Advances in Intelligent Systems and Computing 471

Ewa Piętka
Pawel Badura
Jacek Kawa
Wojciech Wieclawek *Editors*

Information Technologies in Medicine

5th International Conference, ITIB 2016 Kamień Śląski, Poland, June 20–22, 2016 Proceedings, Volume 1

Springer

Advances in Intelligent Systems and Computing

Volume 471

Series editor

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

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

More information about this series at http://www.springer.com/series/11156

Ewa Piętka · Pawel Badura
Jacek Kawa · Wojciech Wieclawek
Editors

Information Technologies in Medicine

5th International Conference, ITIB 2016
Kamień Śląski, Poland, June 20–22, 2016
Proceedings, Volume 1

Springer

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

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

Pawel Badura
Faculty of Biomedical Engineering
Silesian University of Technology
Gliwice
Poland

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

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-319-39795-5 ISBN 978-3-319-39796-2 (eBook)
DOI 10.1007/978-3-319-39796-2

Library of Congress Control Number: 2016940806

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

After many years of research and development in the area of computer-assisted diagnosis and therapy, multimodal medical imaging systems prove the ability to provide quantitative and qualitative information leading to increase the sensitivity and specificity of many clinical procedures. Often the aspect of multimodal analysis is extended to other domains including signals, video recordings, posture information, histopathological image quantification, patient tracking, and retrospective and prospective analyses, all combined with image data or one with another. Mathematical information analysis, computer applications together with medical equipment and instruments have become standard tools underpinning the current rapid progress with developing computational intelligence. However, modern systems still face challenges in performing hybrid data registration, subsequent analysis, and visualization to deliver pertinent quantitative information. We are witnessing a radical change as information 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.

The *Fifth Conference on Information Technologies in Biomedicine* has become a recognized event that helps to bridge the gap between methodological achievements in engineering and clinical requirements in medical diagnosis, therapy, and rehabilitation. In these proceedings, members of academic societies of technical and medical background present their research results and clinical implementations. Many of these areas are recognized as research and development frontiers in employing new technologies in clinical environment.

Image and signal analysis are traditional parts that deal with the problem of data processing, recognition, and classification. Novel approaches to the 2D and 3D image analysis of various pathologically affected anatomical structures, early tumor detection, and stroke diagnosis are introduced. New research studies developed for the processing of heart, brain, and gastric as well as acoustic signals are discussed. Home care, remote rehabilitation assistance, and safety of the elderly require new areas to be explored in telemedicine, telegeriatrics, and biomechatronics. Bioinformatics has become a dynamically developed field of computer-assisted

biological data analysis where information techniques are used in the nanoscale. Solutions of complex problems concerning diagnostics and therapy have been developed in the surgical procedures. A modeling and simulation platform presents also preimplementation studies required at the developmental stage of computer-assisted therapy.

We would like to express our gratitude to the authors who contributed their original research papers as well as to all reviews and *Scientific Committee Members* for their valuable comments.

Ewa Piętka

Scientific Committee Members

E. Piętka (Poland)
M. Akay (USA)
P. Augustyniak (Poland)
A. Bargieła (Great Britain)
R. Bruck (Germany)
K. Capova (Slovakia)
S. Czudek (Czech Republik)
A. Drygajło (Switzerland)
A. Gertych (USA)
D. Greenhalgh (Great Britain)
A. Grzanka (Poland)
M. Grzegorzek (Germany)
M. Gzik (Poland)
A. Hajdasiński (Netherlands)
J. Haueisen (Germany)
Z. Hippe (Poland)
M. Juszczyk (Germany)
E. Krupinski (USA)
M. Kurzyński (Poland)
A. Liebert (Poland)
A. Lipowicz (Poland)
R. Maniewski (Poland)
J. Marciniak (Poland)
M. McNitt-Gray (USA)
A. Mitas (Poland)
A. Napieralski (Poland)
E. Neri (Italy)
A. Nowakowski (Poland)
D. Paulus (Germany)
T. Pałko (Poland)
Z. Paszenda (Poland)

W. Pedrycz (Canada)
I. Provaznik (Czech Republik)
A. Przelaskowski (Poland)
P. Słomka (USA)
E. Supriyanto (Malaysia)
A. Świerniak (Poland)
R. Tadeusiewicz (Poland)
E. Tkacz (Poland)
H. Witte (Germany)
S. Wong (USA)
Z. Wróbel (Poland)

Contents

Part I
Image Processing

A Statistical Dominance Algorithm for Edge Detection and Segmentation of Medical Images

Adam Piórkowski

Abstract This article proposes an algorithm which performs the initial stage of edge detection or segmentation. The algorithm counts the number of pixels with a given relation to the central point of the neighborhood. The output image is a statistical result of the dominance of points over their neighborhoods and allows the classification of these points to be determined (peak, valley, and slope). Therefore, this solution allows the impact of noise or uneven illumination in image results to be reduced. The basic features of the proposed algorithm are considered in this paper. Application of the algorithm is illustrated in the context of image segmentation of a corneal endothelium with a specular microscope, images of specimens of the brain tissue, and hand radiographs.

Keywords Edge detection · Segmentation · Cell counting · Image processing · Preprocessing · Microscopy images

1 Introduction

The analysis of medical images requires appropriate data preprocessing and ends with binarization and/or segmentation. The aim is to analyze the location, size or shape of objects by extracting them in binary form or in the form of their boundaries. The diversity of data often does not allow the use of well-known algorithms to set parameters without an adaptive approach. Binarization methods used, such as the one proposed by Otsu [15], or Top-Hat and edge detection used by Prewitt, Canny [4] does not always yield the required results. In this article a new method is proposed that effectively performs the binarization or supports segmentation of a fairly large class of images.

A. Piórkowski (✉)
Department of Geoinfomatics and Applied Computer Science,
AGH University of Science and Technology, A. Mickiewicza 30 Av., Cracow, Poland
e-mail: pioro@agh.edu.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_1

2 Statistical Dominance Algorithm

This algorithm was established initially in a 3 × 3 version as a neighborhood map for multi-stage image thinning after binarization (application: corneal endothelium) [18]. The presented approach makes a segmentation repeatable, objective, and unambiguous. Research has shown that simply making a *neighborhood map* provides very interesting results for binarization of grayscale images. The most similar algorithm was presented in [24], but it considers fixed levels which are not dependent on neighboring pixels. The Local Binary Patterns algorithm is also related to the algorithm presented here in that it counts the sign of the differences between the central pixel value and its neighbors pixel values, sampled on a circle of given radius [13, 21]. A related issue is also mentioned in geospatial processing, where counting filters are used to count the number of point features, such as houses, wells or landslides [22].

2.1 The Main Idea

The basic idea of the algorithm is to determine for each pixel the number of neighbors in an area with a radius R. The resulting value is the number of pixels whose brightness is greater than or equal to the brightness of the central pixel. The output image is a statistical result of the dominance of points over their neighborhoods and allows the classification of these points to be determined (peak, valley, and slope).

2.2 Options of the Algorithm

The algorithm has the following options:

- two relationships between the central point and points in the neighborhood with equal or greater brightness (>= or >); other relationships (< or <=) which correspond to the first two relations for inverted images,
- additional threshold of the relationship (especially useful for noisy images, e.g. 1–2 % of dynamic range),
- neighborhood reduced to a straight line. A directional version of algorithm with lower computational complexity.

In general cases, assuming any shape and size of the relevant neighborhood, the algorithm is similar to formula 1. Equation 2 represents a constructive approach to the round neighborhood (disc) with a specified radius.

$$p\prime(x,y) = \sum_{p_b \in B(x,y)} \begin{cases} p'(x,y) = p'(x,y) + 1, & p_b(x_b, y_b) \geq p(x,y) + t \\ p'(x,y) = p'(x,y), & \text{otherwise} \end{cases} \tag{1}$$

$$p\prime(x,y) = \sum_{i=-N,j=-N}^{i=N,j=N} \begin{cases} p'(x,y) = p'(x,y) + 1, & p(x+i,y+j) \geq p(x,y) + t, \\ & i^2 + j^2 \leq R^2 \\ p'(x,y) = p'(x,y), & \text{otherwise} \end{cases} \tag{2}$$

where:

- $p(x, y)$—the value of pixel (x, y) in input image,
- $p'(x, y)$—the value of pixel (x, y) in output image,
- $B(x, y)$—a neighborhood of pixel $p(x, y)$,
- $p_b(x_b, y_b)$—the value of p_b pixel of $B(x, y)$ neighborhood (input image),
- R—radius of neighborhood,
- N—size of neighborhood mask, $N = \lceil R \rceil$,
- t—*threshold*—the optional difference to be checked.

2.3 The Code

The Statistic Dominance Algorithm

```
for (x = N; x < SX - N; x++)
  for (y = N; y < SY - N; y++)
  {
    imgout[x,y] = 0;     // initialization

    for (i = -N; i <= N; i++)
      for (j = -N; j <= N; j++)
        if (i * i + j * j <= R * R)  //without calling sqrt()
          if (imgin[x + i, y + j] >= imgin[x, y] + threshold)
            imgout[x, y]++;
  }
```

where:

- *imgin*—input image,
- *imgout*—output image,
- *SX*, *SY*—width and height of input/output image,
- *R*—(floating point) radius of neighborhood,
- *N*—(integer) size of neighborhood mask (also size of mirror margin, not presented here), $N = \lceil R \rceil$,
- *threshold*—*t*—the optional difference to be checked (especially for noisy images).

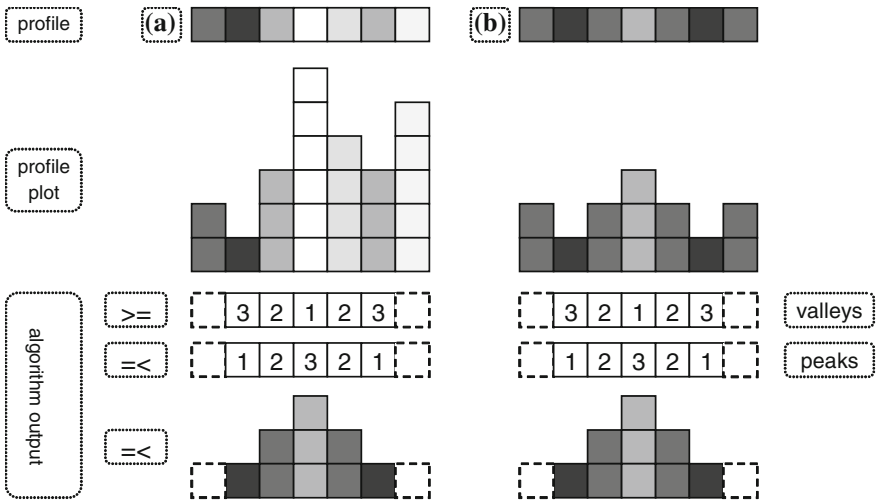


Fig. 1 An example: two profiles with the same course of valleys and peaks but different values, $R = 1$

2.4 Features of the Algorithm

The features of the algorithm are:

- the algorithm is designed for grayscale images,
- the result depends on the objects contained in the image, the selection of the size of radius R is therefore a function of the expected object size,
- the range of values in the output image is strictly defined (maximum value is the number of pixels at radius R),
- the result is mostly independent of differences in the brightness of objects in different areas of a picture as only the selected relation is taken into account, so it is easy to process images with varying levels of signal and noise and with local distortions.

As an example showing the last feature mentioned above an analysis of the two profiles could be provided (directional version assumed for simplicity) that have the same course (but not values) of valleys and peaks, thereby producing the same output from the algorithm (Fig. 1).

3 Examples

3.1 Corneal Endothelium Images from Specular Microscopy

Corneal endothelium images obtained from specular/confocal microscopy are very difficult to preprocess and analyse. There have been a few attempts to carry out proper

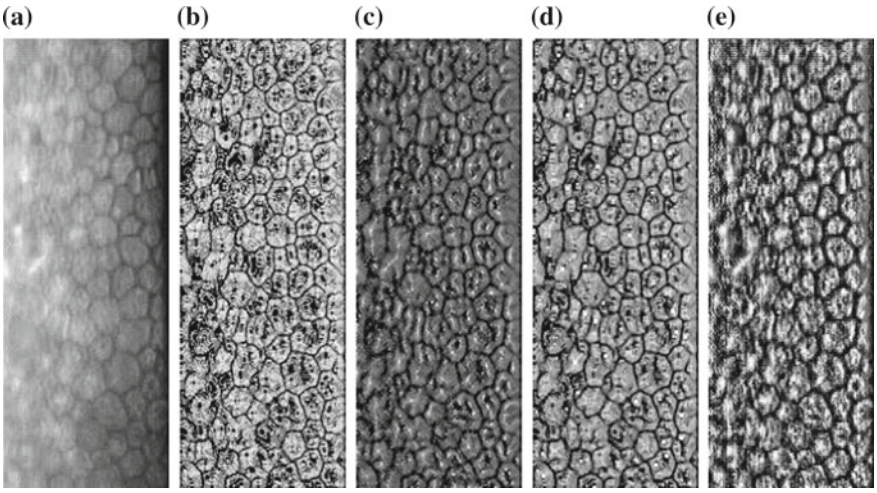


Fig. 2 Corneal endothelium image processing examples. **a** Original, normalized. **b** Max, SDA, $R = 2$. **c** SDA, $r = 3$. **d** Max, SDA, $R = 3$. **e** $R = 9$, preproc. **f** SDA, no

segmentation, but all of them needed manual correction. Watershed algorithms seem to be useless. The most promising method, which considers a stochastic approach to segmentation, is presented in [19]. An interesting approach of active contour use is presented in [5].

An example of a corneal endothelium image is shown in Fig. 2. The original input (Fig. 2a, normalized) contains a clear distinction between cells on the right side, and blurred boundaries between shapes on the left side. The SDA algorithm makes it possible to highlight the appropriate boundaries in both sides. Figure 2b contains SDA output after preprocessing with smoothing and a maximal filter (3×3), and Fig. 2d shows the same combination for a bigger radius (3). The boundaries are clear. Figure 2c contains the same case as Fig. 2b, but without a maximal filter. Pure output of the SDA algorithm without any preprocessing, for big radius ($R = 9$), is shown in Fig. 2e, which also allows highlighting of cell boundaries. The final SDA outputs in Fig. 2 are inverted for printing purposes.

3.2 Microscopic Images of the Brain Sections

Cell counting is a standard procedure of microscopic image analysis [10]. The aim is to count all cells in the selected region of interest (ROI), for example, a sagittal section of the molecular layer in the cerebellum of bank vole's brain is presented and discussed in [8]. Density, area coverage, and distribution of cell sizes are also needed according to the input image and the required outcomes. A shape of objects can be

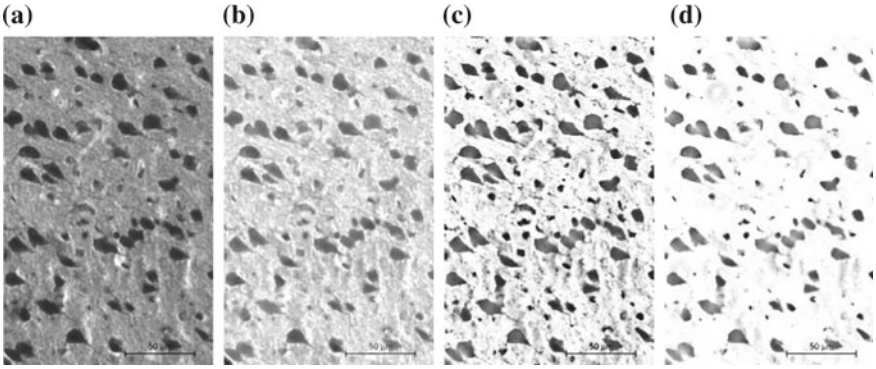


Fig. 3 An example of a sagittal section of the cerebral cortex. **a** Original. **b** Top-Hat, $R = 12$. **c** SDA, $R = 12$, $t = 30$. **d** SDA, $R = 12$, $t = 60$

analysed too [7, 14]. A classic approach to cell counting is the Top-Hat algorithm, but there are also other methods that are based on background removal or a general attempt for cell counting issue [9].

The SDA algorithm can act in similar way to the Top-Hat algorithm. Additionally, the threshold parameter can be taken into account to make the output less sensitive to noise. Figure 3a shows an example of a sagittal section of the cerebral cortex. There is also a Top-Hat output with neighborhood radius of 12, (Fig. 3b) and SDA outputs for thresholds of 30 (Fig. 3c) and 60 (Fig. 3d) and a radius of 12. It can be seen that SDA outputs contain less disturbances than Top-Hat.

3.3 Hand Radiographs

Computer processing of hand radiographs is a common task in numerous studies. In most cases, the analysis requires the extraction of the contours of the fingers or the whole hand and a clear separation of the bones, particularly the phalanges [1, 2, 23]. There is no unified method to this segmentation, so approaches are typically created by authors.

For example, to obtain a separation of lines in fingers' phalanxes the Sobel gradient was used [16]. The preprocessing stage involved background removal based on dynamic thresholding, and erosion with a 3×3 structural element was performed in order to remove all small noisy elements.

In further work [17] the authors presented a more advanced approach to preprocessing. Due to nonuniformity, background removal was performed independently for each highest peak reflecting the soft and bony structure, in relation to the average intensity of the background level. These considerations are not needed when using the SDA algorithm.

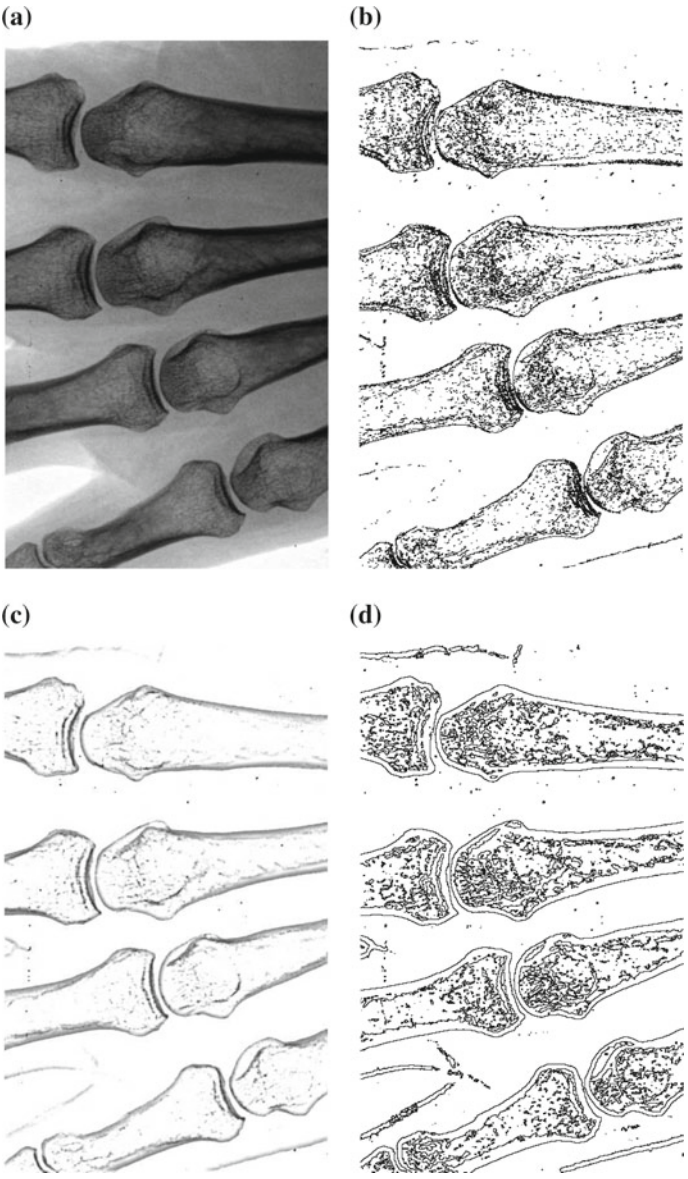


Fig. 4 A part of a hand radiograph. **a** Original, 2712 × 1656, 8 bit. **b** Canny's algorithm output. **c** SDA output, $R = 25$, $t = 25$. **d** SDA output after binarization ($t_{bin} = 1$) and contour extraction

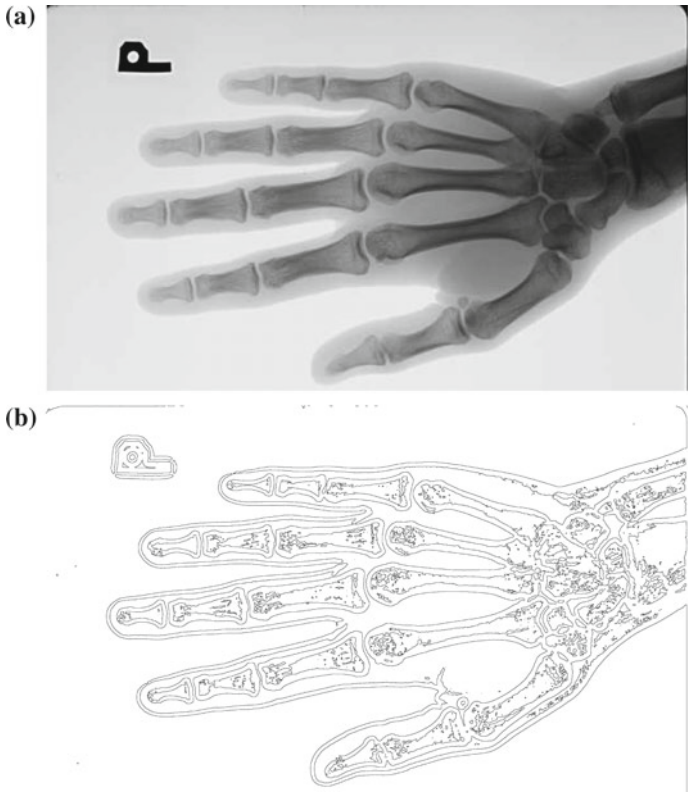


Fig. 5 Full hand radiogram and its SDA output after binarization ($t_{bin} = 1$) and contour extraction. **a** Original, 1425 × 2364, 12 bit. **b** SDA, $R = 15$, $t = 200$

Identification of the borders of the upper and lower surfaces of phalanxes is considered in [3, 25]. The authors proposed a constructive algorithm, based on computing the path running through the borders of both surfaces. Initial border points are used to create a circle with radius equal to the distance from F to E. A special analysis checks neighboring pixels, corresponding to the surface and background. The idea of this approach is related to the SDA algorithm, which checks the content of the round (circle) neighborhood. Another approach is to adapt other algorithms [11] in order to achieve proper hand segmentation [12].

In Figs. 4 and 5 the usage of the SDA algorithm is shown for a hand radiograph. Figure 4 presents contour detection of phalanxes in which all bones are segmented correctly. Figure 5 shows contour detection of a whole hand (size of 2712 × 1656) in which it can be seen that all phalanxes are separated precisely without dynamic thresholding and independent processing of selected regions of interest.

4 An Experimental Comparison of SDA and Top-Hat Algorithms

The effectiveness of algorithms can be evaluated experimentally. For this purpose, cells were counted by comparative segmentation using the standard methods of the Top Hat and Otsu algorithms (Fig. 6). As a reference, the well-known *rice* image was selected [20]. The size of the environment is a parameter common to both algorithms; therefore, a round neighborhood with a radius $R = 12$ was chosen. The SDA algorithm has an additional parameter: thresholding. Based on the selection of this parameter, overlapping results obtained with two methods were analyzed. Figure 7a shows the number of points detected in one of the resulting images for a given threshold. Figure 7b shows the dependence of the Dice Similarity Coefficient for a range of SDA thresholds [6]. Figure 8 shows the number of counted elements; an important parameter for cell counting.

Tests conducted on the *rice* example image show high similarity between the SDA segmentation algorithm and the standard approach (Top-Hat, Otsu) with appropriate threshold adjustment. For the *threshold* $= 53$, both methods achieved the highest compliance rate, reaching a Dice Similarity Coefficient value of 0.9876. Figure 6

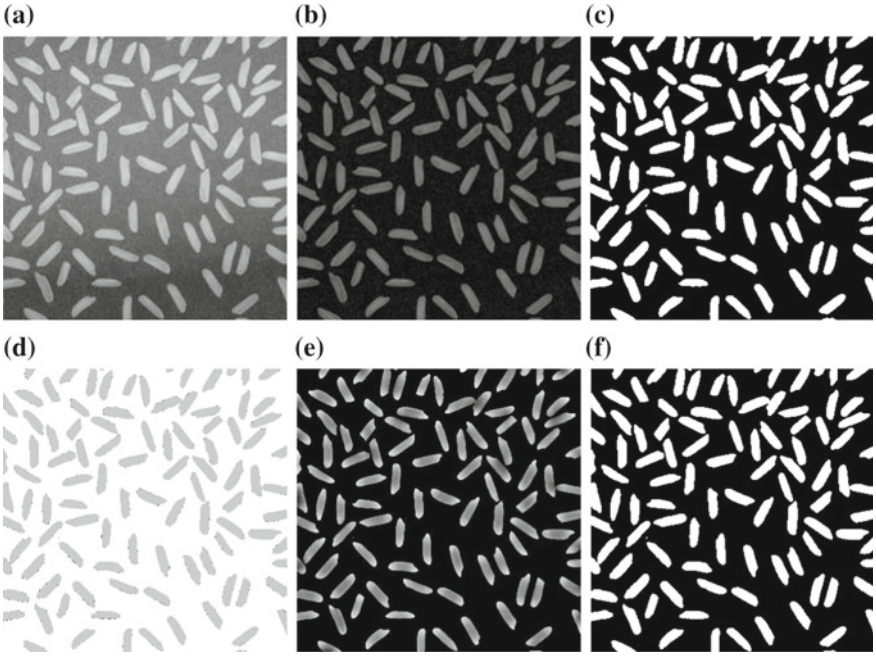


Fig. 6 The image *rice*. **a** Original, 350 × 350, 8bit [20]. **b** Top-Hat ($R = 12$). **c** Top-Hat ($R = 12$) + Otsu output. **d** Top-Hat and SDA differences. **e** SDA output, negated, $R = 12, t = 53$. **f** SDA output after binarisation ($t_{bin} = 1$)

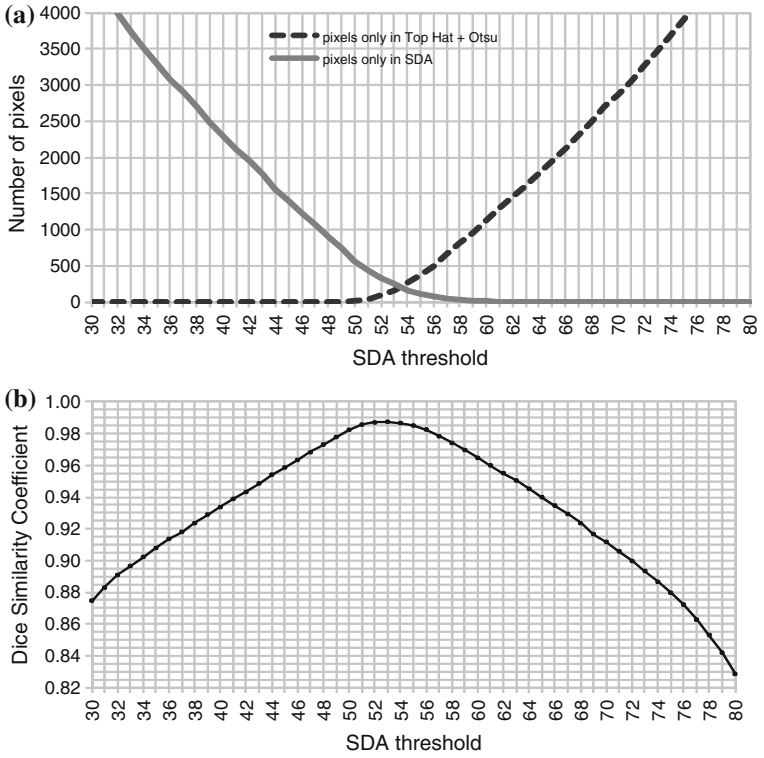


Fig. 7 A comparison of Top-Hat + Otsu and SDA outputs. **a** Number of pixels in output test images. **b** Dice similarity coefficient

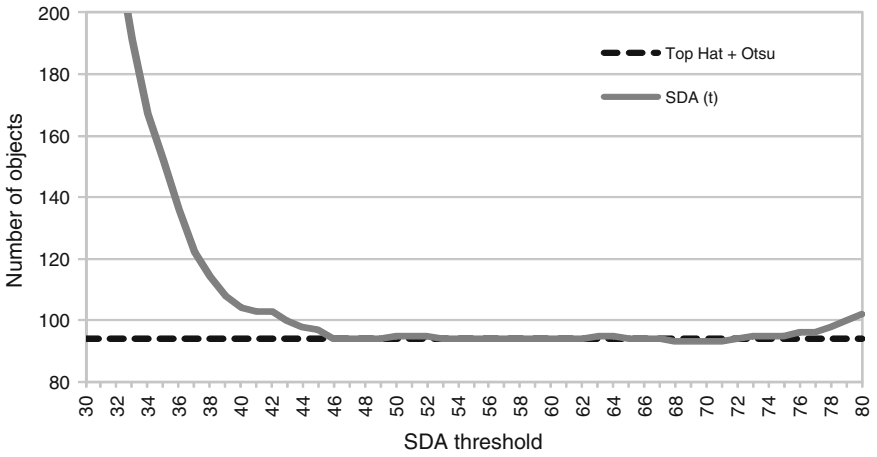


Fig. 8 Number of labeled objects in output test images

shows the difference between the two results. It can be argued that these differences relate mainly to single points on the edges of objects (Fig. 6d, red and blue pixels). When the number of objects is analyzed, it can be shown that with a relatively long thresholding period (46–67), the number of detected objects is almost the same in both output images.

The advantage of the SDA algorithm over the standard approach is the ability to determine the threshold value in order to be able to choose objects with given a difference in brightness. This applies e.g. to microscopic images of nervous tissues, for which it is essential to accurately search for selected elements like glial cells.

5 Summary

The SDA algorithm described in this paper seems to be notable for applications in medical image analysis. Its features are shown with various examples. The most interesting result is the distinct marking of the boundaries of corneal endothelial cells which is clear in the left half and fuzzy in the right. Also, the algorithm for cell counting tasks has been exemplified. The hand radiographs algorithm extracted the contours of phalanges reasonably well, without the need for local processing. Further research and studies should involve full testing of the proposed algorithm in detailed comparison of other well-known methods.

A sample algorithm demo is available on the site:

http://www.home.agh.edu.pl/~pioro/sda/

Acknowledgments This work was financed by the AGH—University of Science and Technology, Faculty of Geology, Geophysics and Environmental Protection as a part of statutory project.

The author would like to thank to Dr. J. Gronkowska–Serafin for corneal endothelium images, Dr. A. Kolodziejczyk for neural tissue images, Prof. M. Korkosz and Dr. M. Bielecka for hand radiographs, and Prof. R. Tadeusiewicz for consulting.

References

1. Bielecka, M., Bielecki, A., Korkosz, M., Skomorowski, M., Wojciechowski, W., Zieliński, B.: Application of shape description methodology to hand radiographs interpretation. In: Computer Vision and Graphics, pp. 11–18. Springer (2010)
2. Bielecka, M., Skomorowski, M., Zieliński, B.: A fuzzy shape descriptor and inference by fuzzy relaxation with application to description of bones contours at hand radiographs. In: Adaptive and Natural Computing Algorithms, pp. 469–478. Springer (2009)
3. Bielecki, A., Korkosz, M., Wojciechowski, W., Zieliński, B.: Identifying the borders of the upper and lower metacarpophalangeal joint surfaces on hand radiographs. In: Artificial Intelligence and Soft Computing, pp. 589–596. Springer (2010)
4. Canny, J.: A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. **6**, 679–698 (1986)
5. Charłampowicz, K., Reska, D., Boldak, C.: Automatic segmentation of corneal endothelial cells using active contours. Adv. Comput. Sci. Res. **11**, 47–60 (2014)

6. Dice, L.R.: Measures of the amount of ecologic association between species. Ecology **26**(3), 297–302 (1945)
7. Jaworek-Korjakowska, J., Tadeusiewicz, R.: Design of a teledermatology system to support the consultation of dermoscopic cases using mobile technologies and cloud platform. Bio-Algorithms Med-Syst. **11**(1), 53–58 (2015)
8. Kolodziejczyk, A., Ladniak, M., Piorkowski, A.: Constructing software for analysis of neuron, glial and endothelial cell numbers and density in histological nissl-stained rodent brain tissue. J. Med. Inform. Technol. **23**, 77–86 (2014)
9. Kuo, T., Byun, J.: ITCN (Image-based Tool for Counting Nuclei). http://www.bioimage.ucsb.edu/automatic-nuclei-counter-plug-in-for-imagej
10. Meijering, E.: Cell segmentation: 50 years down the road [life sciences]. IEEE Signal Process. Mag. **29**(5), 140–145 (2012)
11. Mikrut, Z., Bublinski, Z., Popiela-Mizera, A.: A method of linear star sections applied for object separation in ercp images. In: International Conference on Image Processing, 1996. Proceedings, vol. 1, pp. 363–366. IEEE (1996)
12. Ogiela, M.R., Tadeusiewicz, R., Ogiela, L.: Image languages in intelligent radiological palm diagnostics. Pattern Recogn. **39**(11), 2157–2165 (2006)
13. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on featured distributions. Pattern Recogn. **29**(1), 51–59 (1996)
14. Oszutowska-Mazurek, D., Mazurek, P., Sycz, K., Waker-Wójciuk, G.: Variogram based estimator of fractal dimension for the analysis of cell nuclei from the papanicolaou smears. In: Image Processing and Communications Challenges 4, Advances in Intelligent Systems and Computing, vol. 184, pp. 47–54. Springer (2013)
15. Otsu, N.: A threshold selection method from gray-level histograms. Automatica **11**(285–296), 23–27 (1975)
16. Pietka, E., Gertych, A., Pospiech, S., Cao, F., Huang, H., Gilsanz, V.: Computer-assisted bone age assessment: image preprocessing and epiphyseal/metaphyseal roi extraction. IEEE Trans. Med. Imaging **20**(8), 715–729 (2001)
17. Pietka, E., Gertych, A., Pospiech-Kurkowska, S., Cao, F., Huang, H., Gilzanz, V., et al.: Computer-assisted bone age assessment: graphical user interface for image processing and comparison. J. Dig. Imaging **17**(3), 175–188 (2004)
18. Piorkowski, A., Gronkowska-Serafin, J.: Towards precise segmentation of corneal endothelial cells. Bioinform. Biomed. Eng. Lect. Notes Comput. Sci. **9043**, 240–249 (2015)
19. Selig, B., Vermeer, K.A., Rieger, B., Hillenaar, T., Hendriks, C.L.L.: Fully automatic evaluation of the corneal endothelium from in vivo confocal microscopy. BMC Med. Imaging **15**(1), 13 (2015)
20. Shoelson, B., Tannenbaum, B.: New features for high-performance image processing in matlab. http://www.mathworks.com/company/newsletters/articles/new-features-for-high-performance-image-processing-in-matlab.html
21. Smolka, B., Nurzynska, K.: Power LBP: a novel texture operator for smiling and neutral facial display classification. Procedia Comput. Sci. **51**, 1555–1564 (2015)
22. Spatial data analysis: neighbourhood and connectivity calculations. In: Ilwis 3.0 Academic Users Guide
23. Tadeusiewicz, R., Ogiela, M.R.: Picture languages in automatic radiological palm interpretation. Int. J. Appl. Math. Comput. Sci. **15**(2), 305–312 (2005)
24. Tamburo, R.: An image filter for counting pixel neighbors. Insight J. (2011). http://hdl.handle.net/10380/3267
25. Zieliński, B., Skomorowski, M., Wojciechowski, W., Korkosz, M., Sprężak, K.: Computer aided erosions and osteophytes detection based on hand radiographs. Pattern Recogn. **48**(7), 2304–2317 (2015)

Swarm Intelligence Approach to 3D Medical Image Segmentation

Marta Galinska and Pawel Badura

Abstract In this paper we present a new idea for 3D medical image segmentation based on swarm intelligence and ant colony optimization. The methodology combines selected mechanisms running both mentioned artificial intelligence techniques, e.g. fitness-controlled motion of virtual agents or stigmergy. Foundations of the algorithm are described along with its implementation specification, simulations, results and their analysis also in terms of clarifying the parameterization. Several parameters are introduced and verified in terms of their influence on the method performance. The experiments rely on the segmentation of spleen in computed tomography studies. We also formulate some conclusions on possible ways for the algorithm future development.

Keywords Image segmentation · Swarm intelligence · Ant colony optimization

1 Introduction

Computer-aided diagnosis (CAD) systems designed e.g. for cancer diagnostics or post-treatment monitoring rely on proper analysis of medical images [7]. In most cases they require the expert contours of the region of interest (ROI) and/or the anatomical structure under consideration. The proper delineation and annotation stands for a diagnostic or pre-treatment factor of great importance. There is a large margin of error which depends on the expert when the structure contours are outlined manually [19]. They depend on the expert experience, age and even mood or time of the day. The intra-observer delineation may vary, since the results are considered unrepeatable and the segmentation gold standard is very difficult to detain. Moreover, the process of creating a contour is very laborious. All these factors force a

M. Galinska (✉) · P. Badura
Faculty of Biomedical Engineering, Silesian University of Technology,
Roosevelta 40, Zabrze, Poland
e-mail: marta.galinska@polsl.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_2

need to automate and objectify the segmentation process in diagnostic imaging. Making direct diagnostic decision based on automatic segmentation is, however, hardly acceptable. The semi-automatic methods are more appropriate: a person initiates, improves and controls the results of the algorithm (e.g. localizes the seed points or sets up parameters) but no delineations are made manually only.

Many standards for providing the partial automation of the image segmentation processes have been proposed through the years [9, 18]. None of them can be treated as a solution for all the problems, since each one addresses some specific issue. The topic of traditional 2D images segmentation have been explored particularly deeply, yet with a rapid progress in both, medical imaging data availability and computing power, the 3D processing techniques become more common nowadays. Such a growth of the amount of information brings also new challenges for intelligent analysis, requiring the employment of advanced processing and inference ideas. For that purpose, the artificial intelligence (AI) principles supply the segmentation workflows more often.

Artificial neural networks (ANN) and fuzzy logic (FL) are probably the most widely used AI branches in image segmentation in terms of ANN-based classification [21], fuzzy clustering [12, 25], fuzzy connectedness [5, 23], fuzzy inference systems [1] or hybrid neuro-fuzzy approaches [24]. The presence of evolutionary computation can also be noticed mostly in some specified auxiliary procedures like automated seed points selection [2, 14]. The granular computing have recently been employed for some pre-segmentation organ model definition [10]. The swarm intelligence ideas discovered actually in early 1990s are used in biomedical engineering mainly in bioinformatics [17, 20], yet also gradually appear in the image analysis domain. Multiple variants of the ant colony optimization (ACO) [8] and particle swarm optimization (PSO) [13] are, however, also mainly present as multipurpose procedures helpful at certain stages, not constituting the segmentation methodology itself. That concerns e.g. some sort of PSO-driven classifier training [26] or optimal multilevel threshold selection [15].

The ACO was initially illustrated by the traveling salesman problem [8]. Each virtual ant travels from one town to another and leaves pheromone on the road (stigmergy). The amount of pheromone depends on the road length. Pheromone evaporates over the time as it happens in nature [6]. The probability of choosing the next city on a track is directly proportional to the amount of pheromone on the path between the cities and inversely proportional to the distance among them [22]. The stigmergic information exchange makes the colony smart enough to search for the best solutions in a heuristic manner not only in the graph-defined approaches. The canonic ACO has been willingly modified. For example the ant colony behavior merged with beam search (which is a well-known tree search method) for open shop scheduling has been proposed [4]. The ACO hybridization with constraint programming (CP) and multilevel framework or auxiliary search space is also a well known technique [3].

PSO was designed to imitate social mechanisms observable in swarms or flocks consisting of more or less primitive individuals like birds, fish or insects [13]. Several basic rules control the swarm behavior [16], mainly related to collective searching

for a food and avoiding threats. The particle observes its neighbors, remembers its searching history and explores new spaces driven by some objective function. The PSO mathematical apparatus is very simple since it utilizes only basic arithmetic vector operations with a dose of randomness [13].

The goal of this study is to propose and examine a methodology based directly on swarm intelligence for segmentation of structures within medical imaging. We try to take advantage of possible correspondence of the 2D or 3D image space to the spatially-oriented PSO. The virtual agents exploring the image are stimulated by multiple factors, including the ACO-originating stigmergy. The paper describes implementation of the algorithm, experiments employing abdominal computed tomography (CT) studies, and analysis of the obtained results. Several parameters have been defined and analyzed for the swarm control with the awareness of the fact, that a lot of others are left for possible future formulation.

In the study we first introduce the virtual swarm specification in terms of its members, environment, parameters and auxiliary image processing techniques (Sect. 2). Then, the evaluation protocol is established in Sect. 3, followed by presentation and discussion on the obtained results. Conclusions are drawn in Sect. 4 along with a look at possible directions of the algorithm development.

2 Materials and Methods

The idea of the algorithm is to create a virtual swarm moving in a three-dimensional space related to the topology of the medical image. After the seed point location and parameters initialization, image preprocessing procedures are introduced (Fig. 1) in order to prepare the environment for the swarm operations. The iterative swarm motion is performed in the main loop in terms of the individuals particles movement.

2.1 PSO-Related Particle Movement Mechanisms

Each particle in the PSO algorithm has two main features in a D-dimensional problem space [13]: current position $\mathbf{x} = \{x_i: i = 1 \ldots D\}$ and current velocity $\mathbf{v} = \{v_i: i = 1 \ldots D\}$, both initiated with random values. They are iteratively updated

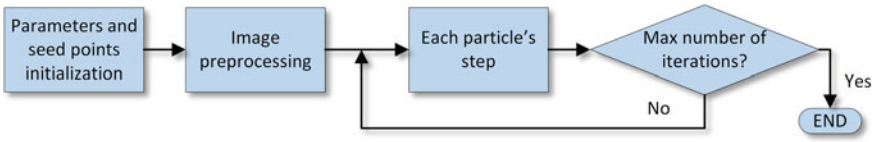


Fig. 1 Block diagram of the algorithm

in accordance with:

$$v_i = wv_i + c_1 r_1(\hat{x}_i - x_i) + c_2 r_2(\hat{s} - x_i), \tag{1}$$

$$x_i = x_i + v_i, \tag{2}$$

where w is the inertia weight, $r_1, r_2 \in [0, 1]$ are random variables, c_1, c_2 are algorithm-specific constants, $\hat{\mathbf{x}}$ is the particle best location so far, and $\hat{\mathbf{s}}$ is the global best location reached by any particle within the swarm. The particles locations are evaluated using a fitness function $f(\mathbf{x})$ subjected to optimization. Once the swarm moves throughout the environment, the communication mechanisms force the particles to search for optimal locations in an intelligent way.

Here, the particle is a sphere of a radius R temporarily centered in $\mathbf{x}$ and moving with velocity $\mathbf{v}$. R is subjected to changes during the exploration with some limitations assumed for its value. The particles are placed in the environment—the image space—according to some predefined seed points, obtained in either, manual or automatic manner. During the exploration of the volume, particles leave certain amount of pheromone in visited places—indicate voxels that should belong to the resulting mask of the structure. In each iteration of the algorithm the attractiveness rate r for several selected voxels $\mathbf{u}_j$ taken from the neighborhood $\mathcal{N}(\mathbf{x})$ of the particle $(\mathbf{x}, \mathbf{v})$ is calculated as:

$$r(\mathbf{u}_j) = a \cdot \frac{f}{\max(F)} + b \cdot \frac{f_d}{\max(F)} + c \cdot \frac{N_n}{N_a} + dm \cdot \frac{z}{||\mathbf{v}||}, \tag{3}$$

where: f denotes the amount of pheromone in $\mathbf{u}_j$, f_d—the amount of pheromone on the path between $\mathbf{x}$ and $\mathbf{u}_j$, F is the matrix containing the distribution of pheromone in the image space, N_n, N_a are the numbers of neighbors and agents defined in the algorithm, respectively, z denotes the distance from the closest obstacle on the path (if present; obstacles are related to the image and m is a multiplier selected randomly from the set $\{-1, 1\}$. The most attractive new location $\hat{\mathbf{u}}$ for the agent is selected:

$$\hat{\mathbf{u}} = \arg \max_{j=1\ldots N_n} \left\{ r(\mathbf{u}_j) \right\} \tag{4}$$

as it moves from $\mathbf{x}$ to $\hat{\mathbf{u}}$.

The following main parameters have been introduced and tested:

Number of agents N_a defines the size of the swarm moving inside the mask. If the swarm is larger, we can expect the object to be segmented faster, yet the leak probability increases.

Pheromone influence a is one of the main parameters in assessing the attractiveness of the area. Its negative value discourages other particles from the already explored area and by this encourages them to explore new places.

Pheromone evaporation e during the exploration of the area particles leave pheromone in visited places. Each particle leaves the amount of pheromone equal

to $(R - dist_{\mathbf{p}})$, where $dist_{\mathbf{p}}$ denotes the distance between voxel $\mathbf{p}$ covered by the agent and its center $\mathbf{x}$. The pheromone successively evaporates; its amount f is reduced at the end of each iteration:

$$f \leftarrow f - e. \tag{5}$$

Search radius s in millimeters determines the size of the area around the particle destination $\mathbf{u}_j$, where all particles are recognized as neighbors. Large number of neighbors decreases the area attractiveness and probability of other particles displacement in this direction. A value of s should be moderate in order to affect the movement.

Additional parameters considered in the method include:

The influence of pheromone on path b Current direction of the particle is also influenced by the amount of pheromone on its way to the potential endpoint. This mechanism favors the previously unexplored paths.

The influence of other particles c The swarm intelligence idea offers some benefits of communication between the particles. Since we want to disperse the swarm to avoid duplication of visited area and expand the explored field, the stimuli should be negative.

The influence of the path length d The swarm has been divided into two parts:

- the part preferring long paths: particles specialized in fast, although imprecise exploration are responsible for the swarm movement and searching for new, large areas;
- the part preferring short paths: particles designed for local exploration are responsible for filling the mask and detecting contours.

Size growth rate The idea of the spherical particle volume increase—when it is in a *safe* place (away from the structure edge)—has a positive impact on the time of filling up the mask. The particle radius R is reduced upon collision in order to help the agent get through to less accessible spaces.

2.2 Image Data Influence on the Particle Movement

The image features: intensity and gradient have also an impact on the particles movement. Figure 2 shows a sample axial slice of the original abdominal CT study and its gradient. They determine restricted areas, which repulse the particle and make it rebound from the boundary of the structure. Two types of restrictions may apply here. On one hand, in case of binary features (e.g., a mask obtained at any pre-segmentation phase), a possible move from one point to another is restricted if an obstacle arises in-between. On the other hand, the real-valued features (intensity, gradient) can stimulate the agent in a fuzzy way, e.g., the closer to the image edge or surface indicated by a high gradient, the more intensive the repulsion affecting the particle. In this study we employed the binary restrictions. The original image data

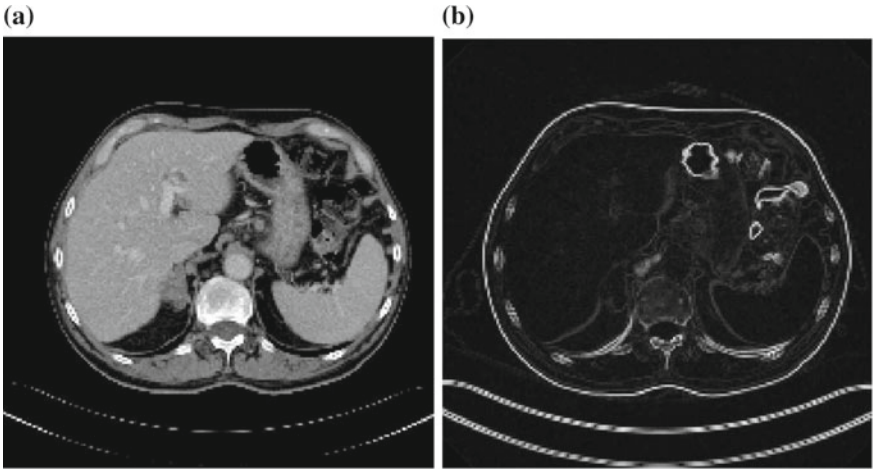


Fig. 2 Sample original slice (**a**) and the gradient image (**b**)

have been initially thresholded in order to obtain the binary volume containing the spleen. The volume is then inverted, so the result indicates areas inaccessible for the particle. The second type of obstacles arise from the gradient image thresholding. Once the particle touches any obstacle, it immediately bounces back without leaving any pheromone.

3 Results and Discussion

The algorithm implemented in Matlab® has been tested using a database of five abdominal CT studies with spleen [11] delineated by an expert.[1] Intensity and gradient thresholds have been determined empirically for this kind of data at 60 Hounsfield units (HU) and 5.5 HU/mm, respectively. All volumes subjected to the analysis have been resampled in order to obtain a topology with a 2 mm linear voxel size in each direction. The results were conducted by repeating the algorithm with different sets of parameters. Each set was tested 5 times and the results were averaged to verify the repeatability. A total of 24 sets were tested (Table 1), yielding 120 runs per study. Seven seed points were selected manually and their location did not change throughout all runs. Two evaluation measures have been used—the sensitivity and Dice index:

$$Sen = \frac{TP}{TP + FN} \cdot 100\,\%, \tag{6}$$

[1] Diagnostic context plays only a supporting role to the main research on the swarm algorithm.

Table 1 Summary of tested parameters

Parameter	N_a	a	e	s
Tested values	{300, 1500, 3000}	{−2, −1}	{0.00, 0.01}	{2, 5}

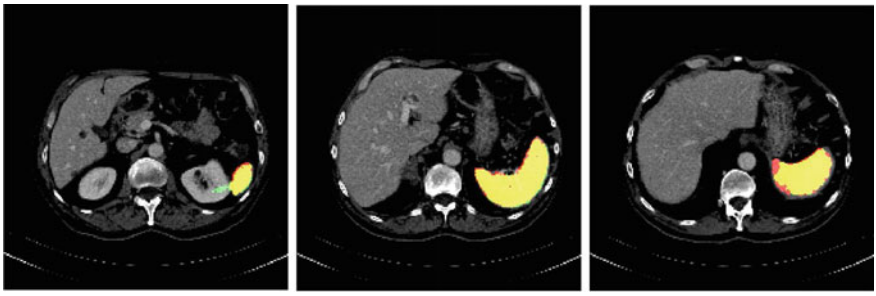


Fig. 3 Illustration of a sample segmentation result. TP voxels in *yellow*, FN—*red*, FP—*green*. Swarm parameters: $N_a = 3000$, $a = -1$, $e = 0$, $s = 5$

$$D = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \cdot 100\,\%, \tag{7}$$

where TP, FP, FN denote the number of true positive, false positive, and false negative voxels, respectively (Fig. 3). Figure 4 shows the pheromone-marked voxels in subsequent iterations, whilst combined sensitivity and Dice index plots in time for different parameter sets are presented in Fig. 5. A great and consistent progress can be observed up to the ca. thirtieth iteration. Later iterations are responsible for the external object surface smoothing and possible leaks. Table 2 presents a summary of optimal parameter values based on the results analysis. As can be seen, the selection of parameters depends on the image. The only *universal* parameter is the number of particles, yet its value may depend on the segmented object size and character. Nonetheless, swarms with 3000 particles worked most efficiently. It is possible that a larger swarm would achieve better results yet, due to the computing time, increasing number of particles is not an optimal solution.

The doubled value of pheromone influence a has a positive impact on the result of segmentation, which forces the swarm members to communicate more extensively in three out of five cases. The difference in evaporation rate e does not significantly affect the algorithm sensitivity. However, in most cases, swarms achieve better results if the evaporation has taken place. This may be due to the fact that during first iterations there is a large randomness in the direction selection; thus, implementation of the evaporation mechanism allows for further correction of non-visited areas. Searching for other particles in some neighborhood also improves the algorithm performance. Selection of the appropriate search radius s depends on the image.

We have to take into account that both, ACO and PSO are heuristic methods, thus there is no guarantee for reaching the best solution. Such algorithms search for the optimal solution with a high level of randomness. The results are unique,

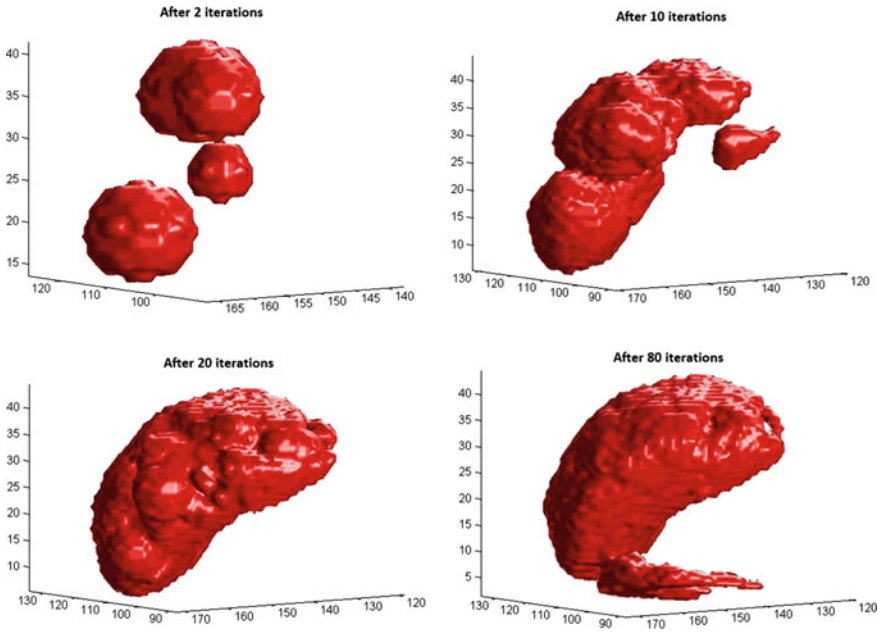


Fig. 4 3D visualization of the temporary segmentation result after 2, 10, 20 and 80 iterations. Obtained by the swarm with $N_a = 3000$, $a = -1$, $e = 0$, $s = 5$

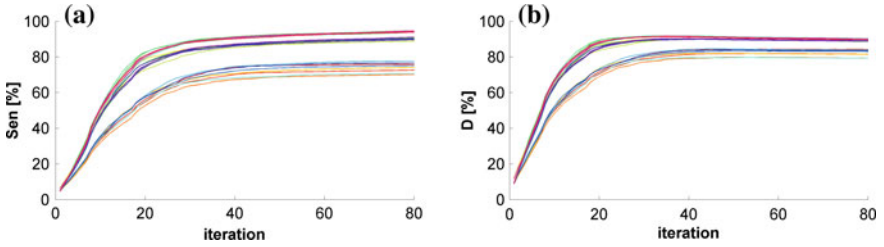


Fig. 5 Combined sensitivity (**a**) and dice index (**b**) plots in time averaged throughout different parameter settings

therefore difficult in verifying. In this study such an unrepeatability has been treated with responsibility: the same sets of parameters have been applied five times. The randomness puts the method application in medicine into question, where repeatability of the results counts above everything else. Since our experiments employed a specific kind of segmentation problem, it is hard to formulate conclusive statements on generalized parameters settings. The development of a swarm intelligence-based approach to the image segmentation should therefore involve various types of imaging techniques and structures at the validation stage.

The use of the proposed solution is probably insufficient for diagnostic purposes in its current form, but it may be an auxiliary tool for the radiologist. It is also possible that the combination of the described method with other applicable AI or

Table 2 Summary of the optimal parameter settings for individual studies

Image number	N_a	a	e	s	*Sen* (%)	D (%)
1	3000	−1	0	2	84	84
2	3000	−2	0.01	5	90	77
3	3000	−1	0	5	94	92
4	3000	−2	0.01	5	96	84
5	3000	−2	0.01	2	75	86

non-AI solutions would give satisfactory results. The method is very easy to modify because of its specificity. Adding new particle movement rules can improve the swarm efficiency. Going one step further we can risk hybridization of the method with evolutionary algorithms. This would reflect the natural Darwinism in the virtual population life, promoting more efficient individuals and removing the weak ones.

4 Conclusion

One of the purposes of this study was a verification if the particle motion randomness improves the efficiency of the performed task (in this case segmentation). The answer is not clear. Tracking the size and shape of segmented object shows that agents disperse from the seed points evenly. This allows for time saving in comparison to visiting all available voxels. There is a risk that at some stage the best direction could be skipped and a particle moves in another direction. Therefore, randomness might have a negative effect on the segmentation efficiency. However, the productivity of the algorithm in terms of time resources usually compensates this drawback, since the difference between the best of randomly drawn directions and the best of all available directions is usually very small. The main advantages of heuristic methods are the easily extensible number of criteria and the objective function adaptability. Although the current implementation operates directly in 3D medical image space which is not the most common case in this domain, one can imagine the swarm acting in a space of a higher dimensionality. That might concern images of different modalities, acquired during different examinations or employing various image features.

Acknowledgments This research was supported by the Polish National Science Center (NCN) grant No. UMO-2012/05/B/ST7/02136.

