Springer International Publishing Switzerland 2014

UHMWPE is a linear, semi-crystalline polymer which contains lamellar crystals and amorphous zone between [5,6]. Polyethylene is a two-phase material, so the ratio of crystalline phase to amorphous phase should be constant. The mechanical properties of the UHMWPE depend on the phase composition, degree of crystallinity and density [5,7]. Despite many advantages, the UHMWPE also has drawbacks, which include: aging, low yield strength, tendency to creep and low friction wear resistance [3,5,7]. Polyethylene acetabular prosthesis after implantation with the time lose their usable properties due to the increase of the crystallinity degree. The increase in the crystallinity degree of the ultra-high molecular weight polyethylene is influenced by factors such as the method of sterilization, as well as the effect of X-rays [6,8]. Too high dose of radiation can cause destruction of the polymer. This may have an impact on the structure, mechanical properties and physico-chemical properties. In a study of Sudol and Czaja on the effects of electron beam irradiation and plastic deformation of UHMWPE, it was found that sterilization has devastating effects, because macromolecules oxidation of the polymer occurs and its degree is dependent on the dose of the radiation. Cross-linking of macromolecules and reduction of molecular weight was also observed [9,10,11,12].

2 Methodology

The key objective of the study was to investigate the influence of ionizing radiation on the stability of the mechanical properties of ultra-high molecular weight polyethylene, used in production of hip endoprostheses. The change of mechanical properties of polyethylene could increase the risk of damage to the endoprostheses.

UHMWPE 1000 (ultra-high molecular weight polyethylene) was material used for testing. On the selected material series of X-ray radiation were conducted. Part of the samples before and during the x-ray was in the Ringer's solution. An additional set of samples was issued to permanent ionizing radiation from CT over a period of four weeks.

Samples needed for the study were in the shape of flats and were prepared in accordance with current standards PN-EN ISO 527-1,2:2012E and PN-EN ISO 178:2011E.

Tests were performed on the following series of samples:

1. after 3 weeks of immersion in Ringer's solution, irradiated 3 times (a) and 10 times (b) by electron beam by varying the value of the dose,
2. without immersion in Ringer solution, irradiated 3 times (a) and 10 times (b) by electron beam with different dose values,
3. placed under the permanent influence of CT over a period of four weeks, a known dose values,
4. reference samples: free from ionizing radiation.

The value of the radiation dose is not constant, as it was selected automatically by X-ray device, adequate for the weight of the patient. The values of the doses

Table 1. Value doses [μGy/m^2], which material have been irradiated

serial number	series Ia (3R)	series Ib (10R)	series IIa (3)	series IIb (10)	series III (CT)
dose value [μGy/m^2]	**674.7**	**4471.5**	**1626.9**	**3796.2**	**466435.5**

received by the sample distinguishing between the series and the number of X-rays are shown in Tab. 1.

Visual Assessment of Surface. Visual assessment of samples surface was to identify potential changes on the surface of samples that may have occurred after exposure to X-rays. Macroscopic observation of the surface was performed using a stereoscopic microscope Zeiss Discovery V8, 25x magnification The samples were observed for a color change, discoloration, and surface defects, such as deformation, cracks, pits or chipping.

Density Measurement - Immersion Method. For the determination of the density the five samples from each batch were selected. Density measurement was performed in accordance with the PN-EN ISO 183-1:2006, by A - immersion method. The analytical weight Mettler Toledo XS105 DualRange with an accuracy of $\pm$ 0.01 mg was used for the test.

The Hardness Measurement - Shore Method. The selected samples were subjected to hardness measurements by elastic rebound (Shore method). For the measurements the Shore hardness tester (type D) was used. Designation of hardness by pressing was made on the basis on the PN-EN ISO 868:2005P standard. Three samples of each series were tested.

Mechanical Properties. Determination of mechanical properties has been carried out using a static tensile test, and the static bending test. MTS Criterion Model 45 testing machine was used for the study. Static tensile test has been carried out based on recommendations of the PN-EN ISO 527-1,2:2012E standard. Static bending test was performed under the specified conditions of the PN-EN ISO 178:20011E standard.

3 Results

Visual Assessment of Surface. The UHMWPE samples tested were free from changes in the quality of the surface, color change or mechanical damage, such as cracks, dents, chipping have not been observed. The surface condition of the test material is shown in Fig. 1.

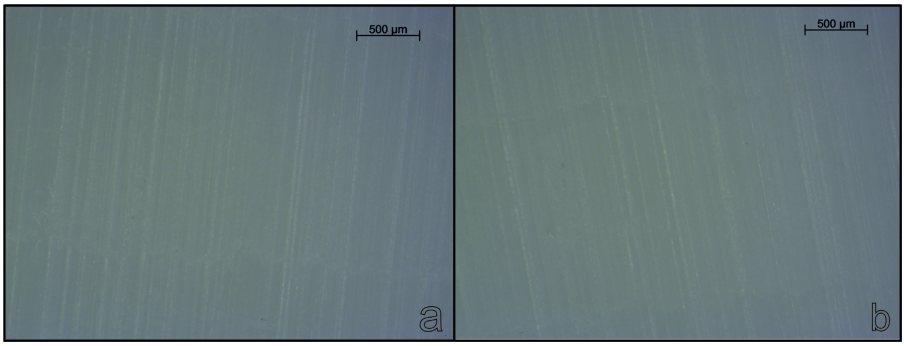


Fig. 1. Macroscopic evaluation of the surface UHMPWE at a magnification of 25x: the reference sample (a), the material was subjected to the action of X-rays (series III - TK) (b)

Table 2. The results obtained for density test of UHMWPE [g/cm^3] - mean

serial number	series Ia (3R)	series Ib (10R)	series IIa (3)	series IIb (10)	series III (CT)	series IV (ref)
density [g/cm^3]	**0.957**	**0.951**	**0.954**	**0.953**	**0.951**	**0.954**

Table 3. Results of hardness measurement, Shore method - mean

serial number	series Ia (3R)	series Ib (10R)	series IIa (3)	series IIb (10)	series III (CT)	series IV (ref)
Shore hardness [D/15:result]	**60.9**	**60.5**	**60.6**	**60.2**	**60.7**	**60.7**

Density Measurement. The results obtained during the determination of the polyethylene density by immersion method of the X-rays treated samples, as well as a comparison with the reference samples are shown in Tab. 2.

Hardness. Results of hardness testing are shown in Tab. 3.

Mechanical Properties. The studies performed for the UHMWPE samples have a comparative nature. Individual series of samples which have been designated for the test did not received the same, constant X-rays dose, thus in order allow them to compare and examine the influence of ionizing radiation on ultra-high molecular weight polyethylene, results were sorted by the dose of radiation and are presented in Tab. 4.