References

1. Badura, P., Pietka, E.: 3D fuzzy liver tumor segmentation. Inf. Technol. Biomed. Lect. Notes Bioinform. **7339**, 47–57 (2012)
2. Badura, P., Pietka, E.: Semi-automatic seed points selection in fuzzy connectedness approach to image segmentation. Comput. Recogn. Syst. Adv. Intell. Soft Comput. **45**(2), 679–686 (2007)

3. Blum, C.: Ant colony optimization: introduction and recent trends. Phys. Life Rev. **2**(4), 353–373 (2005)
4. Blum, C.: Beam-ACO-hybridizing ant colony optimization with beam search: an application to open shop scheduling. Comput. Oper. Res. **32**, 1565–1591 (2005)
5. Czajkowska, J., Badura, P., Pietka, E.: 4D segmentation of Ewing's sarcoma in MR images. Inf. Technol. Biomed. Adv. Intell. Soft Comput. **69**(2), 91–101 (2010)
6. Deneubourg, J., Pasteels, J., Verhaeghe, J.: Probabilistic behavior in ants—a strategy of errors. J. Theor. Biol. **105**(2), 259–271 (1983)
7. Doi, K.: Computer-aided diagnosis in medical imaging: historical review, current status and future potential. Comput. Med. Imaging Graph. **31**(4–5), 198–211 (2007)
8. Dorigo, M., Maniezzo, V., Colorni, A.: Ant system: optimization by a colony of cooperating agents. IEEE Trans. Syst. Man Cybern. Part B **26**(1), 29–41 (1996)
9. Gonzalez, R., Woods, R.: Digital Image Processing. Prentice Hall (2008)
10. Juszczyk, J., Pietka, E., Pycinski, B.: Granular computing in model based abdominal organs detection. Comput. Med. Imaging Graph. **46**(2), 121–130 (2015)
11. Kawa, J., Juszczyk, J., Pycinski, B., Badura, P., Pietka, E.: Radiological atlas for patient specific model generation. Inf. Technol. Biomed. Adv. Intell. Syst. Comput. **284**(4), 69–82 (2014)
12. Kawa, J., Pietka, E.: Image clustering with median and myriad spatial constraint enhanced FCM. Comput. Recogn. Syst. Adv. Intell. Soft Comput. **45**, 211–218 (2005)
13. Kennedy, J., Eberhart, R.C.: Particle swarm optimization. In: Proceedings. IEEE International Conference on Neural Networks, pp. 1942–1948 (1995)
14. Liang, Y., Zhang, M., Browne, W.: Image segmentation: a survey of methods based on evolutionary computation. In: Simulated Evolution and Learning, Lecture Notes in Computer Science, vol. 8886, pp. 847–859 (2014)
15. Maitra, M., Chatterjee, A.: A hybrid cooperative-comprehensive learning based PSO algorithm for image segmentation using multilevel thresholding. Expert Syst. Appl. **34**(2), 1341–1350 (2008)
16. Millonas, M.M.: Swarms, phase transitions, and collective intelligence. In: Artificial Life III. Addison-Wesley (1994)
17. Mohamad, M.S.: An enhancement of binary particle swarm optimization for gene selection in classifying cancer classes. Algorithm Mol. Biol. **8**(1), 1–11 (2013)
18. Pham, D., Xu, C., Prince, J.: Current methods in medical image segmentation. Ann. Rev. Biomed. Eng. **2**, 315–337 (2000)
19. Pietka, E., Kawa, J., Spinczyk, D., Badura, P., Wieclawek, W., Czajkowska, J., Rudzki, M.: Role of radiologists in CAD life-cycle. Eur. J. Radiol. **78**(2), 225–233 (2011)
20. Roseffeld, S.: Critical junction: nonlinear dynamics, swarm intelligence and cancer research. In: IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology, pp. 206–211 (2013)
21. Sharma, N., Ray, A., Sharma, S., Shukla, K., Pradhan, S., Aggarwal, L.: Segmentation and classification of medical images using texture-primitive features: application of BAM-type artificial neural network. J. Med. Phys. **33**(3), 119–126 (2008)
22. Simon, D.: Evolutionary Optimization Algorithms. John Wiley and Sons (2013)
23. Udupa, J., Samarasekera, S.: Fuzzy connectedness and object definition: theory, algorithms, and applications in image segmentation. Graph. Model Image Process. **58**(3), 246–261 (1996)
24. Verma, B., Zakos, J.: A computer-aided diagnosis system for digital mammograms based on fuzzy-neural and feature extraction techniques. IEEE Trans. Inf. Technol. B **5**(1), 46–54 (2001)
25. Zarychta, P.: Features extraction in anterior and posterior cruciate ligaments analysis. Comput. Med. Imaging Graph. **46**(2), 108–120 (2015)
26. Zyout, I., Czajkowska, J., Grzegorzek, M.: Multi-scale textural feature extraction and particle swarm optimization based model selection for false positive reduction in mammography. Comput. Med. Imaging Graph. **46**(2), 95–107 (2015)

Directional Characteristics of Mammographic Spicules in the Complex Wavelet Domain

Magdalena Jasionowska and Artur Przelaskowski

Abstract The subject of this paper is effective recognition of radiating spicules on digital mammograms. The presence of the spicules is the dominant symptom of neoplastic breast lesions called architectural distortions (ADs) or spiculated masses (SMs). The originality of the proposed method lies in the extraction of effective descriptors based on local directional activity of mammographic texture. Additionally, non-directional properties of mammographic findings were used in order to provide complete information about the discussed pathologies. The methodology applied was based on an analysis and constructive modeling of the conditioning of spicules distribution in the complex wavelet domain. The numerical descriptors of local tissue spiculation were calculated in the complex wavelet domain and, next, have been optimized and empirically verified. The experimental study was conducted on the basis of 2516 regions of interests, containing both normal (2091) and abnormal (415) breast tissue (clinically confirmed spiculated findings). Using the feature vector computed in the complex wavelet domain, the accuracy of spicules recognition (both in the case of ADs and SMs) reached over 83 %.

Keywords Computer-aided breast cancer diagnosis · Content-based descriptors · Mammographic spicule recognition · Complex wavelets · Image analysis

1 Introduction

Early diagnosis of breast cancer increases the survival rate and improves the efficiency of treatment [9]. Thus as early as possible detection of breast cancer is a very important challenge for screening programs. The error rate in mammography (commonly used in breast cancer screening [1]) is estimated even up to 30 % for false positives and 20 % for false negatives [19]. This is due to unstable patterns

M. Jasionowska (✉) · A. Przelaskowski
Faculty of Mathematics and Information Science,
Warsaw University of Technology, Koszykowa 75, Warsaw, Poland
e-mail: jasionowskam@mini.pw.edu.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_3

of pathology, variable technological conditioning of medical imaging, and limited expression of diagnostically important symptoms of image features. In radiology, in the decision-making process computer-aided diagnosis (CAD) systems frequently minimize the interpretation error, e.g. caused by evaluation of numerous cases in a short time [7]. However, the efficiency of CAD systems is still insufficient to automatically detect subtle spiculated pathologies. The detection sensitivity of the two commercial CAD systems (R2 Image Checker and CADx Second Look) for architectural distortions (ADs) and subtle spiculated masses (SMs) was less than 50 % and FPs/image was approximately 1.0 [25]. Thus, research into designing descriptors for effective recognition of mammographic spicules is ongoing.

1.1 Mammographic Spicules

Generally, breast tissue in mammographic image is manifested as a directionally oriented image texture—normal breast tissue converges frequently toward the nipple [22]. Mammographic image contains many piecewise-linear differently oriented structures that manifest various elements such as ligaments, ducts, blood vessels, benign findings, as well as the object of our interest—malignant spiculated lesions (both ADs and SMs). The limitations make subtle spiculated pathologies (especially ADs) extremely difficult to differentiate due to their subtlety and a high degree of similarity to other types of spiculated structures. The malignant spiculated findings can be divided into two typical groups:

(a) ADs are typically malignant lesions, but they can be most frequently confused with postsurgical scar. ADs are manifested as relatively short (<1 cm) linear structures radiating from an invisible dense center. The size of these pathological findings is approximately 5–50 mm [6, 9]. Moreover, both the spicule width and length are non-reproducible by radiologists [26].
(b) SMs can be distinguishable in mammogram by lines radiating from a well-visible central mass. These linear structures retain the same width along their entire length. SMs are probably identified with slightly higher number of spicules than ADs. Moreover, in case of SMs, both the size of the spicules and the length of the major axis of the central region are reproducible with high probability [26].

Various mammographic manifestations of spiculated findings are highly unstable and relatively case-dependent. The properties include the number of radiating spicules, their size, angular distribution, linearity, spicule overlapping, and the correlation of pathological findings with the surrounding tissue. Thus, a precise and consequent analysis of oriented-texture patterns seems a truly essential challenge of the CAD system for mammograms. Therefore, many research groups use directionality of spicules (piecewise linear structures) in image processing. For example, the analysis of local oriented edges was used in [16], whereas statistical analysis of a map of pixel orientations featured in [15], and skeleton analysis in [18]. In other cases, top-hat partial reconstruction was used to enhance the spicules images [14], and the

integrated intensity along oriented lines (using generalized DTLO) was described in [3]. Moreover, the extraction of linear structures was realized by estimation of a mean curvature sign and the concentration index [20], Gabor filtering and phase portrait [8, 22] or a curvilinear structure ridge detection [30]. Other approaches concentrated on non-directional properties of spiculated lesions. These include the intensity distribution of pixel context matched to the symptom template by thresholding procedures [21], as well as a fractal texture analysis [4, 17]. Several researchers have recently emphasized the importance of the use of various image representation domains in recognition of subtle piecewise linear structures, e.g. the Radon domain [27], or multiscale domains including discrete tensor wavelets [24], steerable complex filtering [29], Gabor wavelets [10, 23], and dual-tree complex wavelets [5].

1.2 Research Purpose

To the best of our knowledge, no specific directional descriptors of spicules had been designed in the complex wavelet domain. Thus, the purpose of this study is to design and optimize computational descriptors in the complex wavelet domain, efficient in automatic detection of abnormal spicules in mammograms. This paper focuses on originality in three spheres: (a) numerical modeling of local disorders of mammogram texture directionality with the use of both magnitude and phase information in the complex wavelet domain, (b) an extension of a descriptor vector designed to specify both tissue directionality and surrounding background of potential findings, (c) representative database of both normal and abnormal cases with two typical malignant spiculated findings (ADs and SMs).

The complex wavelet domain was selected to study due to our previous experiments, described in [11, 12]. Moreover, according to the results of Anderson et al. [2] or Wong and Scharcanski [31], the phase information of complex wavelets which complements directional distribution of signal energy seems to be used to capture invariantly subtle data dependencies. The proposed diagnostic model of spiculated findings in mammograms provides the background of spicule appearance, spicule profile and convergence that were formalized in the presented set of descriptors, being a complete description of the assumed model of mammographic spiculated findings. Relative to previous study [13], the optimization process of descriptor design and implementation was carried out to maximize specificity and sensitivity of spicule description for detection of both ADs and SMs (not solely ADs), based on earlier experiments and representative database of normal and abnormal cases.

2 Method

The proposed method of the spicule recognition was based on two fundamental assumptions. Firstly, the dominant role seems to be played by a reliable estimate of directional activity in distinction of radiating spiculated structures. Secondly, the

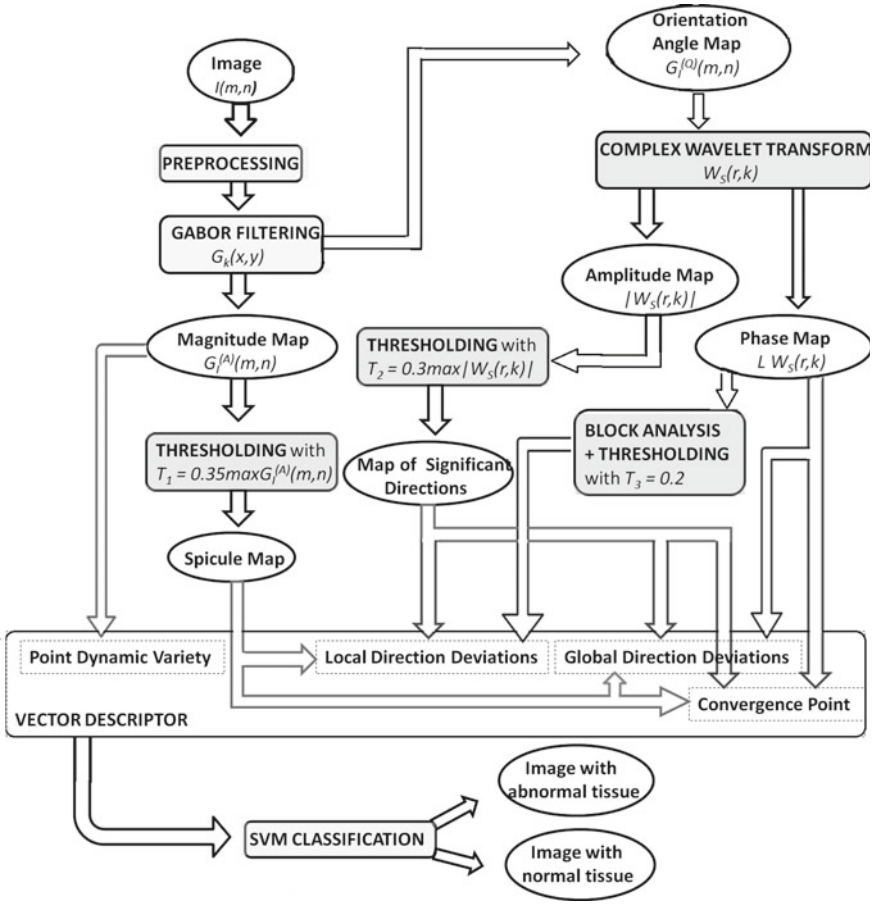


Fig. 1 Feature vector extraction in the complex wavelet domain. Gabor filtering was used to analyze mammograms considering with high amplitudes of the prefiltered image texture, corresponding to potential spiculated structures

properties of the complex wavelets (multi-scalability, sufficient directional resolution in higher dimensions, shift-invariance, information sparsity and selectivity) tend to accurately represent the line-singularity distribution while distinguishing the activity of texture orientation. The algorithm of well-differentiating feature vector extraction (Fig. 1) was implemented on the basis of the following elements of the proposed spicule model:

(a) initial characteristics of spicule orientation, based on a *magnitude map* and an *orientation angle map* as a result of directional Gabor filtering of ROIs;
(b) complementation of the initial characteristics of spicule orientation with the surrounding background (based on the *magnitude map*);

(c) precise description of each spicule in the form of piecewise linear or curvilinear structures modeled with line-singularity (based on spatial and directional structure locations in the complex wavelet domain with the use of the *orientation angle map*);
(d) estimation of directional activity sensitive to locally concentrated spicules of clearly diversified orientation (based on calculation of both local and global direction deviations to estimate distributions of dominant orientation for specific singularities, i.e. abnormal spicules);
(e) basic description of the area of spicule convergence into hypothetical lesion center (based on a local analysis of the distribution of singularity orientation in the complex wavelet domain).

2.1 Texture Orientation Activity

Initial analysis of texture orientation activity for mammographic ROIs was based on maximum magnitudes determination for dominant orientations at each pixel of directional filtered ROIs. Based on Rangayyan and Ayres [22] we found that a bank of Gabor filters of different orientations (adjusted to the approximated model of linear spiculated structures) seems to be practicable to analyze locally oriented piecewise-linear structures in mammogram. Even allowing for image noise and variations in scale, the Gabor image decomposition into the *magnitude map* and *orientation angle map* provides robust performance. The Gabor filter consists of two elements: a Gaussian envelope and a sinusoidal carrier, and it is oriented at angle Θ_k. This function has the form:

$$\mathbf{G}_k(x,y) = \exp\left(-0.5\left(\left(\frac{x'}{S_x}\right)^2 + \left(\frac{y'}{S_y}\right)^2\right)\right)\cos(2\pi f x') \tag{1}$$

where $x' = x\cos\Theta_k + y\sin\Theta_k$ and $y' = y\cos\Theta_k + x\sin\Theta_k$, $\Theta_k \in [-\pi/2, \pi/2]$. The Gabor magnitude map of image I(m,n) contains the highest filtering magnitude response $\mathbf{G}_I^{(A)}(m,n) = \max_k |\mathbf{G}_{(I,k)}^{(A)}(m,n)|$ where $\mathbf{G}_{(I,k)}^{(A)}(m,n) = G_k \times I(m,n)$. The Gabor orientation angle map, respectively, is determined at each pixel as $\mathbf{G}_I^{\Theta}(m,n) = \Theta_{arg\max_k |\mathbf{G}_{(I,k)}^{(A)}(m,n)|}$. A set of empirically optimized parameters includes:

(a) a sufficient number of Gabor filters (which corresponds to the number of Gabor filter orientations) in order to assure enough angular resolution: $k = 1, 2, \ldots, 180$;
(b) $S_x = 4$ and $S_y = 8$ defining two parameters of the Gaussian envelope: width S_x and elongation in the Y-direction (doubled in comparison to the extent in the X-direction), which provides relatively moderate precision of spicule orientations due to the slightly elongated filter shape.

The fixed ratio $\frac{S_y}{S_x}$ is generally sufficient to detect the directional linearity of texture but most frequently it is too vague to identify the nature of spicules. Moreover, the determination rule of the period of cosine term $f = \frac{0.7}{\tau}$ was verified, where $\tau = \frac{2.35}{S_x}$, that is the full-width at half-maximum of the Gaussian term along the X-axis. Consequently, $f = 0.075$ was used to reduce cosine oscillations along the elongated Gaussian slope and, subsequently, to model respective response of Gabor filter (ridge-like formation). It seems worth noting that high magnitude $\mathbf{G}_I^{(A)}(m, n)$ indicates distinctly oriented texture at the pixel (m,n), while low magnitude represents lack of highly oriented structures. Thus, a distinction between local orientation of abnormal and normal breast tissue can be roughly characterized by the relation of minimum to maximum of the *Gabor magnitude map*. This simple normalized scalar feature indicates a point dynamic variety and is defined by the following formula:

$$\mathbf{PDV} = \frac{\min_{m,n} \mathbf{G}_I^{(A)}(m, n)}{\max_{m,n} \mathbf{G}_I^{A}(m, n)}. \tag{2}$$

2.2 Precise Directional Description of Active Texture

The *Gabor orientation maps* of mammogram regions were used for more precise feature extraction and efficient recognition of abnormal directional distortions compared to dominant architecture of the mammary gland. The distribution of angular singularities, determined in locally estimated $\mathbf{G}_I^{\Theta}$, reflects diversity of orientation trends by non-random, possibly abnormal patterns. Multiscale and multidirectional decomposition of the *Gabor orientation angle map* was realized with the dual-tree complex wavelet (DT-CWT) implementation of four scales and six directional subbands. General form of the transformation is

$$\mathcal{W}_s(\mathbf{r}, k) = \int \mathbf{G}_I^{(\Theta)}(\mathbf{r}) 2^{\frac{s}{2}} \psi_c(2^s \mathbf{p} - \mathbf{r}, k) d\mathbf{p}, \tag{3}$$

with complex wavelet ψ_c, scale index $s = 1, \ldots, 4$, subband index $k = 1, \ldots, 6$ and image position $\mathbf{r} = [m, n]$. The applied complex wavelet decomposition was determined with filter banks (FBs) selected empirically: *Near-Symmetric* FB (13, 19 tap filters) for the first tree stage and Q-Shift, i.e. *Quarter Sample Shift Orthogonal* (14, 14) for next decomposition stages, for which the MATLAB implementation of the DT-CWT was used [28]. The number of scales was experimentally adapted to the size of the analyzed images. A 4-level decomposition of the selected ROIs with six directional subbands resulted in 24 complex coefficient matrices $\mathcal{W}_s(m, n, k)$ to be analyzed further.

Analytical capabilities of complex representation determined for the *Gabor orientation map* are mainly due to the invariant expression of properties further clarified as:

(a) directional characteristics of curvilinear singularities in magnitude domain $|\mathcal{W}_s(\mathbf{r}, k)|$ (directional diversity). 24 matrices of the wavelet *amplitude map* were obtained for each scale and all subbands;
(b) detailed description of selected singularity orientation based on block-based division of each subband in the phase transform domain $\angle\mathcal{W}_s(\mathbf{r}, k)$, resulting in more precise orientation correction in local space). Similarly, 24 matrices of the wavelet *phase map* were determined.

Local Direction Deviations of Potential Singularities In order to reduce false positives and diagnostically insignificant image structures, initial selection of potential orientation singularities was carried out in two ways:

(a) the first procedure, based on four binarized matrices created for each scale s by both appropriate subsampling (every 2^s pixel) of the Gabor magnitude map and binarization with experimentally adopted threshold T_1.
As a result, only the dominant coefficients, with amplitudes higher than $0.35 \times \max \mathbf{G}_I^{(A)}(m, n)$ in each subband, were retained for further analysis. The achieved frame of scaled spicule maps indicates the location of potential singularities for feature extraction in $\angle\mathcal{W}_s(\mathbf{r}, k)$ domain.
(b) the other procedure, based on thresholding magnitudes of the CWT, i.e. $|\mathcal{W}_s(\mathbf{r}, k)|$; As a result of fixed T_2 application, significant coefficients with amplitudes higher than $0.3 \times \max |\mathcal{W}_s(\mathbf{r}, k)|$ were retained and normalized in each subband, resulting in the *map of significant directions* of potential spiculated structures.

A proper 16-block analysis to precise directional activity is possible only after the initial selection of relevant representatives. For this purpose, a *map of local direction deviations* **LDD** in the form of 4×4 matrix series based of scaled subdivisions of each of 24 decomposition subbands was calculated. It represents the distribution of the coefficient average phase separately for each subband block across directions and scales.

Firstly, each subband of coefficients for all scales and directions is divided into 16 blocks $B_{\mathbf{p}}$, where $\mathbf{p} = [i, j]$ and $i, j = 1, \ldots, 4$. Therefore, depending on the scales s, each subband is a cell of 4×4 matrix of the block size of 64×64, 32×32, 16×16 or 8×8.

The next stage is to determine the averaged phase of non-zero coefficients for each block $B_{\mathbf{p}}$ of successive directional subbands across scales as $\alpha(s, k, \mathbf{p}) = mean(|\angle\mathcal{W}_s(B_{\mathbf{p}}, k)|)$, where $s = 1, \ldots, 4$, $k = 1, \ldots, 6$, and $\mathbf{p} = [i, j]$, $i, j = 1, \ldots, 4$, normalized to the average phase maximum of all blocks belonging to direction k across scales s.

Further, *active directions* were determined by estimation of phase singularities of specified directional characteristics, which is approximated by thresholded subband blocks. Thus, spatial accuracy is limited by block size in successive scales. For that

purpose, the values of $\alpha(s, k, \mathbf{p})_{norm}$ were binarized with experimentally chosen global threshold $T_3 = 0.2$.

Finally, 16-element matrix $LDD_{[4\times4]}$ was calculated by summing up the indicated blocks of the *active directions* across all scales and directions, taking into account the local structure of 4×4 subband subdivision:

$$\mathbf{LDD}(i, j) = \Sigma_{(s,k)} I(i, j), \tag{4}$$

where indicator function $I(i, j) = 1 \iff \alpha(s, k, [i, j])_{norm} > T_3$ while $I(i, j) = 0$ in other cases. The possible maximum value of the matrix LDD is 24, thus the representative block is active for all the subbands of complex wavelet decomposition. The 16 values of the matrix LDD form a feature vector used to recognize the specific local directional activity of tissue, rendering LDD as the significant factor for the differentiation of ADs or SMs.

Global Direction Deviations of Potential Singularities The obtained map of local direction deviations LDD provides information on the image angular selectivity and image directional resolution. In order to generalize the directional information, a 6-element *vector of global direction deviations* **GDD** was computed as follows:

$$\mathbf{GDD}(k) = \frac{\Sigma_{(s,\mathbf{r})} N_{sel}(\mathbf{r}, s, k)}{\left(\Sigma_{(s,\mathbf{r})} N(\mathbf{r}, s, k)\right)^2} \times \Sigma_{s,\mathbf{r}} |\angle \mathcal{W}_s(\mathbf{r}, k)|, \tag{5}$$

where N_{sel} is a number of selected coefficients of significant directions (based on *spicule maps* and LDD) with N being the number of all coefficients across k direction subbands of all scales. GDD allows for an examination of the size of phase shifts, and thus, for determination of the size of deviations of potential singularities localization in image in relation to the filter support. On this basis, it is possible to clarify the angular selectivity and to determine the localization of potential singularities (here, pathological spiculated structures) with higher directional resolution.

2.3 The Point of Convergence of Spiculated Structures

A final element of the proposed spicule model is the convergence point of spiculated structures, which is the area (not the point) of ADs or SMs and frequently seems to be invisible, especially for ADs. The dominant directions in image are possible to be determined with a high degree of precision through extension of information on the actual image texture orientations by adding the average value of coefficient phase to the generally accepted value of characteristic angles for each subband $\Theta_k = \pm15°, \pm45°, \pm75°$. The average value of phase coefficients that may be of help in the estimation of the convergence point is calculated across directional subbands k of all scales:

$$\mathbf{CP}(k) = \frac{\Sigma_{s,\mathbf{r}} N_{sel}(\mathbf{r}, s, k)}{\left(\Sigma_{s,\mathbf{r}} N(\mathbf{r}, s, k)\right)^2} \times \Sigma_{s,\mathbf{r}} \angle \mathcal{W}_s(\mathbf{r}, k), \tag{6}$$

based on the meaning of previously defined data structures and parameters.

3 Results and Discussion

The paper provides a brief summary of a new approach to abnormal mammographic spicules recognition. With the use of both magnitude and phase information, in order to investigate the potential of complex wavelets to represent meaningful image directional structures, appropriate regions of interest (ROIs 512 × 512 pixels) were indicated in mammograms available in the *Digital Database for Screening Mammography* (image resolution 43.5–50 μm). The ROIs with spiculated distortion of breast tissue were selected manually in order to collect reliable input data, therefore minimizing the number of classification errors associated with wrongly selected input data. As for the ROIs with normal breast tissue, they were automatically selected from the mammograms while omitting the areas of pathological tissue. Moreover, the ROIs with normal breast tissue were selected from both the normal and pathological mammograms to maintain a realistic database. The criteria used to form both training and testing datasets of mammogram ROIs were set to capture the whole image of diagnostically representative lesions as well as the surrounding tissue. The database contains 165 cases with ADs, 260 with SMs, 165 + 260 cases with both pathologies, and, respectively, 165, 260 or 425 normal cases in a balanced test (Table 1), whereas 2091 normal cases were used in a realistic test (Table 2).

Table 1 Balanced test: the results of SVM classification with the QF and RBF, *10-fold* cross validation

Features	Kernel	Features set	No of cases	ACC (%)	SE (%)	SP (%)
LDD	QF	AD+SM/normal	165 + 260/425	61.8	61.3	62.3
		SM/normal	260/260	65.9	66.2	65.6
		AD/normal	165/165	52.3	53.7	50.9
LDD+	QF	AD+SM/normal	165 + 260/425	72.6	70.3	75.0
		SM/normal	260/260	76.6	69.6	84.1
		AD/normal	165/165	59.3	60.2	58.4
LDD	RBF	AD+SM/normal	165 + 260/425	63.7	70.4	56.9
		SM/normal	260/260	65.8	72.3	59.3
		AD/normal	165/165	56.1	61.5	50.8
LDD+	RBF	AD+SM/normal	165 + 260/425	76.0	77.4	74.5
		SM/normal	260/260	77.3	76.9	77.7
		AD/normal	165/165	67.0	70.6	63.5

Table 2 Realistic test: the results of SVM classification with QF and RBF, *10-fold* cross validation

Features	Kernel	Features set	No of cases	ACC (%)	SE (%)	SP (%)
LDD	QF	AD+SM/normal	165 + 260/2091	64.5	65.6	59.1
		SM/normal	260/2091	73.8	75.3	61.9
		AD/normal	165/2091	62.5	64.5	37.6
LDD+	QF	AD+SM/normal	165 + 260/2091	79.6	81.5	70.1
		SM/normal	260/2091	82.9	83.6	77.3
		AD/normal	165/2091	79.4	81.8	49.1
LDD	RBF	AD+SM+normal	165 + 260 + 2091	75.4	80.1	52.2
		SM+normal	260 + 2091	80.1	83.0	56.5
		AD+normal	165 + 2091	82.4	84.6	53.9
LDD+	RBF	AD+SM+normal	165 + 260 + 2091	**83.0**	**85.1**	**72.7**
		SM+normal	260 + 2091	85.5	86.8	75.4
		AD+normal	165 + 2091	87.5	80.7	37.0

The calculated feature vector of LDD, GDD, CP and PVD, described in Sect. 2, was classified with SVM MATLAB classifier with various kernel functions. However, in this paper, the results were solely presented for quadratic kernel and radial basis kernel $r = 3$ (Tables 1 and 2) as, according to the conducted experiments, they proved to be the most effective kernels available in *svmtrain* MATLAB function. Moreover, the *least square* method was used to find a separating hyperplane. Additionally, *K-fold* cross-validation procedure, where experimentally established $K = 10$ ($K = 2, 5, 10, 20, 30$ was tested) was used to reduce the impact of the separation of training and testing data sets on classification results. This process was repeated $K = 10$ times, leaving one separate fold for evaluation each time. The sensitivity (SE), specificity (SP), and accuracy (ACC) of each dataset were calculated in *K-fold* cross-validation tests. Finally, the mean value of these statistical parameters of the computed K values was chosen in order to preserve the independence of the achieved classification results. These averaged values are presented in Tables 1 and 2. Furthermore, in the realistic test, 2091 cases with normal breast tissue were randomly partitioned into n equal-sized subsets—equinumerous to proper datasets for pathology (ADs, SMs, or ADs + SMs dataset, respectively). Thus, in *K-fold* classification procedure, the mean values of SE, SP and ACC were previously calculated for every n subset, and subsequently averaged (Table 2). This approach was meant to ensure independence of the results.

In the balanced test, the efficiency of pathological mammo-spicule recognition tends to be significantly higher in the case of SMs than that of ADs (Table 1). For both quadratic (QF) and radial basis (RBF) kernel function of classification, the ACC of SMs recognition was approx. 10 % higher comparing to that of ADs. Generally, the considerably lower SP was achieved for ADs, which probably resulted from the subtlety of ADs, higher distinct and unambiguous manifestation of SM structures. The lower SP of ADs recognition was also confirmed by the results of the realistic test (Table 2).

The results in Tables 1 and 2 indicate that the ACC of mammo-spicules recognition with the use of RBF kernel in classification process was higher comparing to that of QF (approx. 3–10 %). Therefore, the results of spicule recognition with the use of RBF kernel were considered in greater detail. Based on the results with the use of LDD for the balanced test (Table 1), the ACC was approximately 64 %, where the SP being 57 % for the ADs+SMs dataset. However, the LDD+ extension of the vector of well-differentiating features by addition of descriptors GDD, PVD or CP improved the efficiency of abnormal spicule recognition, resulting in 76 % ACC and the 74 % SP of ADs+SMs recognition. In the realistic test, in case of ADs+SMs (Table 2), the SE increased to approximately 85 %, whereas the SP rose to nearly 73 %. Thus, the recognition ACC of the dataset of both ADs and SMs reached until 83 % for the realistic data, resulting in a lower number of false positives. The representative examples for the proposed spicule recognition are shown in Fig. 2.

The addition of PVD descriptor seems only possible by approximation of the texture characteristics from the center of the mammographic findings and their background. It seems worth noting that the clarified description of the directional characteristics of the analyzed pathological findings allowed, in the majority of cases, to obtain even better recognition sensitivity. The analysis of the directional activity in small blocks as well as the additional estimate of the CP and the GDD define higher

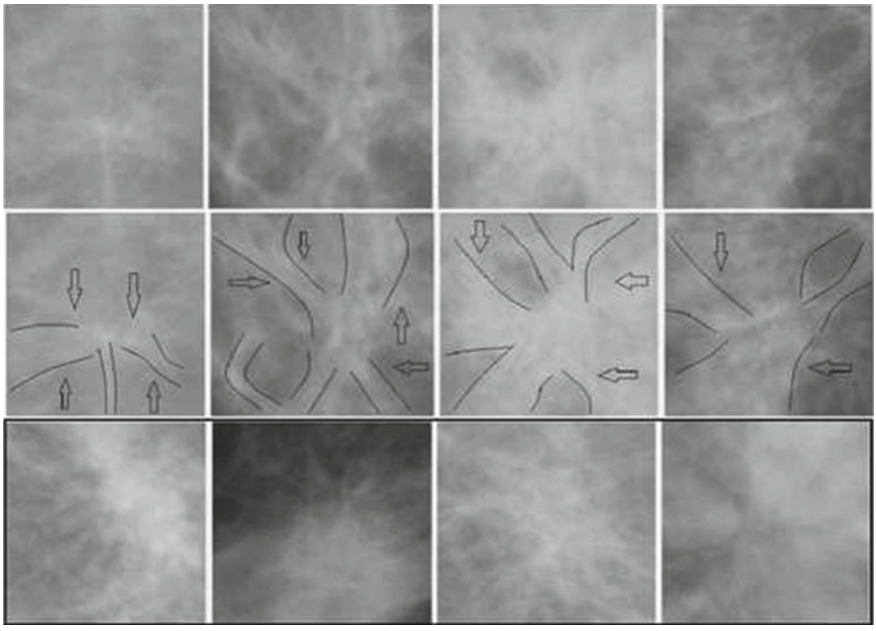


Fig. 2 The selected results of spicule recognition in ROIs with abnormal breast tissue: four false negative cases, indicated causes of unrecognized spicules shown above, and four cases of true positives, from *top* to *bottom* respectively

Table 3 The comparison of results with other research group

Method	Cases: all/spiculated	Database	SE (%)	SP (%)	ACC (%)	FP
LDD+GDD+CP+PVD	**2516/425**	DDSM	**85.1**	**72.7**	**83**	–
Gabor filters+phase portrait [22]	106/36	MIAS	76.5	76.4	–	–
Level sets [3]	60/47	DDSM	–	–	93	3
Neutral networks [4]	4212/262		80	–	–	10.5
Gabor wavelets [10]	80	MIAS	–	–	74.4	–

precision of directional resolution in image and contribute to the increased accuracy, which allows for more accurate recognition of spiculated findings.

The results of the proposed detection method have proved satisfactory and even more effective in comparison to other research groups (Table 3). Because of the different databases, in terms of their size and types, their comparison seems problematic. The most similar database to ours is the one by Banik et al. [4], the similarity lying in the statistics of cases, where the obtained SE reached moderately higher value while having relatively acceptable SP. Furthermore, it is possible to increase ACC of spicule detection procedure. The proposed method is an element of the whole detection method for mammographic pathologies. Improvements of the effectiveness of the proposed procedure in the complex wavelet domain is possible by full-automatic selection of ROIs with potential pathologies (e.g. based on breast tissue density) as well as by more precise description of both the lesion-background relation and the area of spicule convergence.

4 Conclusion

The proposed approach of abnormal spicule recognition assumes: (a) effective use of phase information and energy distribution for more precise and invariant characteristics of spicule recognition, (b) narrowed area of analysis to directional activity centers, designated on the basis of the Gabor amplitude and CWT magnitude maps, (c) increased efficiency of recognition through supplementary information, which describes the distribution of spicules and characteristics of their background in relation to the formulated pathological model-spicule. The obtained results seem to confirm the merits of the applied method. Thus, the usefulness of CWT, directional analysis of mammograms, precise investigation of convergence lines area, and description of the background tend to prove effective in the verification of the suspected abnormalities of active local directionality (in this case, of mammographic subtle spiculated structures). However, the recognition of subtle ADs and SMs in mammograms seems to remain an unsolved problem. It can be assumed the pragmatic difficulty lies in increased number of indicated false positives with simultaneous rise in the sensitivity of spicule detection.

Acknowledgments This publication was funded by the National Science Centre (Poland) based on the decision DEC-2011/03/B/ST7/03649.

References

1. American College of Radiology, ACR Appropriateness Criteria for Breast Cancer Screening, pp. 1–6 (2012)
2. Anderson, R., Kingsbury, N., Fauqueur, J.: Multiscale object features from clustered complex wavelet coefficients. In: IEEE/SP 13th Workshop on Statistical Signal Processing, pp. 437–442 (2005)
3. Ball, J.E., Bruce, L.M.: Digital mammogram spiculated mass detection and spicule segmentation using level sets. In: Proceedings of the 19th Annual International Conference of the IEEE EMBS, pp. 4979–4984 (2007)
4. Banik, S., Rangayyan, R.M., Desautels, J.E.L.: Detection of architectural distortion in prior mammograms. IEEE Trans. Med. Imaging **30**(2), 279–294 (2011)
5. Berks, M., Taylor, C., Rahim, R., Boggis, C., Astley, S.: Modelling structural deformations in mammographic tissue using the dual-tree complex wavelet. LNCS **6136**, 145–152 (2010)
6. Birdwell, R.L., Morris, E.A., Wang, S., Parkinson, B.T.: 100 rozpoznan. Sutek, MediPage (2005)
7. Broeders, M., Moss, S., Nystrom, L., Njor, S., Jonsson, H., Paap, E., Massat, N., Duffy, S., Lynge, E., Paci, E.: The impact of mammographic screening on breast cancer mortality in Europe: a review of observational studies. J. Med. Screen. **19**(Suppl 1), 14–25 (2012)
8. Chakraborty, J., Rangayyan, R.M., Banik, S., Mukhopadhyay, S., Desautels, J.E.L.: Statistical measures of orientation of texture for the detection of architectural distortion in prior mammograms of interval cancer. J. Electron. Imaging **21**(3), 1–13 (2012)
9. Dziukowa, J.: Mammografia w diagnostyce raka sutka. BEL CORP Scientific Publications, Warszawa (1998)
10. Ferrari, R.J., Rangayyan, R.M., Desautels, J.E.L., Frere, A.F.: Analysis of asymmetry in mammograms via directional filtering with Gabor wavelets. IEEE Trans. Med. Imaging **20**(9), 953–964 (2001)
11. Jasionowska, M., Przelaskowski, A.: Multiscale modeling of local directional mammogram findings. J. Med. Inf. Technol. **17**, 183–190 (2011)
12. Jasionowska, M., Przelaskowski, A.: Subtle directional mammographic findings in multiscale domain. Inf. Technol. Biomed. LNCS **7339**, 77–84 (2012)
13. Jasionowska, M., Dobór deskryptorów, P.: według warstw proponowanego modelu spikul w mammogramach. XVIII Krajow Konferencja Biocybernetyki i Inżynierii Biomedycznej, Polska, pp. 118 (2013)
14. Jiang, H., Tiu, W., Yamamoto, S., Iisaku, S.: A method for automatic detection of spicules in mammograms. J. Comput. Aided Diagn. Med. Images **2**, 23–31 (1998)
15. Karssemeijer, N., Te Brake, G.M.: Detection of stellate distortions in mammograms. IEEE Trans. Med. Imaging **15**(5), 611–619 (1996)
16. Kegelmeyer, W.P., Jr.: Evalution of stellate lesion detection in a standard mammogram data set. In: Bowyer, K.W., Astley, S. (eds.) State of the Art in Digital Mammographic Image Analysis, pp. 262–279. World Scientific (1993)
17. Kim, H.J., Kim, W.H.: Automatic detection of spiculated masses using fractal analysis in digital mammography. LNCS, vol. 3691, pp. 256–263. Springer, Berlin (2005)
18. Kobatake, H., Yoshinaga, Y., Detection of stellate distortion in mammograms. IEEE Trans. Med. Imaging, **MI-15**(3), 235–245 (1996)
19. Kolb, T.M., Lichy, J., Newhouse, J.H.: Comparison of the performance of screening mammography, physical examination and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. Radiology **225**, 165–175 (2002)

20. Nemoto, M., Honmura, S., Shimizu, A.: A pilot study of architectural distortion detection in mammograms based on characteristics of line shadows. Int. J. CARS **4**, 27–36 (2009)
21. Ozekes, S., Osman, O., Camurcu, A.Y.: Mammographic mass detection using a mass template. Korean J. Radiol. **6**(3), 221–228 (2005)
22. Rangayyan, R.M., Ayres, F.J.: Gabor filters and phase portraits for the detection of architectural distortion in mammograms. Med. Biol. Eng. Comput. **44**(10), 883–894 (2006)
23. Ragupathy, U.S., Saranya, T.: Gabor wavelet based detection of architectural distortion and mass in mammographic images and classification using adaptive neuro fuzzy inference system. Int. J. Comput. Appl. **46**(22), 37–40 (2012)
24. Rashed, E.A., Ismail, I.A., Zaki, S.I.: Multiresolution mammogram analysis in multilevel decomposition. Pattern Recogn. Lett. **28**, 286–292 (2007)
25. Sampat, M.P., Markey, M.K., Bovik, A.C.: Computer-aided detection and diagnosis in mammography. In: Bovik, A.C. (ed.) Handbook of Image and Video Processing, 2nd edn, pp. 1195–1217. Academic, New York (2005)
26. Sampat, M.P., Markey, M.K., Bovik, A.C.: Measurement and detection of spiculated lesions. IEEE **5747**, 105–109 (2006)
27. Sampat, M.P., Whitman, G.J., Markey, M.K., Bovik, A.C.: Evidence based detection of spiculated masses and architectural distortions. SPIE, Med. Imaging: Image Process. **5747**, 26–37 (2005)
28. Selesnick, I.W., Baraniuk, G., Kingsbury, N.G.: The dual-tree complex wavelet transform. IEEE Signal Process. Mag. **22**(6), 123–151 (2005)
29. Shenk, V.U.B., Brady, M.: Finding CLS using multiresolution oriented local energy feature detection. In: Proceedings 6th International Workshop on Digital Mammography, pp. 64–68 (2002)
30. Wai, L.C.C., Mellor, M., Brady, M.: A multi-resolution CLS detection algorithm for mammographic image analysis. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) LNCS, vol. 3217, pp. 865–872. Springer, Berlin (2004)
31. Wong, A., Scharcanski, J.: SeniorPhase-adaptive superresolution of mammographic images using complex wavelets. IEEE Trans. Image Process. **18**(5), 1140–1146 (2009)

Discrimination of Wheat Grain Varieties Using X-Ray Images

Małgorzata Charytanowicz, Jerzy Niewczas, Piotr Kulczycki, Piotr A. Kowalski and Szymon Łukasik

Abstract A study was conducted so as to develop a methodology for wheat variety discrimination and identification by way of image analysis techniques. The main purpose of this work was to determine a crucial set of parameters with respect to wheat grain morphology which best differentiate wheat varieties. To achieve better performance, the study was done by means of multivariate discriminant analysis. This utilized both forward and backward stepwise procedures based on various sets of geometric features. These parameters were extracted from the digitized X-ray images of wheat kernels obtained for three wheat varieties: Canadian, Kama, and Rosa. In our study, we revealed that selected combinations of geometric features permitted discriminant analysis to achieve a recognition rate of 89–96 %. We then compared the correctness of classification with results obtained by way of employing the nonparametric approach. The discriminant analysis proved effective in differentiating wheat varieties.

Keywords Wheat grain · Morphological features · X-ray imaging · Image processing · Image analysis · Discrimination · Classification · Nonparametric density estimation

M. Charytanowicz (✉) · P. Kulczycki · P.A. Kowalski · S. Łukasik
Systems Research Institute, Centre of Information Technology
for Data Analysis Methods, Polish Academy of Sciences, Newelska 6, Warsaw, Poland
e-mail: mchmat@ibspan.waw.pl

M. Charytanowicz · J. Niewczas
Institute of Mathematics and Computer Science, The John Paul II Catholic
University of Lublin, Konstantynów 1H, Lublin, Poland

P. Kulczycki · P.A. Kowalski · S. Łukasik
Faculty of Physics and Applied Computer Science, Division for Information
Technology and Biometrics, AGH University of Science and Technology,
Mickiewicza 30, Cracow, Poland

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_4

1 Introduction

Recent research in computer-aided data analysis based on image processing, has brought about the high development of accurate and automated systems for the discrimination and classification between wheat grain categories [3, 9, 11]. One of the more well-known and widely used imaging techniques is soft X-ray photography. This is an objective and precise method which provides high quality visualization of the internal kernel structure. Together with photo-scanning procedures, this technique provides sufficient resolution for reflecting distinct features important for the accurate characterization of kernels. Moreover, this is non-destructive and considerably cheaper than other more sophisticated techniques such as magnetic resonance imaging, scanning microscopy or laser technology. It has to be stressed, however, that sole visualization of the kernels, irrespective to the techniques used, does not provide quantitative evaluation of the overall quality of the kernel, e.g., measures of geometric parameters of internal grain features, their correlation and distribution within the structure, etc. In order to carry out an accurate grain quality assessment, specialized image processing procedures need to be employed for the measurement, detection and interpretation of kernels in X-ray images. Various grading systems using different morphological features for the classification of different cereal grains and varieties have been reported in literature [12, 15, 17]. As well as geometric features, the authors used grain surface texture and color to develop a statistical model that effectively employs these variables to classify wheat grain.

The main objective of this work is to investigate whether it is possible to recognize wheat varieties by way of multivariate discriminant analysis as applied to certain basic, critical geometric parameters of kernels that have been extracted from X-ray images. The accomplished study consists of two phases: the construction of a model based on the training set that has been created for cases with known belonging to different groups, and, subsequently, to extend the use of this model for classifying new cases. The classification results given by the classical approach were then compared with nonparametric classification results.

2 Mathematical Preliminaries

Classical methods of data analysis assume that the data are drawn from one of a known parametric family of distributions, determined by their parameters [2, 5]. The density underlying the data could then be estimated by finding from the data estimates of unknown parameters, using optimization criteria. Such attitude requires performing goodness-of-fit tests on the data, in which the null hypothesis states that our data follows a specific distribution. This rigidity can be overcome by nonparametric estimation methods that assume no pre-specified functional form for a density function. In this section, both approaches to discriminant analysis are shortly described.

2.1 Nonparametric Density Estimation

Suppose that there is the n-dimensional random variable X, with a distribution characterized by the density f. Its kernel estimator $\hat{f} : R^n \to [0, \infty)$, calculated using experimentally obtained values for the m-element random sample

$$x_1, x_2, \ldots, x_m, \tag{1}$$

in its basic form is defined as

$$\hat{f}(x) = \frac{1}{mh^n} \sum_{i=1}^{m} K\left(\frac{x - x_i}{h}\right), \tag{2}$$

where $m \in N\backslash\{0\}$, the coefficient $h > 0$ is called a smoothing parameter, while the measurable function $K : R^n \to [0, \infty)$ of unit integral $\int_{R^n} K(x)\mathrm{d}x = 1$, symmetrical with respect to zero and having a weak global maximum in this place, takes the name of a kernel.

The choice of the form of the kernel and the calculation of the smoothing parameter is made most often by way of established optimization criterions. The choice of the kernel has no practical meaning, and, due to this situation, it is possible to take into account primarily the properties of the obtained estimator. In practice, for the one-dimensional case, the function is assumed usually to be the density of a common probability distribution. Most often the standard normal kernel given by the formula

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{3}$$

is used. In the multidimensional case, two natural generalizations of the above concept are employed: radial and product kernels, but from an applicational point of view, the difference is insignificant. The fixing of the smoothing parameter has importance in establishing the quality of estimation. For the one-dimensional case, the effective plug-in method is especially recommended. For the multidimensional case, two natural attitudes are usually used: the plug-in method for product kernels and the cross-validation method for radial kernels.

Broader discussion of this task, as well as additional procedures improving the quality of the estimator obtained, such as modification of the smoothing parameter and linear transformation, as well as general aspects of the theory of statistical kernel estimators are found in [6, 7, 13].

2.2 Discriminant Analysis

Classificatory discriminant analysis is used to allocate new observations into one of the two or more known groups or clusters on the basis of the measured characteristics.

The basic problem of discrimination can be stated as follows. Consider K independent simple random samples containing $m_1, m_2, \ldots, m_K$ elements, respectively drawn from K different populations (classes):

$$\begin{array}{c} x_{11}, x_{12}, \ldots, x_{1m_1} \\ x_{21}, x_{22}, \ldots, x_{2m_2} \\ \ldots \\ x_{K1}, x_{K2}, \ldots, x_{Km_K} \end{array} \tag{4}$$

where $x_{kj} \in R^n$ constitutes the j-multivariate observation from k-th class. These samples form the training set. Given a new observation $z \in R^n$, the main task is to assign that observation to one of K classes.

Suppose, first of all, that observations are drawn from populations having probability density functions $f_1, f_2, \ldots, f_K$. The classical Bayesian approach would then allocate the tested element to the class for which the value

$$c_1 f_1(z), c_2 f_2(z), \ldots, c_K f_K(z) \tag{5}$$

is the largest [4, 8]. The constants $c_k, k = 1, 2, \ldots, K$, are chosen by reference to the probabilities of misclassification, or by considering the prior probability that z comes from the population k and other utilities of correct classification. In practice, the densities $f_k, k = 1, 2, \ldots, K$, cannot be assumed to be known, and so the discriminate rule must be estimated from the training set. One natural approach is to suppose that the unknown densities come from some parametric family and then to estimate the parameters. When the within-class covariance matrices are assumed to be equal, the parametric approach generates a linear discriminant function, otherwise it can be extended to a quadratic form.

When the distribution within each group is not assumed to have any specific distribution or is assumed to have a distribution different from the multivariate normal distribution, nonparametric methods can be used to estimate the densities $f_1, f_2, \ldots, f_K$ and derive classification criteria. A natural description of these distributions allow the specifying of the kernel estimators $\hat{f}_1, \hat{f}_2, \ldots, \hat{f}_K$, constructed under given random samples (4), and consequently, the choosing of the population for which the value $c_k \hat{f}_k(z)$, $k = 1, 2, \ldots, K$, is the greatest. The main concept of the kernel density methodology is described in Sect. 2.1.

3 Methodology

The proposed methodology for wheat variety discrimination and identification by way of image analysis techniques, is summarized as follows:

- selecting wheat grains,
- capturing the wheat grain photograms,

- producing bitmap graphics files for reflecting geometric features of grains,
- rescaling images to standardize the unit of measurement,
- measuring geometric features,
- selecting wheat kernel geometric parameters,
- performing the discriminant analysis and classification process.

3.1 Image Processing and Feature Extraction

The study was undertaken at the Institute of Agrophysics of the Polish Academy of Sciences in Lublin, and conducted using the combine harvested wheat grain of three different varieties: Canadian, Kama, and Rosa, originating from experimental fields. Randomly selected grain samples of these varieties contained 108, 72, and 108 kernels, respectively. A high quality visualization of the internal kernel structure (Fig. 1) was then generated using a soft X-ray technique.

The images were obtained in the form of photograms of size 13 × 18 cm, at the scale of 5 : 1. For each X-ray exposure, 10–12 grain kernels were evenly positioned groove down. The photograms were then scanned by way of an Epson Perfection V700 table photo-scanner that was equipped with a transparency adapter, at 600 dpi resolution and 8 bit gray scale levels. This produced bitmap graphics files with a sufficient resolution for reflecting distinct features important for the accurate characterization of objects. Before taking the measurements of certain geometric features, all images were rescaled to standardize the unit of measurement. Figure 2 presents the X-ray images of these kernels.

Image processing allowed a determination of size measures of wheat grain, as well as a possibility to compute shape coefficients [14, 16]. The complete analysis procedure of the obtained bitmap graphics files was based on the computer software package `Grains`, specially developed for the X-ray diagnostic of wheat kernels [10]. The most important feature of the package `Grains` is that it provides functions for ascertaining the particular characteristics of any selected kernel. In our work,

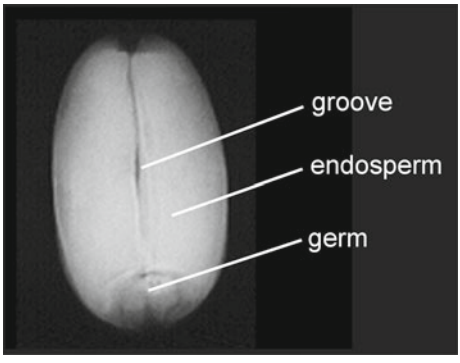


Fig. 1 X-ray image of an individual kernel

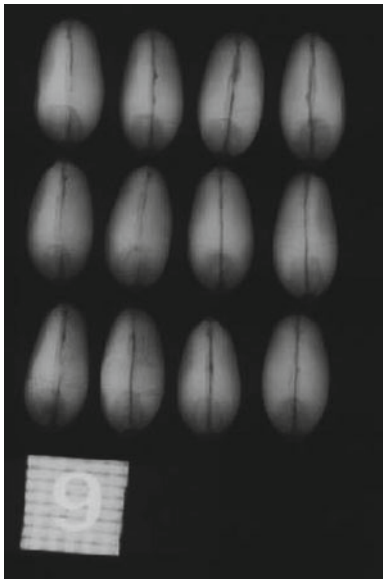


Fig. 2 X-ray photogram (13×18 cm) of kernels

using the commands available in the `Grains` menu analysis, automatic boundary detection, and, subsequently, calculation of main geometric and statistical parameters was accomplished for each individual grain kernel. Figure 3 provides an illustration of the contour indicating boundary of a kernel, as well as the obtainable measurements of the detected features.

To construct the data, the following descriptors were determined for each kernel:

- area A, understood as an area of a grain projection,
- perimeter P, understood as a perimeter of a grain projection,
- compactness,[1] given as $C = 4\pi A/P^2$,
- length of a kernel,
- width of a kernel,
- length of a kernel groove,
- asymmetry coefficient

where the asymmetry coefficient, given as a percentage, is a ratio between two quantities: the absolute value of the difference between areas of the left and right part of a kernel, and the total area of a kernel. Moreover, two parameters were taken by direct measurement using the package `Grains`. These being:

- area of a germ,
- length of a germ.

In addition to these numerical characteristics, the following parameters were obtained indirectly:

[1] The maximum value of the compactness is equal to one and is taken for a circle.

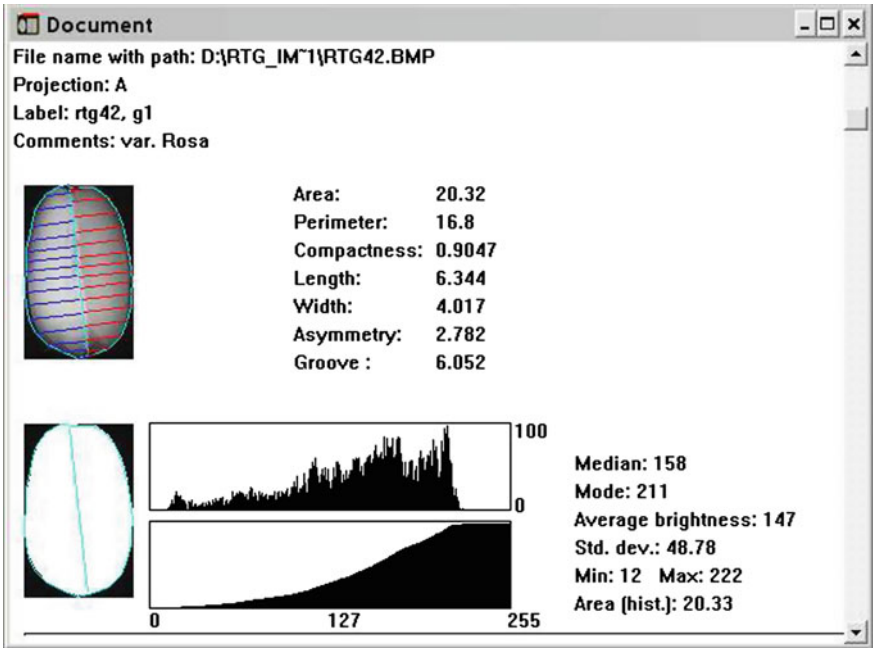


Fig. 3 Document window with geometric parameters of a kernel and statistical parameters of its image (as a unit of measure, millimeters were utilized)

- the ratio between the length of a germ and the length of a kernel,
- the ratio between the area of a germ and the area of a kernel,
- the ratio between the width of a kernel and the length of a kernel,
- the Feret coefficient.[2]

Computer image analysis of whole kernels allowed size and shape descriptors. Finally, the data contained the vectors of real-valued parameters determined for 288 samples belonging to three different varieties of wheat: Canadian, Kama, and Rosa. Randomly selected sets containing 100, 65, and 100 grains respectively, were then used to derive a discriminant procedure. The remaining observations of the test samples, as batches of 8, 7, and 8 grains, respectively, were considered as unknown categories, and, hence, comprised the test data set that was employed in the classification procedure. The level of procedure performance was evaluated by rate of correct classification.

[2]The ratio between the maximum diameter of the kernel in the vertical direction and the maximum diameter of the kernel in the horizontal direction, the measure of elongation.

3.2 Feature Selection

All measurements were made automatically from a total of 288 samples. Grains of the Canadian (the average length of kernels being equal to 5.250 ± 0.203, the average width of kernels being equal to 2.892 ± 0.201) and Kama (the average length of kernels being equal to 5.504 ± 0.230, the average width of kernels being equal to 3.236 ± 0.182) varieties are seen to be less in size than the Rosa (the average length of kernels being equal to 6.246 ± 0.390, the average width of kernels being equal to 3.749 ± 0.245).

In order to ascertain the usability of the determined geometric features in wheat grain discrimination, the distribution of the obtained histograms for each variable was analyzed, and, subsequently, correlation analysis was performed. A histogram distribution for all individual variables was symmetric and unimodal. Due to the significant relationships between some of these variables, these were reduced to the inclusion of the two best six-element sets that minimized correlations (Table 1).

Once the variables were selected, as an input, multidimensional analysis was performed for both data sets of feature vector components. This was done so as to extract the data set that better discriminates the investigated wheat varieties. The analysis was carried out through the separate use of forward and backward stepwise discriminant analysis. Firstly, the assumptions concerned with the discriminant analysis were assessed. Subsequently, multidimensional analysis was undertaken in order to discriminate the varieties. Finally, the classification process employing the discriminant functions, was performed. In so doing, no significant improvement was noticed when forward and backward stepwise discrimination was analyzed. Thus, the results for forward stepwise discriminant analysis will be shown alone.

Table 1 Feature vector components for discriminant analysis

Data set I	Data set II
1. Perimeter of a kernel	1. Area of a kernel
2. Compactness	2. Length of a kernel
3. Asymmetry coefficient	3. Asymmetry coefficient
4. The ratio between the length of a germ and the length of a kernel	4. Length of a germ
5. The ratio between the area of a germ and the area of a kernel	5. Area of a germ
6. The ratio between the width of a kernel and the length of a kernel	6. The Feret coefficient

3.3 Results and Discussion

To compute discriminant functions for the three-group classification, a training set containing 265 observations and two different feature vector components were used. Firstly, for both data sets of geometric features, discriminant analysis allowed the determination of which variables had significant contribution to the discrimination. In the data set I, the compactness was found to be insignificant and was removed from the model. In the data set II, all variables demonstrated significant contribution to the model.

Next, Wilks' Lambda criterion was used to assess feature discriminatory power. In the data set I, the kernel's perimeter, and subsequently, the ratio between the length of a germ and the length of a kernel, and the ratio between the area of a germ and the area of a kernel, were established as being most important in discrimination, while the asymmetry coefficient and the ratio between the width of a kernel and the length of a kernel were seen to be the least important. In the data set II, the length of a germ, and, subsequently, the Feret coefficient, the asymmetry coefficient, the area of a germ and the area of a kernel had the strongest discriminatory power, whereas the length of a kernel had the weakest power.

For both models, two discriminant functions were statistically significant. Thus, the first function discriminated mostly between Rosa, and Canadian and Kama combined. The level of discrimination was equal to 89 % for the data set I, and 88 % for the data set II. The second function was found to discriminate Kama variety, but the level of discrimination was lower and equaled 11 % and 12 %, respectively (Figs. 4 and 5).

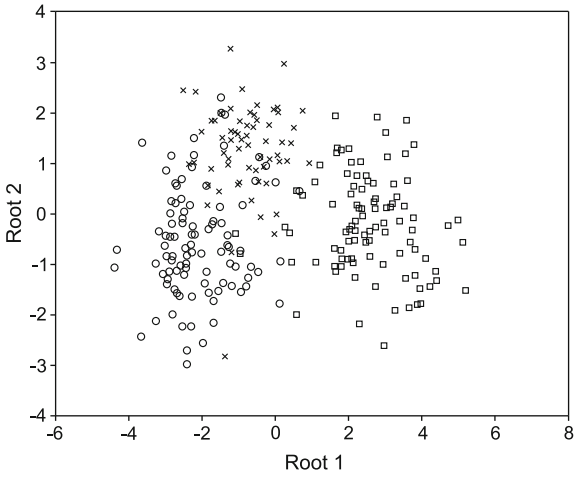


Fig. 4 Results of discrimination by scatterplots of canonical scores for data set I: (○) the Canadian wheat variety, (×) the Kama wheat variety, (□) the Rosa wheat variety

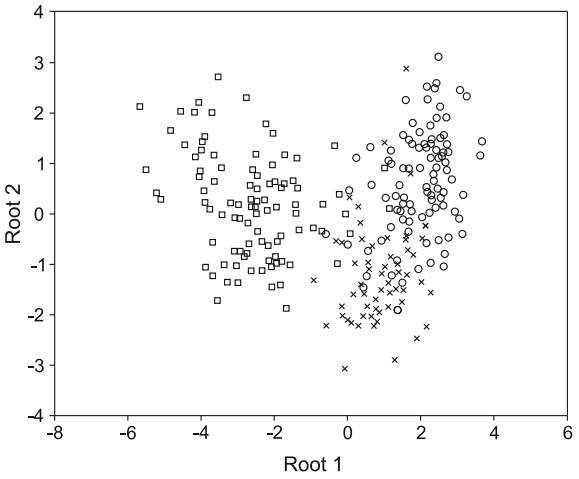


Fig. 5 Results of discrimination by scatterplots of canonical scores for data set II: (○) the Canadian wheat variety, (×) the Kama wheat variety, (□) the Rosa wheat variety

Table 2 The classification correctness for discriminant analysis

Geometric feature	Canadian (%)	Kama (%)	Rosa (%)	Total (%)
Data set I	92	89	96	92
Data set II	89	88	93	90

The classification correctnesses employing both sets of geometric features as an input for computing discriminant functions were summarized in Table 2. Use of the discriminant analysis for data set I resulted in higher proper classification, which, accordingly, equaled 92, 89, and 96 % for Canadian, Kama, and Rosa varieties. For the Rosa variety, 96 of 100 kernels were classified properly, and no kernels of the other two varieties were classified as being of the Rosa variety. For the Canadian variety, 92 of 100 kernels were correctly classified, while 2 kernels of the Rosa variety and 7 kernels of the Kama variety were mistakenly classified, as the Canadian variety. For the Kama variety 58 of 65 kernels were correctly classified, however, 2 kernels of the Rosa variety and 8 kernels of the Canadian variety were mistakenly classified as the Kama variety.

The classification rate was a bit worse when data set II was used. The correct classification results dropped to 89 % for Canadian, 88 % for Kama, and 93 % for Rosa varieties. For the Rosa variety, 93 of 100 kernels were classified properly and none of the other two varieties were misclassified as the Rosa variety. For the Canadian variety, 89 of 100 kernels were correctly classified, while 3 kernels of the Rosa variety and 8 kernels of the Kama variety were mistakenly classified as the Canadian variety. For the Kama variety, 57 of 65 kernels were correctly classified,

however 4 kernels of the Rosa variety and 11 kernels of the Canadian variety were mistakenly classified as the Kama variety.

To test the accuracy of the discrimination, a set of new observations utilizing 23 grains, was subjected to a classification process employing the evaluated discriminant functions. The results confirmed earlier conclusions on the discrimination: proper classification, accordingly, equaled six of 8, five of 7, and eight of 8 kernels for Canadian, Kama, and Rosa varieties. For both geometric feature sets all Rosa kernels were correctly classified, while Kama and Canadian varieties appeared to be misclassified, albeit to a insignificant degree. Herein, two Canadian kernels were classified as being Kama, one Kama kernel was classified as Canadian, and one Kama kernel was misclassified as Rosa. In this assessment a comparable correctness of classification was obtained to that when the nonparametric kernel discrimination was used. The classification rate, of the latter, though, was a bit better: two Kama kernels and only one Canadian kernel, were incorrectly classified. Hence, the nonparametric character of this method provided better fitting to the real data structure.

It is worth noting that the results of the classification confirmed the results obtained by the clustering analysis, which was conducted in the earlier stage of the research study [1].

4 Summary

Recent advances in digital image processing technique provide non-destructive tools for improving insight into seed morphology in terms of image acquisition and automatic feature detection. The accomplished study demonstrated that the image analysis commonly employed in discriminant methods gives reliable results in classifying wheat grain. In such work, selected combinations of geometric features permitted discriminant analysis to achieve a recognition rate of 89–96 %. The Rosa variety is, however, better discriminated, whilst Kama and Canadian varieties are harder to differentiate. Such a summation of results was also confirmed by the results of the subsequent classification of new observations, using both parametric and nonparametric methods. The conducted study confirmed the practical usefulness and effectiveness of the evolved methods in classification practices. Thus, discriminant analysis should be considered as being very effective in separating out wheat varieties.

References

1. Charytanowicz, M., Niewczas, J., Kulczycki, P., Kowalski, P.A., Łukasik, S., Żak, S.: Complete gradient clustering algorithm for features analysis of X-ray images. In: Pietka, E., Kawa, J. (eds.) Information Technologies in Biomedicine, pp. 15–24. Springer, Berlin (2010)
2. Draper, N.R., Smith, H.: Applied Regression Analysis. Wiley, New York (1981)
3. Guevara-Hernandez, F., Gomez-Gil, J.: A machine vision system for classification of wheat and barley grain kernels. Span. J. Agric. Res. **9**, 672–680 (2011)

4. Kowalski, P.A., Kulczycki, P.: Interval probabilistic neural network. Neural Comput. Appl. 1–19 (2015). doi:10.1007/s00521-015-2109-3
5. Krzyśko, M., Wołyński, W., Górecki, T., Skorzybut, M.: Systemy uczące się. Rozpoznawanie wzorców, analiza skupień i redukcja wymiarowości (Learning systems. Pattern recognition, cluster analysis and dimensionality reduction). WNT, Warszawa (2008) (in Polish)
6. Kulczycki, P.: Estymatory jadrowe w analizie systemowej (Kernel Estimators in Systems Analysis). WNT, Warszawa (2005). (in Polish)
7. Kulczycki, P.: Kernel estimators in industrial applications. In: Prasad, B. (ed.) Soft Computing Applications in Industry. Springer, Berlin (2008)
8. Kulczycki, P., Kowalski, P.A.: Bayes Classification for nonstationary patterns. Int. J. Comput. Methods **12**(2), 155008–19 (2015)
9. Majumdar, S., Jayas, D.S.: Classification of cereal grains using machine vision: I. Morphology models. Am. Soc. Agric. Eng. **43**(6), 1669–1675 (2000)
10. Strumiłł o, A., Niewczas, J., Szczypiński, P., Makowski, P., Woźniak, W.: Computer system for analysis of X-ray images of wheat grains. Int. Agrophys. **13**, 133–140 (1999)
11. Utku, H.: Application of the feature selection method to discriminate digitized wheat varieties. J. Food Eng. **46**, 211–216 (2000)
12. Utku, H., Koksel, H., Kayhan, S.: Classification of wheat grains by digital image analysis using statistical filters. Euphytica **100**, 171–178 (1998)
13. Wand, M.P., Jones, M.C.: Kernel Smoothing. Chapman and Hall, London (1994)
14. Wiwart, M., Suchowilska, E., Lajszner, W., Graban, Ł.: Identification of hybrids of spelt and wheat and their parental forms using shape and color descriptors. Comput. Electron. Agric. **83**, 68–76 (2012)
15. Zapotoczny, P.: Discrimination of wheat grain varieties using image analysis and neural networks. Part I. Single kernel texture. J. Cereal Sci. **54**, 60–68 (2011)
16. Zarychta, P.: Features extraction in anterior and posterior cruciate ligaments analysis. Comput. Med. Imaging Graph. **46**, 108–120 (2015)
17. Zhu, M., Hastie, T.J.: Feature extraction for nonparametric discriminant analysis. j. comput. graph. stat. **12**(1), 101–120 (2003)

Face Localization Algorithms as Element of Silhouette Localization Process

Mariusz Marzec, Aleksander Lamża, Zygmunt Wróbel and Andrzej Dziech

Abstract The article presents an algorithm that allows for automatic detection of rectangular masks of human silhouettes that are in the foreground (significant silhouettes) in colour images with complex backgrounds. A very important element of the algorithm was elimination of faces and silhouettes in the background. Typical solutions presented in the literature are based on two approaches of silhouette localization. The first group of methods is based on initial background suppression, which allows for silhouette extraction. The second group of methods is related to direct silhouette localization. According to assumptions the method based on localization of individual silhouette elements has been used. The face, head and torso have been selected as essential silhouette elements and face localization has been used as the main algorithm block. The method presented in the article provides 79 % effectiveness of silhouette localization in different image collections taking into account elimination of silhouettes in the background.

Keywords Face localization · Head localization · Silhouette localization · Silhouette segmentation

1 Introduction

The study presented in this paper aims at preparing an algorithm for automatic detection of a mask covering the area of a human silhouette. It was assumed that the mask should cover the most important silhouette areas. The face, head and torso were adopted as the important areas. The results should allow for automatic addition

M. Marzec (✉) · A. Lamża · Z. Wróbel
Department of Computer Biomedical Systems, Institute of Computer Science, University of Silesia, Będzińska 39, 41-200 Sosnowiec, Poland
e-mail: mariusz.marzec@us.edu.pl

A. Dziech
Faculty of Mining Surveying and Environmental Engineering, AGH University of Science and Technology, Al. A. Mickiewicza 30, 30-059 Kraków, Poland

E. Piętka et al. (eds.), *Information Technologies in Medicine*, Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_5

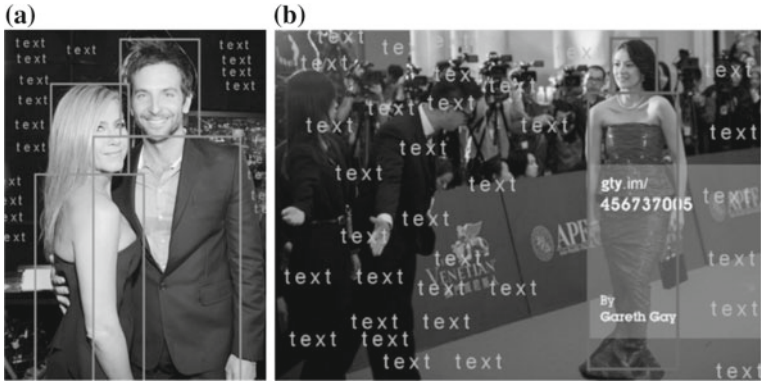


Fig. 1 Example of designating areas and adding text information performed manually by the authors

of text information to the test image in such a way so that it will not distort the information contained therein or it will distort it minimally. Automatically added information (e.g. descriptive text) should be applied to the test image in such a way so as not to cover the most important silhouette elements, i.e. the face or torso of the person in the image.

There arise common problems with which the algorithm has to cope correct detection and removal of the background, detection of the area taking into account the orientation and position of the head and torso, detection of silhouette area homogeneity with regard to lighting, texture, colour and classification of areas in order to select only people in the foreground.

In Fig. 1 mentioned operations were carried out manually in order to present the assumptions and objectives of the study in graphic form. It can be observed that despite the addition of large amounts of descriptive information, the key elements of the silhouettes in the foreground, namely the head and torso, are still clearly visible in the image (they are still recognizable). A high background complexity such as variety of colours, textures and small faces and silhouettes can also be observed.

2 Related Work

2.1 Silhouette Localization Process

A silhouette is understood as the outline of a person or object [5, 25] located in an image. As described in mentioned literature silhouette extraction or segmentation is the process of separating the area containing the silhouette from the background area. The biggest problems in the process of silhouette segmentation are variable lighting, heterogeneity and a high degree of complexity of the background and the

silhouette area [25]. Silhouette segmentation is a very important task in tracking systems, detection of people, detection of activity in an image [5].

In the available literature, there are two attitudes used in the silhouette localization process:

- silhouette localization through background removal [7, 9, 21],
- direct silhouette localization [2, 6, 20, 23].

The first group of methods is based on the assumption that it is possible to prepare a suitable background model which will remove the background leaving only the sought human silhouette. The second group of methods refers to the direct search for the silhouette area. In this case, a suitable model representing the silhouette area is prepared. This group is very interested in point of view this article.

The methods for direct silhouette localization can be divided into two groups:

- localization of a silhouette as a whole [2, 6, 20, 23],
- localization of a silhouette as a set of related parts [3, 14].

The first step in the second group of methods is to determine which areas of the silhouette are important and will be subjected to the localization process. In publication [14], the authors proposed an algorithm that allows for localization of a silhouette as a whole or only the upper torso. The human silhouette model was built as a collection of geometrically related parts which included the head (front and side), the upper part of the body (front and side) and legs. Each of these body parts was described using an *orientation based feature*, i.e. the local histogram of gradient orientations and their local grouping. The first stage involved determination of the dominant gradient orientation in close proximity (3×3 pixels) and in different scales (5 ranges). Then, the individual silhouette parts were localized based on local characteristics. Finally, the whole body was localized based on relationships between silhouette parts and their combination. Publication [3] presents a silhouette model made up of so-called *bubbles* or image portions of small random Gaussian distribution. In typical cases of people in a standing position, the thus obtained masks represented the head and torso, hands and feet. In order to link the obtained masks with each other, geometric relations defining their spatial relationships were used. In order to detect silhouettes made of the designated *bubbles*, ap-ROI detector developed by the authors was applied. The results obtained by the authors gave 40 % efficacy in localization of a silhouette as a whole.

Due to the fact that the face area provides the most information about the person in the picture, face localization block was applied as the main module of the proposed algorithm. It was also established that the most essential silhouette elements besides the face are head and torso. Typical images from the test set of images are presented in Figs. 1, 3, and 8.

2.2 Face Localization

The rich literature on face detection and localization proposes a wide variety of solutions [1, 24, 26]. They can be divided into four basic groups methods based on knowledge about the image, methods based on basic images features, template matching and machine learning methods. In knowledge-based methods, knowledge about the image is described as a set of rules and relationships describing a typical face. The face is described as a set of characteristic points or areas related to each other in terms of distance, mutual position, etc. An example of their application can be found in publication [13], where the authors proposed a face detection method based on geometric relationships of characteristic areas of the face such as the eyes and mouth. Another group of methods (called *feature invariant approaches*) includes algorithms whose task is to look for the basic image features (*invariants*) such as edges, colours, shapes, areas with a particular colour, texture. Examples of solutions based on colour segmentation are presented in publications [1, 8, 18]. The third group of methods relates to template matching. In this group of methods, face detection involves the use of specially prepared templates describing the face as a whole or the individual facial features (i.e. nose, mouth) separately. These templates are prepared by experts based on knowledge about the test image. After matching templates to the image, the correlation between individual templates is calculated. An example of their application is found in publication [10] where the authors propose the use of a deformable model representing the shape of the face and its most important elements to determine the face position. The last group of methods (*appearance-based methods* or *learning-based methods*) are algorithms that use models, templates and patterns resulting from the learning process. Proper operation requires a training set based on which the algorithm learns to recognize and create an appropriate pattern. An important factor in the effectiveness of this group is the right choice of the training set (images containing valid and invalid examples) and methods of learning and classification. An example can be found in publication [15], where the authors propose the use of a unidirectional feed-forward network for face detection based on the skin colour.

Presently used solutions most often belong to the last group of methods and have appeared in the last few years due to the rapid increase in computing power. Learning-based algorithms are gaining an increasingly greater advantage in face localization systems [26] in relation to the 3 previous groups. The most currently implemented solutions are methods based on the Local Binary Pattern [4, 27] and Haar-like features [16, 19, 22]. In publication [4], the authors proposed a face localization method based on image features referred to in the literature as *Local Binary Pattern features*. They provide good results in a wide variety of applications (texture classification and segmentation, image extraction), and therefore they also operate effectively in the case of face localization. In publication [22], the first stage of the algorithm operation involved the use of image features similar to the basic Haar functions, called *Haar-like features*, which enable to encode the difference in intensity between two adjacent rectangular areas of different size as well as extract effectively the

texture regardless of the absolute image intensity. Thus determined features were subjected to classification using AdaBoost. The effectiveness of face localization ranged from 76.1 to 93.9 % depending on the allowable number of false detections. After reviewing the above methods, it has been decided that the method based on the Haar cascade classifier will be used in the face localization block.

3 Problem of the Foreground Silhouette Localization

The proposed method assumes that in each of the test images there is at least one human silhouette and the first elements which will be automatically localized by the system will be the face and head of the person in a still colour image. Based on the analysed literature it was also assumed that the head detection in the first stage and for the test images will be easier than the comprehensive detection of the whole body. In the study, it was assumed that the most important information from the point of view of the realized task is present in the face area and then in the torso area (Fig. 2), so these areas need to be designated in such a way so that the added descriptive content will not distort the information in the image. Another important purpose of algorithm is the need to eliminate *irrelevant* faces so that the designated masks of the face and torso will only represent the person(s) located in the foreground. The image in Fig. 2a demonstrates how the algorithm should behave in accordance with the adopted assumptions and how it could behave without being able to identify the most significant person in the image. Due to the large number of silhouettes located in the background in Fig. 2b, little space was left for adding information, despite the fact that the people in the foreground do not take up so much space.

In view of the above-mentioned assumptions, it was decided that the analysis will be divided into three main steps: significant face localization block will operate first, the second algorithm block will designate the head area and the third torso localization block will be based on the data obtained from the head localization.

It was also assumed that the most important people in the image are in the foreground, face and head sizes in this case are the biggest.

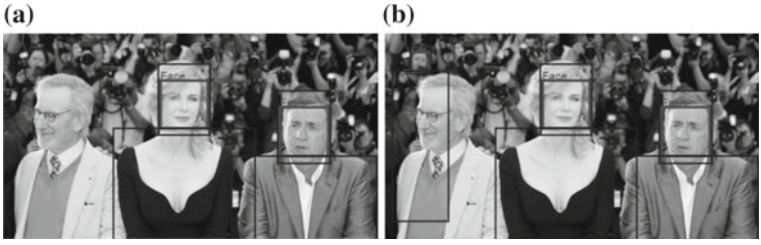


Fig. 2 Illustration of silhouette masks localization and the problem with localizing many people, **a** correct localization, **b** incorrect localization (face in background)

4 Proposed Method for Detection Foreground Silhouettes

The research material was a collection of 137 colour images of varying sizes and content. In each of the images there was at least one person in the foreground. The scale, number and position of silhouettes as well as the background were varied.

The effectiveness of the algorithm was defined as the ratio of correct detections to the number of detections determined by an expert.

The cases when the silhouettes were located correctly in the foreground were treated as correct detections (as in Figs. 2a and 3a—2 silhouettes) and the rectangular mask covered the most significant areas, namely the head and torso.

False-positive detections were these cases when the silhouettes of people in the background, i.e. *irrelevant* ones, or areas that were not actually human silhouettes were found—Figs. 2b and 3b, c. False-negative cases, in turn, were the ones when the silhouette of a person in the foreground was not found. An example of determining the effectiveness for the images with a single silhouette is shown in Fig. 3c—large silhouette of people in the foreground was not localized.

Described algorithms are implemented in the production system which automatically adds text information on images. The algorithm described in the article was implemented in the Java language using OpenCV package [28].

The authors decided that the first step of the proposed algorithm was the face localization block. As mentioned in the introduction, methods for locating the head/face are widely used, and their effectiveness is over 90 %. It is a slightly different approach in relation to methods where the mask covering the test object (e.g. the whole person in the image) is determined directly [2, 6, 20, 23]. First research on using face localization method as a potential element of silhouette localization process was carried by authors in [12]. According to results of face's mask localization and effectiveness authors focused on application this solution in the face localization block.

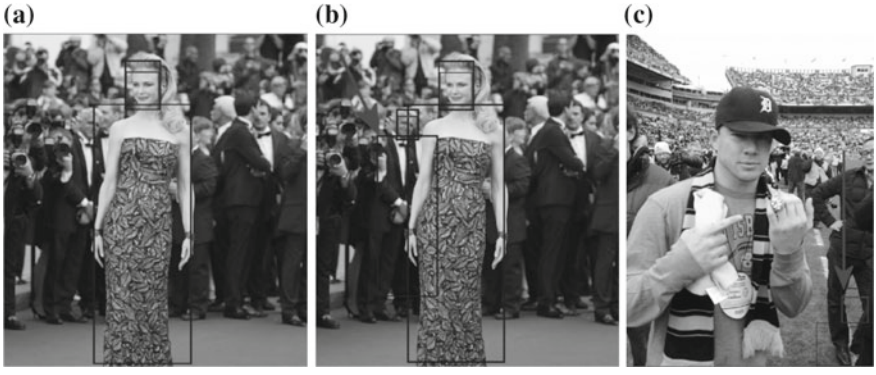


Fig. 3 Localization of silhouettes, **a** relevant, **b** irrelevant, **c** false-positive and false-negative

4.1 Detection of the Face and Head Area

In the face localization block authors used the Haar cascade classifier. The classifier present in OpenCV as Cascade Classifier was applied to implement the algorithm (standard training set was used). The effect of the classifier operation is a list of rectangular masks defining the areas of detected faces. In the next step, they are sorted in descending order of size. In order to eliminate less relevant faces (in the background or smaller ones), filtering with a fixed threshold can be applied. The value of 60 % of the largest localized mask area was selected experimentally based on tests. With such a threshold value, the algorithm correctly located the faces of people in the foreground in the case of images containing both the entire silhouette of a person and also the face with a fragment of the body. This block of the algorithm also involves elimination of cases of facial masks which overlap so as to avoid multiple detection of the area belonging to one silhouette. The result of the first algorithm block is determination of the coordinates of the upper left corner Fx and Fy as well as the width Fw and height Fh of the face area. Examples of the face localization block operation are presented in Fig. 4a, b.

In the next step, the head area is calculated based on the face area obtained above. The simplest and quickest method, namely the geometrical method, was proposed. Based on the analysis of areas returned by the face detector and the anatomical dimensions of the human [11, 17] correction coefficients K were determined allowing for the estimation of the expected head area.

$$K = \left(k_x, k_y, k_w, k_h\right) \tag{1}$$

where k coefficients correspond to: coordinate x, coordinate y of the upper-left corner, width and height of the head.

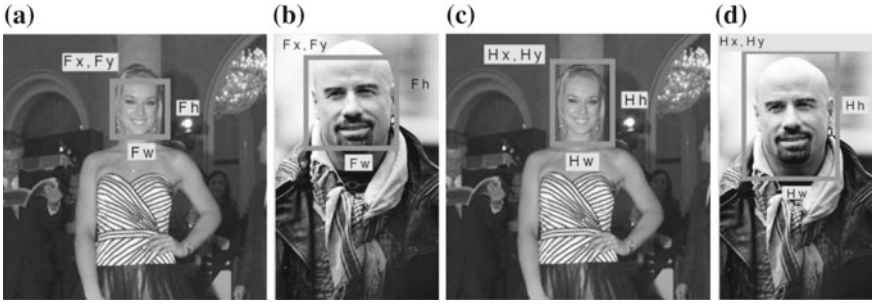


Fig. 4 Examples of correct face localization (**a, b**) and head areas, determined with the geometrical method (**c, d**)

The head area is determined according to the following formulas:

$$\begin{aligned} H_x &= F_x - k_x F_w, \\ H_y &= F_y - k_y F_h, \\ H_w &= k_w F_w, \\ H_h &= k_h F_h. \end{aligned} \tag{2}$$

The performed tests provided optimal values of the coefficient:

$$K = (0.05, 0.35, 1.1, 1.5)\,. \tag{3}$$

Figure 4c, d presents two sample images after applying the method described above. The smaller rectangle is the face area obtained from the face detection block, while the larger rectangle is the designated head area. First example shows face and whole silhouette and second shows face and fragment of silhouette.

4.2 Detection of the Torso Area

The next step of the algorithm operation is to calculate the silhouette area of the person in the test image. Based on the analysis of human anatomical parameters [11, 17] and the results of preliminary tests, it was assumed that the silhouette mask should meet the relations as in Fig. 5c. The height and width of the torso mask were marked B_h and B_w respectively. The coordinates of the upper left corner were marked B_x, B_y—Fig. 5a, b.

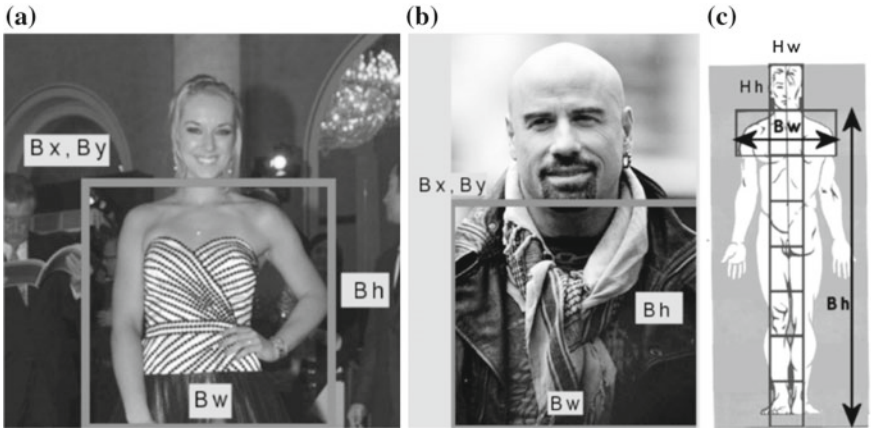


Fig. 5 Examples of the determined torso masks and Geometrical torso mask based on [17]. $B_w = 3H_w$ and $B_h = 7H_h$

Above determined values enabled to mark the whole silhouette in the case of people standing upright and in the case where only part of the silhouette was in the image. In this case, a rectangular silhouette area was limited by the image dimensions. The conducted tests have shown that the geometrical method for determining the torso area is fast but the operation precision is sufficient to avoid significant distortion of the torso area.

5 Experimental Results

The collection of 137 images prepared by the authors was subjected to detailed studies on the effectiveness of determining the forbidden area covering the silhouette of people in the foreground.

In order to compare the performance of the proposed algorithm, the set of images was analysed using 2 selected methods (Table 1 presents the results):

1. The developed method based on the localization of the silhouette elements with the Haar Cascade Classifier [22] (applied in the Faces areas localization block) and then geometrical silhouette detection.
2. The developed method based on the localization of the silhouette elements with the Local Binary Pattern [4] (applied in the Faces areas localization block) and then geometrical silhouette detection.

Figure 3 presents the methodology for interpretation of the results. On this basis, a graph of silhouette localization effectiveness—Fig. 6 and a graph of the number of false negative localizations—Fig. 7 were created. The correct silhouette detections are defined as cases when the forbidden silhouette areas covering the head and torso of a person or people were marked properly. In the results obtained, the number of correct detections of silhouettes in the foreground reached 79 % (149 correct detections out of 188 marked by an expert). The graph in Fig. 6 shows that in a small

Table 1 Comparison of methods used in face localization block

Method	Publication	Applied algorithm	Discussion of results
Method 1	The developed method—the haar cascade classifier as a face detector [22]	Haar like cascade used in the face localization module	Number of studied Images: 137 Positive: 79 % False positive: 12 % False negative: 21 %
Method 2	Local binary pattern as a face detector [4]	LBP is used in the face localization module	Number of studied Images: 137 Positive: 65 % False positive: 17 % False negative: 35 %

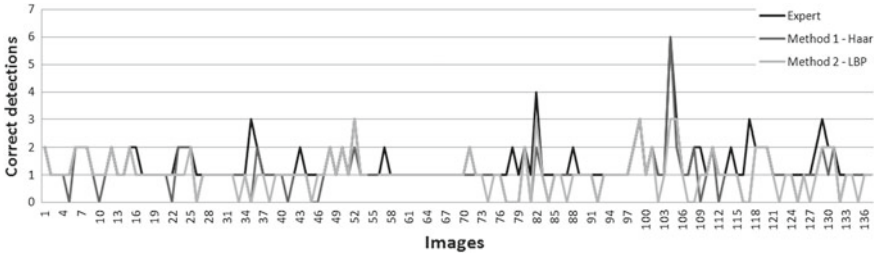


Fig. 6 Number of correct detections of silhouette area

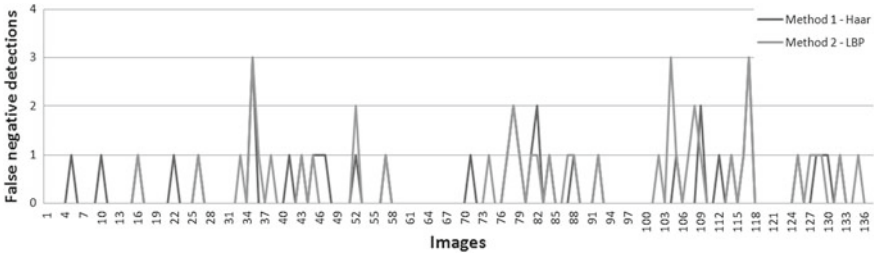


Fig. 7 Number of non-detected silhouettes

group of images, the algorithm did not find the number of silhouettes determined by the expert and there were only few cases when no silhouette was found.

Analysing the results it has been found that problems with the silhouette detection were most often caused by false positive detections in the face detection block. Such situations happened when the algorithm found in the foreground an area which was classified as the face (even though it actually was not the face) and rejected the other faces for which the face area was less than 60 % of the largest found area. Such a case is presented in Fig. 8c and 7 % of these cases occurred in the test set.

Examples of correct detections for different cases of images with unknown number of silhouettes are presented in Fig. 8a, b, d. It can be observed here that despite of typical problems: the complex background, many people in background the algorithm correctly locates important silhouettes defining the forbidden areas.

The number of correct silhouette localizations can be increased by improving the classifier, which will eliminate the cases as in Fig. 8c. After disabling the threshold of elimination of background silhouettes, the results of analysis get better immediately Fig. 8d, but there occur too many false positive detections. The coefficient of correct detections in the test set reached 79 %, with 12 % false positive and 21 % false negative localizations.

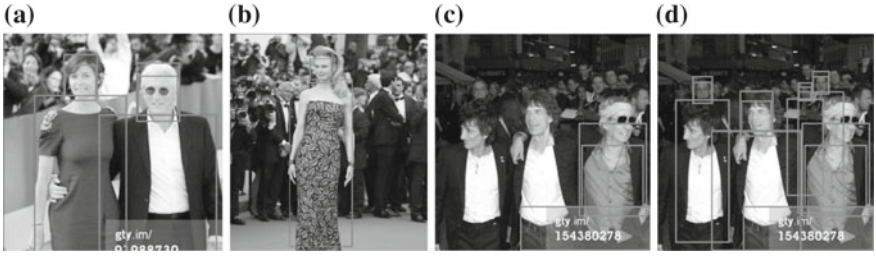


Fig. 8 Examples of the results of silhouette detection algorithm operation

6 Conclusions

The obtained results of the research indicate that the adopted assumptions and methods and their modifications allow to automate of the task of determining masks covering silhouettes of people in the foreground in static images. The analysis and comparison of known, face detection methods used in the literature made it possible to determine the actual capabilities and potential ways of development of the algorithm presented in the article. Attention was also drawn to the decrease in the number of correct detections in the event of a false positive detection with a large face area. The solution to this problem would be to eliminate large *non-faces* that block the search for smaller (appropriate areas of the face) through modifications of the algorithm at the stage of locating the face in the foreground. The elimination of such cases (which accounted for about 8 % of the set of images) would increase the effectiveness of the algorithm. At this stage, 79 % efficacy was obtained in the combined set of images prepared by the authors.

Acknowledgments This work was supported by the UE European Union, Innovative Economy Programme, European Fund for Regional Development *Intelligent Information System for Global Monitoring, Detection and Threat Identification* (INSIGMA), Project number: POIG 01.01.02-00-062/09-00.

References

1. Beigzadeh, M., Mafadoost, M.: Detection of face and facial features in digital images and video frames. In: Cairo International Biomedical Engineering Conference CIBEC, pp. 1–4 (2008)
2. Bolme, D.S., Lui, Y.M., Draper, B.A., Beveridge, J.R.: Simple real-time human detection using a single correlation filter. In: Proceedings of the Twelfth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (2009)
3. Castrillón-Santana, M., Vuong, Q.C.: Combining human perception and geometric restrictions for automatic pedestrian detection. In: Current Topics in Artificial Intelligence, vol. 4177, pp. 163–170. Springer (2006)
4. Chang-yeon, J.: Face detection using IBP features. Technical Report. Stanford University. CS 229 Final Project Report (2008)

5. Chen, X., He, Z., Anderson, D., Keller, J., Marjorie Skubic, M.: Adaptive silhouette extraction and human tracking in dynamic environments. In: Proceedings of the International Conference on Image Processing (2006)
6. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) (2005)
7. Elgammal, A., Harwood, D., Davis, L.: Non-parametric model for background subtraction, computer vision—ECCV 2000, vol. 1843, pp. 751–767 (2000)
8. Hsu, R.L., Abdel-Mottaleb, M., Jain, A.K.: Face detection in color images. IEEE Pattern Anal. Mach. Intell. **24**, 696–706 (2002)
9. Kim, H., Sakamoto, R., Kitahara, I., Toriyama, T., Kogure, K.: Robust silhouette extraction technique using background subtraction. In: Proceedings of the 10th Meeting on Image Recognition and Understand, MIRU (2007)
10. Lanitis, A., Taylor, C., Cootes, T.F.: An automatic face identification system using flexible appearance models. Image Vis. Comput. **13**(5), 392–401 (1995)
11. Loomis, A: Figure Drawing for all it's Worth. Titan Books, Facsimile edn. (2011)
12. Marzec, M., Lamża, A., Wróbel, Z., Dziech, A.: Methods for face localization in static colour images with an unknown background, MCSS 2014, CCSIS 429, pp. 170–181. Springer International Publishing, Switzerland (2014)
13. Miao, J., Gao, W., Chen, Y., Lu, J.: Gravity-center template based human face features detection. In: Proceedings of the Advances in Multimodal Interfaces—ICMI 2000, vol. 1948/2000, pp. 207–214. Beijing, 14–16 Oct. 2000
14. Mikolajczyk, K.,Schmid, C., Zisserman, A.: Human detection based on a probabilistic assembly of robust part detectors. In: Proceedings of the 8th European Conference on Computer Vision, vol. 3021, pp. 69–82. Prague, Czech Republic, 11–14 May 2004
15. Mostafa, L., Abdelazeem, S.: Face detection based on skin color using neural network. In: GVIP 05 Conference. Cairo, Egypt, pp. 53–58 (2005)
16. Padilla, R., Costa Filho, C.F.F., Costa, M.G.F.: Evaluation of Haar cascade classifiers designed for face detection. World Acad. Sci. Eng. Technol. **64** (2012)
17. Peck, S.R.: Atlas of Human Anatomy for the Artists. Oxford University Press (1951)
18. Phung, S.L., Bouzerdoum, A., Chai, D., Kuczborski, W.: A color-based approach to automatic face detection. In: Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology, pp. 531–534 (2003)
19. Ramirez, G.A., Fuentes, O.: Multi-pose face detection with asymmetric haar features. IEEE Workshop Appl. Comput. Vis. WACV **2008**, 1–6 (2008)
20. Schwartz, W.R., Kembhavi, A., Harwood, D., Davis, L.S.: Human detection using partial least squares analysis In: Proceedings of the IEEE 12th International Conference on Computer Vision, 29 Sept. 2009–2 Oct 2009
21. Sulaiman, S., Hussain, A., Tahir, N., Samad, S., Mustafa, M.: Human silhouette extraction using background using background modeling and subtraction. Inf. Technol. J. **7**(1), 155–159 (2008)
22. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proceedings of 2001 IEEE International Conference on Computer Vision and Pattern Recognition, pp. 511–518 (2001)
23. Wren, Ch., Azarbayejani, A., Darell, T., Pentland, A.: Real Time Tracking of Human Body. IEEE Transactions on Pattern Analysis and Machine Intelligence **19**(7), 780–785 (1997)
24. Yang, M., Kriegman, J., Ahuja, N.: Detecting faces in images: a survey. IEEE Trans. Pattern Anal. Mach. Intell. **24**(1), pp. 34–58 (2002)
25. Zeng, G., Quan, L.: Silhouette extraction from multiple images of an unknown background. In: Proceedings of the Asian Conference of Computer Vision (2004)
26. Zhang, C., Zhang, Z.: A Survey of recent advances in face detection. Technical report, Microsoft Research, **577** (2010)
27. Zhang, L., Chu, R., Xiang, S., Liao, S., Li, S.Z.: Face detection based on multi-block LBP representation. Adv. Biometrics Lect. Notes Comput. Sci. **4642**, 11–18 (2007)
28. http://opencv.org/opencv-java-api.html

Imaging of the Anterior Eye Segment in the Evaluation of Corneal Dynamics

Magdalena Jędzierowska, Robert Koprowski and Zygmunt Wróbel

Abstract The imaging of the anterior segment of the eye plays a crucial role in today's ophthalmology. Despite the variety of existing methods, only a few have the possibility to estimate the corneal dynamics. The purpose of this article is to review the currently available methods of imaging the anterior segment of the eye capable of evaluating the dynamics of corneal deformation in response to air pulse. In the paper the Corvis ST, Ocular Response Analyzer (ORA) and optical coherence tomography (OCT) combined with air puff system were described and discussed.

Keywords Anterior segment imaging · Cornea · Corneal dynamics · Optical coherence tomography (OCT) · Ocular response analyzer (ORA) · Corvis ST · Image processing

1 Introduction

There are many methods of imagining the anterior segment of the eye available nowadays. The ideal system should be a non-contact and non-invasive device, providing a real-time information. What is more, it should visualize structures with high quality, contrast and magnification. Besides, it would be easy to use and of a reasonable price [41].

The traditional anterior segment imaging method is still the slit lamp biomicroscopy, first introduced in 1911 [35]. Nevertheless, this technique for quantitative assessment of the anterior structure of the eye is limited—the magnification and image quality are restricted by the slit lamp [18]. Another system is ultrasound biomicroscopy (UBM). Its main advantage is the unique possibility to show the structure behind the iris [37]. Moreover, UBM enables to observe the real-time changes

M. Jędzierowska (✉) · R. Koprowski · Z. Wróbel
Faculty of Computer Science and Material Science, Department of Biomedical Computer Systems, Institute of Computer Science, University of Silesia, ul. Będzińska 39, 41-200 Sosnowiec, Poland
e-mail: magdalena.jedzierowska@us.edu.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_6

in the human eye. Despite this fact and other, not mentioned advantages, it has not become the part of regular clinical practice [26]. A different modality that visualizes the anterior segment structures is slit-scanning imaging (e.g. Orbscan II). It combines two techniques: Placido disk and slit-scanning enabling to reconstruct a three-dimensional image of the cornea [11]. We can also list four devices adopting the Scheimpflug imaging principle [32]. One of them is Pentacam which uses two cameras: a rotating camera (Scheimpflug camera) and a static camera. This system produces a series of topographic data: the anterior and posterior corneal surface and other, such as: corneal pachymetry, anterior chamber depth or corneal wavefront aberrations [18, 32]. Other equipment is anterior segment optical coherence tomography (AS-OCT) which has developed from the retinal OCT. It uses a longer wavelength and consequently has a better penetration through light-retaining tissues such as sclera and limbus. This feature of AS-OCT makes it possible to visualize the cornea, iris, anterior chamber angle and anterior lens [18, 37].

The imaging methods mentioned above delivered a variety of important parameters such as: anterior chamber depth (ACD), central corneal thickness (CCT), anterior corneal curvature (ACC), corneal topography data, etc. However, recently the one of the most desirable abilities is the possibility to present the dynamic changes in the eye.

To date, the technology adopting the high-speed Scheimpflug camera and classical air-puff tonometry leads in assessing the cornea dynamic. Commercially available device based on this method is Corvis ST which records the entire deformation process and provides a series of dynamic parameters, e.g. deformation amplitude (DA) and corneal velocities. Another available system is the Ocular Response Analyzer (ORA, Reichert Technologies) which was the first non-contact tonometer allowing to obtain the biomechanical properties of the cornea. The main parameter produced by ORA is corneal hysteresis (CH) which presents the dynamic response of the cornea to an air puff stimulus [28].

The main purpose of the presented review is to provide an overview of imaging methods capable of assessing the cornea dynamics in the context of its deformation in response to an air puff. The limitations and benefits of these methods will be discussed. Moreover, authors will try to answer why the corneal parameters measured during dynamic displacement are so important and what they indicate.

2 Corvis ST

Corvis ST is a noncontact tonometer which uses a high-speed Scheimpflug camera to record the complete process of cornea displacement in response to the air pulse. Sampling rate of the camera is 4330 frames per second and the horizontal coverage is 8.5 mm [13]. The Corvis ST software supplies several ‘built in’ parameters (Fig. 1) which can describe the dynamic and pseudo-dynamic cornea behavior, like deformation amplitude (DA), length of applanation and velocity of cornea apex during the deformation etc. [31, 45].

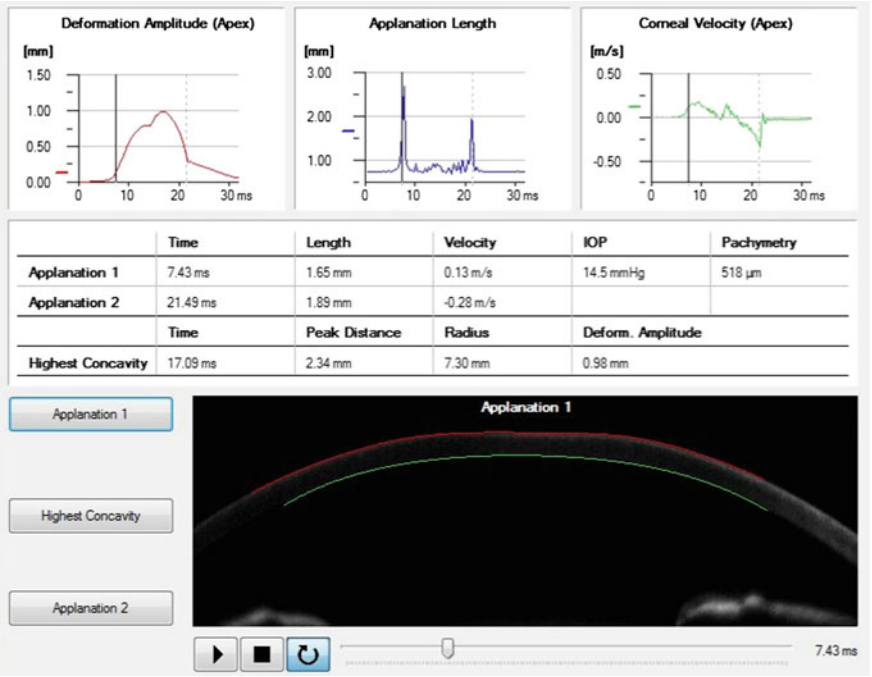


Fig. 1 Screenshot of the Corvis ST display with the available parameters

The repeatability of Corvis ST parameters were tested in [3, 12, 25, 31]. Hon and Lam [12] used a population of 37 healthy adults and pointed out that the deformation amplitude was found repeatable and reproducible. Authors considered it very promising to distinguish the normal corneas from the pathologically changed ones based on this feature. Larger populations were used in publications [25, 31]. Nemeth et al. [31] examined 75 eyes of 75 of healthy subjects, Leung et al. [25] used a much bigger group—180 eyes of 104 subjects (40 normal, 140 glaucoma suspect/glaucomatous eyes). Both confirmed the good repeatability of DA. Most publications agreed that DA is associated with intraocular pressure (IOP) and central corneal thickness (CCT). The higher IOP, the smaller DA [17, 42] and the greater CCT, the smaller DA [12, 42].

Furthermore, deformation amplitude was also investigated in [3] where the authors of the analysis showed that this feature has a potential to identify patients with keratoconus (45 keratoconic and 103 healthy eyes examined). DA was notably greater in keratoconic eyes compared with healthy eyes. However, in this and another work [16] (population of 60 keratoconic and 60 normal eyes) authors pointed out that the results should be still confirmed in further research just to ensure its usefulness in practice.

Based on the collected literature, it is very likely that this dynamic parameter can be one of the most suitable factors for characterizing the dynamics of biomechanical properties of the cornea.

The dynamic changes in cornea deformation registered by Corvis ST were used in paper [36] to explore age related changes in corneal biomechanics in fifty subjects with healthy corneas. Authors used the time-varying Chebyshev polynomial based model and proposed to combine the polynomial coefficient with the subsequent dynamic cornea stages. The first stage: time from the beginning to the maximum concavity of the cornea, the second stage: corneal oscillation phase and the third stage: returning of the cornea to its initial state. Final results showed that age-related differences are mostly seen (statistically significant—Wilcoxon test, $P < 0.05$) in the third stage (which includes the second applanation phase) of the inner corneal contour. It means that the corneal deformation dynamics in that phase changes with age. This outcome confirms the possibility of detecting such discreet relations like dynamic-age dependency with high-speed Scheimpflug camera. However, it should be noted that authors described their results for the 10 % trim of the profile data.

Recently Ji et al. [16] presented a complex paper concerning several aspects of analyzing the dynamic process of corneal deformation based on the video from the Corvis ST. In this paper authors used a group of 40 healthy and 30 keratoconus eyes. Employing their new proprietary method of cornea segmentation and adopting color-appearance-model for image fusion (firstly presented by Li et al. [27]) they proposed a new dynamic parameter: the dynamic curvature topographic map. Dynamic curvature topographic map links the values of anterior cornea curvature (ACC) with the correspondent single frame of Corvis ST video. The values of ACC are identified by colors (form normalized color pallet) what enables in an intuitive way to analyze the dynamic behavior of the anterior corneal surface through the whole cornea response process (Fig. 2). Authors mentioned that the produced map gives an opportunity to distinguish normal from abnormal corneas. The results of the classification of various pathologies of the cornea using topographic maps would be published in further work. In addition, it should be mentioned that the segmentation method proposed by authors has proved to be very accurate and shown better insusceptibility to disruption than Corvis ST built-in method.

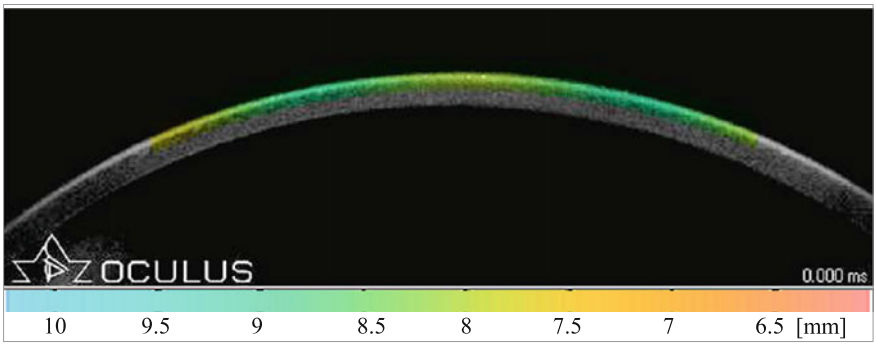


Fig. 2 An example of average curvature topographic map [16]

Another approach to corneal dynamic features available from Corvis ST images is presented in the next paper [19]. In this work the problem of classification of keratoconic and healthy corneas based on corneal vibrations is mentioned. For this purpose a very large number of subjects were tested—269 healthy and 477 keratoconus eyes. Due to advanced image analysis and processing method, authors detected the cornea vibrations based on a series of images from the cornea deformation video. In order to obtain the selective parameters for the classification process, the measurements of vibrations in the characteristic points of the anterior cornea curvature have been taken. Using the decision trees as a classifier, authors obtained very good results—their reported specificity of 98 %, sensitivity of 85 % and accuracy of 92 % for the automatic classification of healthy and keratoconic corneas. These new features of dynamic corneal deformation are very promising to become commonly used in clinical practice. Authors plan to expand their research by trying to modify the algorithm and classify other diseases.

The ultra-high-speed Scheimpflug imaging has been also a very good instrument for investigating the dynamic response of the cornea-model systems [6, 39].

3 Optical Coherence Tomography OCT + Air Puff

Optical coherence tomography (OCT) was firstly developed for visualization and analysis of the posterior part of the eye, but the rapid evolution of this technique and the development of Fourier-domain OCT allows for the acquisition of the anterior eye segment as well [7]. The principle of operation of anterior segment OCT is based on low-coherence interferometry. It works analogically to ultrasound, except that the dispersed factor is not a sound but the light wave [29]. Received signal depends on the properties of the tissue from which light reflects. It is worth mentioning that measurements carried out with this technique have very good reproducibility [30].

Taking into consideration the high quality of pictures available thanks to such technology, imaging of the cornea deformation dynamics with using this method could allow to analyse this phenomena even more accurately. What is more, it could provide clear answers to the questions concerning the links between illness factors and the changes in the dynamic response of the cornea.

Several researchers combined the OCT system with classic tonometry in their papers. In work [4] the swept source OCT (ssOCT) together with an air puff chamber were supposed to image the displacement of the cornea during the air puff. During the measurement, the ssOCT acquired around 1000 A-scans in time (M-scan) at the center of the air puff and showed the anterior and posterior cornea surface (Fig. 3). The obtained images presented the dynamic changes of the detected surface. Authors indicated that such a dynamic behavior is definitely correlated with corneal biomechanics, intraocular pressure and applied force of the air puff.

A similar research was conducted by Dorronsoro et al. in paper [9]. However, this time it was a spectral OCT combined with an air tonometer in a non-collinear configuration. Such a device allows to register the dynamic cornea movements in

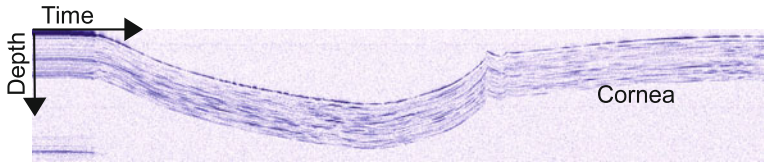


Fig. 3 Part of the M-scan (1300 A-scans) from ssOCT presenting the cornea surface motion [4]

response to the applied force. Full corneal cross-sections can be imaged and what is more, it is possible to obtain additional parameters: the speed of cornea apex, deformation amplitude, diameter and volume of the maximum deformation of the cornea and also the duration of the observed process. The quantitative analysis of deformation parameters was conducted in vivo and in vitro. Differences between corneal responses across species and treatments were seen.

Yet, the above papers demonstrated only novel instruments which are not commercially used and still further investigations are needed to improve their performance.

4 Ocular Response Analyzer (ORA)

ORA is the first noncontact tonometer capable of measuring the biomechanical properties of the cornea [28]. It uses a dynamic, bidirectional applanation process to measure the IOP and additional parameters: corneal hysteresis (CH) and corneal resistance factor (CRF) [8]. A collimated air pulse deforms the cornea which passes through the phase of first applanation, the highest concavity, second applanation and then returns to its normal curvature. Simultaneously, a dynamic infrared electrooptical system records the event of both inward and outward applanations [34]. Disparity in the forces required to flatten the cornea at the certain pressures—P1 and P2 (Fig. 4) called hysteresis [14] reveals the viscoelastic properties of cornea [15] and is caused, among others, by the energy adsorption during the rapid cornea deformation. In other words CH reflects the ability of corneal tissue to absorb and dissipate energy [10]. Going forward with observations based on this dynamic parameter (CH) we could try to investigate the impact of the air-jet on the normal and diseased corneas dynamic behavior.

Hysteresis is commonly applied in the diagnosis of ocular diseases [1, 2]. Many researchers presented CH as a parameter which could be a good classifier in keratoconus [33, 38, 44]. In the work [38] authors used the largest patient group: 207 normal and 93 keratoconic eyes. They found (like in [33, 44]) the hysteresis value higher in normal eyes than in keratoconus ones (a statistically significant difference).

It needs to be mentioned that this equipment has the ability to explore the dynamic properties of the cornea, however they are only the point values presented at the corneal hysteresis figure [16].

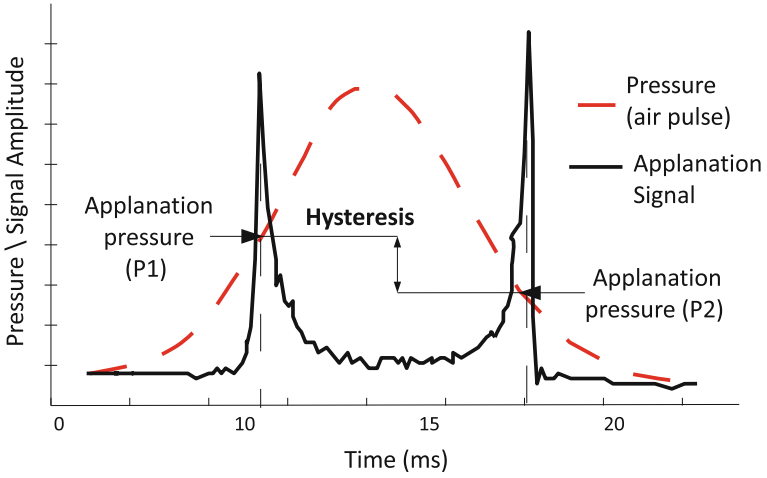


Fig. 4 Corneal hysteresis measured by Ocular Response Analyzer (ORA)

5 Discussion

The imaging methods reviewed in this paper are the only ones with the ability to present the dynamic corneal deformation in response to an air puff.

Over the past 10 years, various studies have shown that the corneal dynamics remains a not entirely known and unexplored phenomenon. The year 2005 when the Reichert launched the first device providing the biomechanical properties of the cornea—The Ocular Response Analyzer (ORA) was the breakthrough in the research concerning dynamics of corneal deformation. With this technique a partial evaluation of the dynamical cornea changes in response to the air stimulus became possible. The biggest advantage of this method is the unique feature—Corneal Hysteresis (CH) which is claimed to be an advanced predictor of glaucoma progression. Unfortunately, ORA cannot visualize the cornea and it delivers only the point values of its dynamics. This limitation may introduce a huge mistake in the assessment of reasons of certain cornea behaviors in response to an air puff.

The 'technical answer' to ORA tonometer was Corvis ST (introduced in 2010). The possibility to record the full reaction of the cornea to a defined air pulse with a high-speed Scheimpflug-camera that takes over 4,300 images per second is the biggest benefit of this equipment. On the basis of this video the series of specialistic parameters can be measured. Currently these features have already been well established and systemized, what is more, they can be a reliable help in clinical practice. They are more repetitive than similar parameters obtained from OCT. In OCT the variability is much higher which makes the parameters less reliable in classification. However, considering the availability in clinical practice (OCT was available in 1995, Corvis ST newly in 2010), such interpretation may be inadequate. In addition, thanks to the complete cornea imaging we are able to observe the cornea reaction in 3D [20].

Competitive systems presented in this paper combine the OCT with the air puff tonometry. The biggest advantage of these systems is the excellent resolution and image quality. Yet, the lack of commercially used equipment based on such conception capable of measuring dynamic response of the cornea is the greatest limitation. Still, this technique is not entirely developed and at this stage cannot be a real competitor in the field of imaging the dynamical changes of the cornea.

An important aspect which also needs to be discussed is the definition of 'dynamics', described in this paper. In a current study authors refer to ocular dynamics as a cornea response to a specific stimulus like the air puff. However, ocular dynamics may be considered in different aspects. Firstly we can discuss dynamics of corneal deformation in response to applied force. Secondly, this term may be related to the ocular pulse. Then some authors identify it with the tear film status [5]. Finally, it is also associated with the dynamical changes of the iridocorneal angle with light- dark adaptation [23, 24]. A new approach to this problem is presented in the paper [40]. Authors used a spectral domain OCT (SD-OCT) to visualize the short-term microfluctuation of the anterior chamber angle. This device captured the sequence of 90 cross-sectional images of the iridocorneal angle (B-scans), which have been consequently analyzed, scan by scan with the use of advanced image processing methods, just to obtain the posterior surface of the cornea and anterior surface of the iris. In such images authors observed a changing in time, variable fluctuation of the temporal position of detected surfaces. They indicated that the power spectra of these movements can deliver an important information concerning individual characteristics of the patient's eye, like the dynamic changes of the inside of the eye globe (blood pulsation in the ocular vessels and muscle reaction) and also the external dynamics like breathing. These individual features may be used in the future (after the detailed research) in classification of the ocular diseases.

It has been confirmed that corneal deformation caused by the air puff allows to measure the dynamic properties of the cornea which are necessary to determine the corneal biomechanics [17]. The cornea's biomechanical properties have an important impact on the intraocular pressure measurements, as well as the progression of corneal diseases [21]. These relations prove the importance of corneal dynamic parameters. A better understanding of these features will be of-substantial help for ophthalmologists who will have a series of new features enabling them to diagnose a variety of eye diseases. It needs to be mentioned that recently several publications proposed new parameters, designed on the basis of images provided by Corvis ST, according to which the classification of ill patients will be possible [19, 20, 43]. The paper [19] showed that the introduced parameter was very accurate, what is very promising when it comes to adopting the new features in the clinical practice. The recent advantages in this field and all available parameters have been summarized in work [22].