Examples of bending and tensile curves are shown in Fig. 2 and Fig. 3. Analyzing the tensile strength (σM) minimal decrease in samples exposed to X-rays was noted. There is a slight difference with respect to the Young's modulus of samples exposed to ionizing radiation, and the reference samples. Elongation

Table 4. The results obtained during the all tests - average value

series / property	tensile strength, σ_M [MPa]	Young's modulus, E_t [MPa]	elongation, A [%]	flexural modulus, E_f [MPa]
CT	**26.5**	**1389**	**129.5**	**1400**
10R	**26.2**	**1373**	**-**	**1492**
10	**26.5**	**1387**	**132.0**	**1362**
3	**26.5**	**1302**	**196.0**	**1464**
3R	**26.3**	**1311**	**-**	**1439**
ref	**26.7**	**1404**	**169.5**	**1350**

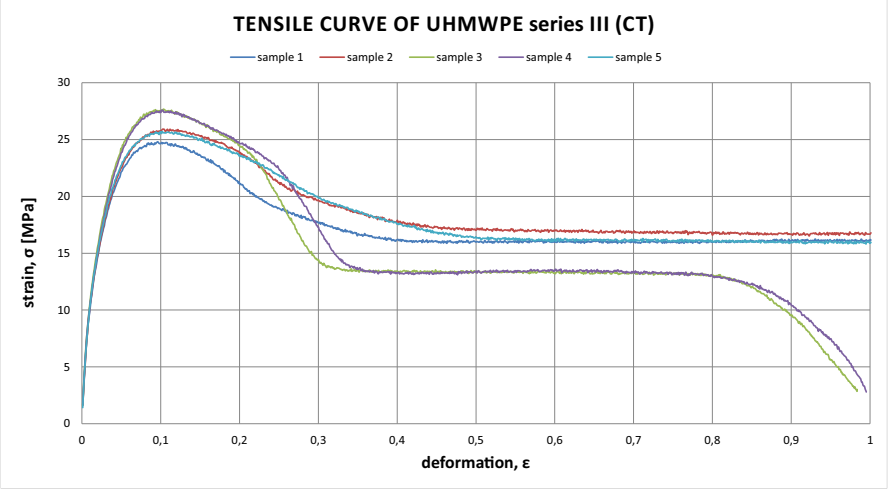


Fig. 2. Example static tensile curves of the samples exposed to CT radiation

is a parameter that cannot be compared, because not all of the samples have been broken. Analysis of the recorded values of flexural modulus Ef showed no influence of the dose of ionizing radiation on the value of the module.

4 Discussion of the Results

The study of the material (UHMWPE) were comparative in character. The different series of samples that were designated for the research did not absorb the same, constant dose of X-rays. The most exposure to ionizing radiation was the series III (CT), which adopted a dose of 466 435.5 μGy/m^2, while the series Ia (3R) took the lowest dose of radiation, namely 674.7 μGy/m^2, irradiated 3 times of immersion in Ringer's solution. Reference material was not subjected to X-ray, a series of IV (ref) was used for comparative purposes. According to the standards for ultra-high molecular weight polyethylene, its density should be within the range of $0.93 \div 0.97$ g/cm^3. The density of the tested polyethylene

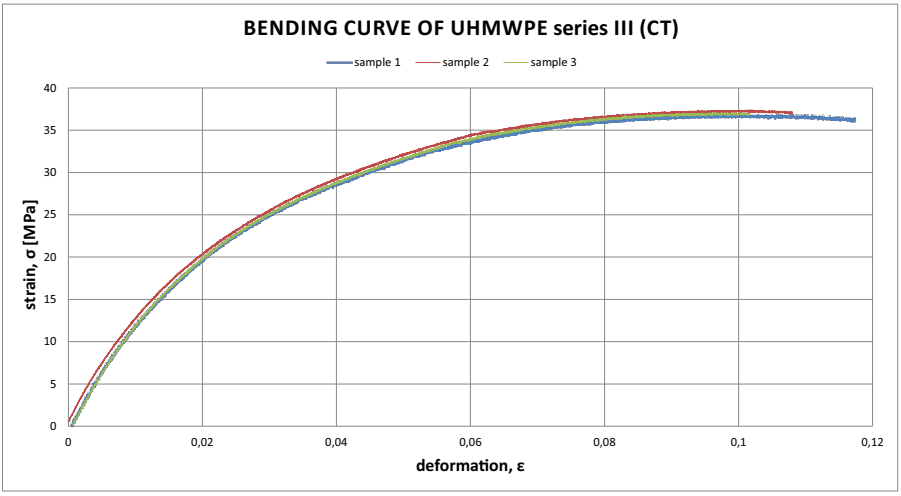


Fig. 3. Example static bending curves for the samples of the series III (CT)

was 0.95 g/cm^3. Change in density is so small that it can be concluded that the quantity of ionizing radiation does not affect the density of the UHMWPE. According to the results obtained during the measurement of a Shore hardness (type D) arranged in 3 it can be seen that the hardness of the test material is in the range D 60÷D 61. It can be concluded that the hardness of all samples of UHMWPE, regardless of whether they have been subjected to ionizing radiation, or does not, have similar values. Analyzing the tensile strength was noted that σ_M the reference material is 26.7 MPa, while the samples treated with X-rays, regardless of the series and the dose falls slightly, reaching a maximum of 26.5 MPa. The Young's modulus of the reference samples also varies slightly with respect to the elastic modulus of the samples exposed to ionizing radiation. Young's modulus for the reference material is greatest, namely 1404.3 MPa, while for the other series is set in a border of 1302.3÷1389.8 MPa. The lowest Young's modulus reached samples of 3R series, 3 times irradiation. The highest value of the flexural modulus of the UHMWPE was determined for the series of 10R, and the lowest modulus value obtained the reference samples.

5 Summary

All values obtained during a studies are in accordance with the standards and did not change significantly. On the basis of density measurements of the ultra-high molecular weight polyethylene, it was found that ionizing radiation has no effect on the density of the tested material. No influence of ionizing radiation on the mechanical properties (Young's modulus, tensile strength, flexural strength, hardness) was observed either. No significant changes in the mechanical properties of the UHMWPE under the influence of ionizing radiation suggest that

patients with endoprostheses whose components are made of the tested material can be subjected to diagnostic radiation without fear of loss of functional properties of these implants.

Acknowledgements The authors wish to thank the Manager of the Department of Diagnostic Imaging Independent Public Provincial Hospital of Traumatology Dr. J. Daaba in Piekary Slaskie, MD Mr. Wojciech Wawrzynek for consultation and exposures of the tested material.

References

1. Kurtz, M.: UHMWPE Biomaterials Handbook, 2nd edn. Ultra-High Molecular Weight Polyethylene in Total Joint Replacement and Medical Devices. Elsevier Inc. (2009)
2. Podrez-Radziszewska, M., Haimann, K., Dudzinski, W.: Zużycie tribologiczne polietylenu UHMWPE w endoprotezach stawu biodrowego i kolanowego. Prace Naukowe Instytutu Konstrukcji i Eksploatacji Maszyn Politechniki Wrocławskiej 87(27), 285–292 (2002)
3. Marciniak, J.: Biomateriały. Wydawnictwo Politechniki Ślaskiej (2013)
4. Nałecz, M.: Biocybernetyka i inżynieria biomedyczna 2000, tom 4 Biomateriały. Akademicka Oficyna Wydawnicza EXIT (2003)
5. Kolczyk, E., Sobczyk, K., Myalski, J., Balin, A.: Badanie wpływu czasu na własciwosci mechaniczne polietylenowych panewek endoprotez stawu biodrowego. Engineering of Biomaterials 11(74), 11–15 (2008)
6. Dobrzanski, L.: Podstawy nauki o materiałach i metałoznawstwo: materiały inżynierskie z podstawami projektowania materiałowego. Wydawnictwo Naukowo-Techniczne (2002)
7. Gierzynska-Dolna, M.: Biotribologia. Wydawnictwo Politechniki Częstochowskiej (2002)
8. Kostrzewa, B., Szewieczek, D.: Wybrane zagadnienia wpływu oddziaływania promieniowania jonizującego na biopolimery. Inżynieria Materiałowa 10(2), 82–85 (1999)
9. Sudol, M., Czaja, K.: Effect of electron beam irradiation and plastic deformation on molar mass of medical grade ultra-high molecular weight polyethylene. Polimery 55(11-12), 890–893 (2010)
10. Burger, N.D.L., de Vaalb, P.L., Meyera, J.P.: Failure analysis on retrieved ultra-high molecular weight polyethylene (UHMWPE) acetabular cups. Engineering Failure Analysis 14(7), 1329–1345 (2007)
11. Gonzalez-Moraa, V.A., Hoffmanna, M., Stroosnijdera, R., Gilb, F.J.: Wear tests in a hip joint simulator of different CoCrMo counterfaces on UHMWPE. Materials Science and Engineering 29(1), 153–158 (2009)
12. Pruszynski, B.: Radiologia: diagnostyka obrazowa RTG, TK, USG, MR i radioizotopy, wyd. II. Wydawnictwo Lekarskie PZWL (2005)