New technologies are not always popular in clinical practice. Moreover, the dynamical properties of the cornea are constantly investigated so the clinicians are still more likely to examine patients with well-known, easy in use noncontact methods. Concerning the above, the OCT imagining has invariably more followers than the newest techniques (like Corvis ST).

The study of the dynamics of the cornea is still the new chapter in the field of imaging in ophthalmology. Therefore, there is a need to develop a new "gold standard" concerning only the dynamic parameters of the cornea e.g. specifying the time range of the first applanation for both healthy and diseased eyes.

6 Conclusion

Currently the dynamics of anterior segment of the eye is the popular topic of research for scientists. The new tonometer Corvis ST gave the new possibilities to visualize the dynamic behavior of cornea. This technique is constantly improved by authors who develop new parameters based on the available images. What is more, the methods which combined OCT with an air puff are a great promise to overcome the current imaging limitations, and hence the knowledge of ocular dynamic and biomechanics will be commonly available.

In conclusion, the sixth corneal layer (Dua's layer) was recently discovered. The properties and impact on corneal dynamic behavior of this layer are still unexplored. Thus, feature research may change the current knowledge regarding the cornea as well as the anterior chamber of the eye.

References

1. Abitbol, O., Bouden, J., Doan, S., Hoang-Xuan, T., Gatinel, D.: Corneal hysteresis measured with the Ocular Response Analyzer in normal and glaucomatous eyes. Acta Ophthalmol. **88**, 116–9 (2010)
2. Agarwal, D.R., Ehrlich, J.R., Shimmyo, M., Radcliffe, N.M.: The relationship between corneal hysteresis and the magnitude of intraocular pressure reduction with topical prostaglandin therapy. Br. J. Ophthalmol. **96**, 254–257 (2012)
3. Ali, N.Q., Patel, D.V., McGhee, C.N.J.: Biomechanical responses of healthy and keratoconic corneas measured using a noncontact scheimpflug-based tonometer. Invest. Ophthalmol. Vis. Sci. **55**, 3651–3659 (2014)
4. 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)
5. Bonatti, J.A., Bechara, S.J., Carricondo, P.C., Kara-José, N.: Proposal for a new approach to corneal biomechanics?: dynamic corneal topography. Arq. Bras. Oftalmol. **72**, 264–267 (2009)
6. Correia, F.F., Ramos, I., Roberts, C.J., Steinmueller, A., Krug, M., Ambrósio Jr., R.: Impact of chamber pressure and material properties on the deformation response of corneal models measured by dynamic ultra-high-speed Scheimpflug imaging. Arq. Bras. Oftalmol. **76**, 278–281 (2013)
7. Denoyer, A., Labb, A., Baudouin, C.: Optical coherence tomography. In: Bernardes, R., Cunha-Vaz, J. (eds.) Optical Coherence Tomography, pp. 125–138. Springer, Berlin (2012)
8. Devi, S.A.: Analyzer, the ocular response. J. Curr. Glaucoma Pract. **3**, 24–27 (2009)
9. Dorronsoro, C., Pascual, D., Pérez-Merino, P., Kling, S., Marcos, S.: Dynamic OCT measurement of corneal deformation by an air puff in normal and cross-linked corneas. Biomed. Opt. Express **3**, 473–487 (2012)

10. Fontes, B.M., Ambrósio, R., Velarde, G.C., Nosé, W.: Corneal biomechanical evaluation in healthy thin corneas compared with matched keratoconus cases. Arq. Bras. Oftalmol. **74**, 13–16 (2011)
11. Gharaee, H., Abrishami, M., Abrishami, M., Mirhosseini, S.M., Bahar, M.M., Eghbali, P.: Anterior and posterior corneal curvature: normal values in healthy Iranian population obtained with the Orbscan II. Int. Ophthalmol. **34**, 1213–1219 (2014)
12. Hon, Y., Lam, A.K.C.: Corneal deformation measurement using scheimpflug noncontact tonometry. Optom. Vis. Sci. **90**, 1–8 (2013)
13. 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)
14. Huseynova, T., Waring, G.O., Roberts, C., Krueger, R.R., Tomita, M.: Corneal biomechanics as a function of intraocular pressure and pachymetry by dynamic infrared signal and Scheimpflug imaging analysis in normal eyes. Am. J. Ophthalmol. **157**, 885–893 (2014)
15. Jedzierowska, M., Koprowski, R., Wrobel, Z.: Overview of the ocular biomechanical properties measured by the Ocular Response Analyzer and the Corvis ST. Inf. Technol. Biomed. **4**, 77–386 (2014)
16. Ji, C., Yu, J., Li, T., Tian, L., Huang, Y., Wang, Y., Zheng, Y.: Dynamic curvature topography for evaluating the anterior corneal surface change with Corvis ST. Biomed. Eng. Online **14**, 53 (2015)
17. Kling, S., Marcos, S.: Contributing factors to corneal deformation in air puff measurements. Invest. Ophthalmol. Vis. Sci. **54**, 5078–5085 (2013)
18. Konstantopoulos, A., Hossain, P., Anderson, D.F.: Recent advances in ophthalmic anterior segment imaging: a new era for ophthalmic diagnosis? Br. J. Ophthalmol. **91**, 551–557 (2007)
19. Koprowski, R., Ambrósio, R.: Quantitative assessment of corneal vibrations during intraocular pressure measurement with the air-puff method in patients with keratoconus. Comput. Biol. Med. **66**, 170–178 (2015)
20. Koprowski, R., Lyssek-Boron, A., Nowinska, A., Wylegala, E., Kasprzak, H., Wrobel, Z.: Selected parameters of the corneal deformation in the Corvis tonometer. Biomed. Eng. Online **13**, 55 (2014)
21. Kotecha, A.: What biomechanical properties of the cornea are relevant for the clinician? Surv. Ophthalmol. **52**(Suppl 2), S109–14 (2007)
22. Lanza, M., Iaccarino, S., Bifani, M.: In vivo human corneal deformation analysis with a Scheimpflug camera, a critical review. J. Biophoton. **14** (2016)
23. Leung, C.K.S., Chan, W.-M., Ko, C.Y., Chui, S.I., Woo, J., Tsang, M.-K., Tse, R.K.K.: Visualization of anterior chamber angle dynamics using optical coherence tomography. Ophthalmology **112**, 980–984 (2005)
24. Leung, C.K., Cheung, C.Y.L., Li, H., Dorairaj, S., Yiu, C.K.F., Wong, A.L., Liebmann, J., Ritch, R., Weinreb, R., Lam, D.S.C.: Dynamic analysis of dark-light changes of the anterior chamber angle with anterior segment OCT. Invest. Ophthalmol. Vis. Sci. **48**, 4116–4122 (2007)
25. Leung, C.K.-S., Ye, C., Weinreb, R.N.: An ultra-high-speed Scheimpflug camera for evaluation of corneal deformation response and its impact on IOP measurement. Invest. Ophthalmol. Vis. Sci. **54**, 2885–2892 (2013)
26. Li, T., Tian, L., Wang, L., Hon, Y., Lam, A.K.C., Huang, Y., Wang, Y., Zheng, Y.: Correction on the distortion of Scheimpflug imaging for dynamic central corneal thickness. J. Biomed. Opt. **20**, 56006 (2015)
27. Li, T., Wang, Y., Chang, C., Hu, N., Zheng, Y.: Color-appearance-model based fusion of gray and pseudo-color images for medical applications. Inf. Fusion **19**, 103–114 (2014)
28. Luce, D.A.: Determining in vivo biomechanical properties of the cornea with an ocular response analyzer. J. Cataract Refract. Surg. **31**, 156–162 (2005)
29. Mansouri, K., Sommerhalder, J., Shaarawy, T.: Prospective comparison of ultrasound biomicroscopy and anterior segment optical coherence tomography for evaluation of anterior chamber dimensions in European eyes with primary angle closure. Eye (Lond.) **24**, 233–239 (2010)

30. Maslin, J.S., Barkana, Y., Dorairaj, S.K.: Anterior segment imaging in glaucoma: an updated review. Indian J. Ophthalmol. **63**, 630–640 (2015)
31. 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)
32. Oliveira, C.M., Ribeiro, C., Franco, S.: Corneal imaging with slit-scanning and Scheimpflug imaging techniques. Clin. Exp. Optom. **94**, 33–42 (2011)
33. Ortiz, D., Piñero, D., Shabayek, M.H., Arnalich-Montiel, F., Alió, J.L.: Corneal biomechanical properties in normal, post-laser in situ keratomileusis, and keratoconic eyes. J. Cataract Refract. Surg. **33**, 1371–1375 (2007)
34. Pepose, J.S., Feigenbaum, S.K., Qazi, 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)
35. Rio-Cristobal, A., Martin, R.: Corneal assessment technologies: current status. Surv. Ophthalmol. **59**, 599–614 (2014)
36. Rogowska, M.E., Iskander, D.R.: Age-related changes in corneal deformation dynamics utilizing scheimpflug imaging. PLoS One **10**, e0140093 (2015)
37. See, J.L.S.: Imaging of the anterior segment in glaucoma. Clin. Exp. Ophthalmol. **37**, 506–513 (2009)
38. 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)
39. Shih, P.-J., Cao, H.-J., Huang, C.-J., Wang, I.-J., Shih, W.-P., Yen, J.-Y.: A corneal elastic dynamic model derived from Scheimpflug imaging technology. Ophthalmic Physiol. Opt. **35**, 663–672 (2015)
40. Siedlecki, D., Kowalik, W., Kasprzak, H.: Optical coherence tomography as a tool for ocular dynamics estimation. Biomed. Res. Int. **2015**, 8 (2015)
41. Silverman, R.H.: High-resolution ultrasound imaging of the eye—a review. Clin. Exp. Ophthalmol. **37**, 54–67 (2009)
42. Tian, L., Huang, Y., Wang, L., Bai, H., Wang, Q., Jiang, J., Wu, Y., Gao, M.: Corneal biomechanical assessment using corneal visualization scheimpflug technology in keratoconic and normal eyes. J. Ophthalmol. **2014**, 8 (2014)
43. Tian, L., Ko, M.W.L., Wang, L.-K., Zhang, J.-Y., Li, T.-J., Huang, Y.-F., Zheng, Y.-P.: Assessment of ocular biomechanics using dynamic ultra high-speed scheimpflug imaging in keratoconic and normal eyes. J. Refract. Surg. **30**, 785–791 (2014)
44. Touboul, D., Roberts, C., Kérautret, 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)
45. Valbon, B.F., Ambrósio 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)

Application of Eye Tracking to Support Children's Vision Enhancing Exercises

Pawel Kasprowski, Michalina Dzierzega, Katarzyna Kruk, Katarzyna Harezlak and Erita Filipek

Abstract The paper presents the results of an initial experiment aimed to check if it is possible to enhance a low vision stimulation intervention with interactive games. There were 35 children with normal vision and various vision deficits examined using three different interactive games. The interactivity was achieved with the usage of an eye tracker—a device that is able to register eye movements and recalculate it to a gaze point—a place on a screen where a child is looking at. The obtained results are encouraging as most of the children were able to fulfill all games' achievements and expressed their engagement.

Keywords Eye tracking · Low vision · Vision stimulation

1 Introduction

Vision plays a crucial role in human life. Vast majority of information that is acquired by human brain is registered by eyes. It is especially important to children when they learn how to handle in their environment and learn various skills like reading and writing. Lack of ability to see the surrounding world properly may cause some serious negative consequences for children not only for cognitive processes but also for psychosocial ones.

Previously, there was a common belief that people with low vision should save their sight and use it sparingly. However, it occurred that intensive use of eyes does not cause damage or decrease the degree of remaining vision [2]. Moreover, intensive eye related exercises may improve vision or at least enable children to use their eyes more efficiently and develop some new visual skills. Therefore, nowadays, it is the

P. Kasprowski (✉) · M. Dzierzega · K. Kruk · K. Harezlak
Silesian University of Technology, Gliwice, Poland
e-mail: pawel.kasprowski@polsl.pl

E. Filipek
Department of Paediatric Ophthalmology, School of Medicine in Katowice,
Medical University of Silesia, Katowice, Poland

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_7

well-established practice to use various vision related exercises during low vision rehabilitation [8]. The main problem of such exercises is that they are boring and cumbersome for children. Many people recall such a treatment from their childhood as a nightmare. Therefore, our idea was to combine vision exercises with fun, by letting children to play games. This idea was possible due to the usage of a simple and affordable eye tracker that enabled to add some interaction during exercises.

1.1 Low Vision Stimulation Intervention

Low vision stimulation intervention should not be confused with vision therapy. Vision therapy is developed to cure specific eye related diseases—e.g. surgically or pharmaceutically—and it is provided by ophthalmologists. In contrast, vision stimulation intervention helps children to use their vision more effectively [8]. It has been proven that it may help to activate brain areas and maximize vision comprehension even without any other medical interventions. It is especially important for brain diseases for which there is no medication or surgery. Cortical Vision Impairment (CVI) is one of the examples of such a disease. CVI is a condition in which the visual pathway and visual processing centers of the brain are damaged and result in reduced vision [6]. There are reports that 95 % of children with CVI developed higher levels of vision within a period of 3.7 years after receiving a program of visual intervention [7].

As there is an evidence that playing computer games may enhance visual skills [1] there are attempts to use it directly during treatment of children with low vision. One of the examples is a virtual-reality (VR)-based display system that facilitates the treatment of amblyopia [4].

1.2 Eye Tracking

Eye tracking is meant as an ability to determine gaze position—the place where the user is looking at. Although there are multiple different eye tracking methods, the most convenient and popular one is a so called videooculography (VOG). Typically VOG utilizes an infrared camera and infrared light sources. Reasons to use infrared light are twofold. Firstly, a pupil (eye center) may be easily localized in infrared light, because it is the darkest region of an image. Secondly, infrared light sources are invisible for participants and therefore not disturbing their vision. A position of reflections of light sources from a cornea—called 1-st Purkinje reflections or glints—may be used as an additional input which makes eye position measurement less sensitive to head movements (Fig. 1).

The data that is taken into account while estimating eye position is a vector between an eye center (center of a pupil) and glints. Obviously, length and direction of such vector depends on a distance from a screen, camera and light sources position

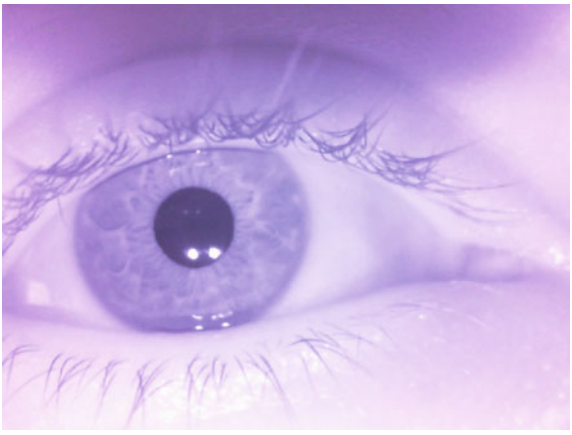


Fig. 1 Eye image with visible glints (1-st Purkinje reflections)

and—last but no least—a corneal curvature, which is person specific. Therefore a calibration procedure must be performed prior to the vector's usage as a source for gaze position. Typically the calibration consists of presentation of some (5–9) points on a screen. The task for participants is to look at these points for some time. Based on the gathered data the system is later able to compute a function that recalculates an eye center and a glint vector to screen coordinates of a participant's gaze.

2 Method

All vision related exercises were conducted with the usage of a computer and an eye tracker device. Participants were seated at a computer display and worked with a specially designed application. Their eye movements were registered with an Eye Tribe eye tracker. The Eye Tribe is an example of VOG devices. It is able to record eye movements with frequency 30 or 60 Hz. The device should be located below a display and there is a prior calibration required before every usage. Based on an eye tracker output it is possible to calculate information about screen coordinates of a gaze point (e.g. a place where the participant is looking at).

The usage of the eye tracker made a gaze contingent interaction between a participant and a computer possible. There were overall seven different visual games prepared for the purpose of the experiment. The aim of each game was to force participant's eye movements to accomplish some tasks (Fig. 2).

Finally, only three of that games were used during experiments. During the first game (named "object game"—OG) child's task was to find objects with a specified color. A child had to look at such an object and click a trigger button. There were always four objects with the specified color on the screen. Figure 3 shows a sample screen from this game.

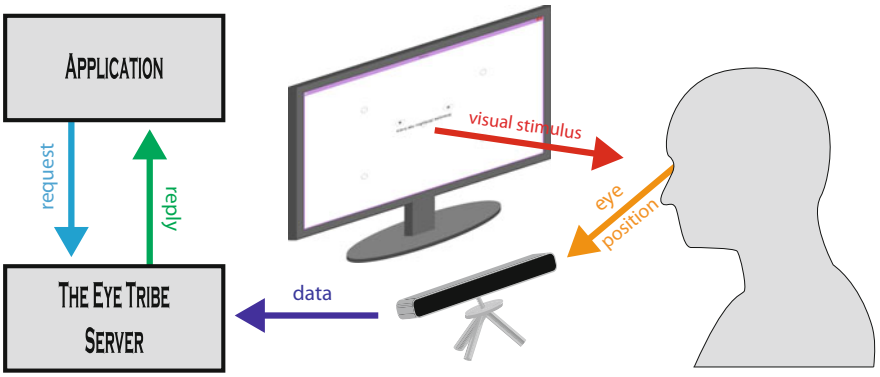


Fig. 2 System architecture

Fig. 3 A sample screen from the first game (OG). There is an initial screen visible on the *left*. The task for a child is to find all *yellow* objects, look at it and click a trigger button. A chosen object disappears. The game screen after pointing all four objects is visible on the right panel. Additionally, *red circles* with lines visualize fixations (places where gaze was focused for some time) and scan paths of the gaze

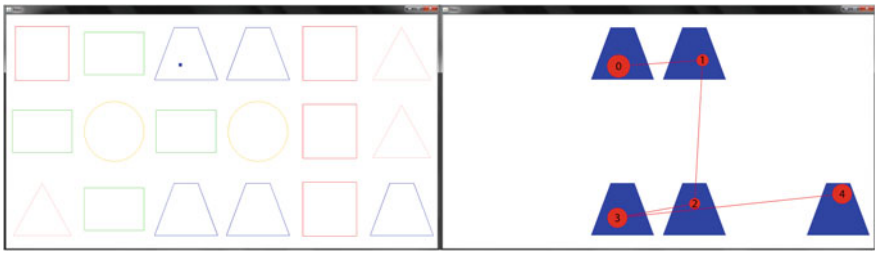


Fig. 4 A sample screen from the second game (FG). There is an initial screen visible on the *left*. The task for a child is to find all trapezoids, look at it and click a trigger button. A chosen object disappears. A sample of recorded fixations (*red circles*) and a scan path are visible on the *right panel*. Additionally, trapezoids have been highlighted, which was not possible during the game

The second game (named "figure game"—FG) was very similar but this time a task was to find geometrical figures. Contrary to the first game, objects were painted in low contrast with white background (Fig. 4).

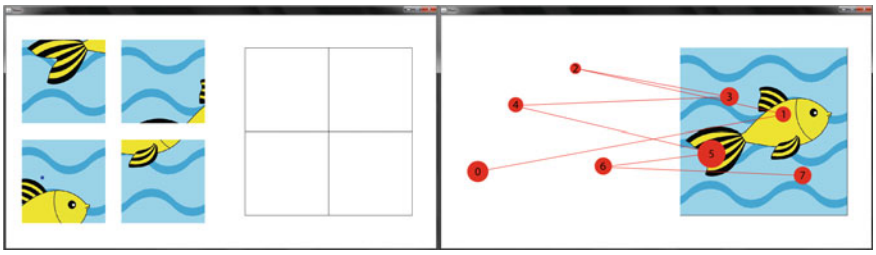


Fig. 5 A sample screen from the third game (PG). There is an initial screen visible on the *left*. The task for a child is to look at one piece and click a trigger button. After that the child should look at the correct place for this piece and once again click the trigger. A sample of recorded fixations (*red circles*) and scan paths after completing the task are visible on the *right panel*

Fig. 6 The four difficulty levels for the puzzle game

The third game was a classic puzzle game (PG). The task for participants was to point with eyes a piece of image on the left side and then point with eyes the correct place on the right side of the screen (Fig. 5).

All games were prepared with four different difficulty levels. A sample of these levels for puzzle game (PG) was presented in Fig. 6.

3 Participants

35 participants aged 5–12 took part in the experiment based on games presented in the previous section. 25 of them were patients of Department of Ophthalmology, Medical University of Silesia, Katowice, Poland. All participants have been divided into four groups:

- HL—children without any diagnosed problems with sight (10 participants),
- GA—children with good visual acuity but with some other dysfunctions (10 participants),
- LD—children with low disorder of sight (10 participants),
- SD—children with significant disorder of sight (5 participants).

All experiments were provided after parent's consent. The purpose of the experiment was to answer the four research questions:

1. Is it difficult for children to work with eye tracker and the application?
2. Do sight disorders influence the game results?
3. Is it more difficult to play games with only one eye?
4. Is such kind of therapy interesting for children?

4 Results

Every experiment started with a calibration of the eye tracker and then a participant played all three games. It is important to mention that there were seven children that did not accomplish at least one of the games. The reason of it could be problems with calibration or just the lack of interest in the game, which resulted in the lack of concentration.

4.1 Timing Analysis

There were three different parameters taken into account when analysing data recorded for OG and FG games:

- Time to find the first element (T1),
- Time to find the next element (TN),
- Total time of the game from the start to the end (TT).

It is visible that the best results are obtained for healthy children (HL). The results for other groups are similar with no significant differences. The total time for group with significant disorder (SD) is higher than for the other groups. What is interesting, the time to find the first element is about 50 % of the total time for the first three groups while it is only 30 % of the total time for the fourth group (Tables 1, 2 and 3).

Although the task was similar to the object game (OG), all groups but the second one (GA) needed more time to accomplish it. The reason was obviously the lower contrast of the image. It was especially problematic for participants from the SD group—they needed three times more time to accomplish the task comparing to the healthy children (HL). Interestingly, children from the GA group finished the second game faster than the first one.

Table 1 Results for the object game (OG) divided into groups

Group (size)	T1 (s)	TN (s)	TT (s)
HL (9)	9.8	3.5	19.5
GA (9)	22.7	9.6	42.7
LD (9)	20.9	8.4	46.2
SD (3)	17.4	12.6	54.5

Table 2 Results for the figure game (FG) divided into groups

Group (size)	T1 (s)	TN (s)	TT (s)
HL (10)	7.5	6.3	26.3
GA (9)	9.3	6.7	27.6
LD (9)	17.6	12.6	55.2
SD (4)	27.8	16.1	79.5

Table 3 Results for the third game divided into groups

Group (size)	T1 (s)	TN (s)	TT (s)
HL (9)	9.8	4.2	40.4
GA (9)	5.8	5.4	65.7
LD (8)	25.6	9.7	93.1
SD (3)	32.9	10.9	210.2

For the third game (puzzle game) there were also three parameters taken into account:

- Time to get the first element (TG),
- Time to put the first element (TP),
- Total time of the game from the start to the end (TT).

Similarly to the previous games, healthy children (HL) were the fastest and children with significant disorders (SD) the slowest ones. The between groups differences are higher, what shows that this game was more challenging for children with low vision.

4.2 Results for both Eyes and One Eye

Eye patching (disclosure) is one of the most popular treatments of children with amblyopia [3]. That is why it was decided to check if it is more difficult to play games with one eye covered. Only the children from HL (healthy children) group took part in the experiment. The children played the same game with two eyes and with one eye covered. The same time related parameters were used to compare both results. It showed no significant difference as it is visible in Figs. 7 and 8 for the first and second games.

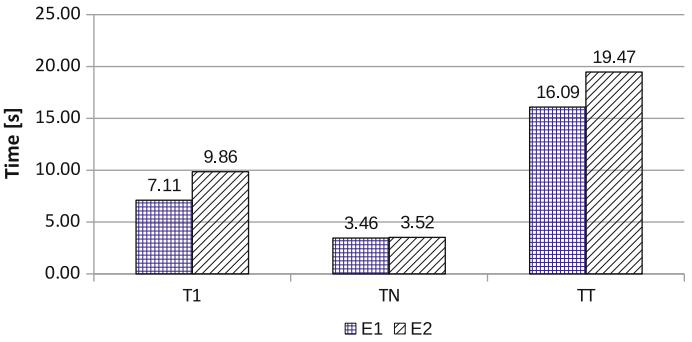


Fig. 7 Results for the HL group playing the first game with one eye covered (E1—*left*) and two eyes (E2—*right*)

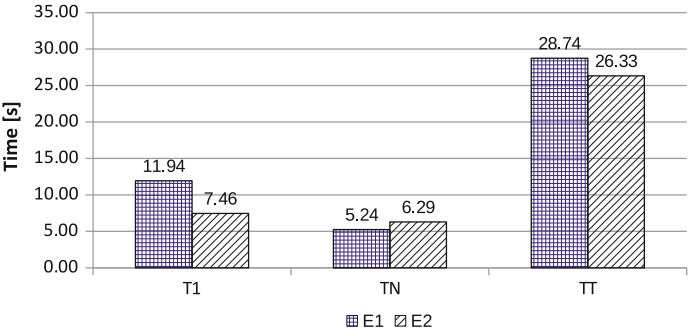


Fig. 8 Results for the HL group playing the second game with one eye covered (E1—*left*) and two eyes (E2—*right*)

4.3 Comparison of Results According to Age

The next step of the analysis concerned the age of children. As the range was quite broad the question was if the results are correlated with age. To have comparable results only children from groups HL and GA were taken into account as these children had the correct vision acuity (visus 5/5). It occurred that the correlation between age and total time is not significant (0.17) what means that age does not have important influence on ability to finish the game.

5 Discussion

Concerning the research question formulated at the beginning of the experiment, if it is difficult for children to work with the eye tracker and application, it occurred that about 87 % of attempts to finish the game were successful. It must be noticed

that in most eye tracking experiments there is some group (approximately 5–10 % of population) that is not able to finish an experiment for different reasons [5]. Children (especially those with low vision) may be regarded as one of the most challenging groups, so, in our opinion, the results are quite encouraging.

There was at least one child in every group which did not accomplish at least one of the games. But it is worth noticing that out of 7 children that had problems, 5 had a diagnosed squint. Overall, only one of six patients with the squint was able to finish all three games. It suggests that the squint may be a serious obstacle while working with gaze contingent interfaces.

Sight disorders influenced the game results. The experiments showed—unsurprisingly—that children with normal vision were faster in accomplishing visual tasks. However, what may be more important, it also proved that even children with significant vision deficits may be successful in playing gaze contingent games. Noticeable, some of such children achieved results comparable to the healthy group.

Our experiments with healthy children showed that eye disclosure (pathing) does not significantly influence the results of the experiment. So it may be concluded that game based therapy may be successfully combined with the traditional patching.

It occurred that such therapy was far more interesting for children than typical vision enhancement exercises. Although some of them had problems with calibration and focusing on the task, most expressed their amusement and asked when they would have an opportunity to improve their results.

6 Summary

The paper presents a preliminary experiment aimed at checking if eye tracker enhanced visual games may be used during low vision stimulation intervention. It showed that such exercises are possible and even children with serious vision deficits are able to take part in them. The next step will be analysis of possibilities to incorporate it into the real low vision developing programs. We hope that this initial idea will be in future extended and it will be possible to evaluate its benefits for children with low vision care.

Acknowledgments We would like to thank the Department of Ophthalmology, Medical University of Silesia, Katowice, Poland and its highly cooperative employees for giving us opportunity to work with their patients. We also greatly acknowledge the support of Silesian University of Technology grant BK/263/RAu2/2016.

References

1. Achtman, R., Green, C., Bavelier, D.: Video games as a tool to train visual skills. Restor. Neurol. Neurosci. **26**(4–5), 435 (2008)
2. Barraga, N.: Increased visual behavior in low vision children. ERIC (1964)

3. Clarke, M.P.: Review of amblyopia treatment: are we overtreating children with amblyopia? Brit. Irish Orthoptic J. **7**, 3–7 (2015)
4. Eastgate, R., Griffiths, G., Waddingham, P., Moody, A., Butler, T., Cobb, S., Comaish, I., Haworth, S., Gregson, R., Ash, I., et al.: Modified virtual reality technology for treatment of amblyopia. Eye **20**(3), 370–374 (2006)
5. Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., Van de Weijer, J.: Eye Tracking: A Comprehensive Guide to Methods and Measures. Oxford University Press (2011)
6. Roman-Lantzy, C.: Cortical Visual Impairment: An Approach to Assessment and Intervention. American Foundation for the Blind (2007)
7. Takeshita, B.: Cortical Vision Impairment: Overview. Southern California CVI Consortium (2012)
8. Takeshita, B.: Developing Your Children's Vision. A Guide for Parents of Infants and Young Children with Vision Impairment. The Center for the Partially Sighted (2012)

Biopsy Needle and Tissue Deformations Detection in Elastography Supported Ultrasound

Joanna Czajkowska, Jan Juszczyk, Bartlomiej Pycinski and Ewa Pietka

Abstract During last decades a fast development of imaging techniques has offered the intra-operative visualization as the integral part of surgical tools. For this, the automated and robust analysis of ultrasound images is required. The paper meets these requirements targeting in detection of tissue deformations caused by biopsy needle inserted in the body. The presented novel technique uses ultrasound data supported by elastography images. In the feature set, the automated detection algorithm introduces Histogram of Oriented Gradients and image entropy. The further classification steps applies Weighted Fuzzy C-Means (WFCM) clustering technique resulting in deformation detection sensitivity and specificity at levels 0.793 and 0.94, respectively.

Keywords Ultrasound · Elastography · Histogram of oriented gradients · Detection · Biopsy needle

1 Introduction

Needle biopsy, widely used as diagnostic procedure, is the injection of skin or mucosa with the needle, usually followed by acquiring the specimen. Injection of the needle has to be as accurate as possible to correctly point at proper place and to prevent injury of contiguous vital structures. Therefore the biopsy often is performed under the control of a medical image, usually ultrasonography (US).

Only a single plane of the body is visualised at a time by an US probe, so the needle might not be completely visible in the image. The best image quality is obtained if a needle guide is used [2]. The guide is fixed to the transducer in the way that prevents the needle from turning out of the scanned plane. The already performed review of needle detection methods is presented in [4].

J. Czajkowska (✉) · J. Juszczyk · B. Pycinski · E. Pietka
Faculty of Biomedical Engineering, Silesian University of Technology,
Roosevelta 40, Zabrze, Poland
e-mail: joanna.czajkowska@polsl.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_8

Ultrasound techniques has grown rapidly over last years. Nowadays not only standard B-mode presentation are widely used, but also other variations have been developed, e.g. Doppler mode, 3D visualisation, advanced measurement systems [1, 17]. One of the newest ultrasound diagnostic methods is elastography.

Since the US recording has to be synchronized with any other signals, defining of the biopsy beginning is very important. An automatic technique is especially desirable for long US recordings. Moreover, for further US data analysis, the detection of tissue deformation area caused by the injected needle places a very important role. This might be a challenge, if the needle is not visible in the image during the injection. In this study both US and elastography images of porcine phantom have been employed to detect the needle and tissue deformations.

The developed algorithm makes it possible to detect and locate the biopsy needle, and the area of tissue deformation caused by the needle insertion. The analysis is performed on the series of ultrasound (US) and elastography images acquired during biopsy procedure. The novel technique, based on the entropy and Histogram of Oriented Gradients (HoG) of the images acquired over the time, enables an accurate detection of: (1) the needle insertion event, (2) the inserted needle, as well as (3) the tissue area deformed by placing the needle. Thanks to the combination of the two used modalities the developed technique copes with the US image data, which is hard to understand for human expert and hard to analyse using other detection techniques.

The paper is organized as follows. In Sect. 2.1 the physical principles concerning acquired image modalities are stated. In Sect. 2.2 the entropy based technique of needle insertion event detection is described. It is followed by HoG features description given in Sect. 2.3. The deformation detection based on both modalities is presented in Sect. 2.4. The obtained results and plans for further development of the algorithms are described is Sects. 3 and 4, respectively.

2 Methods

2.1 Elastography

Elastography is the ultrasound diagnostic method which allows measurement of tissue elasticity parameters. The quasi-static elastography is based on local image deformation analysis. In case of external source of pressure, the deformation of tissue depends on the local tissue elasticity.

The elastography term was proposed in 1991 by Ophir et al. [20]. One of the first results of in vivo studies were presented in 1996 in [19], whereas theoretical basis of elastography can be found in [10, 21, 24].

Elastography visualizes relative elasticity of imaged tissues. Quasi-static elastography is based on minimal movements of tissues generated by an external (outside of tissue) source. In clinical applications the oscillating movements of US probe induces slight changes between consequent images. The relative elasticity is visualized in real time by colour or grey map called elastogram (Fig. 1) [20]. Tissue elasticity model is based on cascaded one dimension spring system, where each regions of tissue in line is represented by spring with the same length but different strain. When the spring system is pressed, the length (after and before compression) varies for different spring [20]. That allows a parameter of relative elastic to be found without information about force that has been applied. Information about tissue elasticity derives from cross-correlation time shift differences among image segments [19]. Clinical application of elastography focuses on tumours in soft tissues detection and qualification. The breast elastography diagnosis was historically the first systematic developed [24]. It has been proven that the lesions in breast tissue have higher strain than normal tissue [11] moreover varies lesions feature a different strain allowing for a differential diagnosis. However others [14] indicate a similar to conventional US diagnostic performance. Liver elastography performed in hepatic cirrhosis and fibrosis diagnostic is another widely expanded study [9, 16, 22, 24]. There are attempts to create a scale for chronic virial hapatis rating based on liver elastography [8]. Currently clinically tested commercial solutions (as dedicated devices) for liver elastography are available [23]. Elastography is applied for the diagnosis of prostate [3, 26] and thyroid [18].

Since the injected needle in US image is hardly or almost invisible in US scan (see Figs. 1 and 2), the elastography is employed.

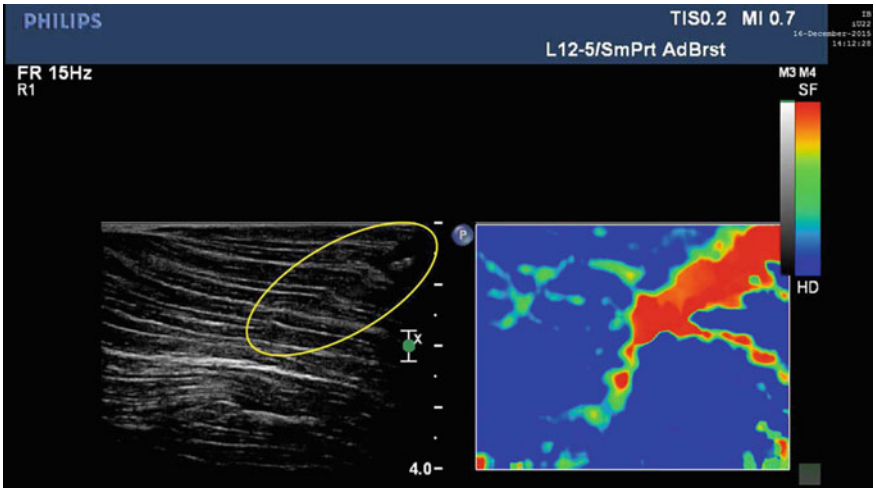


Fig. 1 US image with needle—marked with *yellow* ellipse—injected into the tissue (*left*) and elastography image (*right*)

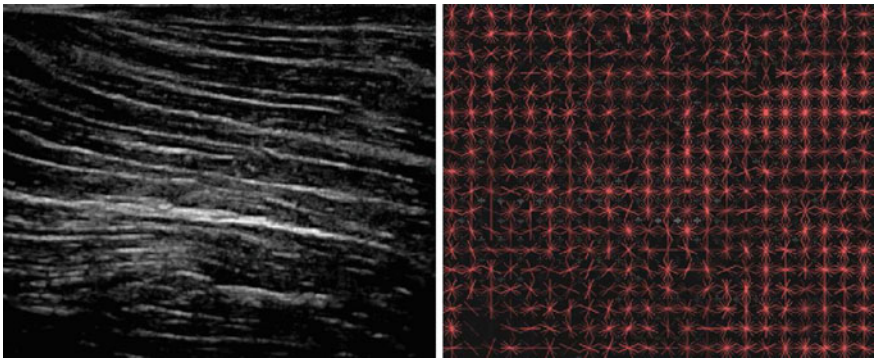


Fig. 2 The HoG differences (*right*) between training data and the input image (*left*). The HoG scale is set to $s = 16$ and the number of analysed directions is 8

2.2 Entropy Based Event Detection

Entropy is a statistical measure referring to disorder or uncertainty in a message [13]. It describes the randomness of the data. In image processing, entropy is defined as the average information generated from all the image pixels and its value is maximum if the pixels are uniformly distributed [15]. Therefore, it can be used to characterize the texture of the input image and is defined as

$$E = -\sum_{k=0}^{L-1} P_k log_2(P_k), \tag{1}$$

where $P_k,\ k = 1, 2, \ldots, L-1$ are probabilities associated with each grey level.

The entropy analysed over the acquired grey scale elastography images is used for an event detection. The grey scale elastography images are obtained as the V channel of RGB to HSV data conversion. The event is one of the biopsy procedure phases: (1) the state before the needle injection, (2) the needle injection. The event being detected by the developed algorithm is the needle insertion. For this, the entropy changes are analysed. The detection algorithm applies Weighted Fuzzy C-Means clustering technique [4, 25] to divide the analysed signal (differences of entropy) into two groups: the state before the needle insertion and needle insertion procedure as well. The acquired image series divided by the event detection step are then used in further processing.

2.3 Histogram of Oriented Gradients

The previous step yields in detection of needle injection beginning. The HoG descriptor [5] is then applied to find the area of tissue deformation caused by the biopsy

needle [4]. The HoG-based approach outperforms existing feature sets by finding lots of applications in different image recognition areas [7, 27] also in the medical image registration field [12]. Our previous study [4] indicates a successful location of a needle by measuring the tissue deformation.

The local appearance of tissue deformations caused by the needle insertion is well characterized by the distribution of local intensity gradients or edge directions [5]. Therefore, it can be detected even without a precise knowledge of the needle edge position or tissue deformation details.

The feature extraction step is then based on analysing local histograms of image gradient orientations. For this, the image is divided into small regions called HoG cells, of the size of $[s \times s]$ pix, where s is the HoG scale, $s \in \{128, 64, 32, 16, 8, 4\}$. The scale selection depends on the image resolution and the object size. A local histogram of gradient directions or edge orientations is calculated over the pixels of each cell. Single HoG cell is processed in five steps. First, the cell image is normalized. As reported in [5] the gamma as well as colour normalisation have an insignificant effect on the classification performance and can be skipped. Next, the gradient image of any single cell is computed using Gaussian smoothing function followed by centred derivative mask $[-1\ 0\ 1]$ at $\sigma = 0$. The edge orientation histogram is then estimated as the weighted vote of pixels within the analysed cell, where the vote is the gradient magnitude at a pixel. To avoid the aliasing problem caused by the orientation bins spaced over 0°–180° the votes are bilinearly interpolated between the neighbouring bins centres in both orientation and position. Finally, to reduce the influence of gradient strengths variations, the local contrast normalization step is introduced [4, 5].

Knowing the HoG features of the analysed US we are able to detect and monitor any changes caused by the injected biopsy needle or the catheter. The HoG differences between training data and the input image presented in Fig. 1 are shown in Fig. 2.

2.4 Data Classification

The integration of both algorithms results in: (1) the detecting of the time moment of biopsy needle injection and (2) HoG features estimation for all the acquired images. The time moment, indicating the beginning of needle injection, makes it possible to automatically divide the data set into two subsets: before and during the needle injection. Thus, the tissue deformation algorithm can be summarized in the steps: (1) training procedure with HoG parameters estimation; (2) deformation area detection.

The training procedure is performed on the US scans labelled as the state before the needle injection (slides $k = 1, 2, \ldots, K$, where $K = 43$), further called I_c. The average image of the acquired US series I_m is then used for HoG parameters definition, which describe the examined tissue. Then, two threshold vectors: $\mathbf{h}^l_{min}$ and $\mathbf{h}^l_{max}$

of HoGs estimated for each image cell l, $l = 1, 2, \ldots, M$, where M is the number of HoG cells in image $I_c(k)$, are given as:

$$\mathbf{h}^l_{min} = \min_{i=k,\ldots,K} |\mathbf{h}^l_m - \mathbf{h}^l_c(k)| \tag{2}$$

and

$$\mathbf{h}^l_{max} = \max_{i=k,\ldots,K} |\mathbf{h}^l_m - \mathbf{h}^l_c(k)|, \tag{3}$$

where $\mathbf{h}^l_m$ and $\mathbf{h}^l_c(k)$ are the HoG features of the average image and kth training frame, respectively. The estimated boundary parameters are then used in the next processing steps. The tissue deformation detection is applied to each frame I creating testing set (labelled as needle injection process). The area of deformation is segmented based on the comparison of HoG parameters between training and testing US scans. For this, the WFCM clustering technique is performed on the feature vectors $\mathbf{v}^l_I$, estimated for each image cell l. It contains values estimated for each direction $d = 1, 2, \ldots, D$ of analysed HoG cells as:

$$\mathbf{v}^l_I = [v^l_I(1), \ldots, v^l_I(D)], \tag{4}$$

where D is the number of HoG directions estimated for every single HoG cell and

$$v^l_I(d) = \begin{cases} 1 & \text{if } |h^l_m(d) - h^l_I(d)| < h^l_{min}(d) \text{ or } |h^l_m(d) - h^l_I(d)| > h^l_{max}(d) \\ |h^l_m(d) - h^l_I(d)| & \text{otherwise} \end{cases}, \tag{5}$$

where $h^l_I(d)$ is the HoG value calculated for the analysed US frame I, image cell l and direction d, $h^l_m(d)$ is the HoG for average image, and $h^l_{min}(d)$, and $h^l_{max}(d)$ are HoG boundary parameters given in (2) and (3), respectively.

To reduce the influence of artefacts occurring in any single image in the US and elastography recording the moving voting algorithm is introduce. For each US scan the detection result is compared with the outcomes of 4 neighbouring images. The final detection area consists of these HoG cells, whose 3 out of 5 cells in consecutive images are assigned into the deformation region. The same moving voting procedure is performed after the needle detection step, leading to the final results.

The above described classification procedure does not include division of the found area into the injected needle and the tissue deformation caused by it. Since the biopsy needle is hardly visible in the registered US scans, the image processing targeting in the needle detection does not provide satisfying or even any interesting results. For this, the elastography data are analysed again.

Since the injected biopsy needle is the stiffest part of the visualized structure and its appearance is mostly visible in a red channel, only this channel will be

analysed in further processing steps (since now called elastography image). First, the elastography image pixels are classified into two groups using WFCM clustering technique. Then, the cluster mask is divided into the regions corresponding to the HoG cells. The final needle detection mask create these cells, within which more than 70 % of pixels are classified as needle in WFCM procedure.

3 Results

3.1 Experimental Set Up

The developed algorithm was evaluated on the set of 7 recordings of biopsy procedure performed on a porcine tissue phantom. Conventional and elastography US images consist of compressed DICOM images (JPEG) of resolution [288 × 352] pix and spacing 0.13 × 0.13 mm. Due to its accessibility and US appearance similar to human body parts the porcine tissue phantoms are often described in literature as reference for US image processing tool evaluation [6]. The experiments were performed using Philips iU22 machine with the linear probe L12-5 and elastography module.

The tissue deformation and needle detection algorithm was tested on the recording consisting of 225 pairs of US and elastography images. It uses the HoG cell size equal to 16 and the number of gradient directions $D = 8$. The number of gradient directions was chosen experimentally as described in [4]. The exemplary result of differential HoG feature extraction step is shown in Fig. 2.

3.2 Results

The evaluation procedure was divided into two parts according to the selection of the algorithm: event or needle detection. The first algorithm correctly indicate the moment of the needle injection beginning in 7 of 7 recordings. The measured entropies of ultrasound and elastography are shown in Fig. 3, whereas their subsequent differences are shown in Fig. 4. The acquired time series contain the images before the needle insertion (slides $k = 1, 2, \ldots, 43$) and the needle insertion process (slides $k = 44, 45, \ldots, 225$).

The accuracy of the needle detection step was evaluated using 45 training and 180 testing images, where the training set was created using the event detection algorithm. The first 45 US image were therefore used for HoG parameters estimation and the actual verification was performed on 180 pairs of US and elastography scans.

The obtained detection results (Fig. 5) were compared with the selection performed manually by three experts, who delineated the needle on each of 180 US

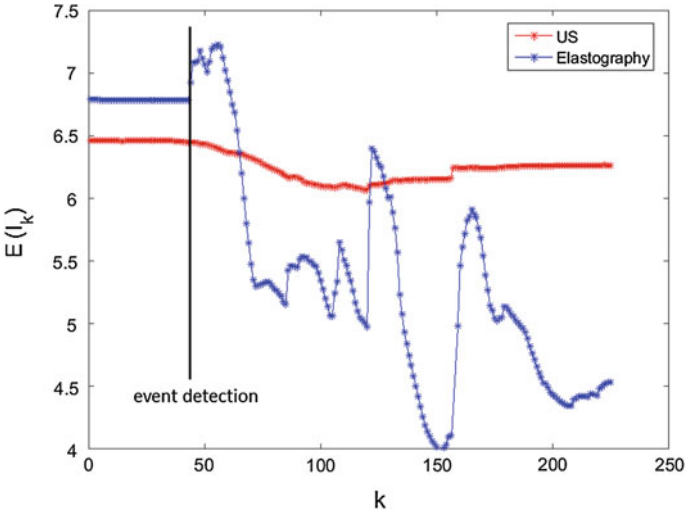


Fig. 3 The entropy values of series of the pair of acquired image data. The k is the index of acquired image pair. The needle insertion begins with $k = 44$

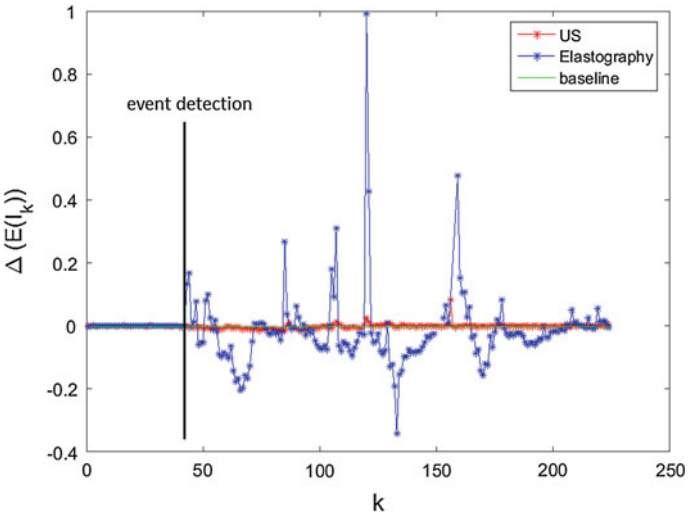


Fig. 4 Differences of subsequent entropy values shown in Fig. 3. The k is the index of acquired image pair. The needle insertion begins with $k = 44$

scans. The resulted images were then divided into HoG cells of the size set at the needle detection step (see Fig. 6). The data evaluation is based on statistical measures of a binary classification performance: sensitivity (TPR) and specificity (SPC), where the HoG cells were treated as the testing probes. If the needle cross the image cell in the reference image (expert selection) it was labelled as a "positive" probe,

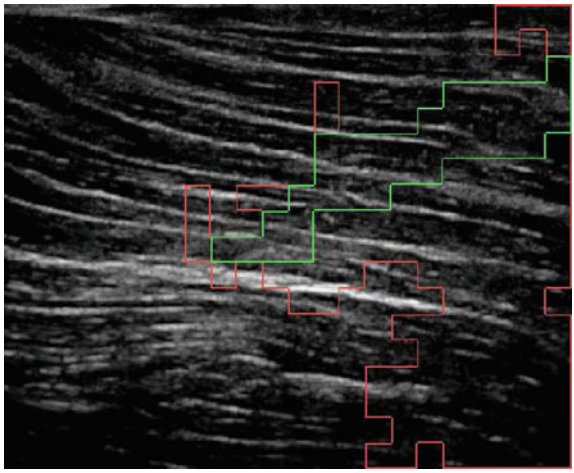


Fig. 5 Needle detection. The *green line* (inner selected area) delineates the biopsy needle, whereas the outer *red line* delineates the deformation area

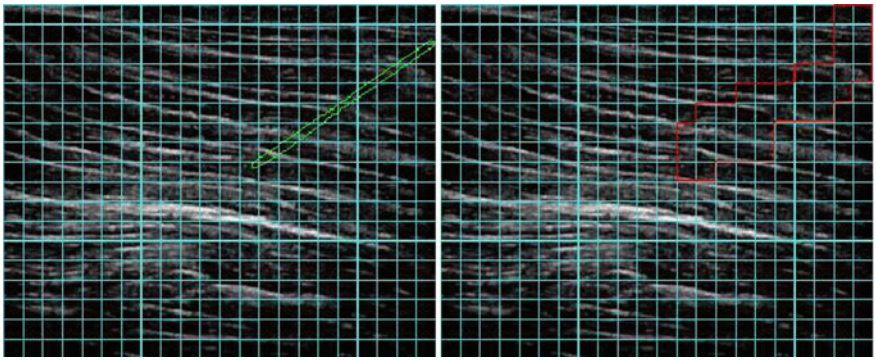


Fig. 6 The exemplary detection results (*right*) compared with manual needle selection (*left*)

otherwise it become a label "negative". The TPR and SPC values calculated for each expert separately as well as the mean value are shown in Table 1. The exemplary deformation area as well as needle detection results are shown in Fig. 5. The red outer area is the deformed tissue and the green inner part is the detected needle. The comparison with the manual needle selection created by the expert is shown in Fig. 6.

For comparison, the same inter-observer experiment was performed. The results are given in Table 2. The experiments show the difficulty of the needle detection, particularly when it cannot be seen in the image. The inter-observer tests show similar results to the automated method evaluation. Moreover, the developed technique is faster than the manual delineation, and repeatable.

Table 1 Sensitivity (TPR) and specificity (SPC) obtained for the automated needle detection compared with the manual needle selection performed by three experts

	TPR	SPC
Expert 1 versus Automated method	0.769	0.936
Expert 2 versus Automated method	0.837	0.937
Expert 3 versus Automated method	0.772	0.947
Mean	0.793	0.94

Table 2 Sensitivity (TPR) and specificity (SPC) obtained from the cross-analysis performed between three experts delineations

	TPR	SPC
Expert 1 versus Expert 2	0.769	0.936
Expert 1 versus Expert 3	0.652	0.978
Expert 2 versus Expert 3	0.805	0.981
Mean	0.742	0.965

4 Conclusions

The paper presents an effective and robust method for automated detection of needle injection moment, tissue deformation area caused by the biopsy needle as well as the needle, in the conventional and elastography US image pairs. The novel approach combining Histogram of Gradients feature applied to US data with entropy based elastography data analysis yields the sensitivity and specificity comparable to the inter-observer results. Currently, the method is tested in the image navigation in minimally invasive surgery set-up.

Acknowledgments This research was supported by the Polish National Science Center (NCN) grant No. UMO-2012/05/B/ST7/02136.

References

1. Badura, P., Pietka, E.: 3D Fuzzy liver tumor segmentation. In: Information Technologies in Biomedicine: Third International Conference, ITIB 2012, Gliwice, Poland, June 11–13, 2012. Proceedings, pp. 47–57. Springer, Berlin (2012)
2. Chapman, G.A., Johnson, D., Bodenham, A.R.: Visualisation of needle position using ultrasonography. Anaesthesia **61**(2), 148–158 (2006)
3. Cochlin, D.L., Ganatra, R., Griffiths, D.: Elastography in the detection of prostatic cancer. Clin. Radiol. **57**(11), 1014–1020 (2002)
4. Czajkowska, J., Pycinski, B., Pietka, E.: HoG feature based detection of tissue deformations in ultrasound data. In: EMBC, pp. 6326–6329. IEEE (2015)
5. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: International Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 886–893 (2005)

6. Duda, R.O., Hart, P.E.: Use of the Hough transformation to detect lines and curves in pictures. Commun. ACM **15**(1), 11–15 (1972)
7. Felzenszwalb, P., Girshick, R., McAllester, D., Ramanan, D.: Object detection with discriminatively trained part-based models. IEEE Pattern Anal. Mach. Intell. **32**(9), 1627–1645 (2010)
8. Friedrich-Rust, M., Ong, M.F., Herrmann, E., Dries, V., Samaras, P., Zeuzem, S., Sarrazin, C.: Real-time elastography for noninvasive assessment of liver fibrosis in chronic viral hepatitis. Am. J. Roentgenol. **188**(3), 758–764 (2007)
9. Friedrich-Rust, M., Ong, M.F., Martens, S., Sarrazin, C., Bojunga, J., Zeuzem, S., Herrmann, E.: Performance of transient elastography for the staging of liver fibrosis: a meta-analysis. Gastroenterology **134**(4), 960–974 (2008)
10. Gao, L., Parker, K., Alam, S., Lerner, R.: Sonoelasticity imaging: theory and experimental verification. J. Acoust. Soc. Am. **97**(6), 3875–3886 (1995)
11. Garra, B.S., Cespedes, E.I., Ophir, J., Spratt, S.R., Zuurbier, R.A., Magnant, C.M., Pennanen, M.F.: Elastography of breast lesions: initial clinical results. Radiology **202**(1), 79–86 (1997)
12. Ghafurian, S., Hacihaliloglu, I., Metaxas, D.N., Tan, V., Li, K.: 3D/2D image registration using weighted histogram of gradient directions **9415**, 94151Z–94151Z-7 (2015)
13. Gonzalez, R.C., Woods, R.E., Eddins, S.L.: Digital Image Processing Using MATLAB. Prentice-Hall Inc., Upper Saddle River (2003)
14. Itoh, A., Ueno, E., Tohno, E., Kamma, H., Takahashi, H., Shiina, T., Yamakawa, M., Matsumura, T.: Breast disease: clinical application of us elastography for diagnosis 1. Radiology **239**(2), 341–350 (2006)
15. Jain, A.K.: Fundamentals of Digital Image Processing. Prentice-Hall Inc., Upper Saddle River (1989)
16. Kirk, G.D., Astemborski, J., Mehta, S.H., Spoler, C., Fisher, C., Allen, D., Higgins, Y., Moore, R.D., Afdhal, N., Torbenson, M., et al.: Assessment of liver fibrosis by transient elastography in persons with hepatitis c virus infection or hiv-hepatitis c virus coinfection. Clin. Infect. Dis. **48**(7), 963–972 (2009)
17. Krysztoforski, K., Krowicki, P., Świątek Najwer, E., Będziński, R., Keppler, P.: Noninvasive ultrasonic measuring system for bone geometry examination. Int. J. Med. Robot. Comput. Assist. Surg. **7**(1), 85–95 (2011)
18. Ma, J., Luo, S., Dighe, M., Lim, D.J., Kim, Y.: Differential diagnosis of thyroid nodules with ultrasound elastography based on support vector machines. In: Ultrasonics Symposium (IUS), 2010 IEEE, pp. 1372–1375. IEEE (2010)
19. Ophir, J., Cespedes, I., Garra, B., Ponnekanti, H., Huang, Y., Maklad, N.: Elastography: ultrasonic imaging of tissue strain and elastic modulus in vivo. Eur. J. Ultrasound **3**(1), 49–70 (1996)
20. Ophir, J., Cespedes, I., Ponnekanti, H., Yazdi, Y., Li, X.: Elastography: a quantitative method for imaging the elasticity of biological tissues. Ultrasonic Imaging **13**(2), 111–134 (1991)
21. Ophir, J., Kallel, F., Varghese, T., Konofagou, E., Alam, S.K., Krouskop, T., Garra, B., Righetti, R.: Imagerie acoustique et optique des milieux biologiques optical and acoustical imaging of biological media. CR Acad. Aci. Paris **2**(4), 1193–1212 (2001)
22. Sandrin, L., Fourquet, B., Hasquenoph, J.M., Yon, S., Fournier, C., Mal, F., Christidis, C., Ziol, M., Poulet, B., Kazemi, F., et al.: Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. Ultrasound Med. Biol. **29**(12), 1705–1713 (2003)
23. Sandrin, L., Fournier, C., Miette, V., Millonig, G., Mueller, S.: Fibroscan in hepatology: a clinically-validated tool using vibration-controlled transient elastography. In: Ultrasonics Symposium (IUS), 2009 IEEE International, pp. 1431–1434. IEEE (2009)
24. Sarvazyan, A., Hall, T.J., Urban, M.W., Fatemi, M., Aglyamov, S.R., Garra, B.S.: An overview of elastography-an emerging branch of medical imaging. Curr. Med. Imaging Rev. **7**(4), 255 (2011)
25. Szwarc, P., Kawa, J., Pietka, E.: White matter segmentation from MR images in subjects with brain tumours. In: Information Technologies in Biomedicine: Third International Conference, ITIB 2012, Gliwice, Poland, June 11–13, 2012. Proceedings, pp. 36–46. Springer Berlin (2012)

26. Tsutsumi, M., Miyagawa, T., Matsumura, T., Kawazoe, N., Ishikawa, S., Shimokama, T., Shiina, T., Miyanaga, N., Akaza, H.: The impact of real-time tissue elasticity imaging (elastography) on the detection of prostate cancer: clinicopathological analysis. Int. J. Clin. Oncol. **12**(4), 250–255 (2007)
27. Vondrick, C., Khosla, A., Malisiewicz, T., Torralba, A.: Hoggles: visualizing object detection features. In: ICCV, IEEE, pp. 1–8 (2013)

Modification of the ICP Algorithm for Detection of the Occlusal Area of Dental Arches

Krzysztof Skabek and Agnieszka Anna Tomaka

Abstract We focused on the the process of finding the occlusal area of dental arches. The algorithm for detection of the occlusion area has been proposed. It is based on the ICP method that iteratively finds the best matching of two sets of points. The introduced method was discussed and compared to manually found occlusions and also to another methods.