EIS Study of SiO_2 Oxide Film on 316L Stainless Steel for Cardiac Implants

Witold Walke[1], Zbigniew Paszenda[1], Marcin Basiaga[1], Paweł Karasiński[2], and Marcin Kaczmarek[1]

[1] Faculty of Biomedical Engineering, Department of Biomaterials and Medical Devices Engineering, Silesian University of Technology, Zabrze, Poland
witold.walke@polsl.pl
[2] Faculty of Electrical Engineering, Department of Optoelectronics, Silesian University of Technology, Gliwice, Poland

Abstract. The correct relationship between implant and tissue environment provides a properly prepared surface of a medical device. The combination of advantages of metal implants with the required biocompatibility is achieved through various methods of surface treatment such as: electrochemical polishing, chemical passivation, preparation of oxide coatings obtained with the sol-gel method. The relatively short history of use of silica layers in biomedical applications transfers into a small amount of literature relating to the electrochemical properties of these layers. Therefore, the authors proposed in the work forming a silica layer of the established technological parameters on surface of the 316L stainless steel. In order to determine the usefulness of this technology electrochemical impedance spectroscopy studies were performed. Furthermore, corrosion resistance tests by means of the potentiodynamic method were also carried out. The obtained results were the basis for optimization of technological conditions for deposition of silica layers on the 316L steel surface.

Keywords: EIS, corrosion resistance, stainless steel, cardiac implants.

1 Introduction

In the past several years there has been a significant reduction in mortality caused by circulatory system diseases. One important factor influencing favorably this process is the progress both in the diagnostic and treatment methods. Particularly important was the introduction and clinical dissemination of vascular stents. However, despite the undisputed advantages resulting from application of these implants, some disadvantage processes which limit the long term efficacy of vascular intervention are observed. It is mainly the process of blood clotting and musculo-fibrous proliferation of vessel's intima (restenosis). As a basic way to reduce these negative processes, development of physical and chemical properties of implants' surface layers is indicated. These properties fundamentally determine the reactivity of the implant surface with the environment

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_38, © Springer International Publishing Switzerland 2014

of the circulatory system. The limited effectiveness of currently proposed technology surface treatment results, among others, from the lack of comprehensive knowledge about mechanisms of initiation and development of blood clotting process and restenosis, and in consequence, vagueness of the already adopted criteria for assessing the final quality of implants. When considering the suitability assessment of the manufactured layers, one must take into account not only the chemical reactivity of the circulatory system, but also electromagnetic compatibility and the used implantation technique. One of the applicable surface treatment method is to use the sol-gel method to form thin oxide coatings based on elements such as: Ti, Si. The relatively short history of the use of silica layers in biomedical applications transfers into a small amount of literature relating to the electrochemical properties of these layers [1,2,3,4,5,6,7,8,9,10]. Therefore, the authors proposed in the work forming a silica layer of the established technological parameters on surface of the 316L stainless steel. In order to determine the usefulness of this technology electrochemical impedance spectroscopy studies were performed. Furthermore, corrosion resistance tests by means of the potentiodynamic method were also carried out.

2 Materials and Methods

Austenitic stainless steel (316 LVM) was used in the study. The steel of the chemical composition shown in Table 1, was in the form of discs with a diameter d = 14 mm and thickness h = 2 mm.

Table 1. Chemical composition of the analyzed steel

Element, % mass	AISI 316L Stainless steel	
	ISO 5832-1:2007	Analyzed steel
C	≤ 0.030	0.022
Si	≤ 1.0	0.59
Mn	≤ 2.0	1.67
P	≤ 0.025	0.015
S	≤ 0.01	0.001
Cr	$17.0 \div 19.0$	17.5
Mo	$2.25 \div 3.0$	2.75
Ni	$13.0 \div 15.0$	14.25

The analyzed samples were subjected to the surface treatment consisting of the following processes: grinding, electropolishing, chemical passivation and deposition of SiO_2 layer using the sol - gel method. Mechanical grinding process was carried out using sand papers with carborundum (SiC) abrasive of 500 grit ($Ra = 0.25\mu m$). The samples were then subjected to the electropolishing ($Ra = 0.14\mu m$), which was carried out in a chromic acid based bath (E-395 POLIGRAT Gmbh), and current density was in the range i = $10 \div 30$ A/cm^2. Subsequently, the samples were subjected to chemical passivation in 45% HNO_3

acid at the temperature of T = 60°C in time of t = 1 h. The final surface treatment step was deposition of the SiO_2 layer by the sol-gel method at various deposition parameters (chemical composition of sols and the temperature of annealing were variables) (1 - v = 2.5 cm/min, T = 400°C, t = 60 min; 2 - v = 2.5 cm/min, T = 430°C, t = 60 min; 3 - v = 2.5 cm/min , T = 460°C, t = 60 min; 4 - v = 2.5 cm/min , T = 490°C , t = 60 min). The silica precursor used in the study was tetraethoxysilane $Si(OC_2H_5)_4$ so called TEOS, and tetramethoxysilane $Si(OCH_3)_4$ so called TMOS. Other starting components include ethyl alcohol (EtOH) and water.

To assess the suitability of the proposed surface modification the authors proposed electrochemical studies in which both potentiodynamic and impedance tests were performed.

Study of corrosion resistance was performed with the use of potentiodynamic method by recording anodic polarization curves [11,12,13,14,15]. The test set consisted of a VoltaLab PGP201 potentiostat, a reference electrode (the NEK KP-113 saturated calomel electrode), an auxiliary electrode (the PtP-201 platinum electrode), an anode (the test sample) and a PC with the VoltaMaster 4 software. Before the study all the samples were cleaned in 96% ethanol using an ultrasonic cleaner (SONICA 1200M for approximately t = 6 min). Corrosion tests initiated at determining the open circuit potential EOCP in currentless conditions. Anodic polarization curves were recorded from the starting potential $E_{init} = E_{OCP}$ - 100 mV. The potential rate was equal to 3 mV/s. After the anodic current density reached 1 mA/cm2 the direction of polarization was changed.

In order to obtain additional information about the physicochemical properties of the analyzed samples' surfaces, the electrochemical impedance spectroscopy test were also performed [12,13]. The measurements were carried out using the Auto Lab measurement system equipped with 302N PGSTAT FRA2 (Frequency Response Analyser) module. The used measuring system has enabled research in the frequency range 104 ÷ 10-3Hz. In the study the impedance spectra were determined and the obtained measurement data were fitted to the equivalent circuit. On this basis the numerical values of resistance R and capacitance C of the analyzed system were determined. Impedance spectra of the analyzed systems were presented in the form of Nyquist diagrams for different values of frequency and in the form of Bode diagrams. The obtained EIS spectra were interpreted after adjusting the least squares method to the equivalent electrical system. All electrochemical studies were carried out in an artificial plasma at the temperature of T = 37±1°C and pH = 7.0±0.2 - Table 2.

3 Results

First, the studies using electrochemical impedance spectroscopy were performed. Impedance spectra recorded for the stainless steel samples of modified surface were shown in Fig. 1 and Table 3.