Keywords 3D segmentation · Surface models · Dental arches · Occlusion detection · ICP algorithm

1 Introduction

In dentistry occlusion is defined as contact between teeth [4]. This contact may occur in many positions during chewing or at rest, but a significant role is attributed to maximum intercuspation—a person's habitual bite, called intercuspation position [6, 9]. As far as occlusion conditions influence the maxilla-mandible relation, and occlusion disorders may cause temporomandibular dysfunctions, evaluation of occlusion is one of the basics examinations in diagnosis of bruxism and other temporomandibular joint disorders [1]. Occlusion evaluation can be done directly by intraoral examination or using dental models, and nowadays it can be performed using computerized virtual dental models or segmented cone beam computed tomography (CBCT). Detection of the intercuspation position is then very important, as techniques of occlusion registration are prone to errors due to finite thickness of a wax bite, difficulties in intraoral scanning of teeth in occlusion and segmentation problems from CBCT with teeth in occlusion.

From a technical viewpoint the detection of intercuspation position is similar to a surface matching task, but in computer vision this matching problem is solved

K. Skabek (✉) · A.A. Tomaka
Polish Academy of Sciences, Institute of Theoretical and Applied Informatics,
BałTycka 5, Gliwice, Poland
e-mail: kskabek@iitis.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_9

for surfaces which have a common overlapping) area, for example scans of the same object from different viewpoints. The most popular technique of registration of such surfaces is the iterative closest point (ICP) algorithm. As the overlapping area is initially unknown the ICP algorithm, assuming the correspondence between matching surfaces as the closest points in actual position, finds the transformation minimizing distance between those points, and iteratively determines the position in which the distance is smallest. In this approach matched surfaces are treated as different views of the same object so intersections are not important. In ICP occlusion detection, however, surfaces of upper and lower teeth represent two complementary objects and direct application of ICP matching algorithm would yield the result in which corresponding points are at the closest position, but surfaces would intersects, which would not correspond to reality due to a collision of real objects.

We prepared technical assumptions for optimal occlusion from observations of tomographic imaging of the natural occlusion. It was stated a prerequisite of keeping the bone structure in non-collision space. Another assumption was also taken that we minimize the contact area on the top surface of dental arches.

In previous work we presented the method for finding the occlusion using a genetic algorithm minimizing a specially prepared energy function [10]. In this chapter we propose another approach based on statistical fitting of geometric structures. We used the ICP algorithm [3] which incrementally approximate the optimal fitting of two point sets. The transformation parameters R and T are estimated. Additionally, we assigned weights for each point of the structure to limit the area of searching.

2 Data Structures

2.1 Mesh and Oriented Point Representation

In the chapter we use mesh structures to represent surfaces. The meshes are obtained using scanning devices (such as Konica-Minolta VI-9i) in form of ordered points constructing the triangle mesh. Such mesh structure consists of a table of vertices V with normals N and topology F:

$$V = \{v_1, v_2 \dots v_k\} = \{(x_1, y_1, z_1) \dots (x_k, y_k, z_k)\}, \tag{1}$$

$$N = \{n_1, n_2 \dots n_k\} = \left\{(x_{n_1}, y_{n_1}, z_{n_1}) \dots (x_{n_k}, y_{n_k}, z_{n_k})\right\}, \tag{2}$$

$$F = \{f_1, f_2 \dots f_l\} = \{(v_1(f_1), v_2(f_1), v_3(f_1)) \dots (v_1(f_l), v_2(f_l), v_3(f_l))\}. \tag{3}$$

The topology of the mesh is stored as a set of triangles F and the triangles f_i are represented by three ordered indices $(v_1(f_i), v_2(f_i), v_3(f_i))$ pointing to the table of vertices V.

In our approach we use vectors normal to the vertices. They are calculated from normal vectors of topologically adjacent faces or, if there is no topological structure given, using PCA method of neighbouring points treated as a tangent plane [5].

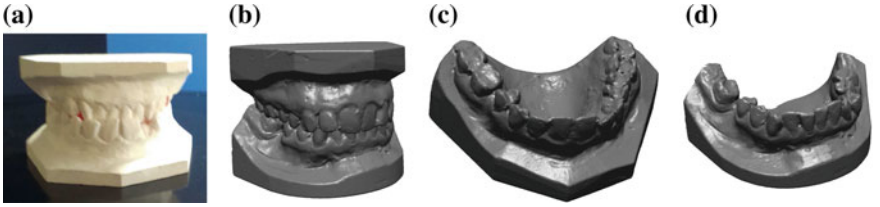


Fig. 1 Data for processing: **a** plaster dental models, **b** setting the correspondence with the wax bite, **c** digitalized model of the jaw, **d** Digitalized model of the mandible

2.2 Dental Models

In our application basic data for processing has the form of 3D triangle meshes which are digitalized 3D images of dental models. Such virtual dental models are obtained by digitizing plaster dental models (Fig. 1a) using a Konica-Minolta 3D scanner. Virtual models of the maxilla (Fig. 1c) and mandible (Fig. 1d) are created. Additional scanning of models with a wax bite enables establishing relation of both models in such a way that maps natural relation of maxilla and mandible (Fig. 1).

The technique of dental model scanning consists of several stages: data acquisition, registration of partial scans from different viewpoint merging them into a model and finally the structure processing [11]. Set of 24 partial scans is taken into consideration, 3 groups of 8 scans per model are acquired. Scans are merged in groups and then integrated. The obtained surface model needs some structure adjusting. Abnormal faces are found and corrected. Holes in the surfaces are also filled and if they are large, additional scanning which covers the artefacts is performed.

3 Algorithm for Finding the Occlusal Area

Our approach is based on the ICP algorithm [2, 3]. The routine minimizes the root square distance as a goodness function for two sets of points: the model point set $\mathbf{Q} = \{q_1, q_2 \ldots q_N\}$ and the input point set $\mathbf{P} = \{p_1, p_2 \ldots p_N\}$. The expression can be formulated as follows:

$$E = \sum_{i=1}^{N} ||\mathbf{R}p_i + T - q_i|\,|^2, \tag{4}$$

where $\mathbf{R}$ and T are rotation and translation of the point set.

The input data is not in correspondence with the model mesh so the first step of ICP algorithm is finding it.

3.1 Stages

There are several stages of the classical ICP algorithm [8]: (1) selection, (2) matching, (3) weighing, (4) rejecting, (5) measuring a distance error, (6) minimization of error metric.

3.2 Point-to-Plane Minimization

We use point-to-plane minimization as an optimization of the given mesh matching [7]. The matching measure is given as follows:

$$E = \sum_{i=1}^{N} ((\mathbf{R}\mathbf{p_i} + \mathbf{T} - \mathbf{q_i}) \cdot \mathbf{n_i})^2, \tag{5}$$

where $\mathbf{n_i}$ is normal in point p_i. We assume that for small angles the rotation matrix can be linearized:

$$\mathbf{R} = \mathbf{R_z}(\gamma)\mathbf{R_y}(\beta)\mathbf{R_x}(\alpha) \approx \begin{bmatrix} 1 & -\gamma & \beta \\ \gamma & 1 & -\alpha \\ -\beta & \alpha & 1 \end{bmatrix}. \tag{6}$$

3.3 Discrimination Measures

In our case there are two structures to be matched: first of them (denoted as $\mathbf{p}$), the upper jaw, is fixed and the second one (denoted as $\mathbf{p}'$), the mandible, is movable. To provide a better selection of structures which are important in the process of matching we introduced weights for the vertices. We establish the assumptions for the relative layout of the matching structures in form of a normalized weight measure. The weight measure consists of three components: $\sigma(\mathbf{p})$—the slope of corresponding vertices with regard to the fixed normal (7), $\rho(\mathbf{p})$—the slope of normals for corresponding vertices (8) and $D(\mathbf{p}, \mathbf{p}')$—the signed distance measure (9). The aim of the measure is to assign scalar values from the interval $[0 \ldots 1]$ to the vertices of the input mesh as weights $W(\mathbf{p})$ that discriminate less important parts in the matching process.

We calculate the dot product $\sigma(\mathbf{p})$ to obtain the direction of occlusion as expression:

$$\sigma(\mathbf{p}) = \vec{\mathbf{p}'\mathbf{p}} \cdot \mathbf{N}_{\mathbf{p}'}, \tag{7}$$

where: $|\mathbf{p}'\mathbf{p}| = d$ and $\mathbf{N}_{\mathbf{p}'}$ is normal vector at the point $\mathbf{p}'$.

Another measure which we considered is the relevance of normals at points $\mathbf{p}$ and $\mathbf{p}'$. This value is calculated from the expression:

$$\rho(\mathbf{p}) = \mathbf{N_p} \cdot \mathbf{N_{p'}}, \tag{8}$$

where: $\mathbf{N_p}$ and $\mathbf{N_{p'}}$ are normal vectors at the corresponding points $\mathbf{p}$ and $\mathbf{p}'$.

Third measure relates to the distance between two directed surfaces. It can be specified as a signed distance as follows:

$$D(\mathbf{p}, \mathbf{p}') = \mathbf{p} \cdot \mathbf{N_p} - \mathbf{p}' \cdot \mathbf{N_p}. \tag{9}$$

The resultant weight $W(\mathbf{p})$ can be specified for each vertex in two ways:

$$W(\mathbf{p}) = \left(1 - D(\mathbf{p})^2\right) \cdot th_0\left(\rho(\mathbf{p})\right) \cdot th_{\alpha > \alpha_0}(\sigma(\mathbf{p})), \tag{10}$$

$$W(\mathbf{p}) = \alpha(1 - D(\mathbf{p})^2) + \beta\, th_0(\rho(\mathbf{p})) + \gamma\, th_{\alpha > \alpha_0}(\sigma(\mathbf{p})), \tag{11}$$

where $th_i(x)$ is a threshold function:

$$th_i(x) = \begin{cases} x & \text{if } x \geq i, \\ 0 & \text{if } x < i. \end{cases} \tag{12}$$

Each component function $D(\mathbf{p})$, $\rho(\mathbf{p})$ and $\sigma(\mathbf{p})$ is normalized in equations (10) and (11). Parameters α, β and γ help to adjust the influence of the component values so the formula (11) is more flexible, but also needs additional adjustments and final normalization.

The influence of the components on particular mesh weight is illustrated in Fig. 2. Vertices with heavy weight are marked red, intermediate are coloured in gray or black, and those with weight of zero are omitted.

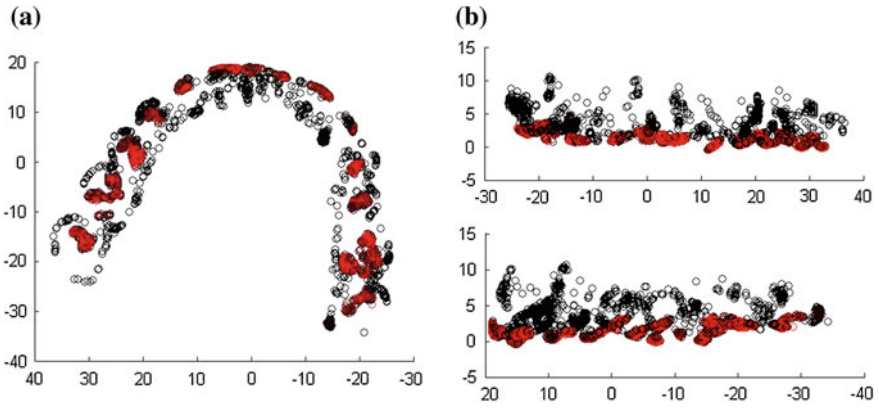


Fig. 2 Weights for ICP

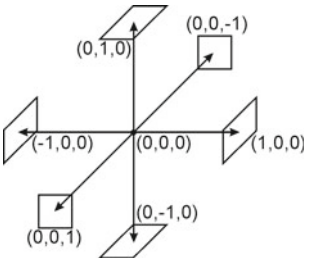


Fig. 3 Unit directions for the final correction

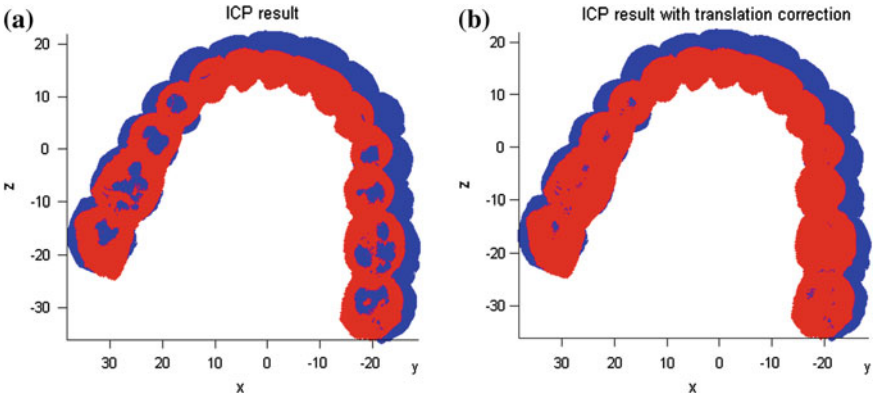


Fig. 4 Occlusion: **a** output of ICP with weights, **b** output of ICP with weights and correction

3.4 Final Correction

The correction ensures physical gap between the maxilla and the mandible. It is done by translating the mandible in 6 unit directions (see Fig. 3) to eliminate the overlapping structures. The set of translation vectors can be also extended to 26 items considering the intermediate directions. The correction is done incrementally assuming the minimum step ($\delta = 0.01$ mm in our case). The adjustment stops when all overlapping structures are eliminated or the current position yields a locally minimal overlapping (Fig. 4).

4 Testing the Occlusions

4.1 Contact Points

In Fig. 5a was shown the surface distribution of contact areas for dental arches within the range 0–1 mm after determination of the occlusal surface with the described

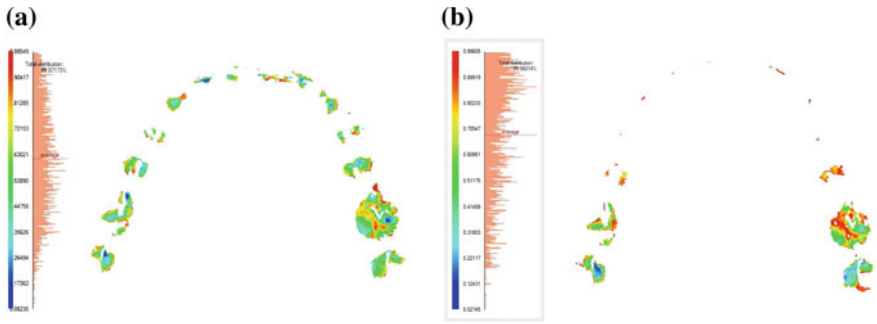


Fig. 5 Occlusion—contact points: **a** determined using the algorithm, **b** determined from CBCT

algorithm. A darker blue color indicates the actual contact places of the teeth. For comparison, Fig. 5b shows the contact area of the teeth in their anatomical position acquired from CBCT imaging.

Contact points and fitting surfaces are also visible in the image of cross sections taken along a circle with center in the middle of the reference structure and the rotation axis OY. The characteristic phases of cross-sections with the contact area are shown in Fig. 6.

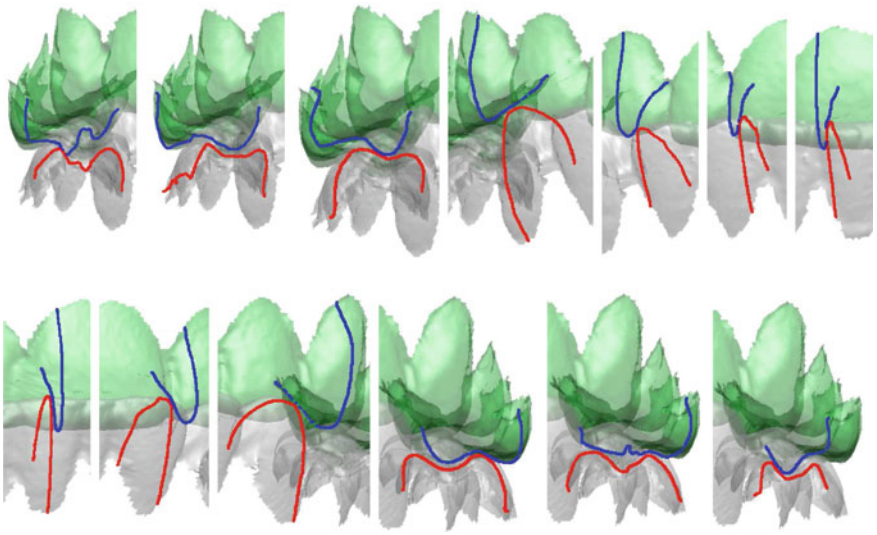


Fig. 6 Occlusion—cross sections

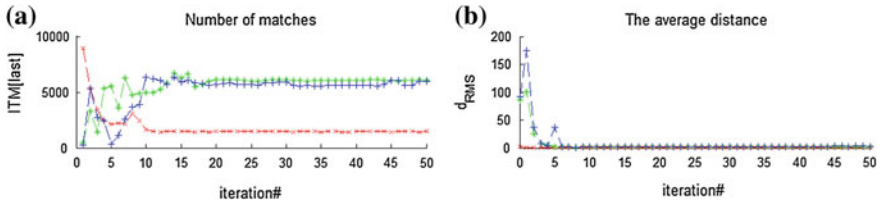


Fig. 7 Stability analysis for three cases (50 iterations): **a** the number of matched items, **b** the average distance

4.2 Stability of the Algorithm

We tested the stability of the aligning algorithm with weights. Experiments were performed in 50 iterations. It turned out that the convergence of the algorithm stabilized after a few iterative steps (Fig. 7). The effect of the initial position on the stability of matches was tested. The fastest convergence was obtained for meshes initially brought into its center of gravity and matched principal components using PCA analysis (case 1 –x– red). Other cases were arranged with structures in the random initial position: translation (case 2 –+– blue), and also translation and rotation (case 3 –*– green).

4.3 Comparison to Practical Cases

We analyzed three kinds of natural occlusions: occlusion from CBCT, occlusion from face scanning, occlusion from dental models.

The best fit was obtained for the 3D scan of plaster models in the fixed position. Several stable points of support can be distinguished in this case. The average distance error with regard to the proposed method was 0.222 mm and standard deviation 0.181.

A similar situation occurred in the case of occlusion obtained from CBCT imaging. The average error distance was, in this case, 0.334 mm and standard deviation 0.632. Comparison of the two cases is shown in Fig. 5.

The most significant differences were obtained for 3D face scan using an occlusal bite. Two distinct points of support were marked on the back molars and there was lack of support of anterior teeth. The results from our fitting method strongly differs from the position of the mandible in this case. The average distance error was 2.66 mm and standard deviation 1.917.

Table 1 Fitting distance errors and standard deviation for ICP and genetic methods with regard to natural alignment from CBCT imaging

ICP1		ICP2		Genetic1		Genetic2	
d_{err}	std_{dev}	d_{err}	std_{dev}	d_{err}	std_{dev}	d_{err}	std_{dev}
0.222	0.181	0.303	0.396	0.217	0.191	0.252	0.187

4.4 Comparison to Other Methods

In the chapter [10] was described another method considering the usage of genetic algorithms with a given optimization formula to find the area of occlusion for two compatible jaws.

After the comparison to the fitting errors obtained by the genetic algorithm, it turned out that the results lead to even smaller differences. The results are given in Table 1.

Moreover, the advantage of genetic method is the application of the objective measure, relevant to the quality of fitting. Unfortunately, disadvantage is the long processing time and much smaller convergence of the optimization process.

5 Conclusions

Finding the occlusal surface is an important part of determining the spatial relationship between the upper and lower dental arch. The process involves the search matches between the natural tooth cavities between the bumps.

The search can be limited to neighborhoods with direct contact friction between the teeth. The chapter selected such surfaces and gave them appropriate weights, greater than the others. In contrast, areas not involved in occlusion were rejected, they were given the importance of zero.

An algorithm for matching the occlusal surface using a method ICP was developed. The algorithm determines the contact surface of teeth by calculating the averaged position of the suitably weighted friction surfaces of the mandible relative to the maxilla. The last step of the algorithm determines the transformation that minimizes the contact area between the two dental arches.

The quantitative convergence of the cases was analyzed by comparing the calculated occlusion area with the positions of the natural occlusion and that obtained by the use of a genetic minimization algorithm.

The next steps were provided to improve the methods of determining the occlusal surface constraints and to optimize the energy function for the genetic algorithm. The genetic approach seems to be promising because of the availability of an objective energy function for surface fitting.

The possible use of the described methods includes the determination of the occlusal surface in the design of occlusal splints for the treatment of bruxism.

Acknowledgments This work was supported in part by the National Science Center, Poland under the research project 2012/07/B/ST6/01238.

References

1. Baron, S.: Bruxism and occlussion. In: Palla, S., Kato, T., Provaznik, I., Tkacz, E., Baron, S. (eds.) Novel Methodology of Both Diagnosis and Therapy of Bruxism, Lecture Notes of the ICB Seminar, 137th Seminar, Warsaw (2014)
2. Besl, P.J., McKay, N.D.: A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell. **14**(2), 239–256 (1992)
3. Chen, Y., Medioni, G.: Object Modeling by registration of multiple range images. In: Proceedings of IEEE Conference on Robotics and Automation (1991)
4. Davies, S., Gray, R.M.J.: Practice: what is occlusion? Br. Dent. J. **191**(5), 235–245 (2001)
5. Hoppe, H., DeRose, T., Duchamp, T., McDonald, J., Stuetzle, W.: Surface reconstruction from unorganized points. Comput. Graphics **26**(2), 71–78 (1992)
6. Mah, J., Hatcher, D., Harrell, W.: Craniofacial imaging in orthodontics. In: Graber, L., Vanarsdall, R., Vig, K. (eds.) Orthodontics: Current Principles and Techniques, ISBN: 032306641, 2012, 5th edn. Mosby, Elsevier (2012)
7. Pomerleau, F., Colas, F., Siegwart, R.: A review of point cloud registration algorithms for mobile robotics. Found. Trends Robot. **4**(1), 1–104 (2015)
8. Rusinkiewicz, S., Levoy, M.: Effcient variants of the icp algorithm. In: Proceedings of Third International Conference on 3-D Digital Imaging and Modeling, pp. 145–152 (2001)
9. Sander, F.: Analiza modeli (Model analysis). In: Diedrich, P. (ed.) Ortodoncja I, Rozwój struktur ustno-twarzowych i diagnostyka (Orthodontics I, Development of Bucco-Facial Structures and Diagnosis), ISBN: 83-89581-85-X, (in Polish). Wydawnictwo Medyczne Urban & Partner, Wrocław (2004)
10. Skabek, K., Tomaka, A.A.: Techniques of processing and segmentation of 3D geometric models for detection of the occlusal area of dental arches. In: Palla, S., Kato, T., Provaznik, I., Tkacz, E., Baron, S. (eds.) Novel Methodology of Both Diagnosis and Therapy of Bruxism, Lecture Notes of the ICB Seminar, 137th Seminar, Warsaw (2014)
11. Tomaka, A.A.: Analiza obrazów wielomodalnych dla potrzeb nieinwazyjnej diagnostyki ortodontycznej (Multimodal Image Analysis for Noninvasive Orthodontic Diagnosis). ISBN: 978-83-62652-46-4, (in Polish), IITIS PAN (2013)

The Fast Matching Algorithm for Rectified Stereo Images

Pawel Popielski, Robert Koprowski and Zygmunt Wróbel

Abstract In their research, the authors focus on the rapid methods for matching rectified images which can be readily implemented on mobile devices. First, the new method for matching images performs binarization of images and transforms them so that they depict edges. The disparity map is created in accordance with the principle that the correct disparity is the minimum distance of the calculated distances between a point in the left image and all the points in the right image in a given row. The method is illustrated on the basis of the authors' own images as well as standard images from the Middlebury library. In addition, the method has been compared with well recognized and commonly used algorithms for matching images, namely variational and semi-global methods.

Keywords Disparity · Stereo correspondence · Stereovision

1 Introduction

Diagnostic methods used in medicine are more or less invasive. Computed tomography, magnetic resonance and other commonly used methods do not remain indifferent to our health. Even laser scanning can be dangerous. The application of a method based on stereovision [16] ensures complete neutrality, because it uses only the reflected light. Other advantages of stereovision include rapid measurement, which is invaluable in the diagnosis of sick people suffering from pain. Stereovision can be used in medicine to examine body posture [7, 8], metabolic diseases [13], or it can be applied in trauma [14] and plastic surgery [4, 5].

Stereovision builds a three-dimensional object model based on the known geometry of two imaging devices and lens parameters. The key step is to look for stereo correspondence between the left and right images of a stereo pair, which enables

P. Popielski (✉) · R. Koprowski · Z. Wróbel
Institute of Computer Science, University of Silesia, Będzińska 39,
41-200 Sosnowiec, Poland
e-mail: pawel.popielski@us.edu.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_10

to compute XYZ ground coordinates of the cloud of points. The authors focused in their study on matching binary images. The method for matching binary images is completely independent of radiometric conditions. Small image complexity ensures fast algorithm operation on tablets and smartphones, and this opens up new possibilities in diagnostics. A physician taking an image of a patient could immediately access a variety of information useful in making a diagnosis.

The section *State of the art* describes the process of determining XYZ ground coordinates from a stereo pair of images. The section *Material* shows the images used in the study, their acquisition and parameters. The section *Proposed algorithm* presents a fast algorithm for matching images. The section *Results* provides the results of the proposed algorithm. The section *Discussion* compares the results with those obtained using other widely recognized methods for matching images. The last section summarizes and outlines further research purposes.

2 State of the Art

Prior to the calibration of imaging devices, all the present optical aberrations must be removed. Modern lenses are now mostly devoid of optical errors, except for distortion. Radial distortion is adjusted in accordance with [6], and tangential distortion in accordance with [10].

The calibration process involves determination of interior orientation elements and then exterior orientation elements [19, 20]. Elements of interior orientation [3], namely the focal length of lenses of a stereo-vision head, as well as the coordinates of the principal points of images are determined in accordance with [25]. Once the elements of internal orientation are known, it is possible to reproduce the position of the beam of rays projecting at the time of registration in relation to the imaging plane. External orientation elements, namely the rotation matrix and translation vector, are calculated according to [26]. The elements of external orientation stabilize the beam of rays projecting in space, enabling the homologous rays of a stereo pair to intersect.

Using the rules of epipolar geometry, images are rectified in accordance with Bouguet's method [2]. The rectification process involves transformation of stereo images onto a common plane in such a way so that the homologous rows of the images are in one line. Now, owing to the fundamental matrix properties [9, 17, 24], the search for stereo correspondence for the left image pixel will only involve searching the same row of the right image. Without rectification, it would be necessary to search the whole right image using a mask, which would disproportionately increase the time complexity of the process of looking for stereo correspondence. The result of the search for stereo correspondence is a disparity map $D(m, n)$, also known as a parallax matrix [12], which contains differences in positions in a given row of homologous points in the left and right images.

Having the disparity map, it is possible to calculate XYZ ground coordinates using [12].

The search for stereo correspondence can be carried out using different approaches. The oldest methods are based on analysis of surroundings of a pixel in the left and right images. When examining the relationship between the mask and searched image, it is the function which determines whether a sufficient level of similarity of areas is achieved. Normalized cross-correlation can be mentioned here as an example. Due to the low effectiveness of these methods and after a preliminary analysis of the results, a decision was made to exclude them from comparative analyses.

Methods based on calculus of variations (VAR) treat a stereo pair of images as a vector field. It involves minimizing the energy functional by solving the corresponding Euler-Lagrange equation, which is represented by a system of partial differential equations. The authors used in their study a variational algorithm as in [22].

The study included also a semi-global method (SG) based on Information Theory, where Mutual Information is used as the cost function [11]. These methods for matching images form the basis for comparative analyses in the research conducted by the authors.

3 Material

It was decided to use for the study both the authors' own images and Middlebury test images [1]. Models of elementary figures, namely a regular cube and a regular pyramid, were made of cardboard. To improve contrast, edges were marked with a black marker.

Additionally, diagonals were drawn in the cube to increase the number of elements for matching. Then, registration was performed using a stereo vision head (Fig. 1) and maintaining an imaging distance of 50 cm. Next, the images were cropped to a resolution of 200×163 pixels for the cube and 200×114 pixels for the pyramid. The images were reduced in such a way so as to preserve the clearest possible figure images and minimize the time spent on manual creation of a ground truth disparity map D_T.

21 datasets were chosen from the Middlebury library which were obtained in 2006 using the technique of [15] and published in [18, 23]. It was decided to choose test images having a height of 370 pixels (Fig. 2). The ground truth disparity maps D_T for these images were also obtained from the Middlebury library.

The stereo vision head used in the study consisted of two Logitech C920 HD Pro webcams which enabled to record a video stream at a resolution of 1920×1080 pixels. The cameras were equipped with the Tessar lens produced by Carl Zeiss Optics. The cameras were attached to a horizontal arm, which enabled to move them away from each other from 10 to 90 cm.

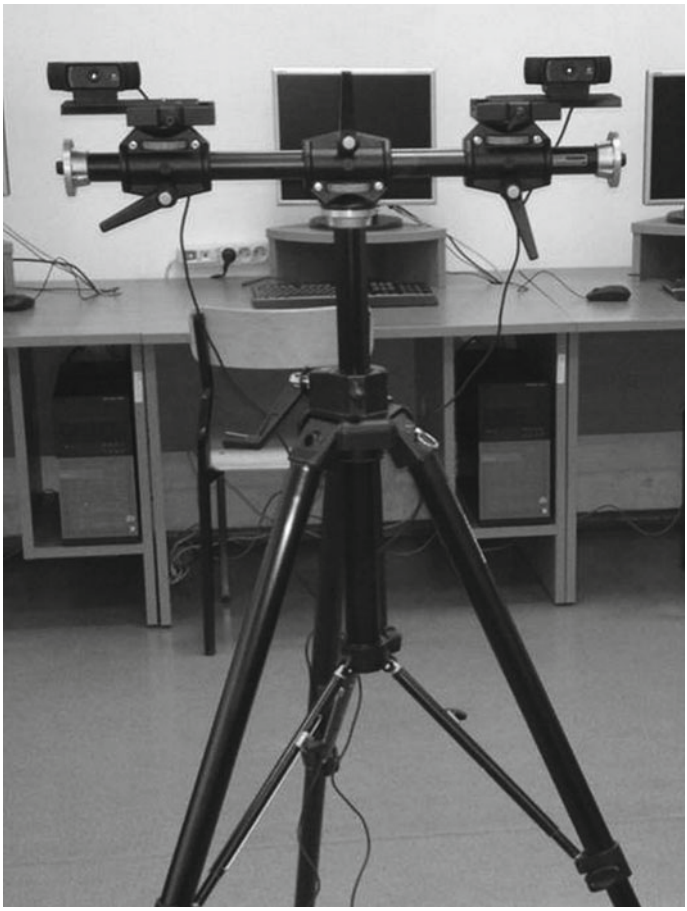


Fig. 1 Stereovision Head used in studies

4 Proposed Algorithm

The rectified left $L_z^{(L)}(m,n)$ and right image $L_z^{(P)}(m,n)$ devoid of aberrations are subjected to binarization using automatic thresholding with the Otsu's method [21]. The resulting binary images $L_b^{(L)}(m,n)$ and $L_b^{(P)}(m,n)$ are the basis for calculating the difference between horizontally adjacent pixels which is calculated for each row (1) (2).

$$L_r^{(L)}(m,n) = L_b^{(L)}(m,n+1) - L_b^{(L)}(m,n)\ , \tag{1}$$

$$L_r^{(P)}(m,n) = L_b^{(P)}(m,n+1) - L_b^{(P)}(m,n)\ , \tag{2}$$

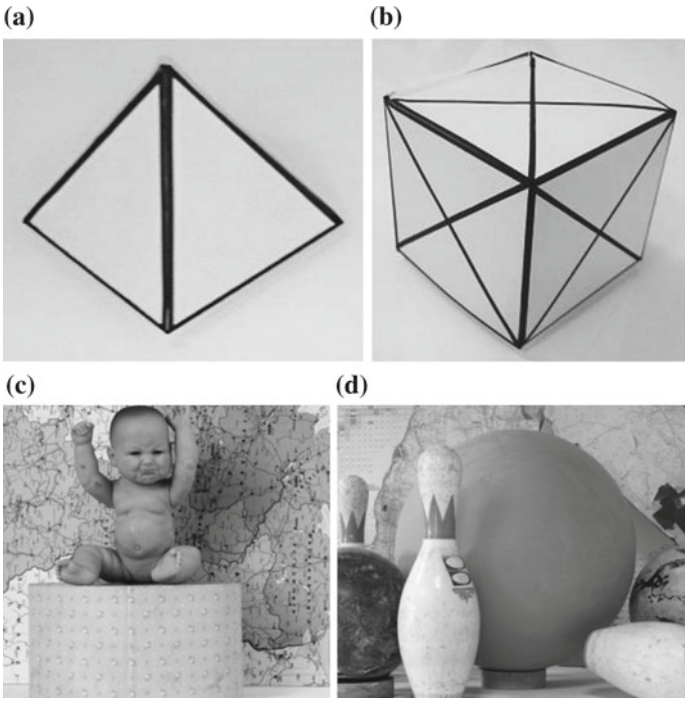


Fig. 2 Sample images used in the study **a** Pyramid, **b** cube, **c** baby1, **d** bowling2

for $n \in (1, N-1)$, where m—number of the next row, n—number of the next column, N—number of image columns.

The resulting images $L_r^{(L)}(m,n)$ and $L_r^{(P)}(m,n)$ show the edges of objects represented by the values $\{-1, 1\}$. The values "+1" are located on the edges resulting from a change in the brightness from light to dark (looking from the left-hand side). The values "−1" are located on the edges resulting from a change in the brightness from dark to light (Fig. 3).

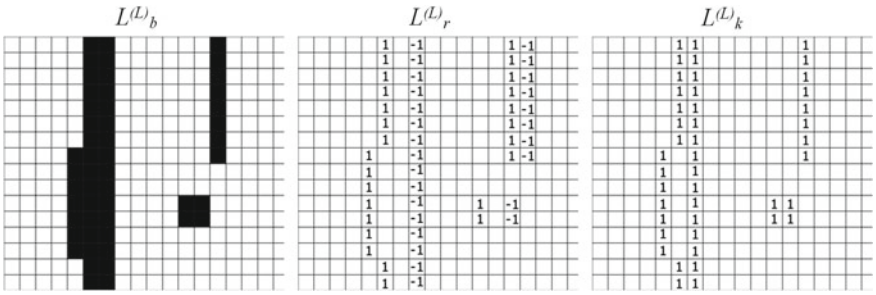


Fig. 3 Binary image $L_b^{(L)}$, image after the subtraction operation $L_r^{(L)}$, and image after shifting $L_k^{(L)}$

Next, the position of value "−1" is shifted by one column to the left and converted to the value "+1", i.e.:

$$L_k^{(L)}(m,n) = \begin{cases} 1 \ for & L_r^{(L)}(m,n) = 1 \\ 1 & L_r^{(L)}(m,n+1) = -1 \\ 0 & L_r^{(L)}(m,n) = 0 \end{cases} \tag{3}$$

$$L_k^{(P)}(m,n) = \begin{cases} 1 \ for & L_r^{(P)}(m,n) = 1 \\ 1 & L_r^{(P)}(m,n+1) = -1 \\ 0 & L_r^{(P)}(m,n) = 0 \end{cases} \tag{4}$$

for $n \in (1, N-1)$.

In the obtained images $L_k^{(L)}(m,n)$ and $L_k^{(P)}(m,n)$, the values "+1" coincide with the inner edges of objects with greater intensity. On this basis, matrices $L_o^{(L)}(m,n)$ and $L_o^{(P)}(m,n)$ are constructed containing information about the location of values "+1" in each row.

$$L_o^{(L)}(m,n) = \begin{cases} n \ for & L_k^{(L)}(m,n) = 1, \\ 0 & L_k^{(L)}(m,n) = 0, \end{cases} \tag{5}$$

$$L_o^{(P)}(m,n) = \begin{cases} n \ for & L_k^{(P)}(m,n) = 1, \\ 0 & L_k^{(P)}(m,n) = 0. \end{cases} \tag{6}$$

To calculate the distance of individual elements from each other in the corresponding rows, the matrices $L_o^{(L)}(m,n)$ and $L_o^{(P)}(m,n)$ are modified in such a way that zero elements of the rows are removed. Then zeros are added at the end of the rows. This addition is related to aligning the number of elements to the row containing the most nonzero elements (Fig. 4). In (Fig. 4) I, K are the numbers of columns of matrices $L_w^{(L)}(m,i)$ and $L_w^{(P)}(m,k)$ respectively and

$$I = \max_m \left(\sum_{n=1}^{N} L_k^{(L)}(m,n) \right), \tag{7}$$

$$K = \max_m \left(\sum_{n=1}^{N} L_k^{(P)}(m,n) \right). \tag{8}$$

Then the difference in distance between a point in the left image and all the points in the right image in a given row is calculated.

$$L_{odl}(m,i,k) = L_w^{(L)'}(m,i,k) - L_w^{(P)'}(m,i,k), \tag{9}$$

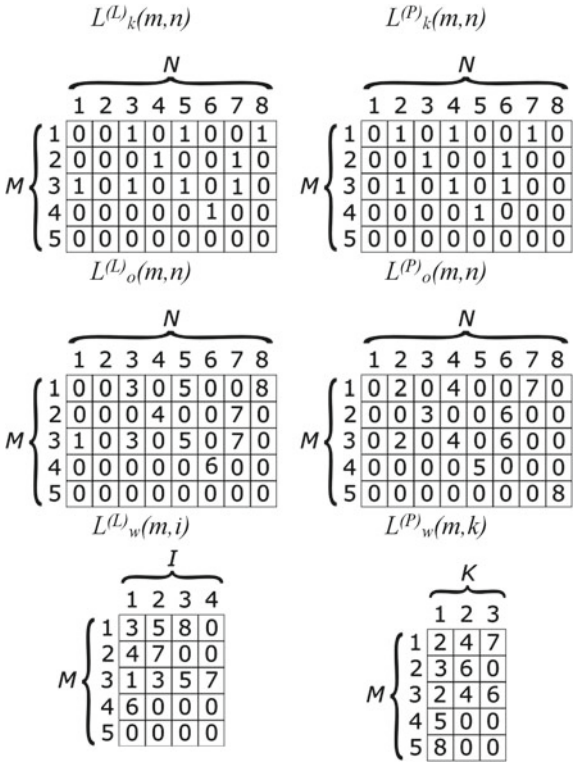


Fig. 4 Method for creating matrices $L_w^{(L)}(m,i)$ and $L_w^{(P)}(m,k)$

where

$$L_w^{(L)'}(m,i,k) = [1,1,\ldots,1]^T_{(1\times K)} \cdot L_w^{(L)}(m,i)\ , \tag{10}$$

$$L_w^{(P)'}(m,i,k) = L_w^{(P)}(m,k)^T \cdot [1,1,\ldots,1]_{(1\times I)}\ . \tag{11}$$

It is assumed that disparity D_m corresponds to the minimum distance calculated between a point in the left image and all the points in the right image for a given row

$$D_m(m,i) = \min_k \left(L_{odl}(m,i,k)\right)\ . \tag{12}$$

A disparity map, whose size is compatible with a stereo pair, is needed for further calculations.

$$D(m,n) = \begin{cases} D_m(m,i) \text{ for } & L_k^{(L)}(m,n) \neq 0, \\ 0 & \text{the others,} \end{cases} \tag{13}$$

for $m \in (1,M)$, $\quad n = L_w^{(L)}(m,i)$, $\quad i \in (1,I)$.

The resulting disparity map $D(m, n)$ enables to calculate the ground coordinates of a cloud of points according to [12].

5 Results

The algorithm was tested on a PC with an Intel Core i5 3.1 GHz. Implementation was made in MATLAB® Version: 8.6.0.267246 (R2015b) Image Processing Toolbox Version 9.3.

The results were evaluated according to three criteria. The first criterion, in accordance with [23], concerns the assessment of the effectiveness of matching images δ_r defined as:

$$\delta_r = \sqrt{\frac{1}{(M \cdot N)} \sum_{m=1}^{M} \sum_{n=1}^{N} |D(m, n) - D_T(m, n)|^2} , \tag{14}$$

where $D(m, n)$—calculated disparity map, $D_T(m, n)$—ground truth disparity map.

The second criterion relates to the percentage of mismatches δ_D defined as

$$\delta_D = \frac{1}{M \cdot N} \sum_{m=1}^{M} \sum_{n=1}^{N} L_D^{(L)}(m, n) , \tag{15}$$

where

$$L_D^{(L)} = \begin{cases} 1 \text{ for } |D(m, n) - D_T(m, n)| > p_D \\ 0 \qquad\qquad\qquad \text{the others} \end{cases}$$

p_D – allowable threshold of mismatches.

The allowable threshold of mismatches p_D was adopted arbitrarily as 2 pixels. This value was adopted taking into account the maximum spatial resolution of an image and a pixel size (for the analysed case the pixel size was 80 μm). The third criterion characterizing the algorithm is its operating time t expressed in seconds for each test image.

The results for the three criteria discussed above for all 21 Middlebury test images (Fig. 2) are available at https://goo.gl/IpNSRT. Some selected representative examples are shown in Table 1.

Although the algorithm was implemented in the interpreted code, the obtained times of the algorithm operation are less than 100 ms (for the computer parameters listed above). The operating time is one of the main criteria for the selection of the image matching algorithm [13].

Table 1 Sample results of matching images with the proposed algorithm

	Pyramid	Cube	Plastic	Lampshade1	Cloth1	Cloth2
$\delta_r\ (pix)$	4	7	15	18	32	32
$\delta_D\ (\%)$	16	31	63	74	97	92
$t\ (s)$	0.01	0.01	0.04	0.04	0.06	0.05

The operating times of the algorithm for the cube and pyramid are 10 ms, while for the images from Middlebury, they are six times longer. This is due to the much higher number of elements which are subject to matching.

The standard deviation reaches the lowest value for the pyramid and is 4 pixels. The highest value of 32 pixels was obtained for the images Cloth1 and Cloth2. This may be due to the low contrast of the images with small, densely laid, repetitive pattern.

The lowest percentage of mismatches is in the case of disparity for the Pyramid, only 16 %. The highest value of 97 % occurs in the image Cloth1. This high value is due to the background, which consists of many small elements. These elements during binarization partly disappear to a different extent in the left and right image.

It should be noted that it is impossible to reach 100 % efficiency for the matching algorithm [18]. This is due, inter alia, to the existence of areas not equally visible in both images, for example, the nose area or object edges [21].

6 Discussion

The evaluation criteria used previously for the proposed algorithm were adopted to compare it with commonly used methods for matching images discussed in the section *State of the art*. The results for all 21 images from the Middlebury library are available at https://goo.gl/IpNSRT. The tables show some selected, representative examples of results obtained for the algorithms.

It should be noted that the VAR and SG algorithms partly use compiled libraries, which significantly accelerates the speed of their operation and distorts the comparison of test results. However, it was decided to publish comparisons involving these algorithms because of the importance of their position among matching algorithms.

In terms of operating time (Table 2), the comparison of algorithms shows the advantage of the proposed algorithm which is the quickest even though two of the compared algorithms use compiled external libraries.

The proposed algorithm also does well in the comparison of standard deviations (Table 3). The lowest values are for the Cube and Pyramid in the case of the proposed algorithm, and for Bowling1, Flowerpots, Mid1 and Mid2 in the case of the SG algorithm. The highest values are for the Cube in the case of the SG algorithm and Mid1 in the case of the proposed algorithm.

Table 2 Comparison of operating times t with other algorithms in seconds

	Pyramid	Cube	Bowling1	Flowerpots	Mid1	Mid2
Proposed algorithm	0.01	0.01	0.04	0.03	0.05	0.04
VAR	0.30	0.42	4.52	4.86	5.11	4.96
SG	0.01	0.02	0.09	0.09	0.09	0.07

Table 3 Comparison of standard deviations δ_r in pixels

	Pyramid	Cube	Bowling1	Flowerpots	Mid1	Mid2
Proposed algorithm	4	7	24	25	30	28
VAR	15	15	13	15	10	10
SG	15	23	8	14	7	8

Table 4 Comparison of percentage of mismatches δ_D

	Pyramid	Cube	Bowling1	Flowerpots	Mid1	Mid2
Proposed algorithm	16	31	79	77	89	86
VAR	100	100	63	49	56	67
SG	99	98	19	21	22	25

In the case of the comparison of the percentage of mismatches (Table 4), the SG method is the best for Bowling1, and the proposed algorithm for the Pyramid and Cube. The highest values of mismatches are observed for VAR and SG methods for the Pyramid and Cube. They reach values up to 100 %, so no disparity was calculated correctly. This is undoubtedly due to the nature of these methods, which are based on the diversity of surroundings of homologous points in both images. In the case of a homogeneous background, which dominates the figure contours, diversity is too small and this causes significant mismatches over the whole surface.

Large areas of mismatches for the proposed algorithm for all Middlebury library images are due to the nature of the method itself, which pairs contours, outlines of the objects resulting from binarization and formation of edges. Pairs are made only for the outlines, ignoring the entire background. The outlines resulting from the use of the proposed algorithm may not coincide with the edges of ground truth disparities D_T, which were performed using other methods by the authors of the Middlebury library images.

7 Conclusions

The proposed new method for matching images copes well with elementary images having a homogeneous background. A special feature of this method is low time complexity making the presented algorithm fast.

Significant areas of mismatches for images from the Middlebury library are a challenge and indicate the need for further work on the algorithm. It may be necessary to introduce in the future additional analyses to the method, in order to improve its performance for complex images with a rich background. This should not increase significantly the time complexity, which is very important having regard to the possibility of implementation on mobile devices with small computational resources and memory.

References

1. 2006 Stereo datasets with ground truth. http://vision.middlebury.edu/stereo/data/scenes2006/
2. Bouguet, J.-Y.: Complete Camera Calibration Toolbox for Matlab (1999). http://www.vision.caltech.edu/bouguetj/
3. Brown, D.C.: Close-range camera calibration. Photogram. Eng. Remote Sens. **37**, 855–866 (1971)
4. D'Apuzzo, N.: Automated photogrammetric measurement of human faces. Int. Arch. Photogramm. Remote Sens. **32**(B5), 402–407 (1998)
5. D'Apuzzo, N.: Measurement and modeling of human faces from multi images. Int. Arch. Photogramm. Remote Sens. **34**(5), 241–246 (2002)
6. Fryer, J.G., Brown, D.C.: Lens distortion for close-range photogrammetry. Photogramm. Eng. Rem. S. **52**, 51–58 (1986)
7. Golec, J., Ziemka, A., Szczygiel, E., Czechowska, D., Milert, A., Kreska-Korus, A., Golec, E.: Photogrametrical analysis body position in hips osteoarthrosis. Ostry Dyzur. **5**, 1–7 (2012)
8. Golec, J., Tomaszewski, K., Maslon, A., Szczygiel, E., Hladki, W., Golec, E.: The assessment of gait symmetry disorders and chosen body posture parameters among patients with polyarticular osteoarthritis. Ostry Dyzur. **6**, 91–95 (2013)
9. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press (2003)
10. Heikkila, J., Silven, O.: A four-step camera calibration procedure with implicit image correction. In: Proceedings of IEEE Computer Society, Conference on Computer Vision and Pattern Recognition. pp. 1106–1112. IEEE Computer Society (1997)
11. Hirschműller, H.: Stereo processing by semiglobal matching and mutual information. IEEE Trans. Pattern Anal. Mach. Intell. **30**, 328–341 (2008)
12. Kraus, K.: Photogrammetry: Geometry from Images and Laser Scans. Walter de Gruyter (2007)
13. Mitchell, H.L.: Applications of digital photogrammetry to medical investigations. ISPRS J. Photogramm. Remote Sens. **50**, 27–36 (1995)
14. Mitchell, H.L., Newton, I.: Medical photogrammetric measurement: overview and prospects. ISPRS J. Photogramm. Remote Sens. **56**, 286–294 (2002)
15. Otsu, N.: Threshold selection method from grey-level histograms. IEEE Trans. Syst. Man Cybern. SMC-9, 62–66 (1979)
16. Patias, P.: Medical imaging challenges photogrammetry. Virtual Prototyp. Bio Manuf. Med. Appl. **56**, 45–66 (2008)
17. Ponce, J., Forsyth, D.: Computer vision: a modern approach (2012)

18. Popielski, P., Wróbel, Z.: An attempt to optimize the process of automatic point matching for homogeneous surface objects. Arch. Fotogram. Kartogr. i Teledetekcji. **22**, 351–361 (2011)
19. Popielski, P., Wróbel, Z.: The Feature Detection on the Homogeneous Surfaces with Projected Pattern. Lecture Notes in Computer Science, vol. 118–128. Springer, Berlin Heidelberg (2012)
20. Popielski, P., Wróbel, Z., Koprowski, R.: The effectiveness of matching methods for rectified images. In: Proceedings of the 8th International Conference on Computer Recognition Systems CORES 2013. pp. 479–489. Advances in Intelligent Systems and Computing, Springer International Publishing (2013)
21. Popielski, P., Wróbel, Z., Koprowski, R.: Object Detail Correspondence Problem in Stereovision. Advances in Intelligent Systems and Computing, vol. 3, pp. 209–222, Springer International Publishing (2014)
22. Ralli, J., Díaz, J., Ros, E.: Spatial and temporal constraints in variational correspondence methods. Mach. Vis. Appl. **24**, 275–287 (2011)
23. Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. Int. J. Comput. Vis. **47**, 7–42 (2002)
24. Shapiro, L., Stockman, G.: Comput. Vis. (2001)
25. Zhang, Z.: Flexible camera calibration by viewing a plane from unknown orientations. Proc. Seventh IEEE Int. Conf. Comput. Vis. **1**, 0–7 (1999)
26. Zhang, Z.: A flexible new technique for camera calibration. IEEE Trans. Pattern Anal. Mach. Intell. **22**, 1330–1334 (2000)

Targeted X-Ray Computed Tomography: Compressed Sensing of Stroke Symptoms

Artur Przelaskowski

Abstract The subject of reported research is model-based compressed sensing applied to CT imaging. Personalized CT examinations were designed according to requirements of CT-based stroke diagnosis in emergency care. Adaptive sensing was optimized to recover more accurately diagnostic information which is partially hidden or overlooked in standard procedures. In addition, limited number of measurements was used to reduce radiation dose. As a result, new paradigm of integrated optimization for CT system was proposed. Formalized diagnostic model is used to improve the relevance of CT imaging in emergency diagnosis. Simulated experiments confirmed a proof of concept realization.

Keywords Model-based compressed sensing · Algebraic iterative reconstruction · Personalized computed tomography · Adaptive sensing and recovery · Computerized stroke modeling · Semantic image processing

1 Introduction

X-ray computed tomography (CT) imaging plays a prominent role in the diagnosis and follow-up of patients. Among others, it is obligatorily used to detect stroke symptoms and confirm ischemic injury [15]. However, its increasing usefulness is associated with relatively high radiation exposure to patients which may result in a non-negligible lifetime risk of cancer [5, 21, 29]. The presented study fits into growing trend towards lower doses increasing the relevance and competitiveness of CT in many diagnostic applications, despite the fact that, in general, the reconstruction quality is proportional to the radiation dose. Most commonly used solutions refer to algebraic iterative reconstruction (AIR) methods. Multiple repetitions approximating current solution converge toward a better recovery of highly undersampled data

A. Przelaskowski (✉)
Faculty of Mathematics and Information Science, Warsaw University of Technology, Koszykowa 75, Warszawa, Poland
e-mail: arturp@mini.pw.edu.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*, Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_11

even with irregularly sampled acquisition positions [4]. The AIR enables to maintain clinically acceptable image quality with reduced number of projections or lowered tube charge (i.e. lowered mAs levels increasing the level of quantum noise on the sinogram). Key elements are more precise modeling of the acquisition system (represented by measurement matrix) including various physical descriptors, compressible signal representations and informative content specificity.

Additional improvement of AIR strategy is possible because of recent achievements of compressed sensing (CS) theory which provides satisfactorily reconstructed signals/images from far fewer measurements comparing the limits Shannon/Nyquist [7]. This means an additional possibility to dose reduction. Ill-defined problem of slice reconstruction is solved as optimization approach constituting typically convex problem with relaxation of important sparsity prior in terms of more computationally tractable norms. Regularization term promoting signal model is minimized subject to a fidelity prior posed by the x-ray projections. Especially, total variation (TV) minimization was investigated as additional regularization term in CT reconstruction with only a few x-ray projections [13, 19, 27, 28, 30]. Minimizing high spatial gradient suppress noise and streaking artifacts. But noticeable disadvantage of uniform penalizing the image gradient irrespective of the informative image structures is to be compensated [31, 35]. However, the theoretical benefits of CS do not simply extend to the conditioning of CT imaging. It is because CS provides theoretical guarantees of accurate undersampled recovery only for certain classes of random measurement matrices while adequate CT system matrix is deterministic. The CS optimization encapsulated in AIR allows for CT reconstruction under additional sufficient-sampling conditions based on parametrized system matrices, necessary to achieve incoherence in designing CT acquisition models [17].

Further dose reduction is associated with the use of a priori knowledge about a specific diagnostic task while the additional benefit may be better conditioning of personalized diagnosis. A concept of interior tomography applies to reconstruct primarily (with high enough quality) region of interest (ROI) which focuses physician's attention. This ROI is a representative of diagnostic model targeting CT imaging. If the ROI is the only intensively covered by X-ray radiation, the radiation exposure to patient is significantly reduced. While the interior problem in 2D is not uniquely solvable, CS tools enable exact and stable reconstruction of a local ROI via the TV minimization while the ROI is piecewise smooth (constant or polynomial), considering statistical nature of local projection data [32, 34]. In addition, model-based CS (also called second-generation or structured sparse modeling) parallels the conventional theory and provides concrete guidelines on how to create model-based recovery algorithms with provable performance guarantees [2]. Ordinary CS assumes sparse or approximately sparse (i.e. compressed) signal models to solve undetermined equations of compressive sampling. Model-based CS uses more realistic signal models that go beyond simple sparsity and compressibility by including dependencies between values and locations of the signal coefficients favoring certain configurations of the informative coefficients. The knowledge of this data dependencies can be exploited in signal stable recovery [8]. A class of model-compressible signals including wavelet trees, block sparsity or separated spikes can be extended

to approximated projections into respective model domain [12]. The effect is substantially decreased number of measurements, i.e. dose reduction without sacrificing robustness of signal reconstruction. Applied models enable better differentiating true signal information from recovery artifacts and uninformative components possibly distorting or masking the required information. However, model-based adaptive sensing means that any measurement may depend on the results of previous measurements while generic CS strategy assumes random observations. Adaptivity means that new measurements are guided by just collected information in order to optimize the gain of new information [14]. Well designed sensing based on true signal model can reliably recover sparse signals at lower SNRs than non-adaptive sensing [11].

Consequently, the problem of targeted imaging can be generalized as the question of possible personalized recovery of primarily high quality information according to the adopted a priori model with reduced dose and increased diagnostic usefulness. A challenging and extremely important application is acute stroke diagnosis. Unenhanced CT is fast, accessible and inexpensive brain imaging test which remains the most practical method of instrumental stroke diagnosis. CT imaging is obligatory used to detect ischemic signs and confirm ischemic injury in correlation with results of neurological recognition of stroke. Reliable recognition of disease allows patients with ischemic stroke to receive timely intravenous fibrinolytic therapy. However, there are serious problems with very early confirmation and assessment of ischemia in classically reconstructed CTs. The direct finding significantly facilitating stroke diagnosis is the extent of cortical/subcortical hypoattenuating (hypodense) area within a vascular territory [15]. But for the acute phase of ischemia CT scans are often normal or subtle tissue density differentiation is almost imperceptible. Therefore, lots of image processing methods and computer-aided systems were applied to improve CT-based stroke recognition with moderate success. Questionable but incidentally possible was estimation of reversible area of ischemic injury, i.e. likely to penumbra [22, 24]. However, none of them has integrated personalized CT imaging with computerized support of stroke diagnosis, according to our knowledge.

2 Problem Formulation

The reported research refers to long term study on computational modeling of acute stroke diagnosis to enhance recognition of ischemia and support the following therapeutic decisions [24–26]. The process of CT imaging was correlated with the results of neurological recognition of stroke. Dealing with challenging demands of acute CT interpretation, iterative reconstruction of CT scans was optimized with semantic model-based criteria. The criteria of dose reduction were integrated with concretized extraction of disease symptoms because of diagnostic signal models embedded into precise modeling of imaging system. The applied diagnostic model integrates the results of neurological examinations with diagnostic protocol of ischemia recognition in a priori semantic model for simultaneous acquisition and compression (CS)

of unenhanced CT imaging with improved recovery of ischemic signs to support diagnosis.

In this paper, the formalized diagnostic model representing structured neurologic information distributed in brain volume has been adapted to the designed targeted CT system in order to improve its diagnostic usefulness. Proposed methodology integrates three principal components: (a) improving stroke diagnosis because of CT targeting the object of interests in sensing and recovery of ischemic extent, (b) lowering doses of CT acquisition scoped to a particular ROI with number-effective measurements, narrowed range of angles and lowered mAs, (c) flexible and feasible CT imaging according to paradigm of personalized medicine (diagnosis).

The proposed idea refers to the CS-AIR algorithms extending CT recovery concept with reliable semantic model used to optimize tomographic sensing and adaptive image recovery. Successive approximation of AIR fashion with CS-inspired model-based optimization offer embedded adaptive models and regularization priors adjusted to diagnostic significance of true signal. A nature of the presentation is proof of concept, done theoretically, algorithmically and experimentally according to target diagnostic criteria. According to our knowledge, resulting algorithms is a first guideline for personalized diagnostic imaging based of adaptive, model-based CS-AIR concept.

2.1 Iterative Reconstruction

Formally, noisy sensing of CT images is defined by a full-rank matrix $\mathbf{A} \in \mathbb{R}^{M \times N}$ modeling imperfectly (e.g. due to discretized space of measurement, geometry limits) the acquisition system. A vector of M observations formulated as $\mathbf{y} = \mathbf{A}\mathbf{x}^{(0)} + \varepsilon$ is used to approximately reconstruct noiseless source $\mathbf{x}^{(0)}$ of N resolution. Small number of measurements $M \ll N$ results from required limits of radiation dose and model-based control of information sensing. The measurement matrix $\mathbf{A}$ is a linear operator of an underdetermined system having infinitely many solutions. A random noise vector ε reflects a statistical nature of the acquired data caused by finite number of the projections and photons, time limits, Poisson-modeled process of X-ray interactions completed with a background Gaussian noise of zero mean [32].

Preferable AIR method is simultaneous algebraic reconstruction technique (SART) greedily performing updates for complete raw data projections. The benefit of SART is best combine of algebraic and simultaneous reconstruction techniques (ART and SIRT based on Kaczmarz' projections) to yield reconstructions of good quality and numerical accuracy in only one iteration [4, 18]. Pixel-oriented weighting of projection rays is changed to more general and flexible representation of pixel-basis (e.g. bilinear elements) approximating the areas of intersection of these ray strips with the pixels. The advantageous effect is less discontinues image model (even continues with high computational cost). Moreover, for a circular reconstruction region, only partial weights are assigned to the first and last picture elements on the individual rays. The noise is additionally reduced because of presumably

considerably smaller inconsistencies with real projection data while the correction terms are simultaneously applied for all the rays in one projection. Instead of sequentially updating pixels on a ray-by-ray basis, the average of the corrections generated by all the rays in a projection is simultaneously applied to a pixel. Pixel updates are computed and stored for each subsequent ray. Next correction update of all rays is added at one to image array. In addition, a longitudinal Hamming window was heuristically adjusted to emphasize the corrections applied near the middle of a ray relative to those applied near its ends.

Fundamental SART algorithm is expressed as follows ($\mathbf{A}_j$ denote the jth raw vector of the array $\mathbf{A}$ and $a_{i,j}$ is the weighting factor representing the contribution of the ijth pixel to the jth ray integral in $k+1$-iteration):

$$\mathbf{x}_i^{k+1} = \mathbf{x}_i^k + \frac{\sum_j \left(a_{i,j} \frac{\mathbf{y}_j - \mathbf{A}_j \mathbf{x}^k}{\sum_i a_{i,j}} \right)}{\sum_j a_{i,j}}. \tag{1}$$

The structure of SART iterative algorithm is strictly related to greedy algorithms of compressed sensing. Prior object information can be used to correct reconstruction process using respective regularization terms. For example, a positivity constraint on the attenuation values, smoothing out differences between neighboring voxels or edge-detecting better fits the measured information [4]. However, parametrization of this procedures according to reliable numerical criteria is not trivial task. Application of additional regularizers, mostly done heuristically in ad-hoc manner, is associated with the risk of unstable quality of reconstruction or unrealistic, unbelievable effect of the reconstruction. Any diagnostic model, i.e. structured model of sensed data can be embedded through adjusting a distribution of the values $a_{i,j}$, controlled distribution of projections or modification of cost function regarding location and extent of known a priori ROI.

2.2 Diagnostic Model of Structured CTs

According to compressed sensing theory, the sampling rate is no longer dependent on the signal spectrum (bandwidth), but it depends on the signal structure, most of all locality and compressibility. One step further is use of true signal model specifying informative signal properties. Prior knowledge formulated in medical domain is transformed to formalized conditions of possibly improved reconstruction and reduced sensing rate.

The most important function of CT-based acute stroke diagnosis is confirming more reliably and in detail neurological recognition of stroke to avoid poor outcomes of patient treatment. A physical and neurological examination establishes the baseline to which all other evaluations are referenced. Among different methods supporting and objectifying neurologic diagnosis, the NIH Stroke Scale (NIHSS) is a well known, validated tool to score the stroke deficit [1, 6]. NIHSS is used to

objectively quantify the impairment caused by a stroke, basing on clinical findings such as symptom intensity and duration. The NIHSS has been repeatedly verified method for assessing stroke severity which is heavily correlated with the brain volume (or extent) affected by the stroke. However, stroke location effects NIHSS prediction of stroke severity, i.e. the acute NIHSS score is significantly higher in case of stroke located within the left cerebral hemisphere, compared with lesions of equal size in the right side [9]. Because of limited correlation of NIHSS scores to the extent or location of the vascular lesion, the Stroke Bricks (StBr) scale used to estimate stroke location and extent was proposed [25], selecting the areas dictated by respective joining of clinical symptoms. The main concept is arbitrary allocation of anatomically reasoned brain map of 15th normalized regions correlated to possible vascular deficits, resulting in respective neurological symptoms. Such a priori predefined map is personally projected to the individual syndromes with activation of the regions recognized in neurological evaluation. In consequence, suggested dysfunctional areas of the brain (ROIs) responsible for the observed symptoms are subjected to a detailed CT image-based analysis aimed at detecting early stroke symptoms conforming disease recognition. Concluded neurological model of stroke is used to support radiologic confirmation of ischemia based on difficult in interpretation CT imaging.

According to proposed concept, prior diagnostic model based on neurological assessment was developed to drive CT measurements and reconstruction procedures more accurately. The adaptive, ROI-oriented sensing verifies hypothesis of interior tomography [32] in a context of the integrated support for hyperacute stroke diagnosis. Moreover, diagnostic stroke model includes other specificity of informative signal. Irreversible ischemic injury is primarily represented by a focal hypodense area, in cortical, subcortical, or deep gray or white matter. A hypodense area is defined as any area in the brain with density lower than normal brain tissues. However, subtle hypodense changes are often masked due to bone artifacts (up to 14 HU), noise level in the range of 4HU and other tissue abnormalities. The hypodense changes are slight, and ischemic area is not well-outlined or contrasted (with slow edges characterized by low-frequency spectrum). Diffusely interspersed changes in gray shade can hardly be distinguished in noisy areas because of low brightness contrast and non-optimum sensing. Sparsifying domains of wavelets and frequency spectrum are often more effective for identifying the signatures of hypodensity.

2.3 Model-Based Adaptive CS

The recovery process in compressive sensing is nonlinear and computationally exhaustive while a variety of algorithms have been proposed including greedy iterative relaxations [3]. On the other hand, a central role play random nonadaptive projections that are incoherent with any fixed sparsifying basis with high probability. So far, research in CS has focused primarily on (a) reducing both the number of measurements M (as a function of N reconstructed data of K-sparse vector

$\boldsymbol{\alpha}$), (b) increasing the robustness, (c) reducing the computational complexity of the reconstruction algorithm. Because solving a convex optimization is computationally expensive, successive sparse approximation is alternatively formulated as a nonconvex problem of

$$\min_{\mathbf{x},\mathcal{I}} \#\left\{\mathcal{I} : \mathbf{y} = \sum_{i\in\mathcal{I}} \mathbf{A}_i x_i\right\}, \tag{2}$$

where $\mathcal{I}$ is a particular subset of selected indices $i = 1, \ldots, N$ and $\mathbf{A}_i$ denotes the ith column of $\mathbf{A}$. Computationally tractable solution of (2) is greedily selecting indices of $\mathbf{A}_i$ (or x_i) and successive iterative forming increasingly better approximations to measurements $\mathbf{y}$.

Required sparsity is most often the result of source signal transform or modeling than natural property of sensed signal. Actually, smooth signals are sparse in the Fourier or cosine basis, while piecewise smooth signals or images are compressible in a scale-space basis like compact wavelets or complex wavelets. Information contained in CT images is typically much smaller than its effective bandwidth applying sparsifying matrices $\boldsymbol{\Phi}$ fashioned with bases, frames or redundant dictionaries, adapted a priori or a posteriori to compressed signal. Applied transforms concentrate signal energy around the edges (wavelets, complex wavelets) or informative frequency components (Fourier, DCT). Resulting sparsity is because of compactness of covered (carried) information, often related to image ROI, frequency spectrum component, specificity of image texture or edge/wedge structure. In other words, implemented information models form compressible signals with underlined structural dependencies between the values and locations of the signal data. The consequence is model-based recovery of sparse signals.

To recover the signal $\mathbf{x}$ from the compressive measurements $\mathbf{y}$, we search for the sparsest coefficient vector $\boldsymbol{\alpha} = \boldsymbol{\Phi}^T\mathbf{x}$ that agrees with the measurements. Appropriate sparsity regularization term, most often used as relaxation of pseudonorm l_0 (i.e. number of nonzeros) is l_1 norm (minimized l_1 means increased sparsity of transformed image resulting usually in preserved edges and low contrast information but also high-gradient noise and artifacts). However, an additional necessary condition of incoherence between a measurement matrix (i.e. a real imaging modality model) A and a sparsifying model $\boldsymbol{\Phi}$ (i.e. a second measurement basis) is required (i.e. no elements of one basis has a sparse representation in terms of the other basis). More formally, the measurement matrix represented in the sparsifying domain has to meet restricted isometry property (RIP) for accurate CT recovery.

Model-based CT extends sparse or compressible signal model with structured sparsity model favoring specific configurations for the magnitudes and indices of the significant coefficients in sparsifying domain. Respective reconstruction algorithms exploit the knowledge-based models of these structures. A useful framework for CS which underlines inter-dependency structure is a union-of-subspaces model [2]. Let's assume that K-sparse signals $\mathbf{x}$ (i.e. model sparse signals) live in $\Sigma_K \subset \mathbb{R}^N$ which is the union of $\binom{N}{K}$ subspaces of dimension K. A signal model $\mathcal{M}_K \subseteq \Sigma_K$ structured basing on a priori knowledge allows only certain subspaces in Σ_K, i.e. $m_k \leq \binom{N}{K}$. It is defined as

$$\mathcal{M}_K = \bigcup_{m=1}^{m_k} \mathcal{X}_m,\ \mathcal{X}_m = \{\mathbf{x} : \mathbf{x}|_{\Omega_m} \in \mathbb{R}^K, \mathbf{x}|_{\Omega_m^C} = 0\}, \tag{3}$$

where each subspace $\mathcal{X}_m$ contains all signals with $\text{supp}(\mathbf{x}) \in \Omega_m$ meaning that the set of possible Ω_m supports the signal model relating to diagnostic ROI in signal domain or specific components in sparsifying domain. K-sparse model is used to reconstruct signal as the best K-term approximation of $\mathbf{x}$ under the model $\mathcal{M}_K$. The RIP constraint on measurement matrix can be relaxed to model-based RIP providing stable recovery [8]. Beneficially, for most greedy algorithms only simple modification of approximation step is necessary. Model-based approximation means that searching only over the m_K subspaces provides the same degree of robust signal recovery. Reconstruction of extracted informative components according to a priori formulated model is possible.

The projection of an arbitrary signal $\mathbf{x}$ to the closest signal to $\mathbf{x}$ that lies in the model was used in proposed targeted CT system of personalized diagnosis.

Model-based CS platform provides reconstruction algorithms adapted to diagnostic models which approximated the optimized solution to ROI-oriented approximation of to specific components reliably represented in carefully selected sparsifying domain.

3 Targeted CT System Design

Distinctive representation of informative signal in reconstructed CTs requires: (a) integration of SART with CS optimization according to formalized diagnostic criteria, (b) use of signal model including respective regularization terms and structure of the data dependencies characterizing informative distribution of tissue density, (c) compact and computationally acceptable implementation approximating iteratively targeted stroke evidence, (d) heuristically optimized parametrization of the integrated algorithm. Hence, the proposed methodology was based on the following activities:

- formulation of a priori knowledge to define model of diagnostic content used to predict case-dependent structure of asymmetric hypodensity distribution in CT volume; the important properties of targeted recovery are clarified distinction, location and size of asymmetric hypodensity (i.e. Ischemic Region of Interests—IROI), including enhanced texture homogeneity and contrast differentiation in a context of appearance; in details, the model was based on the results of neurological examinations projected to responsible areas of impaired brain activity (IROI);
- simulated adaptive sensing of CT data where number of measurements limited because of dose minimization is directed to the important content; incoherent measurement vectors and sparsifying vectors are selected for as sparse as possible representation of sensed information; in details, limited number of angles and raw detectors were adjusted to the IROI—reduced the number of ray photons and

angles outside the IROI was modeled; $\mathbf{\Phi}$ matrix of cosines and smooth wavelets were selected to extract hypodensity components;
- reconstruction of the CT scans with revealed stroke information, clarified in a context of less informative complement; in details, the tomographic reconstruction of projections is SART with embedded CS optimization, where TV-L1 regularizers and quality priors were adjusted to estimates of the IROI.

3.1 CS-Based Optimization

Bregman iteration method is often used to solve a variety of constrained optimization problems, i.e. $\min_{\mathbf{x}} J(\mathbf{x}) : H(\mathbf{y}, \mathbf{Ax}) \leq \varepsilon$. It is done by iteratively solving a small number of unconstrained problems: $\min_{\mathbf{x}} J(\mathbf{x}) + \mu H(\mathbf{y}, \mathbf{Ax})$. $J(\mathbf{x})$ is convex regularization term typically promoting signal sparsity (relaxed with l_1 and respective sparsifying frame or dictionary $\mathbf{x} = \mathbf{\Phi}\boldsymbol{\alpha}$) or degree of smoothness/singularity (Sobolev for uniformly smooth, TV for piecewise constant images with edge discontinuities of small perimeters). $H(\mathbf{y}, \mathbf{Ax})$ is convex quality prior that penalizes the distance between vector of measurements $\mathbf{y}$ and projected reconstruction $\mathbf{Ax}$ with parameter $\mu > 0$ (typically LS of Lasso: $H(\mathbf{y}, \mathbf{Ax}) = \frac{1}{2}\|\mathbf{Ax} - \mathbf{y}\|_2^2$ or alternatively MSE or Dantzig selector). Stable and fast Bregman iterations are due to two fundamental steps:

$$\begin{aligned} \mathbf{x}^{k+1} &= \arg\min_{\mathbf{x}} J(\mathbf{x}) + \frac{\mu}{2}\|\mathbf{Ax} - \mathbf{y}^k\|_2^2, \\ \mathbf{y}^{k+1} &= \mathbf{y}^k + \mathbf{y} - \mathbf{Ax}^{k+1}. \end{aligned} \tag{4}$$

Mixed regularization in a form of $J(\mathbf{x}) = TV(\mathbf{x}) + \lambda\|\mathbf{\Phi x}\|_1, \lambda > 0$ was often applied for medical imaging [20, 23] but also for low dose reconstruction of CT scans [35]. The important cause substantial also in the reconstruction of stroke symptoms is regarding data dependencies in acquired signal. Specificity of embedded information for stroke diagnosis is distinction of piecewise constant regions with depicted edges. Therefore, TV norm of reconstructed noise-free information as additional regularization term specific to the symptoms of ischemia was implemented. Additionally, interior problem of CT was also solved basing on CS-TV optimization [34]. Sparsity of $\boldsymbol{\alpha}$ representation in model-adjusted $\mathbf{\Phi}$ (best if representing non-informative content non-sparsely) enables dominant reconstruction of required information. Therefore, the optimization procedure for reconstruction of CT images was formulated as follows:

$$\min_{\mathbf{x}} TV(\mathbf{x}) + \lambda\|\mathbf{\Phi}^T\mathbf{x}\|_1 + \frac{\mu}{2}\|\mathbf{Ax} - \mathbf{y}\|_2^2 \tag{5}$$

extending set of regularization terms. Useful formula of general Bregman is susceptible to flexible extend resulting in reconstruction speed up for compound nature of regularization. Splitting iteration was proposed as a solution of the general

l1-regularized optimization problem for CS applications [10]. Especially, split Bregman can simplify two-term regularization of (5) by splitting l_1 and TV components. Introduced auxiliary variables $\mathbf{u}$ and $\mathbf{r}$ approximate noiseless vector sensed according to $\mathbf{A}$ and iteratively updated noisy errors, respectively. Equation (5) is transformed to both fundamental steps of Bregman (4) as follows:

$$(\mathbf{x}^{k+1}, \mathbf{u}^{k+1}) = \arg\min_{\mathbf{x},\mathbf{u}} TV(\mathbf{A}^T\mathbf{u}) + \lambda\|\mathbf{\Phi}^T\mathbf{x}\|_1 + \frac{\mu}{2}\|\mathbf{u}-\mathbf{y}\|_2^2 + \cdots$$
$$\cdots + \frac{\beta}{2}\|\mathbf{u}-\mathbf{A}\mathbf{x}-\mathbf{r}^{k+1}\|_2^2 \qquad (6)$$
$$\mathbf{r}^{k+1} = \mathbf{r}^k + \mathbf{A}\mathbf{x}^{k+1} - \mathbf{u}^{k+1}$$

Iterative minimization with respect to $\mathbf{x}$ and $\mathbf{u}$ can be performed separately because of splitting l_1 and TV components in (6) [10]. The resulting algorithm is summarized in the following steps:

(a) initialize $\mathbf{r} = 0, \mathbf{u} = \mathbf{y}$
(b) while change of $\mathbf{x}$ is essential

- $\mathbf{x} = \arg\min_{\mathbf{x}} \lambda\|\mathbf{\Phi}^T\mathbf{x}\|_1 + \frac{\beta}{2}\|\mathbf{u}-\mathbf{r}-\mathbf{A}\mathbf{x}\|_2^2$
- $\mathbf{u} = \arg\min_{\mathbf{u}} TV(\mathbf{A}^T\mathbf{u}) + \frac{\mu}{2}\|\mathbf{u}-\mathbf{y}\|_2^2 + \frac{\beta}{2}\|\mathbf{u}-\mathbf{r}-\mathbf{A}\mathbf{x}\|_2^2$
- $\mathbf{r} = \mathbf{r} + \mathbf{A}\mathbf{x} - \mathbf{u}$

The computationally tractable implementation of split Bregman iterations (6) called RecPF [33] was encapsulated into SART procedure of tomographic data reconstruction (Sect. 2.1). Slightly corrected realization of target signal approximation occurred highly efficient, stable, robust and very fast because of a small number of iterations. As a fundamental concept of solution, a classic approach for optimization problems with separable variables was proposed. Calculation of gradients for TV extended to high-order context of iso- and anisotropic models was used.

3.2 Formalized Diagnostic Model for CT Recovery

Neurologically defined model based on Stroke Bricks method (see Sect. 2.2) was used to indicate the IROI as formal model for CS-ART sensing and recovery (Sect. 2.3). The IROI definition controls sensing procedure through simulated parametrization of angles distribution (number and uniformity), photon distribution (energy concentration in IROI with exponentially sloping distribution around) and region-oriented weighting of projection rays. Optimized recovery is modified according to a priori defined subspace of possible solutions formed by region-tailored components of cost function (quality prior and TV) in the optimization procedure according to (6).

Moreover, primarily recovered tissue property of ischemic extent is asymmetric hypodense distribution in reconstructed CT images. Such hypodense signs are modeled as low-frequency change, slightly distinguished, not well-outlined, weakly

contrasted with poor, blurred outline. Such characteristics imposes specified vectors of sparsifying matrix in CS optimization procedure (regularizer of l_1 norm) according to model-based paradigm.

Concretising, formalized model $\mathcal{M}_{IROI}$ of stroke symptoms includes:

1. the defined supports Ω_x and Ω_α of possible $\mathbf{x}$ narrowing the space of possible solutions to subspace $\mathcal{X}_{IROI}$; it represents the diagnostic model in the union-of-subspaces (3) used in recovery procedure (6); it was realized by
   - unsharp and redundant representation of IROI in signal domain with support Ω_x, which was used in sensing and recovery procedures; parametrized distribution of angles, photons and region-oriented weighting of projection rays we simulated to optimize restoration quality with possibly reduced X-ray dose; next, quality prior of $\|\mathbf{Ax}-\mathbf{y}\|_2^2$ and $TV(\mathbf{x})$ term were modified with increased weights of the respective data differences and gradients calculated in Ω_x;
   - indexing of selected frequency components representing informative tissue density change and poor outlines, which were enhanced in adjusted sparsifying matrices; the bases of DCT and optimized smooth wavelets were applied, respectively, to modify the regularizer of $\|\alpha\|_1$;
2. balanced parametrization of regularization terms and quality prior in (5) including weighting of respective formulas according to adaptive implementation of $\mathcal{X}_{IROI}$.

3.3 General Algorithm

The proposed method of targeted image sensing and recovery is as follows:

(a) simulated sensing of parallel projections in tomographic acquisition system according to real physical model; respective linear measurements are represented by $\mathbf{A}$ according to SART improvements;
(b) main loop of recovery (while the stop criterion is not met)
   - initialization: $\mathbf{x} = 0$, iteration index $k = 1, \ldots$, parametrization of diagnostic model $\mathcal{M}_{IROI}$ basing on neurologically reasoned and determined IROI projected to formalized Ω_x and $\mathbf{\Phi}$ with respective Ω_α
   - a single SART iteration of restoration (1): update of $\mathbf{x}^{k+1}$ with primarily sensed IROI according to adaptive sensing scenario adjusting distribution of angles, photons and weights of projections
   - positivity constraint: $\mathbf{x}^{k+1} = \max(\mathbf{x}^{k+1}, 0)$
   - update of model Ω_α basing on actual spectrum of $\mathbf{\Phi}^T\mathbf{x}^{k+1}$ (extraction of informative coefficients)
   - $\mathbf{x}^{k+1} = \text{RecPF}(\mathbf{x}^{k+1})$ with split Bregman engine (6) and models of Ω_x and Ω_α, respectively
   - Positivity constraint: $\mathbf{x}^{k+1} = \max(\mathbf{x}^{k+1}, 0)$

Such general algorithm was heuristically optimized and experimentally verified. The indications about the most effective configurations of the tested parameters were discussed in the next section.

4 Experiments and Discussion

Studies on the effects of optimized CT acquisition protocols influencing diagnostic quality of CT images cannot be performed on volunteers because legislation and ethical considerations do not allow such practices in the case of CT. Therefore, simulated expositions of a low dose CT with limited number of projections and controlled noise level in raw projections were used to mimic image quality of low dose CT acquisitions similarly to [16]. Moreover, the imaged objects are both numerical phantom and real CT scans representative for stroke diagnosis. The reported results of optimized sensing and recovery of CT according to adaptive model of informative signal and diagnostic criteria of acute stroke diagnosis are not equivalent in literature according to our knowledge.

Simulated sensing of data in a form of tomographic parallel projections was simply parametrized to model real systems. The initial assumption was that the vector $\mathbf{y}$ of M measurements is sensed in CT system with a limited number of raw detectors and angles distributed uniformly in range of 0°–360°. It is accepted that rotation center distance is 40 cm and the detector to rotation center is 40 cm. This relates to acquired image array of $25.6 \times 25.6\,\text{cm}^2$ while the distance from the first sensed ray to the last equals 402. The detector whose length is 60 cm is modeled as straight-line array of 402 uniform bins what means that number of parallel rays for each angle is also 402. Next, the number of angles and their distribution was minimized to reduce radiation dose in CT examinations while maintaining or even improving the diagnostic accuracy of the recovery. In more detail, distribution of angles was reduced uniformly verifying diagnostic usefulness of the recovered scans. Moreover, increased concentration of simulated X-ray in IROI assuming Gaussian sloping of region importance was verified. Other optimized option was lowered tube charge simulated by increased noise level (i.e. standard deviation of Poisson-like noise) added to the projection data sensed in uninformative background of IROI. The number of reduced projections, detectors and especially noise level is directly proportional to dose reduction. All of them depends on a priori estimated, case-dependent diagnostic model, conditioning of real acquisition procedures and clinical context of emergency diagnosis.

Simulated phantom of Shepp-Logan (SP), extended with additional small penumbra and infarct regions to depict asymmetric distribution of tissue density (see Fig. 1) was used in the numerical experiments. Moreover, the results for recovered real phantom based on diagnostic CT examination were reported to make the experiment more reliable. Used CT scan (Fig. 1) is a part of real examinations representative for challenging problem of effective stroke diagnosis. Selected examples of test reconstructions were also presented in Fig. 1 while summarized results were given in Table 1. The following numerical metrics were used to assess the recovery as the first

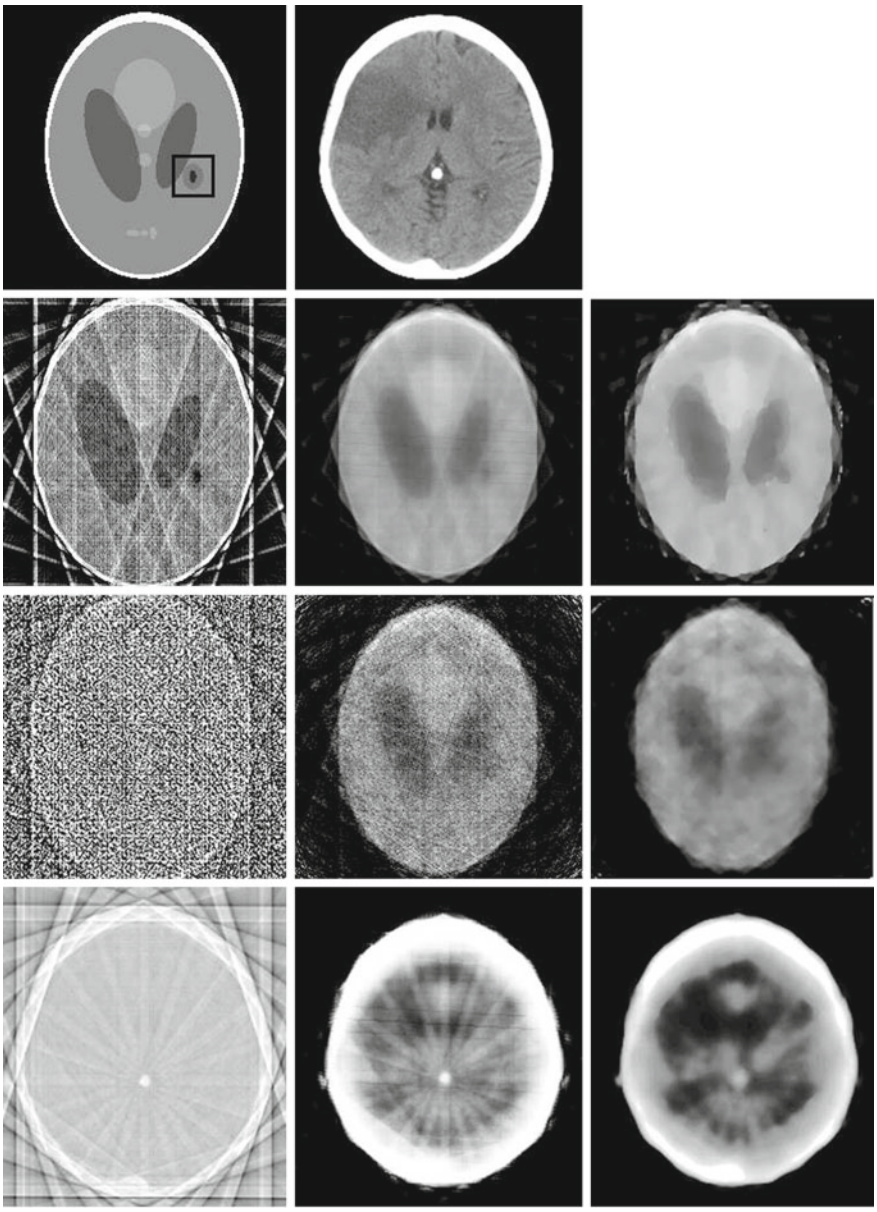


Fig. 1 The exemplary results of the image recovery in CT system. Stroke phantom modeled on Shepp-Logan procedure and real CT image with stroke evidence sensed by simulation (*top left* and *right*, respectively) were the objects of numerical tomographic imaging. Two reference methods of FBP (first column of next raws) and SART (second column) were used to verify proposed CS-SART algorithm. Second raw contains SP reconstructions for noiseless case, third one is for recovery of noisy SP projections while the bottom one is for real phantom

Table 1 The results of simulated experiments due to reconstruction of stroke symptoms in personalized CT system

		Reconstruction method								
		fbp(i)	fbp(b)	fbp(r)	sart(i)	sart(b)	sart(r)	cs-sart(i)	cs-sart(b)	cs-sart(r)
Noiseless SP projections	rmse	86.30	37.47	41.86	28.72	11.30	16.29	30.43	10.00	16.32
	ad	56.50	29.87	33.19	11.71	8.22	11.32	12.57	6.91	10.87
	snr	1.31	10.87	7.01	11.87	119.55	46.33	10.57	152.78	46.12
	psnr	9.41	16.66	15.70	18.97	27.07	23.89	18.46	28.13	23.88
Noisy SP projections	rmse	150.62	124.44	126.41	53.84	48.12	46.18	31.80	15.26	18.55
	ad	118.61	120.69	120.60	33.03	38.55	37.12	12.70	11.08	13.14
	snr	0.43	0.99	0.77	3.38	6.59	5.76	9.68	65.58	35.73
	psnr	4.57	6.23	6.09	13.51	14.48	14.84	18.08	24.46	22.77
Real projections	rmse	132.01	116.01	108.69	30.75	44.22	14.85	33.11	44.78	12.34
	ad	87.90	103.64	95.59	13.81	29.33	9.32	14.86	29.60	6.97
	snr	0.42	0.58	0.53	7.67	3.96	28.37	6.62	3.86	41.08
	psnr	5.72	6.84	7.41	18.37	15.22	24.70	17.73	15.11	26.30

Standard FBP and SART methods were compared to optimized model-based CS-SART for recovered noiseless and noisy SP stroke phantoms and real CT phantom. The applied metrics define a distance between input phantoms and reconstructions in range of a whole image (i), only brain area (b) and pointed out simulated IROI (r). Only 20 angles of the projections were used in reported experiment. The used quality metrics were defined in Sect. 4

criterion of the comparative tests in reported experiment of targeted CT optimization: root mean square error ($rmse$), average difference (ad) signal to noise ratio (snr), and peak signal to noise ratio ($psnr$):

$$\text{rmse} = \sqrt{\frac{\sum e_i^2}{N}}, \quad \text{ad} = \frac{\sum e_i}{N}, \quad \text{snr} = \frac{\sum \left(x_i^{(0)}\right)^2}{\sum e_i^2}, \quad \text{psnr} = 10 \log \frac{N \cdot \max\left(\mathbf{x}^{(0)}\right)^2}{\sum e_i^2}$$

where $\mathbf{e} = |\tilde{\mathbf{x}} - \mathbf{x}^{(0)}|$.

Significant improvement of CS-sart reconstruction is possible especially in real noisy SP case. However, the most significant improvement in the efficiency of targeted imaging was achieved in case of the real CT phantom in terms of extraction of stroke symptoms. Diagnostically significant properties of ischemia have been reconstructed with significantly improved reliability in both the numerical assessment, as well as primarily subjective assessment of imaging results.

5 Conclusions

Proposed approach refers to a development of stroke confirmation paradigm in emergency application of acute stroke care. Recently developed computer-aided diagnosis (CAD) have been based on quality improvement, normalization and analysis of processed CT scans while presented methodology of targeted CT imaging extends the image processing concept with preceding clinical data analysis (inter alia neurological assessment) and integrated optimization of the acquisition and reconstruction in CT system. This means that conventional radiological assessment may be integrated with neurological syndromes related to brain areas, structures and vascularization dictated by clinical symptoms. Therefore, CT-based diagnosis of acute stroke was focused on model-based acquire and regularized reconstruction of the informative signal to extract the earliest phases of the evolutioning ischemic extent. More reliable stroke recognition and precise characteristics of pathological symptoms in neurologically reasoned areas lead to accurate diagnostic decisions. The proposed way is optimized sensing and recovery of CT scans to accurately confirm or undermine neurological observations. Anatomically defined areas related to respective neurological symptoms, correlated to radiologic pragmatism of reported assessment formulates semantic model of potential ischemia to be accurately recovered for diagnosis of acute stroke basing on optimized CT examinations.

The presented studies confirmed the validity of proposed and realized targeted CT imaging for stroke demonstrating it in principle, verifying its feasibility and potential usefulness to suggest the real potential of being used. The obtained results are an incentive for further research, which should cover most of all: a more complete modeling of the entire diagnostic process and physical experiments on phantoms with the possibility of flexible adaptation of acquisition parameters.

Acknowledgments This publication was funded by the National Science Centre (Poland) based on the decision DEC-2011/03/B/ST7/03649.

References

1. Adams, H.P., Davis, P.H., Leira, E.C., et al.: Baseline NIH stroke scale score strongly predicts outcome after stroke: a report of the trial of org 10172 in acute stroke treatment (TOAST). Neurology **53**(1), 126–131 (1999)
2. Baraniuk, R.C., Cevher, V., Duarte, M.F., Hegde, C.: Model-based compressive sensing. IEEE Tran. Inf. Theory **56**(4), 1982–2001 (2010)
3. Baraniuk, R., Davenport, M.A., Duarte, M.F., Hegde, C. et al.: An introduction to compressive sensing. http://cnx.org/content/col11133/latest/ (2011)
4. Beister, M., Kolditz, D., Kalender, W.A.: Iterative reconstruction methods in X-ray CT. Physica Med. **28**, 94–108 (2012)
5. Brenner, D.J., Hall, E.J.: Computed tomography—an increasing source of radiation exposure. N. Engl. J. Med. **357**, 2277–2284 (2007)
6. Brott, T., et al.: Measurements of acute cerebral infarction: a clinical examination scale. Stroke **20**, 864–870 (1989)
7. Donoho, D.: Compressed sensing. IEEE Trans. Inf. Theory **52**, 1289–1306 (2006)
8. Duarte, M.F., Cevher, V., Baraniuk, R.C.: Model-based compressive sensing for signal ensembles. In: Proceedings of the 47rd Allerton Conference on Communication, Control and Computer, pp. 244–250 (2009)
9. Fink, J.N., et al.: Is the association of National Institutes of Health Stroke Scale scores and acute magnetic resonance imaging stroke volume equal for patients with right- and left-hemisphere ischemic stroke? Stroke **33**, 954–958 (2002)
10. Goldstein, T., Osher, S.: The split Bregman method for L1 regularized problems. SIAM J. Imaging Sci. **2**(2), 323–343 (2009)
11. Haupt, J., Castro, R., Nowak, R.: Improved bounds for sparse recovery from adaptive measurements. IEEE ISIT, 1563–1567 (2010)
12. Hegde, C., Indyk, P., Schmidt, L.: Approximation-tolerant model-based compressive sensing. arXiv:1406.1579 (2014)
13. Huang, J., Zhang, Y., Ma, J. et al.: Iterative image reconstruction for sparse-view CT using normal-dose image induced total variation prior. Plos One **8**(11) (2013)
14. Iwen, M.A., Tewfik, A.H.: Adaptive compressed sensing for sparse signals in noise. In: Proceedings of the ASILOMAR Conference on Signals, Systems, and Computers, pp. 1240–1244 (2011)
15. Jauch, E.C., Saver, J.L., Adams, H.P., Bruno, A., et al.: Guidelines for the early management of patients with acute ischemic stroke. Stroke **44**(3), 870–947 (2013)
16. Joemai, R.M., Geleijns, J., Veldkamp, W.J.: Development and validation of a low dose simulator for computed tomography. Eur. Radiol. **20**, 958–966 (2010)
17. Jorgensen, J.S., Sidky, E.Y., Pan, X.: Quantifying admissible undersampling for sparsity-exploiting iterative image reconstruction in X-Ray CT. IEEE Tran. Med. Imag. **32**(2), 460–473 (2013)
18. Kak, A.C., Slaney, M.: Principles of computerized thomographic imaging. Society of Industrial and Applied Mathematics, chapter 7 (2001)
19. Liu, Y., Ma, J., Fan, Y., Liang, Z.: Adaptive-weighted total variation minimization for sparse data toward low-dose x-ray computed tomography image reconstruction. Phys. Med. Biol. **57**, 7923–7956 (2012)
20. Michailovich, O., et al.: Spatially regularized compressed sensing for high angular resolution diffusion imaging. IEEE Tran. Med. Imag. **30**(5), 1100–15 (2011)

21. National Council on Radiation Protection and Measurements (NCRP) Report 160. Ionizing radiation exposure of the population of the US. NCRP, Bethesda (2009)
22. Oliveira, M.S., Fernandes, P.T., Avelar, W.M., et al.: Texture analysis of computed tomography images of acute ischemic stroke patients. Braz. J. Med. Biol. Res. **42**(11), 1076–1079 (2009)
23. Patel, V.M., et al.: Gradient-based image recovery methods from incomplete Fourier measurements. IEEE Trans. Image Process. **21**(1), 94–105 (2012)
24. Przelaskowski, A.: Recovery of CT stroke hypodensity—an adaptive variational approach. J. Comp. Med. Imag. Graph. **46**, 131–141 (2015)
25. Przelaskowski, A., Ciszek, B., Jozwiak, R., Domitrz, I., Sobieszczuk, E.: Stroke bricks—the segments of interests to analyze early stages of the disease evolution. Report submitted to International Journal of Medical Informatics (2016)
26. Przelaskowski, A., Sklinda, K., Bargieł, P., et al.: Improved early stroke detection: wavelet-based perception enhancement of computerized tomography exams. Comp. Biol. Med. **37**(4), 524–533 (2007)
27. Qi, H., Chen, Z., Zhou, L.: CT image reconstruction from sparse projections using adaptive TpV regularization. Comp. Math. Met. Med. (2015)
28. Sidky, E.Y., Pan, X.: Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. Phys. Med. Biol. **53**(17), 4777–4807 (2008)
29. Smith-Bindman, R., Lipson, J., Marcus, R., et al.: Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer. Arch. Intern. Med. **169**, 2078–86 (2009)
30. Tang, J., Nett, B.E., Chen, G.-H.: Performance comparison between total variation (TV)-based compressed sensing and statistical iterative reconstruction algorithms. Phys. Med. Biol. **56**(18), 5949–5967 (2011)
31. Tian, Z., Jia, X., Yuan, K., et al.: Low dose CT reconstruction via edge-preserving total variation regularization. Phys. Med. Biol. **56**(18), 5949–5967 (2011)
32. Xu, Q., Mou, X., Wang, G., Sieren, J., Hoffman, E.A., Yu, H.: Statistical interior tomography. IEEE Tran. Med. Imag. **30**(5), 1116–1128 (2011)
33. Yang, Y., et al.: A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data. IEEE J Sel. Top. Sig. Proc. **4**(2), 288–297 (2010)
34. Yu, H., Wang, G.: Compressed sensing based interior tomography. Phys. Med. Biol. **54**(9), 2791–2805 (2009)
35. Zhu, Z., et al.: Improved compressed sensing-based algorithm for sparse-view CT image reconstruction. Comp. Math. Met. Med. (2013)

Simple Atlas Selection Strategies for Liver Segmentation in CT Images

Dominik Spinczyk and Agata Krasoń

Abstract The paper present two atlas selection strategies: segmentation with a fixed, single individual atlas and segmentation with the best atlas for liver CT images. These two strategies was implemented and results were compared using DICE similarity coefficient, mean surface and Hausdorff distance. The average mean surface distance for single individual atlas equals 3.49 and 3.08 mm for Sum of Squared Differences and Mutual Information respectively. Average Hausdorff distance for the similarity measures mentioned above, which measure outliers, equal to 6.02 and 4.65 mm. The average DICE similarity coefficient are 0.49 and 0.59. The better results were obtained for Mutual Information similarity measure.

Keywords Atlas based segmentation · Liver segmentation · Atlas selection strategies

1 Introduction

Liver is the biggest organ in the abdomen and its segmentation in the different medical image modalities is a significant challenge for contemporary medicine. The most popular medical image modality used for liver diagnosis is Computed Tomography (CT). Different methods were proposed to this task: region growing, active contours, level set, graph cuts, clustering and threshold based methods, deformable models [1–4].

One of the proposed segmentation method is the atlas based segmentation. Okada et al. proposed a probabilistic atlas as an initialization step for multi-level statistical shape model and reported average distance between 4–2 mm depending on the level of statistical shape [5]. Pieter et al. used the Free Form Deformation algorithm and reported average overlap error of 10.4 % and an average RMS distance of 5.0 mm

D. Spinczyk (✉) · A. Krasoń
Faculty of Biomedical Engineering, Silesian University of Technology,
Roosevelta 40, Zabrze, Poland
e-mail: dspinczyk@polsl.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_12

[6]. Linguraru et al. proposed the mean organ model registration to a new case, then geodesic active contour as the first improvement and normalized probabilistic atlas corrected for precise computation of the organ volume. They reported DICE similarity coefficient of 96.2 [7].

The goal of the paper is to compare the implementation of two simple atlas selection strategies: single individual atlas and the best atlas for liver segmentation in CT images. The evaluation is based on DICE similarity coefficient, mean surface and Hausdorff distance. The Materials and Methods section include the description of mentioned above two simple atlas selection strategies, similarity measures and evaluation criteria. The Results and Conclusions section presents obtained results.

2 Materials and Methods

2.1 Steps of Building Atlas

The atlas development process requests manual semi-automatic segmentation of selected set of cases, verified by medical expert. The data consists of two images: a patient intensity image mask and the corresponding binary segmentation mask, which is generally created manually. In the next step the atlas-based segmentation method propagates the segmentation of an atlas image using the image registration technique (Fig. 1).

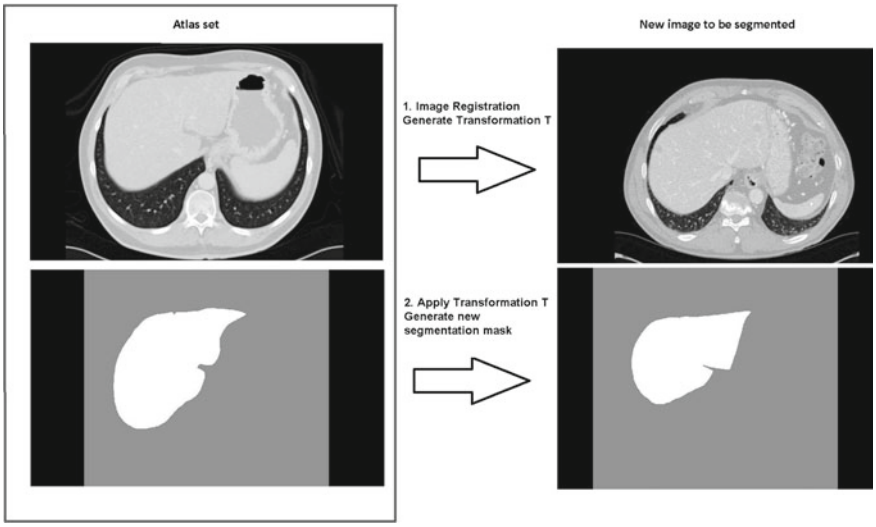


Fig. 1 Concept behind the single atlas-based segmentation

In the image registration, a spatial transformation between two images, the fixed (target) image and a moving (source) image is determined [8]. Mathematically, it is commonly written as an optimization process which finds the transformation $T(\mathbf{x}) = \mathbf{x} + u(\mathbf{x})$ that relates the two images such that the transformed moving image $I_M(T(\mathbf{x}))$ spatially matches the fixed image $I_F(\mathbf{x})$ at every position of $\mathbf{x}$, where a metric $M(I_F(\mathbf{x}), I_M(T(\mathbf{x})))$ is used to quantify the quality of the match and $u(\mathbf{x})$ is a function that represent affine and b-spline transform.

The user can select one of the following transformation types:

- rigid transformation,
- affine transformation,
- B-spline transformation,
- non-parametric transformation (e.g. fluid, elastic registration).

In this paper the B-spline transformation will be used. A rectangular grid $G = K_x \times K_y \times K_z$ is superimposed on the image (size $N_x \times N_y \times N_z$, $K_x \ll N_x$, $K_y \ll N_y$, $K_z \ll N_z$) which gets deformed under the influence of the control points. The dense deformation is given as a summation of tensor products of univariate splines. The displacement field u($\mathbf{x}$) is given as:

$$T(\mathbf{x}) = \sum_{l=0}^{3} \sum_{m=0}^{3} \sum_{n=0}^{3} B_l(\mu_x) B_m(\mu_y) B_n(\mu_z) d_{i+l, j+m, k+n}, \tag{1}$$

where
$i = \lfloor x/N_x \rfloor - 1$, $j = \lfloor y/N_y \rfloor - 1$, $k = \lfloor z/N_z \rfloor - 1$,
$\mu_x = x/N_x - \lfloor x/N_x \rfloor$, $\mu_y = y/N_y - \lfloor y/N_y \rfloor$, $\mu_z = z/N_z - \lfloor z/N_z \rfloor$,
B_l—represents lth basis function of the B-spline and d denotes the displacement.

Then registration refers to as a selection the set of the parameters:

$$\hat{\mu} = \arg\min C(\mu; I_F, I_M), \tag{2}$$

where
$C(\mu; I_F, I_M)$—is the cost function related to the similarity metrics.

The advantage of an atlas based segmentation is the ability to segment the image with no well defined relation between regions and pixels intensities. Another important advantage is applicability in clinical practice, for computer aided diagnosis, whereas, they are often used to measure the shape of an object or detect morphological differences between patient groups [9]. The reader can realize that different sophisticated methods whave been proposed to improve the results. However, the complexity prevents them from implementation. Thus in these paper we focus on basic strategies for atlas' selection. Generally four different selections can be distinguished [10]:

- segmentation with a fixed, single individual atlas,
- segmentation with the best atlas for an image,

- segmentation with an average shape atlas,
- multi-atlas segmentation: a classifier approach.

In this article first two strategies will be evaluated: segmentation with a fixed, single individual atlas and segmentation with the best atlas for an image.

Segmentation with a Fixed, Single Individual Atlas The most straight forward strategy for selection of an atlas is to use one individual segmented image. The selection can be random, or based on heuristic criteria such as image quality, lack of artefacts, or normality of the imaged subject. This strategy is by far the most commonly used method for creating and using an atlas [11]. It requires only one atlas, which greatly reduces the preparation effort as compared to the more complex methods described below.

Segmentation with the Best Atlas for an Image This case assumes that many images exist in atlas set. For each new image to be segmented there is one atlas that will produce the best segmentation accuracy among all available atlases. Two criteria could be helpful to select the best atlas:

- image similarity measure applied after image registration step,
- magnitude of the deformation required to map the coordinates of the image onto that of the atlas.

Different modalities images can be segmented using the atlas based approach and then Mutual Information is used as similarity measure:

$$MI(\mu; I_F, I_M) = \sum_{f \in L_F} \sum_{m \in L_M} p(\mu; f, m) \log_2 \left(\frac{p(\mu; f, m)}{p_F(\mu; f) p_M(\mu; m)} \right), \quad (3)$$

where

- L_F and L_M—sets of regularly spaced intensity bins of histograms of the fixed image and the moving image respectively,
- p—the discrete joint probability of the image intensities of the two images,
- p_F and p_M—the marginal discrete probabilities of the fixed image and the moving image intensities respectively.

As the magnitude of deformation, usually two measures are used:

- average deformation of the atlas over all voxels—after non-rigid registration, the magnitude of the deformation between the raw image and each individual atlas is computed and averaged over all voxels,
- maximum deformation of the atlas over all voxels—this criterion is identical to the previous one, except that it uses the maximum deformation over all voxels rather than the average.

In our study the cost function is defined as:

$$C(\mu; I_F, I_M) = -MI(\mu; I_F, I_M) + \alpha R(\mu), \tag{4}$$

$$R(\mu; I_M) \frac{\Delta}{=} \frac{1}{\sum_{\mathbf{x}} c(\mathbf{x} + u(\mathbf{x}))} \sum_{\mathbf{x}} c(\mathbf{x} + u(\mathbf{x})) \times \left\{ c_{AC} \sum_{k,i,j} AC_{kij}(\mathbf{x})^2 + \right.$$
$$\left. + c_{OC} \sum_{i,j} OC_{ij}(\mathbf{x})^2 + c_{PC} PC(\mathbf{x})^2 \right\}, \tag{5}$$

where

- $R(\mu)$—regularization term which constrains the non-rigid deformation,
- $c(x) \in [0, 1]$—rigidity coefficient,
- c_{AC}, c_{OC}, c_{PC}—weights,
- α—a weight term, which balances the similarity metric and regularization term.

Additionally Sum of Squared Differences measure was used:

$$SSD(\mu; I_F, I_M) = \sum_{x_i \in \Omega_F} \left(I_F(x_i) - (I_M(T_\mu(x_i))) \right)^2 \tag{6}$$

where

- Ω_F—domain of the fixed image.

and

$$C(\mu; I_F, I_M) = SSD(\mu; I_F, I_M) + \alpha R(\mu) \tag{7}$$

Configuration of the Registration Step Data set consists of fifteen 3D non-contrast enhanced abdominal CT studies, with liver mask manually segmented by radiologists, which is divided into two sets: atlases and new image to be segmented. Insight Toolkit Image Registration Framework is used as registration environment. Multiresolution strategy is used with three levels of resolution. As regularization term a local rigidity term is used (see (5)) [12].

2.2 *Quantifying Segmentation Accuracy*

Computing the accuracy of a segmentation requires a gold standard or ground truth. It is commonly accepted as a gold standard to use a manual segmentation by a human expert.

As similarity index the following measures are used:

- Dice similarity coefficient (DICE) defined as:

$$DICE\left(I_{F-Seg}, I_{M-Seg}(T)\right) = 2\frac{|I_{F-Seg} \cap I_{M-Seg}(T)|}{|I_{F-Seg}| + |I_{M-Seg}(T)|}, \tag{8}$$

where
I_{F-Seg}, I_{M-Seg} present two segmentations and |.| denominates number of voxels inside the segmentation,

- Mean surface distance (MSD) defined as:

$$MSD(I_{F-Seg}, I_{M-Seg}(T)) = \frac{1}{n_X + n_Y}\sum_{i=1}^{n_X} d_i \sum_{j=1}^{n_y} d_j, \tag{9}$$

where
n_X, n_Y represent the number of voxels on the two segmentation surfaces respectively, and d are the closest distances from each voxel on the surface to the other surface. The value of this measure helps to identify distance between registered surfaces.

- Hausdorff distance defined as:

$$H(I_{F-Seg}, I_{M-Seg}(T)) = \max\left\{\sup_{x\in N_X}\inf_{y\in N_Y} d(x,y), \sup_{y\in N_Y}\inf_{x\in N_X} d(x,y)\right\}, \tag{10}$$

where
N_X, N_Y represent the voxels on the two segmentation surfaces correspondingly, and d_i and d_j are the closest distances from each voxel on the surface to the other surface.

- Mean deformation field defined as:

$$MDF = \frac{1}{n}\sum_{i=1}^{n} |v_i|, \tag{11}$$

where
$v_i = [x_i, y_i, z_i]$—deformation vector for specific i-th voxel,
$|v_i| = \sqrt{x_i^2 + y_i^2 + z_i^2}$—magnitude of deformation vector for specific i-th voxel,
n—number of voxels for which the deformation field was calculated.
The value of this measure helps to identify maximum distance between registered surfaces.

Presented strategies were tested on a dataset consisted of fifteen 3D CT volumes of abdominal cavity. Images were separated into training set that consists of 9 volumes and 6 new cases for automatic liver segmentation.

3 Results and Conclusions

For the two atlas selection strategies specified and the defined above similarity measures atlas-based segmentations were performed. Following results were obtained:

- similarity index (Tables 1, 2, 3, 4, 5 and 6),
- the example of the best and the worst registrations with a fixed, single individual atlas for Sum of Squared Differences similarity measure for selected axial projection (Fig. 2—see Table 7: atlas seg 6 row).

The average mean surface distance for a single individual atlas equals 3.49 and 3.08 mm for Sum of Squared Differences and Mutual Information, respectively. Average Hausdorff distance, which measures outliers, equal to 6.02 and 4.65 mm, respectively. The average DICE similarity coefficients are 0.49 and 0.59, respectively. The better results were obtained for the Mutual Information similarity measure. The mean surface distance is similar to results in [5, 6]. The Dice similarity coefficient is worse than in [6, 7].

Comparing two evaluated simple atlas selection strategies: fixed single, individual atlas and the best atlas (using the additional criteria like mean magnitude of the deformation), the average distance between registered surfaces decreases significantly (see comparison in columns of Table 2 with best and worst case in corresponding columns of Table 8; the best case in column is bold).

Table 1 Mean surface distance in atlas registration for sum of squared differences similarity measure (mm)

Atlas case	New case 1	New case 2	New case 3	New case 4	New case 5	New case 6
Atlas seg 1	0.55	0.64	3.39	1.49	3.21	3.22
Atlas seg 2	10.04	1.45	5.75	3.48	0.67	8.7
Atlas seg 3	3.42	0.82	2.16	3.07	0.98	2.24
Atlas seg 4	1.3	3.95	0.18	0.21	1.37	0.61
Atlas seg 5	12.5	9.01	2.37	4.05	1.51	12.52
Atlas seg 6	10.8	4.12	1.44	5.4	0.25	5.15
Atlas seg 7	1.7	0.97	2.29	6.7	2.85	0.17
Atlas seg 8	0.85	6.13	0.1	0.08	2.36	0.46
Atlas seg 9	12.3	1.3	3.43	4.8	0.16	9.58
Average value	5.94	3.15	2.35	3.25	1.48	4.74

Table 2 Mean surface distance in atlas registration for mutual information similarity measure (mm)

Atlas case	New case 1	New case 2	New case 3	New case 4	New case 5	New case 6
Atlas seg 1	1.75	4.23	1.45	1.6	1.59	3.15
Atlas seg 2	0.25	2.2	0.59	0.49	4.27	1.07
Atlas seg 3	1.01	12.28	1.73	1.14	3.22	1.24
Atlas seg 4	2.7	10.15	2.75	6.69	7.15	2.17
Atlas seg 5	0.36	0.18	0.98	0.62	0.75	0.7
Atlas seg 6	0.41	8.45	2.95	0.47	0.75	1.72
Atlas seg 7	15.25	0.93	0.87	1.64	8.88	0.72
Atlas seg 8	2.8	8.45	8.18	1.91	11.45	1.15
Atlas seg 9	0.88	0.93	2.25	0.77	5.43	0.78
Average value	2.82	5.31	2.42	1.70	4.83	1.41

Table 3 Hausdorff distance in atlas registration for sum of squared differences similarity measure (mm)

Atlas case	New case 1	New case 2	New case 3	New case 4	New case 5	New case 6
Atlas seg 1	0.55	1.28	5.61	2.57	6.42	5.63
Atlas seg 2	13.69	2.9	11.05	5.66	1.31	14.32
Atlas seg 3	6.2	1.65	4.32	6.13	0.98	2.48
Atlas seg 4	2.6	7.95	0.18	0.37	2.67	1.11
Atlas seg 5	18	16.32	4.67	8.25	3.03	15.72
Atlas seg 6	15.14	6.36	2.98	8.17	0.77	8.23
Atlas seg 7	3.51	1.93	2.57	13.04	5.71	0.34
Atlas seg 8	1.75	12.25	0.1	0.16	4.71	0.91
Atlas seg 9	24.06	2.6	6.83	9.62	0.32	19.38
Average value	9.5	5.91	4.26	6.0	2.88	7.57

The main reason for the worse performance of DICE coefficient is a big shape variation of liver shape in patients' population. The registration algorithm based on Mutual Information similarity measure is a good initial step for liver surface segmentation but additional correction is desirable using the knowledge of shape variation in population.