Table 2. The chemical composition of the artificial plasma

Compound	Concentration, g/dm^3
NaCl	6.800
$CaCl_2$	0.200
KCl	0.400
$MgSO_4$	0.100
$NaHCO_3$	2.200
Na_2HPO_4	0.126
NaH_2PO_4	0.026

To analyze the experimentally determined impedance spectra of corrosive systems for the analyzed stainless steel, an electric equivalent circuit, shown in Fig. 2, was applied.

Impedance spectra obtained for the samples were interpreted by comparison to an equivalent electrical circuit, which indicates the presence of a double layer (two time constants shown on the plot), where R_s is the electrolyte resistance, R_p - the resistance of the electrolyte in the pores, and CPE_P - capacitance of double layer (porous surface), while R_{ct} and CPE_{dl} - resistance and capacitance of the oxide layer. The use of two Constant Phase Elements in the equivalent electrical circuit positively influenced the fit quality of the curves determined experimentally.

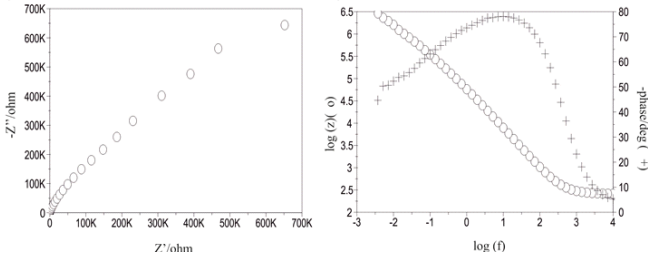
Mathematical impedance model of the above mentioned system presents the equation (1):

$$Z = R_s + \frac{1}{\frac{1}{R_p} + Y_{01}(j\omega)^{n_1}} + \frac{1}{\frac{1}{R_{ct}} + Y_{02}(j\omega)^{n_2}} \tag{1}$$

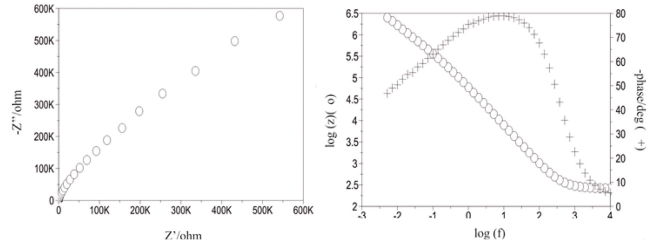
Table 3. The results of electrochemical impedance spectroscopy

Surface treatment	E_{OCP}, mV	R_s, Ωcm^2	R_p, kΩcm^2	CPE_p Y_{01}, $\Omega^{-1}cm^{-2}s^{-n}$	n_1	R_{ct}, kΩcm^2	CPE_{dl} Y_{02}, $\Omega^{-1}cm^{-2}s^{-n}$	n_3
(1)	-102	17	93	0.1882E-4	0.89	1480	0.2438E-4	0.89
(2)	-95	15	85	0.1853E-4	0.90	1274	0.2219E-4	0.90
(3)	-91	16	193	0.1013E-4	0.91	1677	0.1931E-4	0.90
(4)	-111	15	308	0.2141E-4	0.97	1623	0.9943E-5	0.86

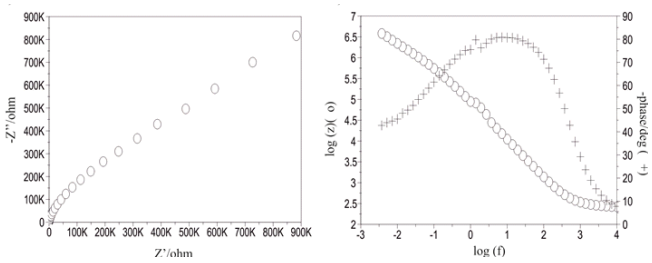
Supplementary potentiodynamic tests was also carried out by recording anodic polarization curves. On the basis of the test it was found that the corrosion potential for the samples subjected to the process of electropolishing and chemical passivation (initial state) reached E_{cor} = -108 mV (mean). Polarization of the samples caused an increase in the anodic current for the potentials in the range E_b = +1354 mV. When the anodic current density reached 1 mA/cm^2 the direction of polarization was changed, which allowed to register the return curve.



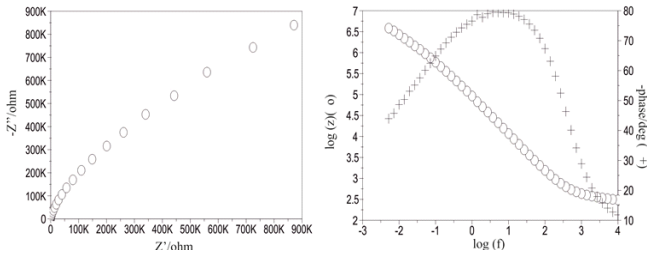
(a) Sample (1): Nyquist plot (left), Bode diagram (right)



(b) Sample (2): Nyquist plot (left), Bode diagram (right)



(c) Sample (3): Nyquist plot (left), Bode diagram (right)



(d) Sample (4): Nyquist plot (left), Bode diagram (right)

Fig. 1. Impedance spectra

Thus obtained anodic polarization curves were characterized by the presence of hysteresis loops, indicating the course of pitting corrosion. In addition, based on the Stern method a polarization resistance was determined, which was equal to $R_p = 550\ k\Omega cm^2$. A corrosion current density for all the samples was also determined and was equal to $i_{cor} = 0{,}012\ \mu A/cm^2$ - Table 4.

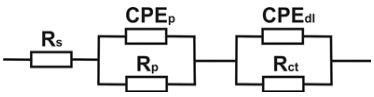


Fig. 2. Electrical equivalent circuit model for the stainless steel - oxide layer - artificial plasma system

Table 4. The results of potentiodynamic test for the analyzed samples

Surface treatment	E_{cor}, mV	E_b mV	R_p, kΩcm^2	i_{cor} μA/cm^2
Initial state	-108	+1354	550	0,012
1	-146	+1603	285	0.091
2	-137	+923	479	0.054
3	-131	+816	310	0.083
4	-121	+1610	385	0.067

Next, the samples with the SiO_2 layer deposited by the sol-gel method at different annealing temperatures were analyzed. Based on the study insignificant differences in the values of corrosion potential E_{cor} compared to the samples in the initial state were found - Table 4. Anodic polarization curves indicate a wide range of a passive region. The rapid increase of the anodic current was observed for the potentials in the range $E_b = +816 \div +1610$ mV. These values for the samples (1) and (4) were greater than the values obtained for samples after the process of electropolishing and chemical passivation. On the other hand, for the samples (2) and (3) a negative decrease of the breakdown potential, with reference to the initial state, was observed - Table 4. Furthermore, for these samples slight differences in the values of the polarization resistance were also found, and the corrosion current density i_{cor}, compared to the samples without the layer, were also found - Table 4.

4 Conclusions

As a basic way to reduce blood clotting and restenosis the appropriate physicochemical properties of implants' surface layers are indicated. One of the most progressive methods to improve the properties of the surface layers of metal biomaterials is the sol-gel method. The relatively short history of use of silica layers in biomedical applications transfers into a small amount of literature relating to the physical and chemical properties of these layers. Therefore, the authors of the work proposed a surface treatment of the stainless steel with a SiO_2 layer using the sol-gel method with the assumed process parameters in order to improve its hemocompatibility.