3.1 Conclusions

In the paper elements of atlas selection strategies have been presented and atlas-based liver segmentation for fixed, single individual atlas and the best atlas have

Table 4 Hausdorff distance in atlas registration for mutual information similarity measure (mm)

Atlas case	New case 1	New case 2	New case 3	New case 4	New case 5	New case 6
Atlas seg 1	1.5	5.74	3.9	1.88	3.4	4.2
Atlas seg 2	0.44	2.93	1.09	0.91	4.9	0.79
Atlas seg 3	1.51	19.95	4.3	1.72	5.7	1.8
Atlas seg 4	6.02	6.06	3.8	9.19	8.7	4.4
Atlas seg 5	0.41	0.21	1.57	0.71	0.79	1.34
Atlas seg 6	1.11	13.91	4.05	0.89	13.8	2.41
Atlas seg 7	21.3	1.55	1.74	2.18	1.73	0.9
Atlas seg 8	3.5	9.44	12.8	6.14	15.53	4.7
Atlas seg 9	1.33	5.55	4.09	1.35	10.02	1.2
Average value	4.12	7.26	4.14	2.77	7.17	2.41

Table 5 DICE similarity coefficient for sum of squared differences similarity measure

Atlas case	New case 1	New case 2	New case 3	New case 4	New case 5	New case 6
Atlas seg 1	0.73	0.68	0.52	0.63	0.41	0.49
Atlas seg 2	0.26	0.54	0.37	0.44	0.63	0.24
Atlas seg 3	0.5	0.65	0.52	0.47	0.65	0.53
Atlas seg 4	0.52	0.43	0.59	0.65	0.62	0.66
Atlas seg 5	0.18	0.18	0.55	0.37	0.54	0.23
Atlas seg 6	0.25	0.48	0.54	0.36	0.73	0.3
Atlas seg 7	0.6	0.67	0.48	0.23	0.47	0.5
Atlas seg 8	0.63	0.44	0.65	0.7	0.56	0.61
Atlas seg 9	0.24	0.58	0.39	0.44	0.71	0.1
Average value	0.43	0.52	0.51	0.48	0.59	0.41

Table 6 DICE similarity coefficient for mutual information similarity measure

Atlas case	New case 1	New case 2	New case 3	New case 4	New case 5	New case 6
Atlas seg 1	0.72	0.52	0.59	0.67	0.63	0.57
Atlas seg 2	0.75	0.54	0.69	0.61	0.47	0.67
Atlas seg 3	0.62	0.21	0.56	0.68	0.46	0.62
Atlas seg 4	0.51	0.54	0.53	0.62	0.79	0.54
Atlas seg 5	0.79	0.83	0.56	0.75	0.27	0.69
Atlas seg 6	0.73	0.23	0.53	0.72	0.68	0.59
Atlas seg 7	0.25	0.71	0.68	0.7	0.68	0.74
Atlas seg 8	0.55	0.47	0.71	0.52	0.31	0.57
Atlas seg 9	0.73	0.54	0.52	0.73	0.42	0.73
Average value	0.63	0.51	0.6	0.67	0.52	0.63

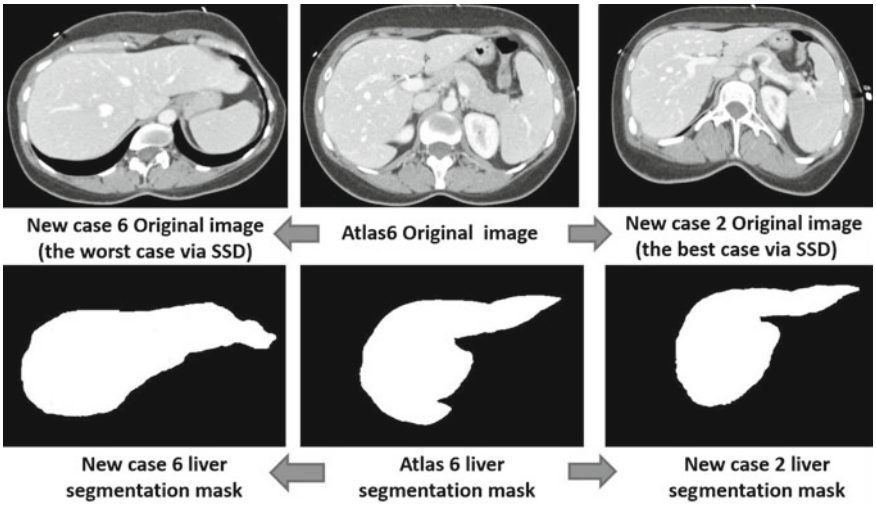


Fig. 2 The best and the worst registrations via SSD with a fixed, single individual atlas for Mutual Information similarity measure

Table 7 Sum of squared differences image similarity measure for the best atlas selection (bold font is the best atlas for specific new case)

Atlas case	New case 1	New case 2	New case 3	New case 4	New case 5	New case 6
Atlas seg 1	11178.43	3138.55	**4139.75**	15458.3	2374.38	3655.57
Atlas seg 2	15987.92	2967.74	4304.61	21132.64	2047.69	2855.2
Atlas seg 3	10140.52	2034.43	4849.66	17005.21	1392.71	3836.22
Atlas seg 4	9727.43	3376.71	4821.52	18980.12	3335.68	15464.35
Atlas seg 5	**1656.15**	1639.02	8198.71	9380.82	13614.2	8340.55
Atlas seg 6	2124.77	**863.03**	7979.474	**2502.6**	**600.91**	7609.31
Atlas seg 7	12704.72	2988.98	3643.52	13104.75	2486.52	**1758.4**
Atlas seg 8	17964.09	2279.81	4304.4	19242.13	1856.02	11797.34
Atlas seg 9	19781.91	3288.17	11123.07	30757.19	2434.46	8486.11
Average value	11251.77	2508.49	5929.41	16395.97	3349.17	7089.23

been implemented. Better results were obtained for the Mutual Information image similarity measure than for Sum of Squared Differences. As regards the selection of the best atlas, two similarity measures were used: Sum of Squared Differences image similarity measure, which is applied after the image registration step and mean magnitude of the deformation, that is required to map the coordinates of the image onto that of the atlas. As main evaluation criterion mean surface distance between registered surfaces was used. Additional evaluation criteria like: Hausdorff distance, DICE similarity coefficient and magnitude of deformation, measure the geometric

Table 8 Mean magnitude of the deformation for the best atlas selection (bold font is the best atlas for specific new case)

Atlas case	New case 1	New case 2	New case 3	New case 4	New case 5	New case 6
Atlas seg 1	58.17	91.23	19.62	64.62	151.2	20.53
Atlas seg 2	56.17	23.01	23.54	178.54	**35.85**	23.23
Atlas seg 3	**39.26**	41.06	19.18	**16.93**	141.35	23.07
Atlas seg 4	86.58	178.03	104.74	106.15	231.34	18.13
Atlas seg 5	103.62	113.43	**8.7**	76.87	57.64	**12.29**
Atlas seg 6	55.76	39.73	13.01	66.28	79.09	13.71
Atlas seg 7	92.91	186.01	178.89	97.3	166.26	135.15
Atlas seg 8	99.73	209.77	142.84	115.68	134.78	113.27
Atlas seg 9	69.6	**21.5**	20.2	68.59	39.6	26.57
Average value	73.53	100.42	58.97	87.88	115.23	42.88

similarity, were used as measures for geometric similarity. Moreover the Sum of Square distance was used which in contrast to the previous measures presents image content similarity.

Acknowledgments The study was supported by National Science Center, Poland, Grant No. UMO-2012/05/B/ST7/02136.

References

1. Rathore, S., Iftikhar M., Hussain M., Jalil A.: Texture analysis for liver segmentation and classification: a survey. In: Frontiers of Information Technology, pp. 121–126 (2011)
2. Mharib, A., Ramli, A., Mashohor, S., Mahmood, R.: Survey on liver CT image segmentation methods. Artif. Intell. Rev. **37**(2), 83–95 (2012)
3. Priyadarsini, S., Selvathi, D.: Survey on Segmentation of liver from CT images. In: IEEE International Conference on Advanced Communication Control and Computing Technologies (2012)
4. Punia, R., Singh, S.: Review on machine learning techniques for automatic segmentation of liver images. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **3**(4), 666–670 (2013)
5. Okada, T., Shimada, R., Sato, Y., Hori, M., Yokota, K., Nakamoto, M., Chen, Y., Nakamura, H., Tamura, S.: Automated segmentation of the liver from 3D CT Images using probabilistic atlas and multi-level statistical shape model. Med. Image Comput. Comput.-Assist. Intervention **4791**, 86–93 (2007)
6. Slagmolen, P., Elen, A., Seghers, D., Loeckx, D., Maes, F.: Haustermans K.: Atlas based liver segmentation using nonrigid registration with a B-spline transformation model. Proceedings of MICCAI Workshop on 3D Segmentation In The Clinic: A Grand Challenge, pp. 197–206 (2007)
7. Linguraru, M., Sandberg, J., Li, Z., Pura, J., Summers, R.: Atlas-based automated segmentation of spleen and liver using adaptive enhancement estimation. Med. Image Comput. Comput.-Assist. Intervention **5762**, 1001–1008 (2009)

8. Wyawahare, M., Patil, P., Abhyankar, H.: Image registration techniques: an overview. Int. J. Signal Process. Image Process. Pattern Recogn. **2**(3), 11–28 (2009)
9. Kalinić, H.: Atlas-based image segmentation: a survey. Department of Electronic Systems and Information Processing, Universiy of Zagreb, pp. 1–7 (2008)
10. Rohlfing, T., Brandt, R., Menzel, R., Russakoff, D., Maurer, C.: Quo Vadis, Atlas-Based Segmentation? Handbook of Biomedical Image Analysis, pp. 435–486 (2005)
11. Kikinis, R., Shenton, M., Iosifescu, D., McCarley, R., Saiviroonporn, P., Hokama, H., Robatino, A., Metcalf, D., Wible, C., Portas, C., Donnino, R., Jolesz, F.: A digital brain atlas for surgical planning, model-driven segmentation, and teaching. IEEE Trans. Visual Comput. Graphics **2**(3), 232–241 (1996)
12. Staring, M., Klein, S., Pluim, J.: A rigidity penalty term for nonrigid registration. Med. Phys. **34**(11), 4098–4108 (2007)

Automatic 3D Segmentation of Renal Cysts in CT

Pawel Badura, Wojciech Wieclawek and Bartlomiej Pycinski

Abstract A fully automatic methodology for renal cysts detection and segmentation in abdominal computed tomography is presented in this paper. The segmentation workflow begins with the lungs segmentation followed by the kidneys extraction using marker controlled watershed algorithm. Detection of candidate cysts employs the artificial neural network classifier supplied by shape-related 3D object features. Anisotropic diffusion filtering and hybrid level set method are used at the fine segmentation stage. During the evaluation 23 out of 25 cysts delineated by an expert within 16 studies were detected correctly. The fine segmentation stage resulted in a 92.3 % sensitivity and 93.2 % Dice index combined over all detected cases.

Keywords Image segmentation · Renal cyst · Abdominal computed tomography · Artificial neural network · Level sets

1 Introduction

Medicine pays more attention to the image analysis techniques due to their applicability in computer-aided diagnosis and therapy (CAD). Multiple approaches are designed to support the diagnostic decision using various imaging modalities, like computed tomography (CT) [11, 24] or magnetic resonance imaging (MRI) [6, 26]. Large amount of data requires advanced processing techniques and sufficient amount of resources in terms of time and processing equipment. Decision yielded by CAD system has to be approved by the physician, yet a higher level of processing automation should be provided.

Renal cyst is a fluid reservoir inside the kidney. The most common type is a simple cortical cyst, which is benign, non-congenital and usually single. The presence of cortical cysts increases with the age, reaching 50 % at 50-year-old population [4].

P. Badura (✉) · W. Wieclawek · B. Pycinski
Faculty of Biomedical Engineering, Silesian University of Technology,
Roosevelta 40, Zabrze, Poland
e-mail: pawel.badura@polsl.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_13

Usually a treatment is not needed and the cysts are periodically monitored using medical imaging. If the diameter of cysts increase or another ones appear, the more complex diagnosis ought to be considered. Differential diagnosis of e.g. cystic carcinoma should be also performed in case of internal inhomogeneity, thicken blurred wall or presence of septa. The density of simple cortical cysts in CT studies is near pure water (10–20 Hounsfield Units, HU) and does not increase significantly after the contrast agent is administrated intravenously. The diameter of simple cysts may spread from about 5 mm to even 15 cm.

Even so CT imaging is widely used in renal cysts diagnosis, segmentation and classification methods of the kidney pathologies still require much effort. In [22] an interactive marker-controlled watershed algorithm is implemented for kidney segmentation. Renal lesions (defined as cysts and solid tumours) are identified by thresholding. Another interesting automatic supervised kidneys classification method assigns the kidney as pathological (i.e. containing at least one cyst) or non pathological [5]. Other semi-automatic approach uses a combination of fast marching and geodesic active contour level sets for cysts and tumour detection [13]. A semi-automatic segmentation algorithm from [3] enables measurements of simple kidney cysts from CT images. More recently, a fuzzy C-means clustering was used for renal cyst segmentation [20].

In this paper two major goals have been defined. First, a reliable cyst outlining procedure has been specified and implemented using 3DSlicer open source software [8]. The procedure consists of three stages: initial delineation on axial slices and possible corrections in sagittal and coronal projections. Secondly, a fully automated system has been designed for the cysts detection and segmentation. The system consists of a sequence of segmentation procedures: lungs segmentation, kidneys masks definition, detection and selection of candidate cyst objects, and fine segmentation of indicated cysts. Various 2D and 3D image processing techniques have been merged. A comprehensive analysis and evaluation of the methodology is possible using the reference database.

The paper is organized as follows. Materials and methods are presented in Sect. 2. That concerns the image database description, the expert delineation procedure specification and presentation of the entire automatic system for renal cysts detection and segmentation. The experimental results of consecutive segmentation stages are presented in Sect. 3 and discussed in Sect. 4. The paper is concluded in Sect. 5.

2 Materials and Methods

2.1 Materials

The clinical data (DICOM image study together with radiological reports) have been selected from a PACS (Picture Archiving and Communication System). For this research 16 patient studies with a total of 25 renal cysts have been chosen [12]. The

exams have been registered with a GE LightSpeed16 machine yielding mean image resolutions of $0.791 \times 0.791 \times 2.5$ mm.

All the studies include 4 series: the first one before the iodinated contrast agent injection and three others 20–25 s, 60–70 s and 10 min after the injection, respectively. Due to the renal cortex appearance in the image, in the current study only the second contrast-enhanced series is employed. The renal cortex intensity comes up high and homogeneous in this phase.

2.2 Expert Outlining

The contours of renal cysts are outlined on the images displayed in the abdominal intensity window (window level = 40 HU, window width = 350 HU). These values ensure high contrast between the cyst and the renal cortex, e.g. if the cyst mean value is equal to 10 HU and the cortex—160, the gray intensities are then equal to 41.4 % and 84.3 %, respectively. Common visualisation settings among all studies ensure the repeatability of classification of the borders, because with different setting, the cyst size may vary (Fig. 1). Moreover, the radiologists read CT studies mostly using the same intensity window.

The process of outlining the cysts is performed using the 3DSlicer open source software [8]. The image is zoomed so that the kidney covers over a half of the screen. Therefore each CT pixel is displayed on a box of dimension at least 8×8 pixels. Image pixels gray levels are linearly interpolated using four neighbouring voxels. The outlinings are drawn with a one-screen-pixel-width pencil and the image pixel is marked if at least 50 % of its surface is located inside the border.

The cysts are outlined in original axial planes, slice by slice. After that, the slices are reoriented into a sagittal projection with the results of previous marks displayed. The cysts borders are verified among all slices and corrected, if necessary. Finally, the same procedure is performed in coronal slices (Fig. 2). The outlining on all

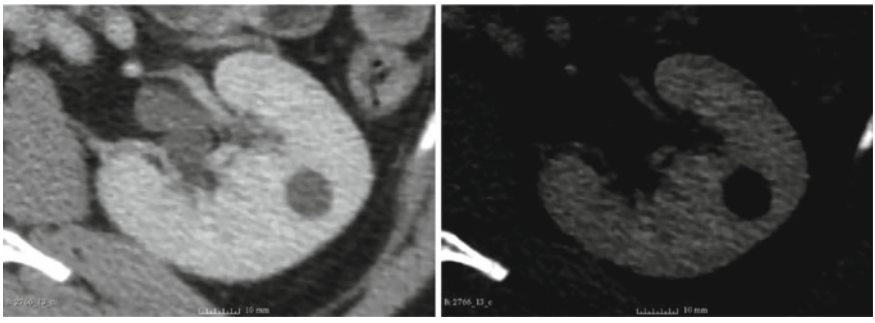


Fig. 1 Brightness of anatomical structures with respect to intensity window values (WL, WW). *Left* (40, 350) [HU], *right* (130, 110) [HU]

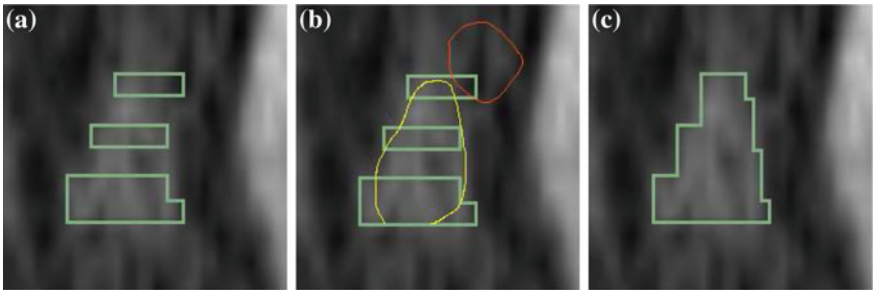


Fig. 2 Manual outlining of the cyst in coronal planes: axial masks reoriented to coronal planes (**a**), outline correction of the cyst (*yellow*) and background (*red*) (**b**), the final mask (**c**)

three orthogonal plane orientations is used in order to improve the delineation 3D reliability. If only axial slices are used, then the image would be wrongfully treated as the set of independent 2D slices. In the process of outlining cysts, a priori knowledge about their typical shape (mentioned earlier) is used.

2.3 Automatic Detection and Segmentation of Renal Cysts

The renal cyst detection and segmentation algorithm is fully automated and performed according to the scheme presented in Fig. 3. With each subsequent procedure the spatial and logical accuracy increases, from a rough lung regions location to precisely delineated renal cyst area.

Lungs Segmentation The lungs segmentation algorithm helps the kidneys to be located. It takes into consideration selected principles of the lungs appearance in abdominal CT. It consists of a series of automatic operations:

1. Otsu thresholding of the entire original volume [17] yielding a binary volume containing the lungs among other dark objects.
2. Axial slice-by-slice removal of 2D connected components attached to the slice borders, erasing the region outside the patient body.
3. Elimination of all 3D connected components not intersecting the uppermost axial slice. Since the abdominal CT studies employed during the examination do not cover the entire thorax, such a procedure deletes all the abdominal objects (intestines, stomach content etc.) while keeping the lung regions.

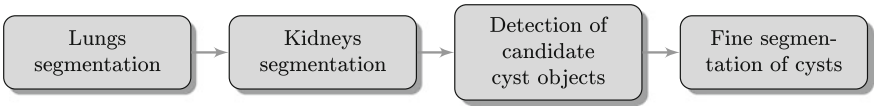


Fig. 3 The automatic detection and segmentation of renal cysts scheme

4. If there are more than two objects left at this stage, an additional condition is investigated: only at most two 3D connected components of an assumed minimum volume remain.
5. Finally, a set of morphological corrections is performed in order to remove holes within the lungs (including vasculature or potential nodules) as well as smooth their surfaces [21].

Kidneys Segmentation Segmentation of the right and left kidney is implemented separately, starting with the right one (Fig. 4). Two assumptions have been made. The first one refers to the kidney intensity, which normally (without contrast agent) equals about 30 HU for kidneys. Since the analysis concerns contrast-enhanced studies, a higher value range $[0, 200]$ HU has been assumed. The preprocessing step remains voxels with the intensity within this range, whilst all others assigns to zeros. The second condition is an anatomical constraint associated with geometric kidney location and significantly limits the search space: the kidneys are supposed to be located inside the corridor formed by the axial lungs shadow (Fig. 4, after preprocessing).

The next stage is a preliminary kidney segmentation consisting of two procedures. First, the sharpening procedure is performed employing the average spatial filtering and 3D grayscale geodesic morphological closing. Then, bright objects are detected using a 3D extended H-maxima transform [21]. The resulting binary volume is subjected to the mask-border cleaning (removing all objects attached to the red box

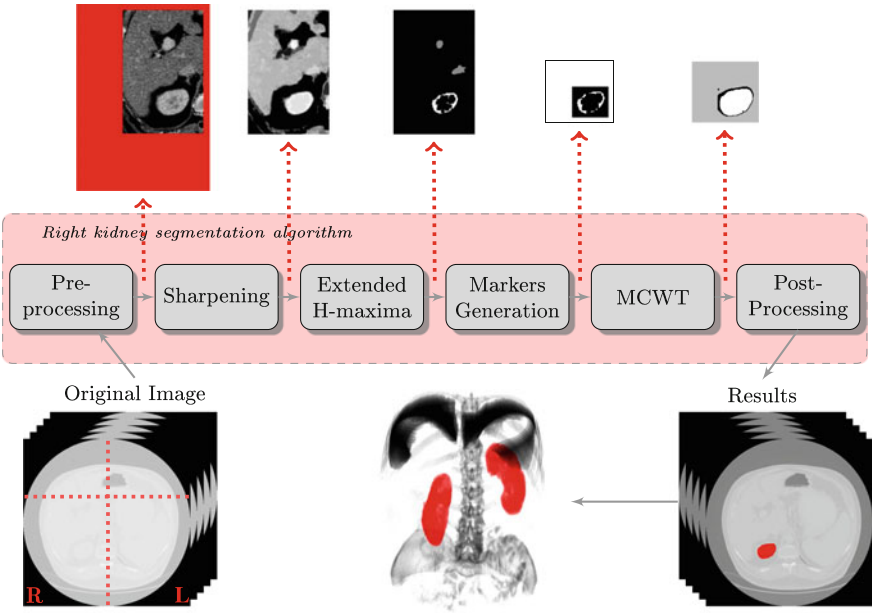


Fig. 4 The kidneys segmentation algorithm

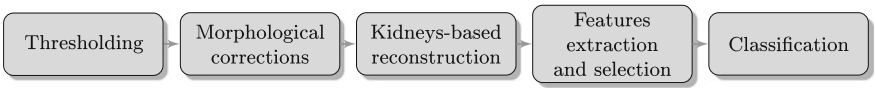


Fig. 5 The detection of candidate renal cyst objects scheme

in Fig. 4) and selection of the largest object left. This object is supposed to be the undersegmented right kidney (Fig. 4, white structure after the extended H-maxima procedure) and serves as a seed area for the next stage.

The final stage employs the 3D MCWT (3D marker controlled watershed algorithm) [15, 18]. Two types of markers are prepared: the background and the object markers [1]. The seed area from the previous stage is used as the object marker; the background marker is defined as the area outside the 3D bounding box of the undersegmented kidney (Fig. 4, after markers generation). The watershed dams are searched between the markers using the 3D gradient magnitude image. Since gradient computation methods are noise sensitive, the average spatial filtering and morphological opening are performed beforehand. The area between the markers covers the external kidney surface, being overlapped by the watershed dams with a high accuracy.

The left kidney segmentation is performed in an analogous manner. The only significant difference is the initial mask generation stage. The right kidney occurrence mask is obtained and mirrored to the left side. Then, all [0, 200] HU intensity objects intersecting the mask are identified. The following border-cleaning, morphological reconstruction and largest object selection procedures yield the undersegmented left kidney. It is then subjected to a MCWT algorithm.

Detection of Candidate Cyst Objects At the beginning of the automatic cyst detection (Fig. 5), the CT volume is thresholded within a predetermined range. The thresholds values rely on the statistical intensity analysis of both, kidneys and annotated cysts. Final values used for volume binarization have been set to $[-15, 60]$. This range covers most of the cyst volume and excludes surrounding tissues, including contrast-enhanced kidneys.

Since the resulting binary volume contains multiple false positive regions spread throughout the patient body, a sequence of correcting operations is employed. First, a morhological opening and closing occurs with a 3D spherical structuring element of a 2 mm radius,[1] followed by filling of 3D holes in objects. Then, a morphological reconstruction is performed using the kidneys mask. To avoid the influence of possible kidneys undersegmentations, the mask is dilated with a spherical structuring element of a 5 mm radius before being used for reconstruction.

The reconstruction yields many 3D candidate connected components of a different size and shape. Apart from the intensity characteristics, the renal cysts feature some size-independent shape properties. The most important one used here is their close-to-spherical character. Consequently, many candidate objects might be easily rejected

[1] In case a metric size is used, the structuring element depends on the CT voxel size.

based on some rules, however the direct definition of a discriminant hyperplane in the feature space is a challenging task. Therefore, an artificial neural network (ANN) [10] classifier has been implemented. In order to avoid overfitting and overtraining [23], the main assumption is to construct the network of a structure as simple as possible, supplied by only a few well matched features. Thus, the multilayer perceptron (MLP) has been used with 3 inputs, one output, 2 hidden layers and a total of 6 active neurons only (3-2-1 units in consecutive layers). The hyperbolic tangent activation function has been applied in hidden layers' neurons, whereas the output neuron has a linear form $\phi(x) = x$ saturated at 0 and 1. Such an output has been used since there are only two classes (0 or 1) defined for candidate objects. The MLP output is thresholded at 0.5 to produce the binary decision.

A set of 13 shape features has been defined for each analyzed 3D connected component. All of them take the metric voxel size into account while being computed. Note also, that all candidate objects that occur in just a single projection (slice) have been rejected before the features extraction and selection. The appropriate object features have been selected using a Fisher's linear discriminant (FLD) [9] at a confidence level $p = 0.05$. The reference set was constructed from all candidate objects available after reconstruction. Three features have been indicated and defined as the classifier inputs:

1. Sphericity S computed as a ratio of radii R^{in}_{max}, R^{out}_{min} of two spheres: the largest one inscribed in, and the smallest one circumscribed on the current object, respectively [2]:
$$S = \frac{R^{in}_{max}}{R^{out}_{min}}. \tag{1}$$
Its value reaches 1.0 for a perfect sphere, ca. 0.4-0.8 for a cube, cuboid or ellipsoid, and well below 0.3 in cases of irregular, thin structures.
2. Normalized object sphericity S_n:
$$S_n = \frac{R^{out}_{min} - R^{in}_{max}}{R_{eq}}, \tag{2}$$
with R_{eq} being the radius of a sphere of the volume equal to the object volume V:
$$R_{eq} = \sqrt[3]{\frac{3V}{4\pi}}. \tag{3}$$
The higher the S_n value, the less spherical the object under consideration.
3. The mean distance $\overline{DT}$ of the object voxels to its external surface computed as an average of the object distance transform [7]. $\overline{DT}$ is small for non-spherical objects.

The Levenberg-Marquardt backpropagation has been used for the classifier training [14].

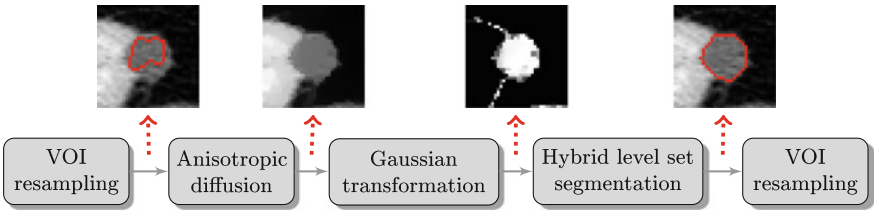


Fig. 6 The renal cyst fine segmentation scheme

Fine Segmentation of Cysts Once a candidate object passes the classification, it is subjected to a fine segmentation procedure. It consists of 5 steps (Fig. 6).

First, the volume of interest (VOI) is determined around the candidate object [25]: its bounding box extended by a margin od 10 mm in each projection is extracted from the entire study and then three-dimensionally resampled in order to obtain a cubic voxel shape of a 1 mm side. Then, the 3D anisotropic diffusion filtering is applied to smooth homogeneous regions while preserving the edges [2, 19].

The diffused volume is subjected to a Gaussian intensity transformation $T(m, \sigma)$:

$$I_T = e^{-\frac{(I-m)^2}{2\sigma^2}}, \tag{4}$$

where I denotes diffused voxel intensity and m, σ are transformation parameters. The mean intensity and standard deviation within a region of nondiffused volume indicated by the initial mask have been chosen as m and σ. The transformation yields a well separated region with the expected cyst intensity enhanced over both, lower and higher HU values (Fig. 6).

The final processing step is performed using a hybrid level set (LS) algorithm [27]. The LS segmentation is designed for the extraction of structures using the evolution of a deformable model ϕ [16]. Here, the evolution is caused by two stimuli: (1) the region component, resulting from the expansion/contraction caused by inner/outer regions and (2) the geodesic active contour component with two factors: the boundary-attraction and the curve smoothing control. The minimized ϕ functional has been defined as [27]:

$$\mathcal{E}(\phi) = -c_\alpha \int_\Omega (I_T - \mu) H(\phi) d\Omega + c_\beta \int_\Omega g \, |\nabla H(\phi)| \, d\Omega, \tag{5}$$

where: c_α and c_β weigh the influence of both components of (5), Ω denotes the spatial volume domain, I_T stands for the voxel intensity, $H(\phi)$ is a Heaviside step function of ϕ, g is a decreasing function called boundary feature map, related to the image gradient, and μ is a parameter indicating the expected lower intensity bound. The hybrid LS segmentation is performed over the transformed I_T volume with the initial cyst mask determining the zero level set. Since the expected cyst shape is rather spherical and the object to be extracted is homogeneous and well enough isolated

from the surroundings, the shape component has a significantly greater impact during the evolution of ϕ ($c_\beta \gg c_\alpha$).

The resulting binary volume is interpolated to the original VOI space and inserted into the full study.

3 Experimental Results

A set of 16 CT studies with 25 renal cysts have been subjected to the evaluation analysis. All cysts have been delineated by an expert. The accuracy of lung and kidneys segmentation, cysts detection and their fine segmentation have been evaluated independently.

3.1 Lungs and Lidneys Segmentation

The lungs segmentation algorithm has not missed any lung region from the employed database. The pleural walls have been properly reflected and no significant undersegmentations have been registered. In a single case a ca. 6.5 cm^3 leak occurred from the left lung into the intestine area due to a partial volume averaging causing a slice-to-slice pleura disappearance. The influence of this fact is however limited, since the segmented lung regions produce a corridor for kidneys segmentation (Fig. 4) and the leak region is still covered by the lung shadow.

The kidneys segmentation algorithm has been evaluated using the entire available data set. A single CT study contained a scan of a patient after the left kidney amputation. In all segmentation cases the kiedneys regions have been correctly identified. A total of two kidneys have been considered undersegmented, two other—oversegmented. Undersegmentation is caused by the large cysts appearance. The oversegmentation are related to a slight leak into the neighboring liver area. Neither of these cases cause any problem for the cyst detection and segmentation. Figure 7 presents 3D visualizations of sample lungs and kidneys regions against the skeleton.

3.2 Renal Cysts Detection

The renal cysts detection validation (Fig. 5) relies on traditional sensitivity and specificity measures:

$$Se_d = \frac{TP_d}{TP_d + FN_d} \cdot 100\,\%, \tag{6}$$

$$Sp_d = \frac{TN_d}{TN_d + FP_d} \cdot 100\,\%, \tag{7}$$

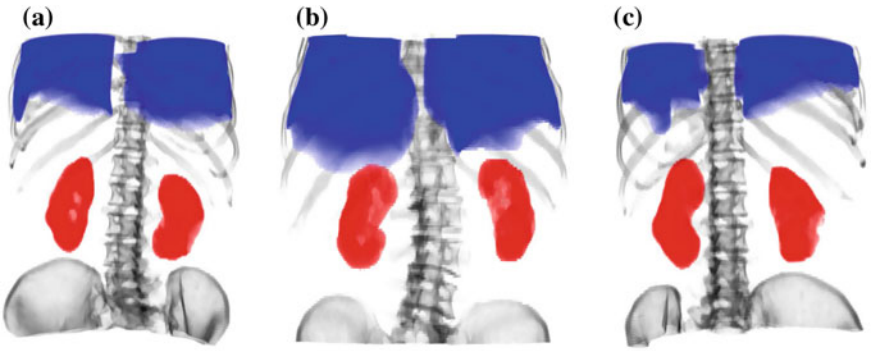


Fig. 7 Illustration of sample lungs (*blue*) and kidneys (*red*) segmentation results

where TP_d, FN_d, TN_d, FP_d denote the number of true positive, false negative, true negative and false positive cyst detections, respectively. In order to reflect the false positive rate, the number of false positives per study $fpps_d$ has also been determined. The k-fold cross validation scheme has been used with $k = 8$ in terms of the studies partition instead of the candidate cysts. With a total of 16 studies, in each of 8 verification procedures 14 studies have been used for training, with the remaining 2 passed as a testing set. Table 1 presents the detection statistics obtained over all 16 cases and 25 cysts annotated by an expert. Two cysts have been missed at the thresholding and morphological operations stage mainly due to their small size and relatively high intensity (Fig. 8). Single-slice true negative objects (668) have been removed before classification. Therefore, the classifier has been supplied by a total of 372 candidate objects, yielding only 5 misclassifications.

It has to be noted, however, that another experiment has been performed, with a single classifier training using the entire set of 372 candidate objects. This ended with

Table 1 Renal cysts detection efficiency

	TP_d	FP_d	TN_d	FN_d	Se_d (%)	Sp_d (%)	$fpps_d$
ANN classifier	20	2	347	3	87.0	99.4	0.12
Entire detection	20	2	1038	5	80.0	99.8	0.12

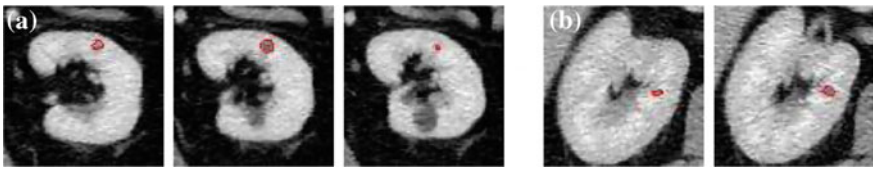


Fig. 8 Cysts missed during detection at the thresholding and morphology stages

a successful separation of the expected positive and negative indications in the three-dimensional feature space: a classifier has been found yielding 100 % sensitivity and specificity levels. Thus, the following verification of the fine segmentation was possible for all 23 cysts still present after the thresholding & morphology stages.

3.3 Renal Cysts Fine Segmentation

The fine segmentation evaluation employed the expert delineations of the true positive cyst detections. Validation relies on voxel-wise sensitivity and Dice index measures:

$$Se_s = \frac{TP_s}{TP_s + FN_s} \cdot 100\,\%, \tag{8}$$

$$D_s = \frac{2 \cdot TP_s}{2 \cdot TP_s + FP_s + FN_s} \cdot 100\,\%, \tag{9}$$

where TP_s, FN_s, FP_s denote the number of true positive, false negative, and false positive voxel detections with respect to the expert delineation, respectively. Due to the true negative voxels dominance over the other groups, the specificity has not been employed for evaluation. Columns 1–3 of Table 2 show the results averaged over a set of all 23 cases. The rightmost column presents the sensitivity and Dice index computed once for the totals of TP_s, FN_s, FP_s accumulated over all cases. The overall metrics reach higher values mainly due to a noticeable relation between

Table 2 Renal cysts fine segmentation efficiency

	Mean (%)	Worst case (%)	Best case (%)	Overall (%)
Se_s	84.2 ± 11.7	55.8	97.6	92.3
D_s	84.8 ± 9.0	70.4	95.4	93.2

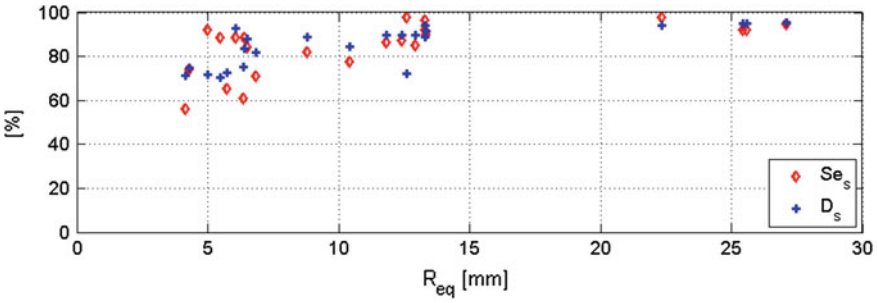


Fig. 9 Fine segmentation efficiency versus the cyst equivalent radius R_{eq}

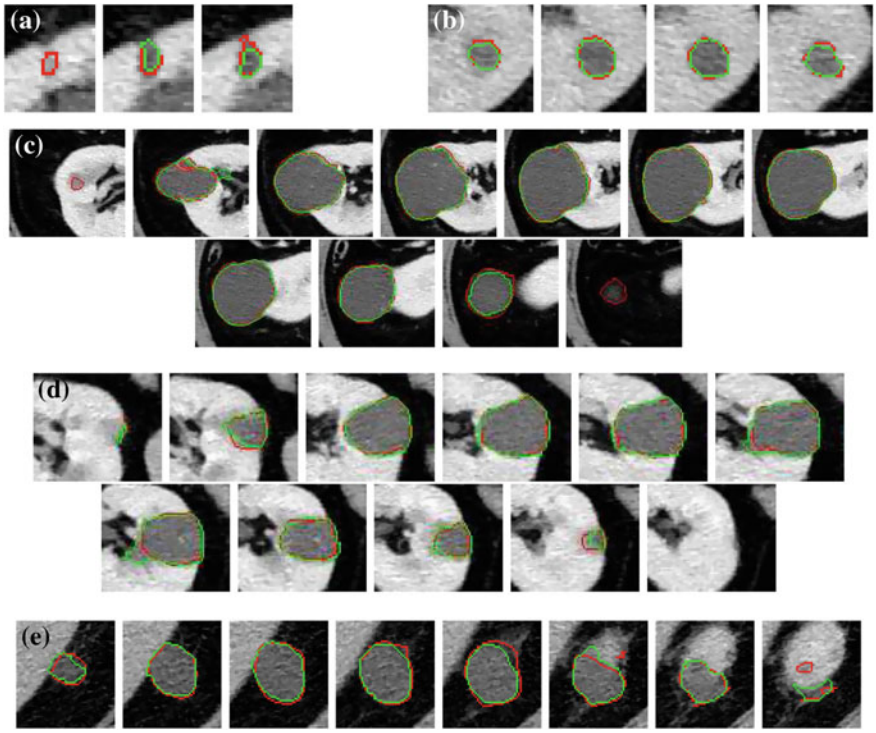


Fig. 10 Sample renal cysts segmentation results with the expert (*red*) and system (*green*) outlines. The worst case in terms of sensitivity Se_s is shown in (**a**). **a** $Se_s = 55.8\,\%$; $D_s = 71.1\,\%$. **b** $Se_s = 88.3\,\%$; $D_s = 92.6\,\%$. **c** $Se_s = 92.1\,\%$; $D_s = 94.8\,\%$. **d** $Se_s = 96.4\,\%$; $D_s = 94.2\,\%$. **e** $Se_s = 87.2\,\%$; $D_s = 89.8\,\%$

the segmentation efficiency and the cyst size. Small objects are susceptible to a higher level of missegmentation; the large ones are extracted more accurately (Fig. 9).

Several segmentation results are shown in Fig. 10.

4 Discussion

Methodology described in this paper enable a full automation of the detection of pathological structures of a certain type. Since its first stages: the lungs and kidneys segmentation do not require extremely high accuracy, eventual slight under- or oversegmentations are considered acceptable. Both procedures properly set up consecutive milestones on the logical path from the abdominal CT data to potential renal cysts locations.

In the experiments conducted over 16 CT studies each actual cyst has been covered by the rough kidneys VOI. The automatic detection encountered, however, some problems, mainly related to small cysts. On one hand, such cysts appear in a small number of slices and may vanish during procedures designed for the initial elimination of potential false positive objects of a few voxel size. On the other hand, such small cysts generally feature higher HU intensities, which pushes them out of the optimal range. Nonetheless, in this study such a problem concerns only 2 objects of a ca. 6 mm diameter. For all larger cysts some consistent relations exist between their intensity distribution and shape properties: it is possible to detect them as candidates and define mainly shape-related features able to discriminate the positive and negative matches. The heuristic ANN classifier of intensionally limited architecture produced well and repeatable responses during research. In one of the experiments a proper discrimination hyperplane has been found between the 23 actual cysts and hundreds of fake candidates.

The high level of a cyst sphericity assumed in the study implied the level set approach for fine segmentation. To make the segmentation more reliable, the image data have been preprocessed for a better cyst isolation from the background using nonlinear combination of anisotropic diffusion filtering and Gaussian intensity transformation. Both, AN filtering and level set segmentation are highly time consuming when applied on large data; here, the initial VOI cropping makes this problem negligible. The level set algorithm gets the well prepared data and finalizes the segmentation process. Possible local sphericity deviations are responsible for most of missegmentations and this process is the more apparent, the smaller the extracted object.

Full automation and high efficiency are the main advantages of the presented methodology. Possible future research paths might comply its generalization into processing of other phases of abdominal CT, including non-contrast studies.

5 Conclusion

Automatic 3D image segmentation methodology described in this paper offers high detection and segmentation efficiencies at each stage. Various data processing techniques at different levels of the analysis have been employed and verified using expert annotations over the clinical CT data.

Acknowledgments This research was supported by the Polish National Science Center (NCN) grant No. UMO-2012/05/B/ST7/02136.

References

1. Badura, P., Pietka, E.: Semi-automatic seed points selection in fuzzy connectedness approach to image segmentation. In: Computer Recognition Systems, Advances in Intelligent and Soft Computing, vol. 45(2), pp. 679–686 (2007)
2. Badura, P., Pietka, E.: 3D fuzzy liver tumor segmentation. In: Information Technologies in Biomedicine, Lecture Notes in Bioinformatics, vol. 7339, pp. 47–57 (2012)
3. Battiato, S., Farinella, G.M., Gallo, G., Garretto, O., Privitera, C.: Objective analysis of simple kidney cysts from CT images. In: 2009 IEEE International Workshop on Medical Measurements and Applications, pp. 146–149 (2009)
4. Bosniak, M.A.: The current radiological approach to renal cysts. Radiology **158**(1), 1–10 (1986)
5. Boukerroui, D., Touhami, W., Cocquerez, J.P.: Automatic regions of interest identification and classification in CT images: application to kidney cysts. In: IEEE First Workshops on Image Processing Theory. Tools and Applications, pp. 1–8 (2008)
6. Bugdol, M., Czajkowska, J., Pietka, E.: A novel model-based approach to left ventricle segmentation. In: Computers in Cardiology Series, vol. 39, pp. 561–564 (2012)
7. Danielsson, P.E.: Comput. Vis. Graph **14**(3), 227–248 (1980)
8. Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., Buatti, J., Aylward, S., Miller, J.V., Pieper, S., Kikinis, R.: 3D slicer as an image computing platform for the quantitative imaging network. Magn. Reson. Imaging **30**(9), 1323–1341 (2012)
9. Fisher, R.A.: The use of multiple measurements in taxonomic problems. Ann. Eugen. **7**(2), 179–188 (1936)
10. Haykin, S.: Neural Networks: A Comprehensive Foundation, 2nd edn. Prentice Hall PTR, USA (1998)
11. Juszczyk, J., Pietka, E., Pyciński, B.: Granular computing in model based abdominal organs detection. Comp. Med. Imag. Graph. **46**, Part 2, 121–130 (2015)
12. Kawa, J., Juszczyk, J., Pycinski, B., Badura, P., Pietka, E.: Radiological atlas for patient specific model generation. In: Information Technologies in Biomedicine, Advances in Intelligent Systems and Computing, vol. 284(4), 69–82 (2014)
13. Linguraru, M.G., Yao, J., Gautam, R., Peterson, J., Li, Z., Linehan, W.M., Summers, R.M.: Renal tumor quantification and classification in contrast-enhanced abdominal CT. Pattern Recognit. **42**(6) (2009)
14. Marquardt, D.W.: An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Ind. Appl. Math. **11**(2), 431–441 (1963)
15. Mittal, U., Anand, S.: Modified watershed segmentation with denoising of medical images. Int. J. Innov. Res. Sci. Eng. Technol. **2**(4), 982–987 (2013)
16. Osher, S., Sethian, J.A.: Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. J. Comput. Phys. **79**(1), 12–49 (1988)
17. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. **9**(1), 62–66 (1979)
18. Parvati, K., Prakasa Rao, B.S., Das, M.M.: Image segmentation using gray-scale morphology and marker-controlled watershed transformation. Discret. Dyn. Nat. Soc. **2008**, 1–8 (2008)
19. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. **12**(7), 629–639 (1990)
20. Piao, N., Kim, J.G., Park, R.H.: Segmentation of cysts in kidney and 3-D volume calculation from CT images. Int. J. Comput. Graph. Animation **5**(1), 1–16 (2015)
21. Soille, P.: Morphological Image Analysis: Principles and Applications. Springer (1999)
22. Summers, R., Agcaoili, C., McAuliffe, M., Dalal, S., Yim, P., Choyke, P., Walther, M., Linehan, W.: Helical CT of von Hippel-Lindau: semi-automated segmentation of renal lesions. In: Proceedings of 2001 IEEE International Conference on Image Processing, vol. 2, pp. 293–296 (2001)
23. Tetko, I.V., Livingstone, D.J., Luik, A.I.: Neural network studies. 1. comparison of overfitting and overtraining. J. Chem. Inf. Comp. Sci. **35**(5), 826–833 (1995)

24. Wieclawek, W., Pietka, E.: Fuzzy clustering in segmentation of abdominal structures based on CT studies. In: Information Technologies in Biomedicine, Advances in Intelligent and Soft Computing, vol. 47, pp. 93–104 (2008)
25. Zarychta, P.: Posterior cruciate ligament—3D visualization. In: Computer Recognition Systems, Advances in Intelligent and Soft Computing, vol. 45, pp. 695–702 (2007)
26. Zarychta, P.: Cruciate ligaments of the knee joint in the computer analysis. In: Information Technologies in Biomedicine, Advances in Intelligent Systems and Computing, vol. 283(3), pp. 71–80 (2014)
27. Zhang, Y., Matuszewski, B., Shark, L., Moore, C.: Medical image segmentation using new hybrid level-set method. In: 2008 Fifth International Conference BioMedical Visualization: Information Visualization in Medical and Biomedical Informatics, pp. 71–76 (2008)

The Importance of the Features of the Posterior Cruciate Ligament in Diagnosis

Anna Zarychta-Bargieła and Piotr Zarychta

Abstract The main aim of this research was finding the proper features of the posterior cruciate ligaments helpful in diagnosis. These features were determined based on the extracted structures of the posterior cruciate ligament and based on the T1-weighted MRI images of the knee joint. On the basis of these features a feature vector has been built. This feature vector has been used in computer diagnosis of the posterior cruciate ligaments.

Keywords Feature vector of the posterior cruciate ligaments · Cruciate ligaments · Fuzzy c-means · Fuzzy connectedness

1 Introduction

The knee joint is the largest, the most complex and the most complicated joint of the human body. This anatomical structure is sensitive, fragile and unfortunately frequently susceptible to injuries [1]. Inside this joint two elements being particularly vulnerable to injuries. They are the cruciate ligaments: anterior cruciate ligaments (ACL) and posterior cruciate ligaments (PCL) [2].

The anterior and posterior cruciate ligaments (Fig. 1) fulfill many functions in the knee joint. The elementary function of both ACLs and PCLs is a passive stabilization of the knee joint in three planes: sagittal, coronal and transverse. ACLs and PCLs ensure proper kinematics of the knee joint. They take care of the smooth movement of the knee joint and protect the articular cartilage. As the knee moves, two activities

A. Zarychta-Bargieła
Hospital in Dąbrowa Górnicza, ul. Szpitalna 13, 41-300 Dąbrowa Górnicza, Poland

P. Zarychta (✉)
Faculty of Biomedical Engineering, Silesian University of Technology,
Roosevelta 40, Zabrze, Poland
e-mail: piotr.zarychta@polsl.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_14

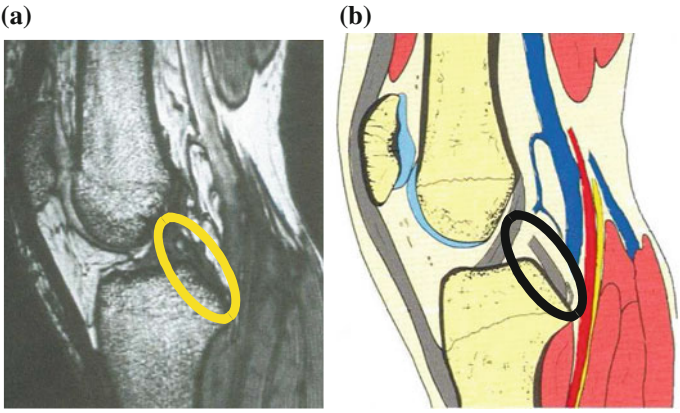


Fig. 1 Posterior cruciate ligament **a** T1-weighted MRI sequence, **b** illustration [5]

can be distinguished. The first one is called active motion and during this motion of flexion, extension and rotation ACLs and PCLs resist translations within the joint caused by the action of muscles and reduce shear forces. The second one is called passive motion and during this motion ACLs and PCLs help to change rolling into sliding movements. Preliminary ACL tension at the movement of flexion initiates tension of PCL and vice versa. Cruciate ligaments always remain partially strained as a result of non-uniform shape and unequal length of their fibers [3].

Nowadays the primary method for diagnostics of the knee joint especially cruciate ligaments injuries is magnetic resonance imaging (MRI). This is a completely non-invasive method. MRI gives the possibility to assess the shape and internal structure and allows the visualization of anterior and posterior cruciate ligaments throughout their length and in the following planes: sagittal, coronal and transverse. Usually MRI of the knee joint is performed on T1-weighted sequences (the anatomy of the muscles, ligaments, tendons is well visible) and various T2-weighted sequences (very sensitive to damage, swelling, inflammation of soft tissues). Typically the posterior cruciate ligament (Fig. 1a) is clearly visible (oval, thick structure, characterized by uniform black color) in the sagittal plane on 4–6 slices of the T1- or T2-weighted MRI sequences of the knee joint. In healthy case, PCL has a slightly arcuate shape and smoothly goes from attachment to the femur to attachment to the tibia [1].

In case of PCLs being damaged, the structure of ligament may take different forms, but often the fibers of the ligament tear apart over its entire length giving the effect of ligament swelling. For this reason, the swelling of the ligament is the fundamental criterion in the diagnosis of posterior cruciate ligaments injuries, and the primary diagnostic method is MRI. Therefore, in this paper the computer analysis of PCLs is performed on the example of MRI of the knee joint [1, 2, 4].

In medical practises the success of the treatment of damaged ligaments depends on many factors, primarily on accurate diagnosis. The accurate diagnosis, especially in complex cases is based on a combination of information from multiple sources.

In experts' (radiologists, orthopedists) opinion, particularly important information in addition to MRI of the knee joint is provided by: 3D structure of the cruciate ligaments, 3D visualization of bony structures in the knee joint and especially the data flowing from a properly constructed feature vector. In experts' opinion and based on the preliminary studies, it can be prove, that properly constructed feature vector is able to contribute to improve this diagnostic procedure.

2 Methodology

In this study have been implemented following procedures:

- the procedure finding 2-dimensional region of interest (2D ROI) including PCLs—this procedure has been based on the entropy measures of fuzziness and fuzzy c-means (FCM) algorithm with median modification. The main aim this procedure is reduction of the computational complexity and to increase of the efficiency an automatical procedure of the extracting PCLs structures. It is worth to remember, that the analysis of the T1-weighted MRI slices in a sagittal series (the resolution of the T1-weighted MRI slices of the knee joint is usually 256×256 or 512×512 pixels) and after this procedure the processed area has been reduced meanly three-five times (the average length of the PCL 2D ROI is about 100×100 pixels);
- the procedure finding 3-dimensional region of interest (3D ROI) including PCLs—this procedure is obtained by mapping the 2D ROI on all of the T1-weighted MRI slices of the knee joint.
- the procedure of the extraction of the PCLs structure from T1-weighted MRI series)—this procedure is based on a fuzzy connectedness concept. The generalized definition of fuzzy connectedness [6–8] is based on the fuzzy affinity relation and introduces an iterative method permitting fuzzy connectedness to be determined with respect to a chosen image pixel—seed point. In order to increase the efficiency of computational procedures the fuzzy connectedness method has been limited to 3D ROI [9]. In this point, the Live-Wire approach has been used as a reference method in the segmentation process. An exhaustive description of this approach can be found in [10, 11].

In this article above presented procedures are not widely described. The exhaustive description can be found in [12]. This article discusses the selection of the feature vector and their impact on the diagnostics of the posterior cruciate ligaments lesions. Preliminary results of a pilot study dedicated to the selection of the feature vector and their impact on the diagnostics of the PCLs lesions [13, 14] received positive opinions from the reviewers encouraging the author of this paper to present a broader description of this methodology.

3 Discussion

This article discusses the selection of the features and their impact on the diagnostics of the posterior cruciate ligaments lesions. In experts' opinion, the most important criteria in the cruciate ligaments diagnostics are as follows:

- the surface area of posterior cruciate ligament structures (PCLa);
- the perimeter of posterior cruciate ligament structures (PCLp);
- the skeleton of the extracted PCL structures (A-length and B-length);
- the intercondylar distance of the head of the tibia bone (ICD_{TIBIA});
- the intercondylar distance of the head of the thigh bone(ICD_{FEMUR}).

The information about value of the surface area of the PCL structures in the direct form is useless. This information should be independent of size of the knee joint. For diagnosis, useful is ratio of the surface area of the extracted ligament to the surface area of the 2D ROI including this ligament (PCLa/ROIa ratio). Such ratio makes the feature independent of the patient's size as the region of interest is dependent on the anatomy of the knee joint of the patient (Fig. 2).

Quite similar like in case of surface area, the information about value of the perimeter of the PCL structures in the direct form is useless. From a practical point of view, useful is the information, which is independent of size of the knee joint. Therefore for diagnosis, useful is ratio of the perimeter of the extracted ligament to the perimeter of the 2D ROI including this ligament (PCLp/ROIp ratio). Such ratio makes the feature independent of the patient's size, like in case of surface area, as the region of interest is dependent on the anatomy of the knee joint of the patient (Fig. 3).

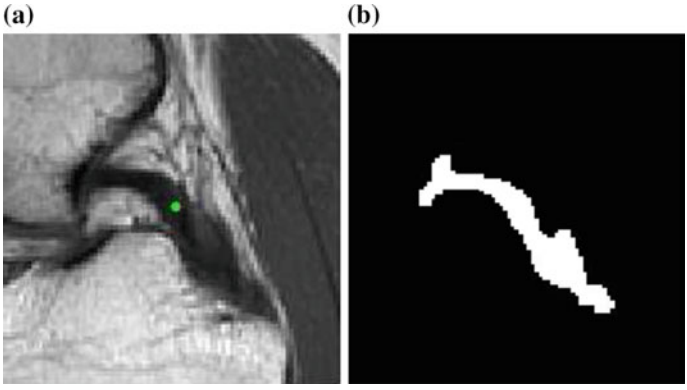


Fig. 2 2D ROI including the PCL structure (**a**) with marked surface area of the extracted PCL structure (**b**)

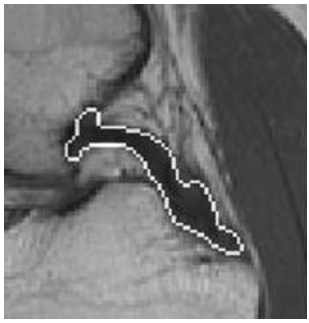


Fig. 3 2D ROI including the PCL structure with marked perimeter of the extracted PCL structure

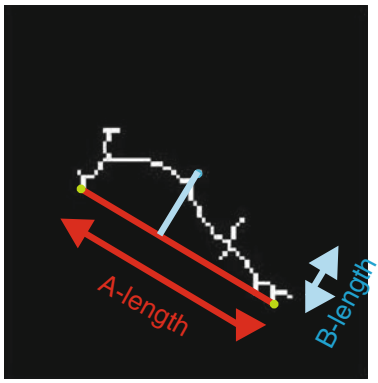


Fig. 4 The skeleton of the PCL with marked A- and B-lengths

In experts' opinion, the skeleton of the extracted PCL structure provides very important information about this structure. In computer aided diagnosis of posterior cruciate ligaments two lengths play very important role. They are: A-length—the maximum distances between a straight line combining the start and end point of the skeleton (Fig. 4) and B-length—its top and bottom edges (Fig. 4). However, as in the case of the surface area and perimeter of the extracted structures, it is not possible to use these values directly for diagnostic purposes. In medical practise the B/A ratio is ideal.

The next important criterion in the cruciate ligaments diagnostics is based on the intercondylar distance of the head of the tibia bone (ICD_{TIBIA}). Quite similar like in case of surface area and perimeter, the information about value of the intercondylar distance may not be used directly for diagnostic purposes. However, the B/ICD_{TIBIA} ratio is very good diagnostic tool, since it takes into account both the shape and size of the patient's knee joint, as well as the structure of the cruciate ligament.

The same situation is in case of the intercondylar distance of the head of the thigh bone (ICD_{FEMUR})—the direct information about value of the intercondylar distance is useless. The situation is changing in combination with the B-length. The B/ICD_{FEMUR} ratio is very good diagnostic tool. This ratio takes into account both the shape and size of the patient's knee joint, as well as the structure of the cruciate ligament.

4 Results

The research has been tested on 97 clinical T1-weighted MRI studies of the knee joint. This group consists of 61 healthy and 36 pathological cases of the PCLs (Fig. 5a). The selection of the features and their impact on the diagnostics of the posterior cruciate ligaments lesions have been done for 86 cases, for which properly the PCL structures have been extracted. The evaluation has been performed by three independent experts (two radiologists and one orthopedist), and it was based on a comparison of the original T1-weighted MRI slices of the knee joint with the segmented PCLs structures.

In experts' opinion, a very important criterion in the cruciate ligaments diagnostics is the surface area of cruciate ligament structures (Fig. 5b). For diagnosis, useful is

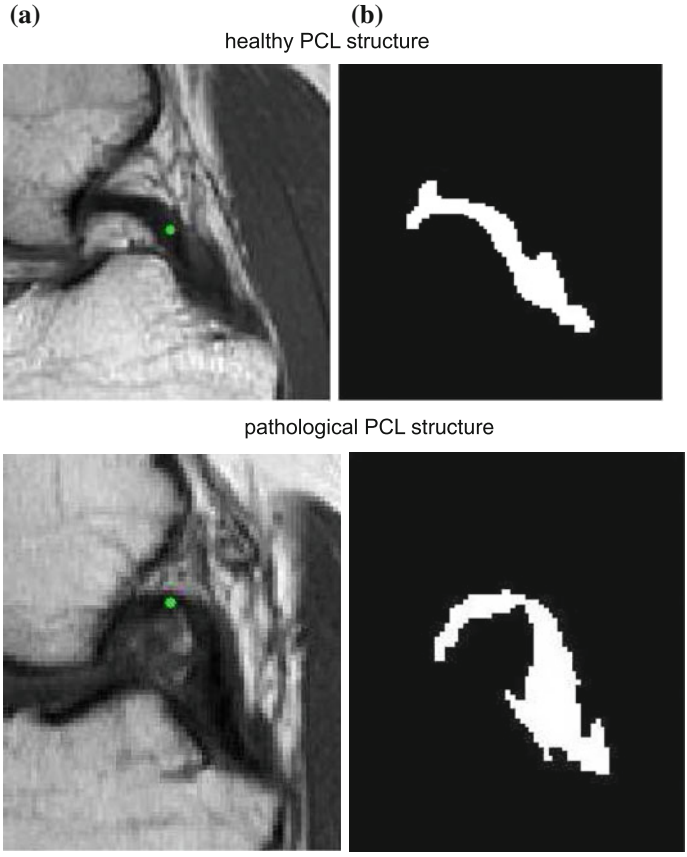


Fig. 5 Healthy and pathological structure of the posterior cruciate ligament **a** 2D ROI including the PCL structure, **b** surface area of the extracted PCL structure

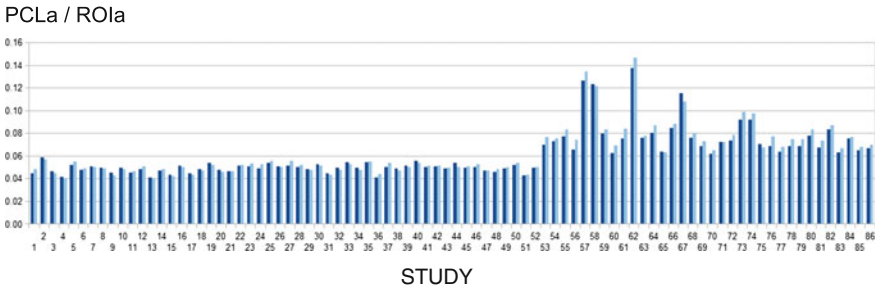


Fig. 6 PCLa/ROIa ratio for all 86 PCL studies where surface areas of the extracted anatomical structures have been marked by the expert (*light bars*) and automatically determined (*dark bars*)

ratio of the surface area of the extracted ligament to the surface area of the 2D ROI including this ligament (PCLa/ROIa ratio).

A comparison of PCLa/ROIa ratio for healthy and pathological structures of the posterior cruciate ligaments (for corresponding slices—the largest surface area of the extracted PCL in each case) on the basis of 86 clinical T1-weighted MRI studies of the knee joint, gives the results as follows. For healthy PCLs structures this ratio is lower or equal than 0.059. For pathological PCLs structures PCLa/ROIa ratio is greater than 0.059.