One of the methods to verify the usefulness of the surface modification are electrochemical studies. For this purpose, the authors proposed the impedance test and additionally potentiodynamic tests. The analyzed material was subjected to the following surface treatments: grinding, electropolishing, chemical

passivation and deposition of SiO_2 layers under controlled conditions by varying the chemical composition of the sols and the value of the annealing temperature in the range of 400°C to 500°C. The work is a continuation of research conducted by the authors regarding the applicability of the sol-gel method for surface modification of metal biomaterials with SiO_2/TiO_2 oxide used for implants in contact with blood [16,17,18].

The results of impedance measurements confirm the significant effect of surface treatment on the corrosion resistance of stainless steel. It was found that in all cases the impedance of the analyzed corrosive systems decreases with increasing frequency. In addition, it was also found that the phase angle also changes with frequency. The recorded impedance spectra for the analyzed samples indicate the electrochemical stability of the anodic layer formed on the stainless steel surface (electropolishing, chemical passivation and sol-gel method). Resistance of the oxide layer R_{ct} reaches large values indicating the appropriate properties of the layer forming material, its relatively high thickness, and a suitable phase composition which provides correct chemical structure. The presence of the surface layer with a high surface development on the surface of the oxide layer causes the additional barrier protecting the steel from the corrosive effects of the environment. In turn, the appearance of the surface layer is the result of the reaction with the oxide layer depleted corrosive environment , in this case an artificial plasma. This phenomenon may be the beginning of the initiation of the conduction of ion channels. This can cause a lack of layer tightness and further uncontrolled digestion, and consequently a breakdown during the polarization potential, for example, generated in the cells. Therefore, the potentiodynamic studies aimed at determining the values of the breakdown potential for the modified surface of the stainless steel. The results clearly show that the values of the potentials causing breakdown of the surface layer, occurring for each surface treatment variant, are much higher than those generated in tissues (max. up to +90 mV). This phenomenon should be considered very beneficial. In summary, the use of the sol-gel technology for coating the stainless steel used for blood contacting implants with SiO_2 is fully justified and useful for improving the safety of use.

Acknowledgement. The project was funded by the National Science Centre allocated on the basis of the decision No. 2011/03/B/ST8/06499.

References

1. Areva, S., Aaritalo, V., Tuusa, S., Jokinen, M., Linden, M., Peltola, T.: Sol-gel-derived TiO_2-SiO_2 implant coatings for direct tissue attachment. Journal of Materials Science, Materials Medicine 18, 1633–1642 (2007)
2. Huan, Z., Chang, J.: Study on physicochemical properties and in vitro bioactivity of tricalcium silicate-calcium carbonate composite bone cement. Journal of Materials Science, Materials Medicine 19, 2913–2918 (2008)
3. Patel, A., Knowles, J.C.: Investigation of silica-iron-phosphate glasses for tissue engineering. Journal of Materials Science, Materials Medicine 17, 937–944 (2006)

4. Zhao, W., Chang, J., Wang, J., Zhai, W., Wang, Z.: In vitro bioactivity of novel tricalcium silicate ceramics. J. Materials Science, Materials Medicine 18, 917–923 (2007)
5. Karasiński, P., Gondek, E., Drewniak, S., Kityk, I.V.: Nano-sized blue spectra shift in sol-gel derived mesoporous titania films. Journal of Sol-Gel Science and Technology 61, 355–361 (2012)
6. Klein, L.C.: Sol-gel optics, processing and application. Kluwer Academic Publishers (1994)
7. Brinker, C.J., Scherer, G.W.: Sol-gel science. Academic Press, Incorporated, San Diego (1990)
8. Karasiński, P.: Sol-gel derived optical waveguide films for planar sensors with phase modulation. Optica Applicata 34(4), 467–475 (2004)
9. Biswas, R.G., Sanders, R.D.: The Effects of a CeO_2 Coating on the Corrosion Parameters of Type 304 Stainless Steel. Journal of Materials Engineering and Performance 7, 727–732 (1998)
10. Xiao-kui, Y., Qing, L., Jun-ying, H., Xian-kang, Z., Shi-yan, Z.: The electrochemical corrosion behavior of sealed Ni-TiO_2 composite coating for sintered NdFeB magnet. Journal Applied Electrochemistry 40, 39–47 (2010)
11. ASTM F746-04: Standard Test Method for Pitting or Crevice Corrosion of Metallic Surgical Implant Materials (2009)
12. Przondziono, J., Walke, W.: Influence of the type of medical sterilisation on electrochemical properties of passive layer created on X10CrNi 18-8 steel. Solid State Phenomena 212, 141–144 (2014)
13. Kajzer, W., Kajzer, A.: Badania potencjodynamiczne i impedancyjne implantów do leczenia zniekształceń przedniej ściany klatki piersiowej. Przegląd Elektrotechniczny 12, 275–279 (2013)
14. Kajzer, W., Chrzanowski, W., Marciniak, J.: Corrosion resistance of Cr-Ni-Mo steel intended for urological stents. Int. J. Microstructure and Materials Properties 2(2), 188–201 (2007)
15. Dziekońska, M., Ziębowicz, A., Ziębowicz, B., Dobrzański, L.A.: Corrosion resistance of neodymium composite materials reinforced with metal powders. Archives of Materials Science and Engineering 58, 137–145 (2012)
16. Basiaga, M., Popik, P., Paszenda, Z., Pochrząst, M.: The research of electrochemical properties of silica layers used in medical devices intended for contact with blood. Przegląd elektrotechniczny (Electrical Review) 12(R 89), 374–377 (2013)
17. Basiaga, M., Walke, W., Paszenda, Z., Karasiński, P.: Research on electrochemical properties SiO_2 layer, intended for contact with blood, deposited by sol-gel method. European Cells and Materials 26(6), 157 (2013)
18. Walke, W., Paszenda, Z., Basiaga, M., Karasiński, P.: SiO_2 layer deposited by sol-gel method on the AISI 316L for contact with blood. Journal of Achievements in Materials and Manufacturing Engineering 56(2), 75–82 (2013)

Electrochemical Impedance Spectroscopy and Corrosion Resistance of SiO_2 Coated CpTi and Ti-6Al-7Nb Alloy

Marcin Basiaga[1], Zbigniew Paszenda[1], Witold Walke[1], Paweł Karasiński[2], and Jan Marciniak[1]

[1] Faculty of Biomedical Engineering, Department of Biomaterials and Medical Devices Engineering, Silesian University of Technology, Zabrze, Poland
marcin.basiaga@polsl.pl

[2] Faculty of Electrical Engineering, Department of Optoelectronics, Silesian University of Technology, Gliwice, Poland

Abstract. One of the ways to increase the hemocompatibility of surface of titanium and its alloys is to apply sol-gel method to form oxide films based on such elements as Ti or Si. However, the literature suggests existence of group of undefined phenomena associated with oxide layer formation on surfaces of metallic biomaterals in the presence of silicon. Therefore, the study involves development of conditions suitable for production of surface layers that have phisycal and chemical properties adequate for application in cardiovascular system. To assess the usefulness of proposed modifications, the Authors carried out electrochemical tests involving potentiodynamic and impedance measurements. The tests were performed in artificial serum in temperature T = 37±1°C in order to simulate real-life conditions. The obtained results show diverse properties of SiO_2 layer that depend on process parameters of its formation.

Keywords: EIS, corrosion resistance, cpTi, Ti-6Al-7Nb.

1 Introduction

The human body is an extremely complex combination of many organ systems. One of the most important systems, which ensures proper function of others, is the cardiovascular system. If any of its functions are disturbed, this fact affects also other systems and as a result operation of whole body is disregulated. Due to this, nowadays, when there is significant awareness of meaning of blood circulation for maintaining proper health, there is an increasing effort to discover new and develop existing methods of preventive healthcare, as well as new methods for cardiology treatments. Long-term observations of cardiovascular system functioning mechanisms, as well as numerous studies carried out by physicians in close cooperation with engineers allowed to model structures and biological processes related to blood circulation. This in turn allowed to achieve broad possibilities of performing new tests of whole system functioning and simulation of pathological situations, based on which it was possible to develop series of

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 4,* Advances in Intelligent Systems and Computing 284,
DOI: 10.1007/978-3-319-06596-0_39, © Springer International Publishing Switzerland 2014

implants and surgical methods restoring normal health conditions or as near to them as possible. The proper adhesion between implant and tissue environment is ensured by properly treated surface of medical product. Combination of advantages of metal implant structures with biocompatibility, essential in medical application, can be achieved by applying different methods of surface treatments. One of the surface treatments used for titanium and its alloys is the application of sol-gel method to form oxide films based on such elements as: Ti, Si. However, the literature suggests existence of series of undefined phenomena associated with oxide layer formation on surfaces of metallic biomaterals in the presence of silicon. These phenomena are confirmed by numerous publications in international journals (mainly in medical journals) [1,2,3,4,5,6,7,8]. However, they usually present only partial results of the studies, results that does not allow fully to assess usefulness of produced surfaces. Just as often, the publication Authors tend to omit issues concerning influence of processing parameteters of sol-gel method on electrochemical properties of silica coatings [7,8]. Therefore, this study aims to analyse elektrochemical properties of SiO_2 layers deposited on surfaces of titanium Grade 4 and Ti-6Al-7Nb alloy under various technological conditions. In order to evaluate electrochemical properties of deposited films, potentiodynamic and impedance measurements were performed.

2 Materials and Methods

In the study, the tests were performed on test pieces of Ti-6Al-7Nb alloy and titanium (Grade4) extracted from rod of diameter d = 14 mm [9,10]. The surfaces of test pieces were then modified in different ways, namely: grinding, electrolytic polishing and subsequent deposition of SiO_2 film by means of sol-gel method. The mechanical grinding was performed using waterpapers with carborundum abrasive of grain size 500 ($Ra = 0.25\mu m$). The test pieces were subsequently electrolytically polished ($Ra = 0.15\mu m$) in a bath based on chromic acid (POLIGRAT Gmbh E-395) with current density i = 10÷30 A/cm^2. In the final step SiO_2 layer was deposited by sol-gel method for various heating temperatures (1 - v =2.5 cm/min, T = 400°C, t = 60 min, 2 - v =2.5 cm/min, T = 430°C, t = 60 min, 3 - v =2.5 cm/min, T = 460°C, t = 60 min, 4 - v =2.5 cm/min, T = 490°C, t = 60 min). In the studies, tetraethoxysilane $Si(OC_2H_5)_4$, TEOS and tetramethoxysilane $Si(OCH_3)_4$, TMOS were used as silica precursors. The other initial compounds were ethyl alcohol (EtOH) and water.

For assessment of usefulness of proposed surface modification, the Authors have proposed electrochemical tests involving potentiodynamic and impedance measurements [11,12,13,14,15,16,17].

Corrosion resistance was tested using potentiodynamic method by recording anodic polarization curves. The measurement system consisted of potentiostat VoltaLab PGP201, reference electrode (saturated calomel electrode NEK KP-113), auxiliary electrode (platinum electrode PtP-201), anode (tested test piece) and PC with software VoltaMaster 4. Prior to the tests, the surfaces of all test pieces have been cleaned in ultrasonic washer in 96% ethyl alcohol for approx.

t = 6 min. The corrosion tests have been started from determination of open circuit potential E_{OCP} under no current conditions. The anodic polarization curves were recorded starting from initial potential $E_{init} = E_{OCP}$ - 100 mV. The potential was changed in anodic direction with rate 3 mV/s. After obtaining anodic current density equal to 1 mA/cm^2, the direction of polarization was reversed.

In order to obtain additional information on physical and chemical properties of examined test piece surfaces, the tests involving electrochemical impedance spectroscopy (EIS) were also performed. The measurements were carried out using measurement system Auto Lab PGSTAT 302N equipped with module FRA2 (Frequency Response Analyser). The measurement system allowed to perform tests in frequency range $10^4 \div 10^{-3}$Hz. The voltage amplitude of sinusoidal signal was 10 mV. The tests allowed to determine system impedance spectra and to fit obtained measurement data to auxiliary system. Based on that the values of resistance R and capacitance C of examined systems were determined. The impedance spectra for examined system have been presented in form of Nyquist diagrams for different frequencies and as Bode diagrams. The obtained EIS spectra have been analysed after fitting using least squares method to auxiliary electric system. All electrochemical tests were carried out in synthetic serum of temperature T = 37±1°C and pH = 7.0±0.2 - Table 1.

Table 1. Chemical composition of synthetic serum

Chemical compound	Concentration, g/dm^3
NaCl	6.800
$CaCl_2$	0.200
KCl	0.400
$MgSO_4$	0.100
$NaHCO_3$	2.200
Na_2HPO_4	0.126
NaH_2PO_4	0.026

3 Results

The test results of pitting corrosion resistance of cpTi titanium test pieces and Ti-6Al-7Nb test pieces with differently treated surfaces are presented in Table 2 and on Figures 1 and 2. The performed tests allowed to determine influence of test piece surface treatment on its corrosion resistance.

Based on the obtained results, it was concluded that the corrosive potential for test pieces in initial state (after electrolytic polishing) has mean value of E_{cor} = -168 mV (cpTi Grade 4) and E_{cor} = -269 mV (Ti-6Al-7Nb alloy). The curves of anodic polarization determined this way have suggested the existence of passive range up to value of potential E = +4000 mV. In this case the rapid increase of anodic current has not been observed, which otherwise would suggest initiation

of pitting corrosion. Additionally, using Stern method the value of polarization resistance was determined as R_p = 159 kΩcm^2 (cpTi Grade 4) and R_p = 369 kΩcm^2 (Ti-6Al-7Nb alloy). Moreover, the value of corrosive current density was determined as equal to i_{cor} = 0.163 μA/cm^2 (cpTi Grade 4) and i_{cor} = 0.070 μA/cm^2 (Ti-6Al-7Nb alloy).

Subsequently, the test pieces with SiO_2 film deposited by sol-gel method for different heating temperatures were examined. Based on the performed tests it was found that regardless of used version of deposition (various heating temperatures) and type of base this process caused advantageous increase of polarization resistance R_p in comparison with initial test pieces. This is proven by values of parameters characterizing their corrosive resistance – Table 2, Figures 1, 2. Moreover, the advantageous decrease of corrosive current density and the minor decrease of corrosive potential E_{cor} were observed.

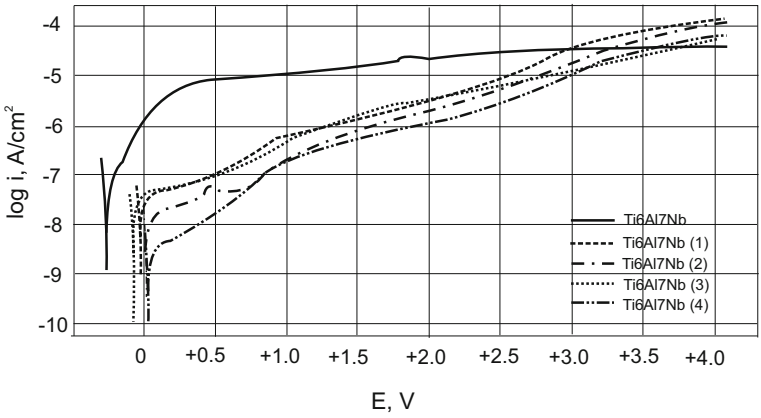


Fig. 1. Anodic polarization curves for Ti-6Al-7Nb titanium alloy test pieces

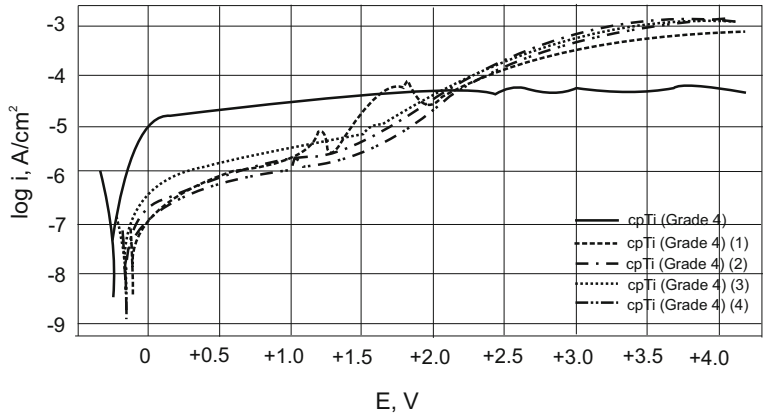


Fig. 2. Anodic polarization curves for cpTi (Grade 4) test pieces

Table 2. Results of potentiodynamic tests of analysed test pieces

Method of test piece surface treatment		E_{cor}, mV	i_{cor}, $\mu A/cm^2$	R_p, $k\Omega cm^2$
cpTi Grade 4	Initial state	-168	0.163	159
	1	-133	0.083	312
	2	-178	0.055	468
	3	-194	0.076	342
	4	-172	0.042	606
Ti-6Al-7Nb Alloy	Initial state	-269	0.070	369
	1	-33	0.040	586
	2	+17	0.017	1460
	3	-82	0.036	709
	4	+26	0.003	8280

In order to identify the nature of generated oxide film on surface of selected biomaterials, measurements using electrochemical impedance spectroscopy were performed in the subsequent stage of studies.Typical impedance spectra recorded for test pieces of titanium Grade 4 and Ti-6Al-7Nb alloy with electrolytically polished surface and for surface with deposited SiO2 film are presented on Fig. 3, 4. While obtained electric properties determined on the basis of recorded spectra are summarized in Table 3.

For analysis of experimentally determined impedance spectra of corrosive systems: cpTi and Ti-6Al-7Nb alloy, the auxiliary electric system was used presented on Figure 5.

Impedance spectra for test pieces cpTi and Ti-6Al-7Nb have been analysed by comparison with auxiliary electric system that suggests presence of double layer, where R_s corresponds to electrolyte resistance, R_p - resistance of electrolyte in pores and CPE_p - capacitance of double layer (porous, interfacial), while R_{ct} and CPE_{dl} - resistance and capacitance of oxide layer. The use of two constant phase elements in electric auxiliary circuit have favourable affected fit of curves determined experimentally [18,19]. The mathematical model of impedance of above system can be described by equation (1):

$$Z = R_s + \frac{1}{\frac{1}{R_p} + Y_{01}(j\omega)^{n_1}} + \frac{1}{\frac{1}{R_{ct}} + Y_{02}(j\omega)^{n_2}} \tag{1}$$

Only for Ti-6Al-7Nb alloy test piece (4) the fitting was done by approximation of experimental data using different model of electric auxiliary circuit - Fig. 6.

In electric auxiliary system the resistor R_{ct} and Constant Phase Element CPE represents ion transition resistance and capacitance of passive oxide layer deposited on alloy surface, respectively. While resistor R_s reflects resistance of synthetic serum, in which the tests were performed. The mathematical model of system impedance is described by equation (2):

$$Z = R_s + \frac{1}{\frac{1}{R_{ct}} + Y_0(j\omega)^n} \tag{2}$$

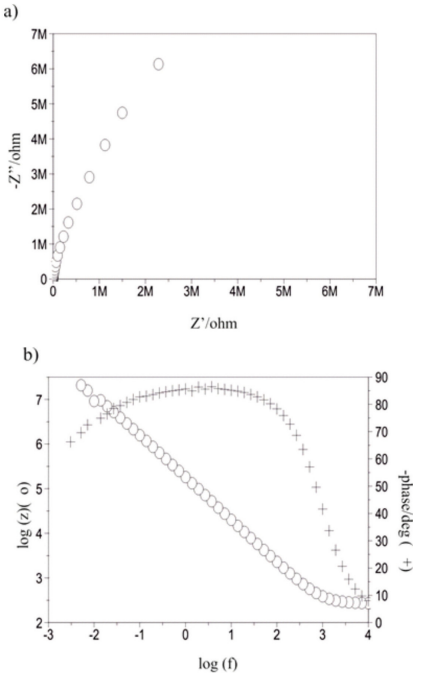


Fig. 3. Impedance spectra for Ti-6Al-7Nb titanium alloy test piece (4): a) Nyquist diagram, b) Bode diagram

Table 3. EIS results for analyzed test pieces

Type of surface modification	E_{OCP}, mV	R_s, Ωcm^2	R_p, $k\Omega cm^2$	CPE_p Y_{01}	n_1	R_{ct}, $k\Omega cm^2$	CPE_{dl} Y_{02}	n_3
				cpTi				
1	-144	17	162	0.6326E-5	0.92	2103	0.1200E-4	0.91
2	-119	16	188	0.6605E-5	0.93	1590	0.2065E-4	0.88
3	-84	17	162	0.5103E-5	0.92	1506	0.1684E-4	0.88
4	-151	189	189	0.5499E-5	0.92	1118	0.1957E-4	0.89
				Ti-6Al-7Nb				
1	-127	16	115	0.6225E-5	0.97	9580	0.7959E-5	0.91
2	-191	16	53	0.9823E-5	0.98	9440	0.5294E-5	0.93
3	-56	17	170	0.9231E-5	0.95	27170	0.5032E-5	0.94
4	-70	15	-	-	-	14580	0.3255E-5	0.94

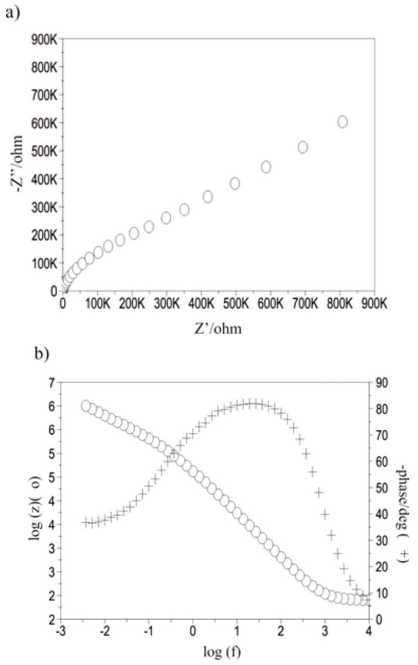


Fig. 4. Impedance spectra for cpTi test piece (4): a) Nyquist diagram, b) Bode diagram

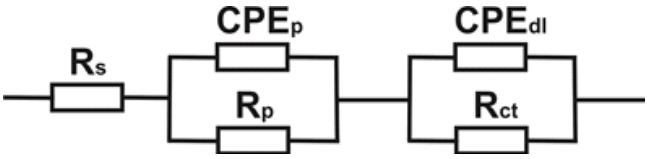


Fig. 5. Model of electric auxiliary circuit for system: cpTi/Ti-6Al-7Nb - oxide layer - synthetic serum

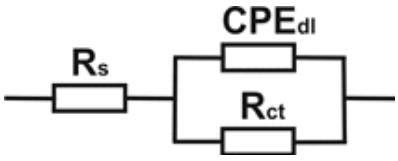


Fig. 6. Model of electric auxiliary circuit for system: Ti-6Al-7Nb - oxide layer - synthetic serum

4 Conclusions

The techniques of surface modification play very important role in tailoring physical and chemical properties of titanium and its alloys, as well as

improving their hemocompatibility. One of the ways to increase the hemocompatibility of titanium-based biomaterials is application of sol-gel method for formation of oxide films based on such elements as Ti and Si. Therefore, the study proposes modification of surfaces of pure titanium Grade4 and Ti-6Al-7Nb alloy with SiO_2 using sol-gel method. In order to assess usefulness of proposed modification the Authors suggested electrochemical tests, involving potentiodynamic an impedance measurements. The formation of SiO_2 was obtained under controlled conditions, for different chemical compositions of sols and various parameters of deposition process. This study is a continuation of research performed by Authors concerning usefulness of application of sol-gel method for modification of surfaces of Ti-6Al-7Nb alloy and titanium cpTi (Grade 4) with SiO_2 used as material for implants for contact with blood [20,21].

The performed electrochemical tests of oxide films deposited using sol-gel method under technological conditions proposed by the Authors have shown differences in corrosive resistance of studied biomaterials in contact with blood. In the case of titanium cpTi, the porous oxide layer of large surface area was found deposited on the biomaterial surface. This layer have formed as a result of partial degradation of oxide film during its contact with synthetic serum. This porous layer is additional barrier protecting material from corrosive environment. The determined values of ion transition resistance for specific variants prove that oxide layers formed in lower temperatures are more stable. In turn, for Ti-6Al-7Nb alloy, the identical nature of surface layer was observed for heating temperatures lower than T = 460°C. In this case the inverse relation was found - the higher the heating temperature, the higher value of ion transition resistance suggesting good properties ensuring protection against synthetic serum. Different nature of surface layer was found for heating temperature T = 490°C. The tests have shown presence of single oxide layer with relatively good electrical parameters that allow to protect alloy from synthetic serum. The high value of ion transition resistance obtained in this case, results in high stability of layer and no reaction with synthetic serum when in contact, which is advantageous.

While potentiodynamic measurements for all cases studied here (regardless the used base), have shown no hysteresis in range up to +4V, which would be a sign of pitting corrosion. It has been confirmed by EIS measurement, where no Warburg elements has been found which presence would mean possible initiation of corrosive phenomena.

In conclusion, the conditions proposed here for SiO_2 layer deposition have advantageous effects on improvement of corrosive resistance. These layers have high tightness, thus preventing direct contact of synthetic serum with surface of titanium or its alloy. Moreover, it was found that the heating temperature of oxide layer affects its nature and electrochemical properties that directly affect corrosive resistance. For cpTi, the temperature increase had no advantageous effect on oxide layer electrochemical properties, while for Ti-6Al-7Nb the higher the heating temperature, the better properties protecting from corrosive environment.

Acknowledgement. The project was funded by the National Science Centre allocated on the basis of the decision No. 2011/03/B/ST8/06499.

References

1. Karasiński, P., Jaglarz, J., Reben, M., Skoczek, E., Mazur, J.: Porous silica xerogel films as antireflective coatings - Fabrication and characterization. Optical Materials 33, 1989–1994 (2011)
2. Areva, S., Aaritalo, V., Tuusa, S., Jokinen, M., Linden, M., Peltola, T.: Sol-gel-derived TiO2-SiO2 implant coatings for direct tissue attachment. Journal of Materials Science, Materials Medicine 18, 1633–1642 (2007)
3. Karasiński, P.: Influence of technological parameters on the properties of sol-gel silica films. Optica Applicata 26, 253–263 (2005)
4. Brinker, C.J., Scherer, G.W.: Sol-gel science. Academic Press, Inc., San Diego (1990)
5. Klein, L.C.: Sol-gel optics, processing and application. Kluwer Academic Publishers (1994)
6. Brinker, C.J., Scherer, G.W.: Sol-gel science. Academic Press, Incorporated, San Diego (1990)
7. Shibli, S., Mathai, S.: Development and bio-electrochemical characterization of a novel TiO2-SiO2 mixed oxide coating for titanium implants. Journal of Material Science 19, 2971–2981 (2008)
8. Huan, Z., Chang, J.: Study on physicochemical properties and in vitro bioactivity of tricalcium silicate-calcium carbonate composite bone cement. Journal of Materials Science, Materials Medicine 19, 2913–2918 (2008)
9. ISO 5832-2:1999 Implants for surgery - Metallic materials - Part 2: Unalloyed titanium
10. ISO 5832-11:2007 Implants for surgery - Metallic materials - Part 11: Wrought titanium 6-aluminium 7-niobium alloy
11. ASTM F746-04: Standard Test Method for Pitting or Crevice Corrosion of Metallic Surgical Implant Materials (2009)
12. Kajzer, W., Kajzer, A.: Badania potencjodynamiczne i impedancyjne implantów do leczenia zniekształceń przedniej ściany klatki piersiowej. Przegląd Elektrotechniczny 12, 275–279 (2013)
13. Kiel, M., Szewczenko, J., Walke, W., Marciniak, J.: Application of EIS method for evaluation of physicochemical properties of modified ti-6al-4v ELI alloy. Przegląd Elektrotechniczny (Electrical Review) 88(12B), 232–235 (2012)
14. Szewczenko, J., Jaglarz, J., Basiaga, M., Kurzyk, J., Paszenda, Z.: Optical methods applied in thickness and topography testing of passive layers on implantable titanium alloys. Optica Applicata 43(1), 173–180 (2013)
15. Dziekońska, M., Ziębowicz, A., Ziębowicz, B., Dobrzański, L.A.: Corrosion resistance of neodymium composite materials reinforced with metal powders. Archives of Materials Science and Engineering 58, 137–145 (2012)
16. Pochrząst, M., Marciniak, J., Basiaga, M., Walke, W.: Comparison of electrochemical properties of surface modified Ti-6Al-4V Eli and Ti-6Al-7Nb alloys used in medicine. Przegląd elektrotechniczny (Electrical Review) 12(R 89), 310–313 (2013)
17. Basiaga, M., Popik, P., Paszenda, Z., Pochrząst, M.: The research of electrochemical properties of silica layers used in medical devices intended for contact with blood. Przegląd Elektrotechniczny (Electrical Review) 12(R 89), 374–377 (2013)

18. Kaczmarek, M., Walke, W., Paszenda, Z.: Application of electrochemical impedance spectroscopy in evaluation of corrosion resistance of Ni-Ti alloy. Przegląd Elektrotechniczny (Electrical Review) 87(12B), 74–77 (2011)
19. Szewczenko, J., Pochrząst, M., Walke, W.: Evaluation of electrochemical properties of modified Ti-6Al-4V ELI alloy. Przegląd Elektrotechniczny (Electrical Review) 87(12B), 177–180 (2011)
20. Walke, W., Paszenda, Z., Basiaga, M., Karasiński, P.: SiO_2 layer deposited by sol-gel method on the AISI 316L for contact with blood. Journal of Achievements in Materials and Manufacturing Engineering 56(2), 75–82 (2013)
21. Basiaga, M., Walke, W., Paszenda, Z., Karasiński, P.: Research on electrochemical properties SiO_2 layer, intended for contact with blood, deposited by sol-gel method. European Cells and Materials 26(6), 157 (2013)

Author Index