The PCLa/ROIa ratio for all 86 PCLs studies is shown in Fig. 6. On the basis of this ratio it is possible to determine both cases: healthy and pathological, respectively. In the group of healthy PCLs structures are following cases: 1–52. In the group of pathological PCLs structures are the remaining cases: 53–86.

Figure 6 also presents the quantification of the results. The results of the PCLa/ROIa ratio obtained for two specific cases have been compared. In the first case, the PCLa/ROIa ratio has been determined automatically, whereas in the second case it has been based on the contour made by an expert. The results are coincident. They allow for the indication of healthy and pathological structures of the extracted PCLs. The differences in the values of the PCLa/ROIa ratio are due to the lack of possibility of precise and unequivocal description of the segmented PCLs structures.

The second important criterion in the PCLs diagnostics is the perimeter of these structures (Fig. 7b). For diagnosis, useful is ratio of the perimeter of the extracted ligament to the perimeter of the 2D ROI including this ligament (PCLp/ROIp ratio).

A comparison of PCLp/ROIp ratio for healthy and pathological structures of the PCLs (for corresponding slices—the largest surface area of the extracted PCL in each case) on the basis of 86 clinical T1-weighted MRI studies of the knee joint, gives the results as follows. For healthy PCLs structures this ratio is lower or equal than 0.41. For pathological PCLs structures PCLp/ROIp ratio is greater than 0.41.

The PCLp/ROIp ratio for all 86 PCLs studies is shown in Fig. 8. As in the case of the surface area, on the basis of this ratio it is possible to determine both cases: healthy and pathological, respectively. In the group of healthy PCLs structures are following cases: 1–52 (the same structures as for the PCLa/ROIa ratio). In the group

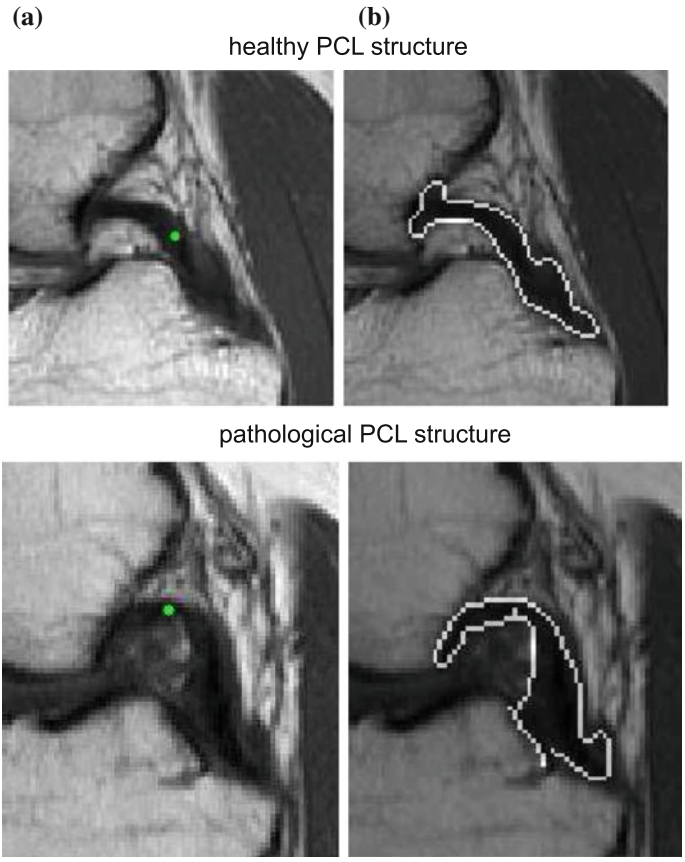


Fig. 7 Healthy and pathological structure of the posterior cruciate ligament **a** 2D ROI including the PCL stucture, **b** perimeter of the extracted PCL structure

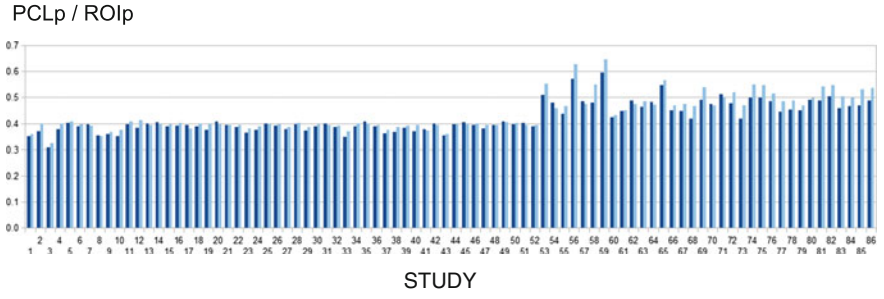


Fig. 8 PCLp/ROIp ratio for all 86 PCL studies where perimeter of the extracted anatomical structures have been marked by the expert (*light bars*) and automatically determined (*dark bars*)

of pathological PCLs structures are the remaining cases: 53–86 (the same structures as for the PCLa/ROIa ratio).

Figure 8 also shows the PCLp/ROIp ratio in two cases. In the first case, the PCLp/ROIp ratio has been determined automatically, whereas in the second case it has been based on the contour made by an expert. These results have been compared, and found that, they are coincident. They allow for the indication of healthy and pathological structures of the extracted PCLs and the differences in the values of the PCLp/ROIp ratio are due to the lack of possibility of precise and unequivocal description of the segmented PCLs structures (quite similar as in the case of the surface area).

The experts prove that the most important criterion in the PCLs diagnostics is the shape of these structures (Fig. 9b). A good tool to evaluate the shape of the ligament and for diagnosis, is the skeleton of the extracted PCL structures (B/A ratio).

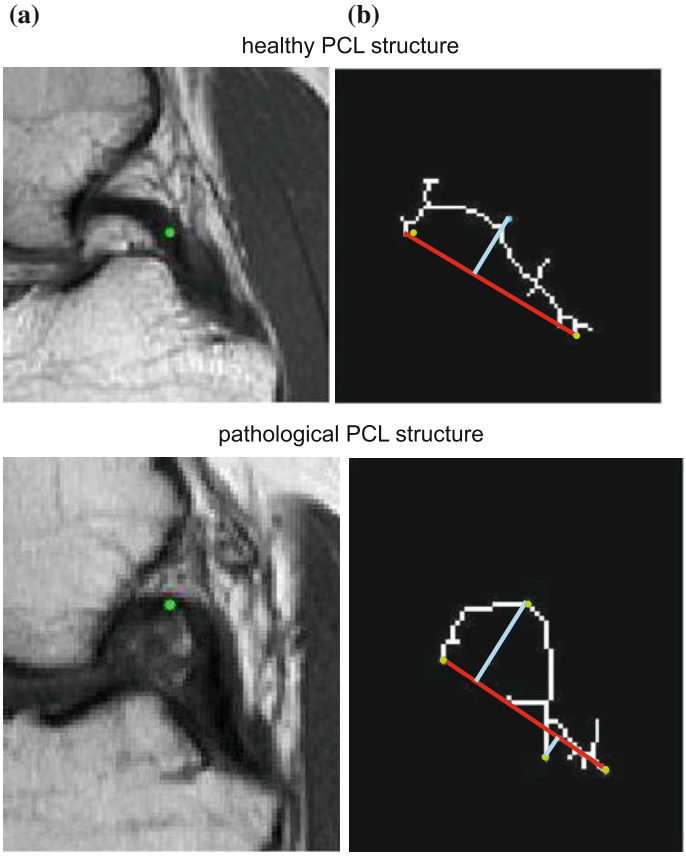


Fig. 9 Healthy and pathological structure of the posterior cruciate ligament **a** 2D ROI including the PCL structure, **b** skeleton of the extracted PCL structure

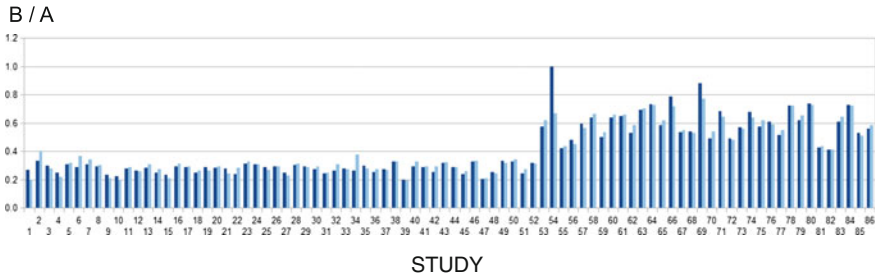


Fig. 10 B/A ratio for all 86 PCL studies where B- and A-lengths have been marked by the expert (*light bars*) and calculated on the basis of skeleton (*dark bars*)

A comparison of B/A ratio for healthy and pathological structures of the PCLs (for corresponding slices—the largest surface area of the extracted PCL in each case) on the basis of 86 clinical T1-weighted MRI studies of the knee joint, gives the results as follows. For healthy PCLs structures this ratio is lower or equal than 0.4 and for pathological PCLs structures is greater than 0.4.

The B/A ratio for all 86 PCLs studies is presented in Fig. 10. As in the previous cases, on the basis of this ratio it is possible to determine both cases: healthy and pathological, respectively. In the group of healthy PCLs structures are following cases: 1–52 (the same PCLs as previously). In the group of pathological PCLs structures are the remaining cases (the same PCLs as previously).

Figure 10 also shows the quantification of the results obtained for the B/A ratio in two cases. In the first case, the B/A ratio has been determined on the basis of the skeleton and points, which have been calculated automatically, and in the second case on the basis of the selection made by the expert. These results have been compared, and found that, they are coincident. Quite similar as in the previous cases, they allow for the indication of healthy and pathological structures of the extracted PCLs and the differences in the values of the B/A ratio are due to the lack of possibility of precise and unequivocal description of the segmented PCLs structures.

In experts' opinion it is also possible to use a ratio that includes the intercondylar distance of the head of the tibia bone—ICD_{TIBIA} (Fig. 11b) or the intercondylar distance of the head of the thigh bone—ICD_{FEMUR}. For diagnosis, useful are both ratio: the B/ICD_{TIBIA} and B/ICD_{FEMUR}.

A comparison of both ratios: B/ICD_{TIBIA} and B/ICD_{FEMUR}, respectively for healthy and pathological structures of the posterior cruciate ligaments (for corresponding slices—the largest surface area of the extracted PCL in each case) on the basis of 86 clinical T1-weighted MRI studies of the knee joint, gives the following results. For healthy PCLs structures B/ICD_{TIBIA} ratio is lower or equal than 0.24 and for pathological PCLs structures this ratio is greater than 0.24. For healthy PCLs structures B/ICD_{FEMUR} ratio is lower or equal than 0.33 and for pathological PCLs structures this ratio is greater than 0.33.

The B/ICD_{TIBIA} ratio for all 86 PCLs studies is shown in Fig. 12 and B/ICD_{FEMUR} ratio for the same studies is shown in Fig. 13. As in the previous cases, on the

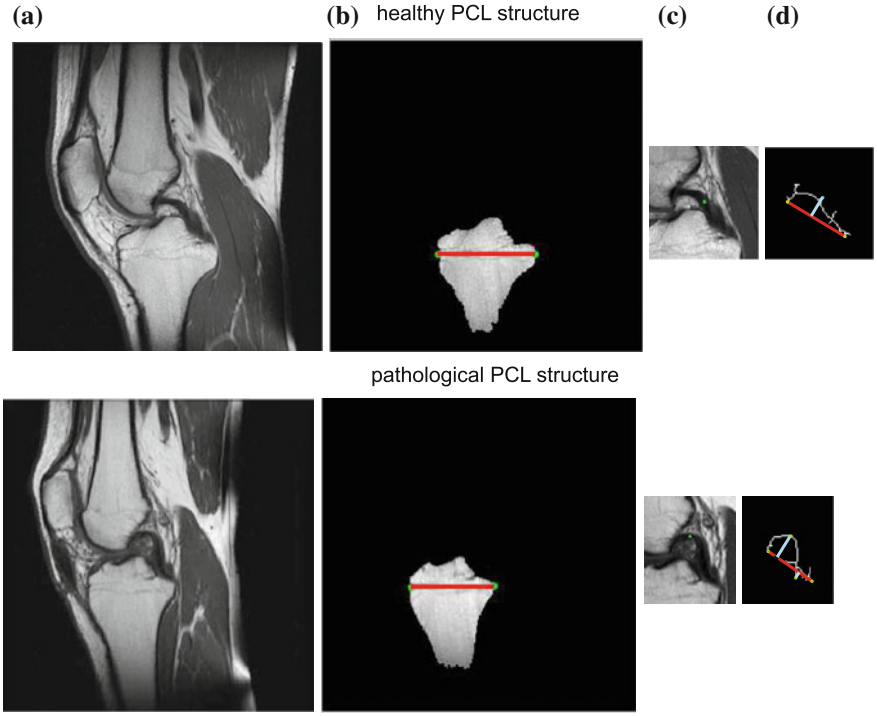


Fig. 11 Healthy and pathological structure of the posterior cruciate ligament **a** selected slice of the T1-weighted MRI series, **b** extracted tibia bone with marked intercondylar distance, **c** 2D ROI including the PCL structure, **d** skeleton of the extracted PCL structure

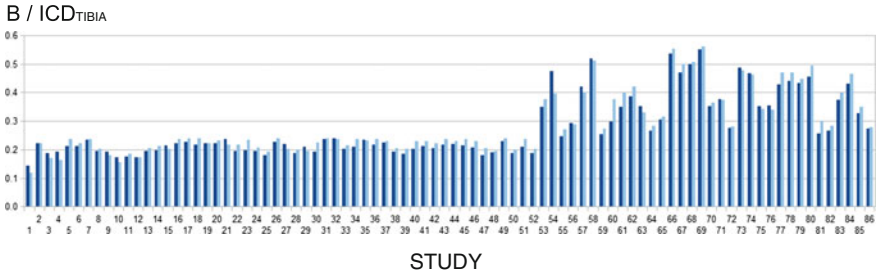


Fig. 12 B/ICD$_{TIBIA}$ ratio for all 86 PCL studies where B- and ICD$_{TIBIA}$-lengths have been marked by an expert (*light bars*) and automatically calculated (*dark bars*)

basis of these ratios it is possible to determine both cases: healthy and pathological, respectively. In the group of healthy PCLs structures are following cases: 1–52, and in the group of pathological PCLs structures are the remaining cases (the same structures as previously).

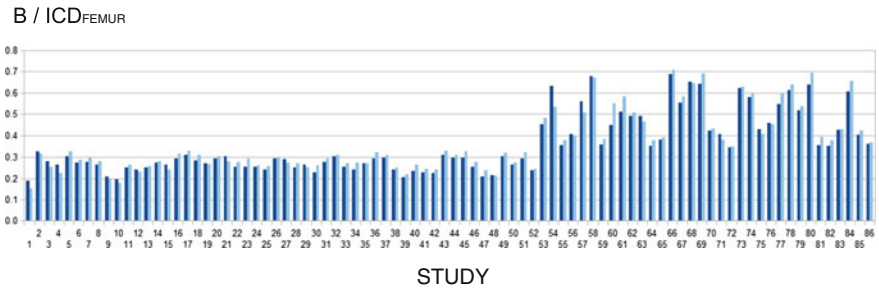


Fig. 13 B/ICD$_{FEMUR}$ ratio for all 86 PCL studies where B- and ICD$_{FEMUR}$-lengths have been marked by an expert (*light bars*) and automatically calculated (*dark bars*)

Figures 12 and 13 also show the quantification of the results obtained for both ratios in two cases. In the first case, the ICD$_{TIBIA}$ and ICD$_{FEMUR}$ values have been determined on the basis of the extracted heads of the tibia and femur, respectively, and points (the left and right condyle edges), which have been calculated automatically, and in the second case on the basis of a selection made by an expert. These results have been compared, and found that, they are coincident. Quite similar as in the previous cases, they allow for the indication of healthy and pathological structures of the extracted PCLs and the differences in the values of the B/ICD$_{TIBIA}$ ratio and B/ICD$_{FEMUR}$ ratio are due to the same cause as before.

Although both the B/ICD$_{TIBIA}$ and B/ICD$_{FEMUR}$ ratios are well suited for practical tasks, nevertheless, the fundamental disadvantage of using them for the posterior cruciate ligaments diagnostics is the need to perform calculations on the entire (256×256 pixels or 512×512 pixels) T1-weighted MRI slice of the knee joint. Therefore, a good solution in this case is a modification of the two expressions. This modification consists in replacing the intercondylar distance (ICD$_{TIBIA}$ and ICD$_{FEMUR}$) by the width of the 2D ROI.

5 Conclusions

The described concept is an innovative approach, which consists in using medical images processing in order to create feature vector. In this case, on the basis of MRI of the knee joint the feature vector of the posterior cruciate ligament has been created. According to experts, all of the above features (PCLa/ROIa ratio, PCLp/ROIp ratio, B/A ratio, B/ICD$_{TIBIA}$ and B/ICD$_{FEMUR}$) are helpful in diagnosis of the posterior cruciate ligament. The use of the feature vector of the posterior cruciate ligament ensures a proper and faster diagnosis of those anatomical structures. The current results seem to be quite good, given the fact that this is a prototype solution, which is constantly tested. The future work should focus on optimizing and simplifying the methods of images processing and the extension of the feature vector to other elements, which may improve the diagnostic process.

References

1. Bochenek, A., Reicher, M.: The Human Anatomy. PZWL, Warsaw (1990)
2. Dziak, A.: Uszkodzenia więzadeł krzyżowych kolana (Injuries of the cruciate ligaments of the knee joint). Acta Clinica **4**(1), 271–274 (2001) (in Polish)
3. Pasierbiński, A., Jarząbek, A.: Biomechanika więzadeł krzyżowych (Biobiomechanics of the cruciate ligaments). Acta Clinica **4**(1), 284–293 (2001) (in Polish)
4. Voos, J.E., Mauro, C.S., Wente, T., Warren, R.E., Wickiewicz, T.L.: Posterior cruciate ligament. Anatomy, biomechanics and outcomes. Am. J. Sports Med. **40**(1), 222–231 (2012)
5. Moeller, T.B., Reif, E.: Pocket Atlas of Sectional Anatomy CT and MRI. Head, Neck, Spine, and Joints, vol. 1. Thieme, Stuttgart (2000)
6. Badura, P., Pietka, E.: Semi-automatic seed points selection in fuzzy connectedness approach to image segmentation. In: Kurzynski, M. et al. (eds) Conference on Computer Recognition Systems. Advances in Intelligent and Soft Computing, vol. 45, pp. 679–686. Springer, Heidelberg (2007)
7. Zarychta, P., Konik, H., Zarychta-Bargieła, A.: Computer assisted location of the lower limb mechanical axis. In: Pietka, E., Kawa, J. (eds.) Information Technologies in Biomedicine. Lecture Notes in Bioinformatics, vol. 7339, pp. 93–100. Springer (2012)
8. Badura, P., Pietka, E.: 3D fuzzy liver tumor segmentation. In: Pietka, E., Kawa, J. (eds.) Information Technologies in Biomedicine. Lecture Notes in Bioinformatics, vol. 7339, pp. 47–57. Springer (2012)
9. Zarychta, P.: Features extraction in anterior and posterior cruciate ligaments analysis. Comput. Med. Imaging Graph. **46**(2), 108–120 (2015). doi:10.1016/j.compmedimag.2015.03.001
10. Wieclawek, W., Pietka, E.: Fuzzy clustering in segmentation of abdominal structures based on CT studies. In: Pietka, E., Kawa, J. (eds.) Information Technologies in Biomedicine. Advances in Intelligent and Soft Computing, vol. 47, pp. 93–104. Springer (2008)
11. Wieclawek, W., Pietka, E.: Numerical complexity reduction in Live-Wire algorithm. In: Kurzynski, M. et al. (eds.) Conference on Computer Recognition Systems. Advances in Intelligent and Soft Computing, vol. 45, pp. 140–147. Springer, Heidelberg (2007)
12. Zarychta, P.: Posterior cruciate ligament—3D visualization. In: Kurzynski, M. et al. (eds.) Conference on Computer Recognition Systems. Advances in Intelligent and Soft Computing, vol. 45, pp. 695–702. Springer, Heidelberg (2007)
13. Zarychta, P.: ACL and PCL of the knee joint in the computer diagnostics. In: Napieralski, A. (ed.) 21st International Conference Mixed Design of Integrated Circuits and Systems MIXDES, pp. 489–492 (2014)
14. Zarychta, P.: Cruciate ligaments of the knee joint in the computer analysis. In: Pietka, E. et al. (eds.) Information Technologies in Biomedicine. Advances in Intelligent and Soft Computing, vol. 283, pp. 71–80. Springer, Heidelberg (2014)

Multimodal Image Registration for Mandible Motion Tracking

Agnieszka Anna Tomaka, Michał Tarnawski and Dariusz Pojda

Abstract The analysis of the condyle position in temporomandibular joint is crucial issue in differential diagnosis of bruxism and centric relation/maximal intercuspidation discrepancy or other temporomandibular joint disorders. The matter of importance are both static position of condyles during maximal intercuspidation and dynamics of mandible movement. The paper presents the method of evaluation of the dynamics of temporomandibular joint using the multiregistration of multimodal images: CBCT scans, virtual dental models and dynamic 3D photos of patient during the function. The mandible movements acquired with the optic 3D scanner are applied to the images gathered from CBCT, yielding the movement tracking of condyles in temporomandibular joint. Therefore low invasive 3D photo sequences registered with single high invasive CBCT yields the simulation of the high invasive imaging repetition. The anatomical evaluation of both the condyle position and occlusion condition during the movement may lead to the broaden TMJ diagnostics.

Keywords Multimodal image registration · Mandible movement acquisition · Temporomandibular disorder diagnosis

1 Introduction

The gnathology investigates two main classes of the masticatory system overload conditions: bruxism and the centric relation/maximal intercuspidation discrepancy. Bruxism is a neurological condition, characterised by involuntary and periodical teeth grinding or clenching [1]. Bruxism leads to orofacial pain, teeth wear, attrition and abfractions, periodontal recessions and temporomandibular joint disorders (TMD) [2, 3, 10].

A.A. Tomaka (✉) · M. Tarnawski · D. Pojda
Polish Academy of Sciences, Institute of Theoretical and Applied Informatics, Gliwice, Poland
e-mail: ines@iitis.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*, Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_15

Centric relation/maximal intercuspidation discrepancy (CR/MI) is characterized by improper position of the condyles in relation to articular fossa when contacts between the opposite teeth are the most abundant. It is a relation which usually occurs during the saliva swallowing. The contacts of the opposing dental arches lead to disclocation of the condyles from the neutral neuromuscular position [12–14]. Due to proprioceptive reflexes of masticatory system it leads to the disorders similar to ones caused by bruxism [3, 10, 12].

Diagnosis of bruxism and CR/MI discrepancy requires verification of centric relation during maximal intercuspidation, which can be done on the basis of clinical and instrumental examination, confirmed by the medical imaging. The most reliable one is the MRI, which allows for both the condyle and the articular disc position examination. CBCT and X-rays are considered as a valuable tool for verification of the CR, symmetry of the condyles position and the uniform articular space.

On the other hand, examination of the dynamics of the jaws movements is very important in proper restoration of dentition and evaluation of pathology. There exist commercial motion tracking systems like Axiographs [11], Cadiax and Zebris which enable some registration of movement of the mandible. However, the registration is performed as a graphical registration of positions of points forming Bonwille triangle [11] or as the 3D movement of the average mandible as the function of time.

Some latest papers describe attempts for the anatomical mandible movement simulation based on CT scans and the dynamic markers tracking, but the protocol was adjusted to hospital grade CT, not of dental CBCT's [5], or the registration is based on enforced head positioning [17].

The aim of our work is to combine the results of multimodal examinations: the morphology from the CBCT reconstructions, dental models scans and mandibular motion acquisition as the sequence of 3D photos. The essential difference to the other authors works is putting the stress on avoiding artifact physiological influences like head bracketing, and on the precise teeth surface reconstruction and registration, facilitating simultaneous TMJ and dental contacts examination.

Therefore, this multimodal fusion of high- and not- (X-ray) invasive imaging yields the simulation of the high invasive imaging repetition, while only the non invasive imaging is repeated, moreover the examination can be led in a way close to the clinical guidelines.

2 Integration of Multimodal Image Data

Due to technical development, sophisticated medical imaging methods become customary tools used in supporting the medical diagnosis. Different imaging modalities quantifies various aspects of human morphology, physiology and pathology transforming them into digital information gathered in computer systems. Such examinations yield results hard to imagine some decades ago, but quite new perspectives are achieved by fusion of different multimodal patient data. Integration of multimodal images requires a registration—bringing the data into the same, common coordinate

system [8]. Once registered, data set contains additional information concerning the relations between data from multimodal examinations.

The main prerequisite of registration process is the assumption of existence of common corresponding features which are recognized during registration and used to determine the suitable transformations. The kind of the transformation depends on the process of forming of an image and on the used image representation. This work focuses on the representation in a form of triangular mesh, the transformation belongs to the group of rigid body transformations, such that:

$$\mathbf{T}(\mathbf{X}) = \mathbf{RX} + \gamma \mathbf{1}_{\mathbf{k}}^{\mathbf{T}},$$

where **T**—transformation, **X**—matrix of points' coordinates, **R**—orthogonal rotation matrix, 1_k^T—vector of ones, γ—translation vector, k—number of points.

For known correspondence the transformation that minimizes distances between corresponding points is obtained using least squared method, known as Procrustes method [7]. For unknown correspondence the iterative closest point (ICP) algorithm can be applied [4]. ICP assumes temporary correspondence of analyzed points in the first image as the closest point in the second image. This temporary correspondence is used to determine parameters of the transformation. After the transformation the procedure of temporary correspondence establishing and determining of transformation parameters is repeated iteratively until the sum of distances of corresponding points reaches the minimum. The result of the registration depends on the area selected for correspondence matching and on initial positions of images.

3 Initial Condyle Position, Fusion of the Images

The diagnosis of stomatognathic system is supported by the following multimodal imaging techniques: CBCT scans, 3D photos, and virtual dental models, all of them acquired in the device dependent coordinate system (Fig. 1).

Analysis of CBCT yields 3D surface reconstructions obtained with different thresholds: reconstructions of soft tissues, bones and teeth. Further segmentation of bone's layer separates skull and mandibular models. Reconstructed in a coordinate

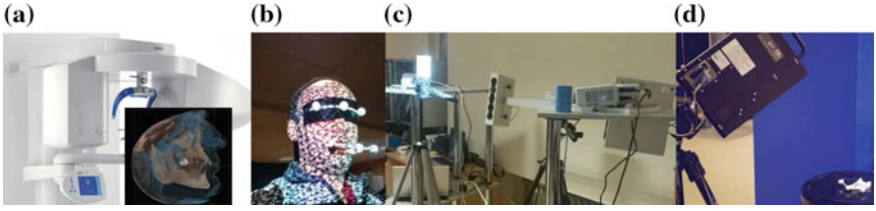


Fig. 1 Acquisition of **a** 3D surfaces from CBCT; **b**, **c** mandible movement using modified 3dMD scanner; **d** virt. dental models with arches from Minolta VI-9i scanner

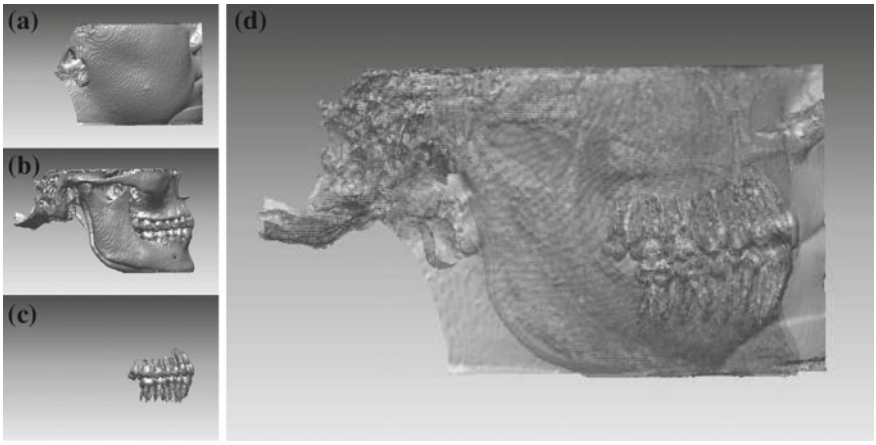


Fig. 2 CBCT reconstructions **a** soft tissues, **b** bones, **c** teeth, **d** mutual relation

system connected with CBCT device, models positions reflects the relation between different patient's tissues. Especially, positions of maxillary and mandibular models reflects the real static patient occlusal and joint conditions during the examination (Fig. 2).

3D photos and virtual dental models are 3D images of outer surfaces of respectively patient's face and his dental models, acquired with 3D scanners. Unluckily, the images haven't got a common part. For disjoint images, without overlapping areas, the usage of additional reference objects is suggested and so special transferring arches were designed. The arch is an easy to be scanned object printed by means of 3D printer and fixed to the patients' teeth with dental acrylic. In fact, this technique introduces additional scanning of a face with arch and virtual models with arch, but provides the relations between 3D photo and virtual dental models (Fig. 3) [15].

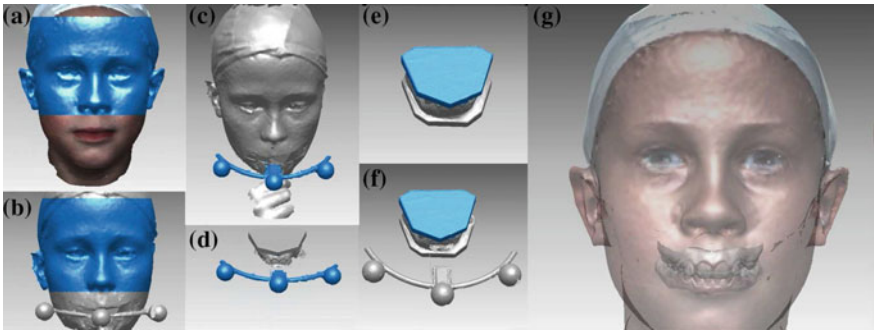


Fig. 3 Common surfaces for registering **a**, **b** 3D photo and 3D photo with arch; **c**, **d** 3D photo with arch and virtual dental models with arch; **e**, **f** virtual dental models and virtual dental models with arch; **g** yielded registration

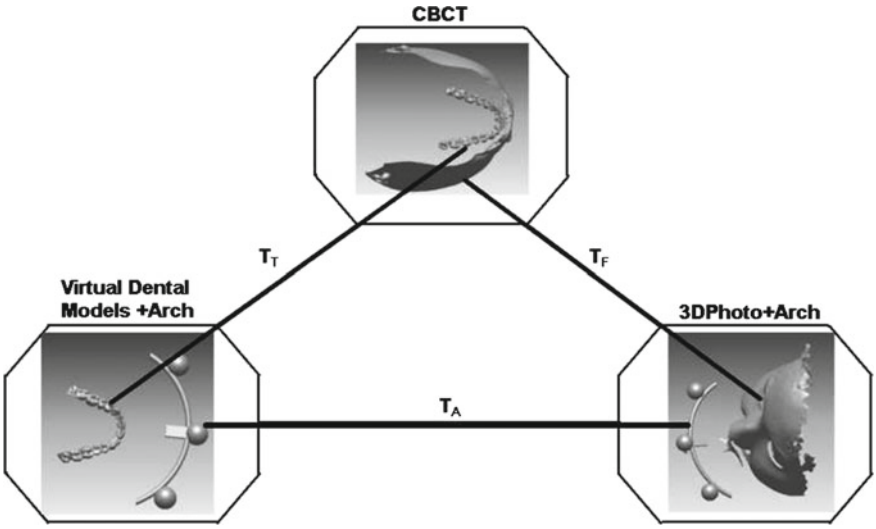


Fig. 4 Common areas for CBCT, 3D photos and virtual dental models registration

For the registration of vision-based and X-ray derived images the following common parts were assumed: surfaces of teeth, upper parts of face and arch surfaces. The idea of registration is shown in Fig. 4. For each matching the mathematical transformation describing rigid body motion is sought.

There are two possible registration strategies: pair wise registration, iteratively finding particular transformations, or simultaneous registering of all common parts, which finds all transformations $\mathbf{T_T}$, $\mathbf{T_A}$ and $\mathbf{T_F}$ at the same time and which distributes registration error equally on each matching.

Transformations determined for common areas can be applied to all images coming from the same source and in this way all images are registered in the same coordinate system.

While CBCT is usually performed in maximal intercuspation, repeating the matching process for lower teeth arch yields the registration of initial maxilla—mandible relations with 3D photos containing mandible arch.

4 Mandible Movements Acquisition

4.1 The Idea of Measurement

Evaluation of temporomandibular joint requires the separation of mandibular movement from overall head movements, and relays on measuring the maxillo-mandibular relations. Both mandible with its teeth and maxillo-facial complex with upper teeth

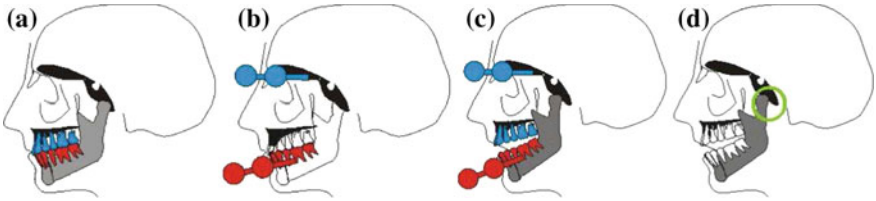


Fig. 5 **a** CBCT maxillo-mandibular models; **b** CBCT versus 3dPhotoAndArches registration; **c** acquisition of the movement; **d** transferring the movement on the models

are not subject to deformations in this movement, and are considered to be rigid bodies, and the movement—rigid body transformation. Since the soft tissues associated with the mandible are not tightly bound to the underlying bones and may be subjected to deformation, it is difficult to acquire mandible movement basing only on 3D photos. Therefore transfer-kinematic mandibular arch was rigidly attached to the lower patient's teeth. The assumption that movement of the mandible with the kinematic arch is a also rigid body transformation means that the deformation of soft tissues and slight bone deformation are not taken into account.

To eliminate overall head movement the initial registration of upper arch is performed. Soft tissues of upper face, after selection of the most stable areas, can be used as upper arch marker.

The overall idea of the measurement relays on the registration of the static maxillo-mandibular models acquired in CBCT (Fig. 5a), with the initial position of mandibular arch (Fig. 5b), using the registration method described above. Next the relation change between arches (Fig. 5c) is measured and yielded transformation of mandibular arch is applied upon the mandible model (Fig. 5d).

4.2 Medical Strategies of Scanning for Static 3D Scanner

Tracking of the mandible movement with the usage of a static 3D scanner it is not possible, but the acquisition of some chosen mandible positions can be performed. The acquisition protocol included the following mandible position: bite in the static habitual occlusion, incisal edge to edge position, maximal protrusion with teeth contacts, medium laterotrusions to both sides, maximal laterotrusions with teeth contact, than maximal laterotrusions in a medium mouth opening, symmetrical medium and maximal mouth opening (Fig. 6).

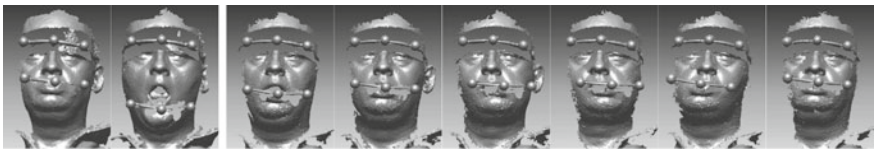


Fig. 6 3D photos with arches for chosen mandibular positions

4.3 Visualization of the Results

Basing on the described method of multimodal images registration and movement acquisitions the experiment of mandible movement evaluation was performed. For a patient with CBCT data, dental models were scanned, and 3D photos with arches attached to upper head and to lower teeth were acquired. All calculations were done in our application which uses RapidForm2006 environment for 3D visualization, and uses some of its functions especially for surface registering. The input data for the program consisted of CBCT reconstructions od maxilla and mandible, virtual dental models with arches and the face scans with arches for mandible extreme articulations. First the registration of the static position were performed for all data. Next in each scan of with mandible movement (Fig. 6), the surface of lower arch was determined.

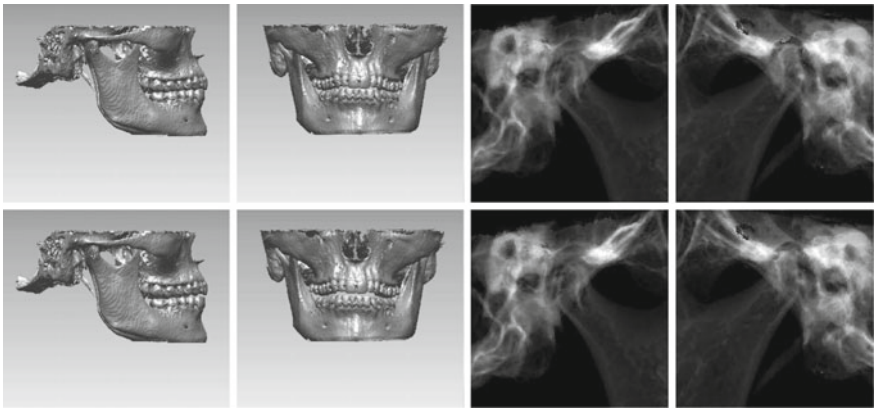


Fig. 7 Static habitual occlusion and resting jaw position

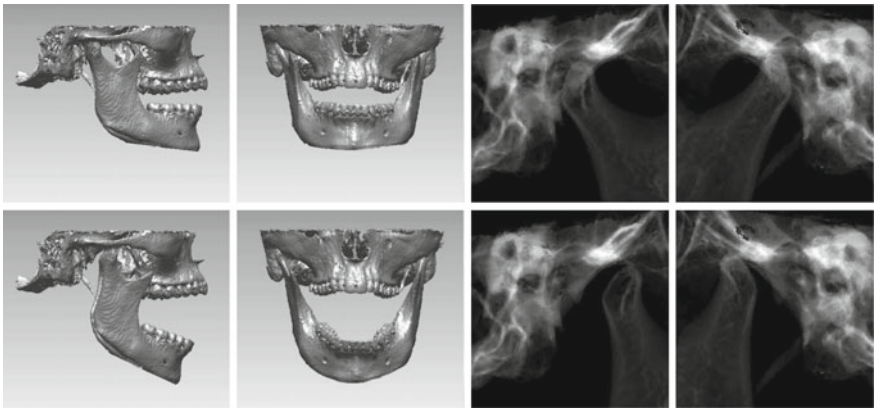


Fig. 8 Symmetrical medium and maximal mouth opening

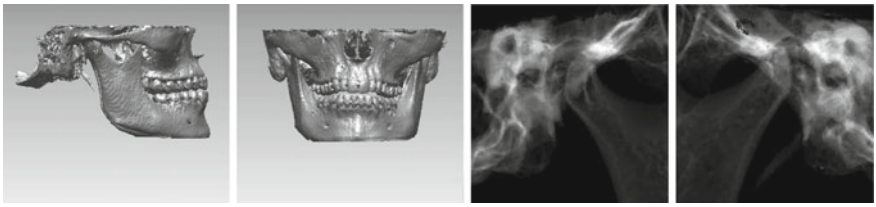


Fig. 9 Incisal edge to edge position

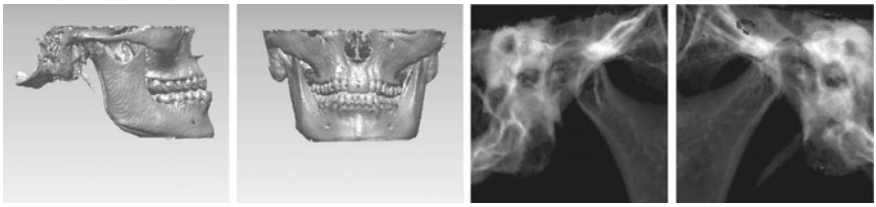


Fig. 10 Incisal edge to edge position and maximal protrusion

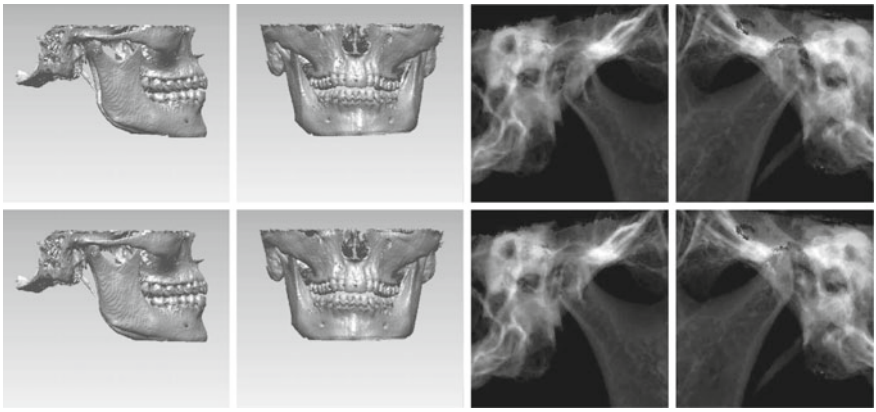


Fig. 11 Laterotrussion to the *left* and to the *right*

The transformation of lower arch regarding its initial position was established and used to transform all images connected with mandible. Maxillo-mandibular relation modified for each 3D photo, can be analyzed and visualized in 3D graphics applications using different views, translucency or displaying profiles (Figs. 7, 8, 9, 10, 11 and 12).

As far as many experienced practitioners have long practice of evaluating the position of the TMJ by means of 2D X-ray images, the reconstructed models in recorded maxillo—mandibular relation can be treated with virtual X-ray simulator [16]. The simulation is based on the measuring, for assumed geometry of X-ray imaging system and patient positioning, and for each point of the detector, the lengths of the

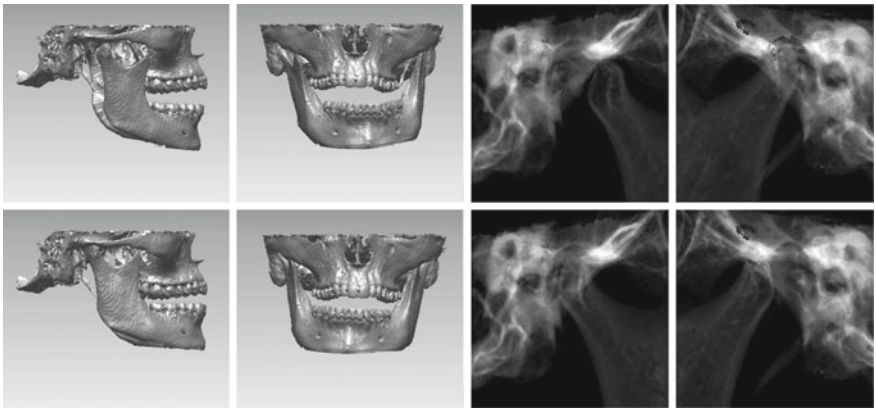


Fig. 12 Maximal laterotrusions to the *left* and to the *right* in medium mouth opening

way of the rays in the models of maxillary and mandibular structures. This approach assumes the constant X-ray attenuation coefficient in each point of the model. It assumes also models as water-tight surfaces. Additionally, to separate images of both condyles, models for the simulation are virtually cut according to the midsagittal plane. Obtained projections contain the images of condyles in chosen mandible positions (Figs. 7, 8, 9, 10, 11 and 12). For each relation then virtual xrays were performed yielding dynamic X-rays of TMJ.

5 Mandible Motion Tracking

Although the chosen articulations can be acquired with the usage of static 3D scanner, as far as time of flash charging is quite long, there is no possibility to trace movement trajectory. The application of interpolation technique was tested, but unfortunately in obtained mandible positions there occured the collision between maxillary and mandibular models. Therefore interpolation techniques cannot be used for mandible movement evaluation.

5.1 3D Dynamic Scanners

A movement of 3D objects can be acquired using dynamic 3D scanner. Three-dimensional scanning of dynamic scenes (temporal-3D, 4D) is a relatively new issue in the field of 3D scanning. Currently applied approaches uses the stereo matching method with structural light for two or multiple synchronized cameras (3dMDdynamic Systems), or *time of flight* cameras like Kinect for Xbox One.

3dMDdynamic Systems: Face 3dMDdynamic System or 3dMDdynamic Body System are equipped with the scanning head consisting of a pair of stereo cameras, a structured light projector and a color camera recording the texture of an object being scanned. Depending on the application the scanners use two (in scanning of a human face) up to twenty-two (scans of the whole body) scanning heads. The scan rate is 60 scans per second with linear accuracy range of 0.5 mm or less [6]. Unfortunately high cost of the scanner exceeded the project budget.

In time-of-flight (ToF) cameras the 3D scene reconstruction is resolved measuring the time-of-flight of a light signal between the camera and the subject for each point of the image. This technology initially applied as slow point-by-point scanning(LIDAR systems), now can capture the entire scene at each laser or light pulse providing up to 60 images per second. Microsoft Kinect for Xbox One scanner with a 512×424 resolution of ToF camera and 1920×1080 resolution video camera can operate in real time [9]. The accuracy of the system is about ± 1 mm, which seems to be insufficient.

Our solution used static 3dMD scanner, but the technique of image acquisition was modified. The calibrated monochromatic cameras of both heads of 3dMD scanner with additional continuous structural light were used and the synchronized sequences of images were acquired. 3D reconstructions were done using 3dMD reconstruction software. This solution enables the acquisition of up to 7 images per second, but as far as the reconstruction was done by the 3dMD software, the accuracy seems to be near the accuracy of original 3dMD scanner.

5.2 Medical Strategies of Mandible Motion Tracking

The examination of the motoric schemes of the mastication system consists of at least several tests. Initially, the maximal intercuspidation is examined. The freeway space is taken into the consideration. Then open-close movement is described, with special regard for the symmetry and the smoothness of the movement. The patient is asked to make protrusion, keeping the teeth in the light contact. The laterals excursions from the maximal intercuspidation are examined. The contacts of the teeth are described during these movements (in articulation). Articulation has the big impact on the muscle-joint play, so in order to watch the muscle and joint free movements, the teeth should be not in contact moreover, to avoid some reflexes of *anticipation of contact* the dental arches should be in certain distance. Again the smoothness of the movement, symmetry and the range of the borderline positions of the mandible, together with all disturbances like joint clicking, or movement constraint are taken into the consideration. This scheme is to be recapitulated during the 3D mandible tracking.

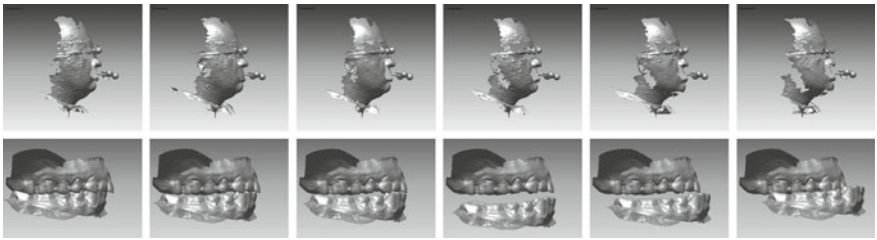


Fig. 13 Selected 3D scans during protrusion movement and responding teeth relations

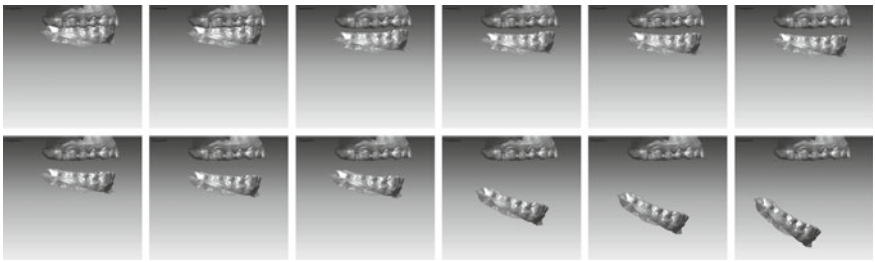


Fig. 14 Selected teeth relations for mouth opening movement

5.3 Visualization of the Results

The acquisition of sequence of 3D photos with arches were done using the modified 3dMD scanner while the mandible motion was executed in accordance with the recommendations of the medical expert.

From technical point of view for each movement there were about several dozen 3D photos with arches to be analyzed instead of one as in Sect. 4.3, but overall idea was the same. The change of lower arch position in subsequent 3D photos was small, therefore registration were more robust. As the result the sequence of transformations was calculated for each movement sequence. These transformations put upon registered images of mandible yield subsequent mandible positions and make tracing of the movement possible (Figs. 13, 14).

Sequence of transformations can be a basis for further analysis of the motion giving possibility to distinguish the rotational from translational movements.

6 Conclusions

The fusion of morphology obtained from CBCT and the sequences of 3D optical scanner photos (acquired during mandible movements) makes the simulation of the changing maxillo-mandibular relations possible. This simulation can be visualized in different ways, for example as animated sequences of static 2D or 3D images.

The paper is focused on the visualization aspects, the resulting data allows for the easy quantitative analysis of the joint and teeth positions and movements in any 3D program. Described method creates a new tool for evaluation of joint and occlusion useful in temporomandibular disorders diagnosis. But as far as the method was tested on the patients with CBCT examination, who volunteered to have some additional, noninvasive records—facial scans and facial scans with the mandible position tracking, further clinical tests have to be performed.

However, it must be remembered that dynamic TMJ imaging is still a simulation, therefore its accuracy depends on the accuracy of CBCT segmentation, accuracy of 3D scans, their registration and on the accuracy of X-rays simulation. All these elements and steps can be improved by applying better segmentation and 3D reconstruction algorithms, better 3D dynamic scanner and less simplified X-ray simulations, but we prove that fusion of multimodal image can bring a new quality in analysis being the basis for diagnosis process. From the point of view of clinician, the visualizations are surprisingly high convincing. The presented paper showed the novel way of evaluating the medical information. This type of image fusion gives the new possibilities for the understanding the clinical facts, although present, impossible to be totally examined by human. The paper shows out that there is still place for looking for the information in the medical records which contains more data that we could anticipate.

Acknowledgments This work was supported in part by the National Science Center, Poland under the research project 2012/07/B/ST6/01238.

References

1. Baron, S.: Bruxism and Occlusion. Lecture Notes of ICB Seminar, 137th ICM Seminar, pp. 3–14 (2014)
2. Baron, S., Tarnawska, B., Karasinski, A.: Analiza zwarcia ekscentrycznego z zastosowaniem systemu T-Scan II - badania własne (The excentric occlusion analysis by the means of T-scan II system—Individual research). Mag. Stom. **3**, 14–17 (2003). (in Polish)
3. Baron, S., Tarnawska, B., Tarnawski, M.: Wpływ zaburzeń zwarcia ekscentrycznego na wybrane elementy układu stomatognatycznego (The influence of the excentric occlusion disturbances on the selected elements of masticatory system). Mag. Stom. **5**, 62–67 (2004). (in Polish)
4. Besl, P., MvKay, N.: A method for registration of 3-D shapes. IEEE Trans. PAMI **144**, 239–256 (1992)
5. Chang, A., Han, J., Kim, D., Yi, W., Hwang, S.: Evaluation of intra-articular distance narrowing during temporomandibular joint movement in patients with facial asymmetry using 3-dimensional computed tomography image and tracking camera system. J. Craniomaxillofac. Surg. **43**(3), 342–348 (2015)
6. Dynamic 3dMD scanner. http://www.3dmd.com/3dmd-systems/dynamic-surface-motion-capture-4d/
7. Dryden, I., Mardia, K.: Statistical Shape Analysis. Wiley (1998). ISBN 0-471-95816-6
8. Hill, D., Batchelor, Ph.: Registration methodology: concepts and algorithms. In: Hajnall, J., Hill, D., Hawkes, D. (eds.) Medical Image Registration. CRC Press (2001). ISBN 0-8493-0064-9

9. Kinect for Xbox One. https://dev.windows.com/en-us/kinect/hardware
10. Kleinrok, J., Kleinrok, M.: Jatrogenny bólowy zespół dysfunkcji narządu żucia po leczeniu protezami stałymi (Latrogenic mandibular pain dysfunction syndrome after prothetic treatment with fixed dentures). Prot. Stom. **5**, 265–273 (2001). (in Polish)
11. Koeck, B.: Zaburzenia czynnościowe narządu zucia (The masticatory system dysfunctions). wyd polskie red. T. Maślanka, Elsevier Urban i Partner, 121, 126 (1997) (in Polish)
12. Majewski, S.: Relacje międzyszczękowe w świetle nowego rozumienia okluzji (Intermaxillary relations, new concept of occlusion). Mag. Stom. **3**(96), 15–17 (1996) (in Polish)
13. Olivo, S., Bravo, J., Magee, D., Thie, N., Major, P., Flores-Mir, C.: The association between head and cervical posture and temporomandibular disorders: a systematic review. J. Orofac Pain. **20**(1), 9–23 (2006)
14. Palaskar, J., Murali, R., Bansal, S.: Centric relation definition: a historical and contemporary prosthodontic perspective. J. Indian Prosthodont. Soc. **13**(3), 149–154 (2003)
15. Tomaka, A., Tarnawski, M., Luchowski, L., Liśniewska-Machorowska, B.: Digital Dental Models and 3D Patient Photographs Registration for Orthodontic Documentation and Diagnostic Purposes. Computer Recognition Systems 2, Advances in Soft Computing, vol. 45, pp. 645–652 (2007)
16. Tomaka, A.: Analiza obrazów wielomodalnych dla potrzeb nieinwazyjnej diagnostyki ortodontycznej (Multimodal image analysis for noninvasive orthodontic diagnosis). ITAS PAS (2013) (in Polish)
17. Yuan, F., Sui, H., Li, Z., Yang, H., Lü, P., et al.: A method of three-dimensional recording of mandibular movement based on two-dimensional image feature extraction. PLoS One **10**(9), e0137507 (2015). doi:10.1371/journal.pone.0137507

Breast Ultrasound Tomography: Preliminary In Vivo Results

Krzysztof J. Opieliński, Piotr Pruchnicki, Marcin Jóźwik, Jarosław Majewski, Tadeusz Gudra, Mariusz Bułkowski and Włodzimierz Roguski

Abstract The paper presents the preliminary results of examinations of biological structures and healthy breast tissue in vivo obtained with the developed ultrasound tomography model. This device uses an innovative combination of ultrasound transmission tomography (UTT) and ultrasound reflection tomography (URT) which enable to obtain three-dimensional images in the same way as in the classical computer tomography (CT). The huge advantage of these methods is not exposing the patient to ionizing radiation. Compared to conventional breast ultrasonic echography (US), UTT and URT methods enable to acquire a much larger amount of information whose clinical significance has not been well documented yet. Based on the preliminary results, it is anticipated that the ultrasound tomography may contribute to the creation of a new standard of diagnosis and treatment of breast cancer. Currently, a Polish manufacturer of ultrasound scanners aims at getting an ultrasonic tomography device for examining women's breasts to market.

Keywords Ultrasonic tomograph · Ultrasound medical imaging · Breast tissue · Cancer diagnosis

K.J. Opieliński (✉) · P. Pruchnicki · T. Gudra
Chair of Acoustics and Multimedia, Faculty of Electronics, Wroclaw University of Technology, Wroclaw, Poland
e-mail: krzysztof.opielinski@pwr.edu.pl

M. Jóźwik
Faculty of Medical Sciences, Department of Gynecology, Gynecologic Endocrinology and Obstetrics, University of Warmia and Mazury, Olsztyn, Poland

M. Jóźwik
Department of Gynecology, Scanmed St. Rafal Hospital, Krakow, Poland

J. Majewski
Faculty of Telecommunications, Computer Science and Electrical Engineering, University of Technology and Life Sciences, Bydgoszcz, Poland

M. Bułkowski · W. Roguski
Dramiński S.A. Ultrasound Scanners, Owocowa 17, Olsztyn, Poland

E. Piętka et al. (eds.), *Information Technologies in Medicine*, Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_16

1 Introduction

Breast cancer is the most common women's cancer in Poland and the second, after bronchial cancer, most common cause of death in women due to cancer. Early detection affects the survival, since early cancer can be removed before the metastasis develops. Breast cancer can be detected early as a lump by self-examination of the patient. However, not all breast cancers produce early symptoms or arouse the suspicion of breast cancer by causing skin stretch or discharge from the nipple. The consulting physician can order imaging in addition to palpation. Visualization of the change, particularly in cases of weakly palpable changes, enables to perform a biopsy for the histopathological verification. In young women, the imaging tests include breast ultrasonography in the first place, because of the density of glandular tissue. In older women, a higher detection rate is observed in the case of mammography, which, in addition to the risks associated with exposure to X-rays [1], is painful sometimes [8]. Both of these methods are complementary, but do not replace each other. In addition, the result of mammography is affected by the pre- or post-menopausal status, and taking hormone replacement therapy [7]. There is also controversy over the initial age and frequency of mammography, because mammography does not reflect properly dense breasts, which are characteristic of most young women, and breast implants. There is also a potential risk of accumulation of the effects of ionizing radiation and for this reason, doctors do not recommend mammograms more than every 2–3 years and for women under 40. Some breast cancers are not indicated in breast ultrasonography or mammography and in these cases one can use magnetic resonance imaging (MRI) [11]. Breast MRI has recently been recognized as the preferred method of screening tests for women at high risk of developing cancer and as very significant examinations supporting mammography and ultrasonography in order to estimate the size and spread of breast cancer [9, 19]. The use of MRI in the diagnosis of breast cancer is important because of the high sensitivity (the ratio of true positives to the sum of true positives and false negatives) and moderate specificity (the ratio of true negatives to the sum of true negatives and false positives) for the tumor masses larger than 5 mm, including ductal carcinoma in situ (DCIS) [20]. With the ability to analyze the breast morphology, MRI provides qualitative and quantitative data about the tumor vasculature, which allows for the differentiation of benign mass from carcinogen mass [21]. However, MRI scanners are quite expensive because of the price, the need to maintain them in special rooms, servicing, and hiring a qualified staff. Furthermore, because of the requirement to detect details, intravenous contrast agent is very often required here that can penetrate into the extravascular space or cause allergic reactions. The administration of contrast agents is not recommended in patients with hepatic or renal diseases. These disadvantages, as well as the price of MRI limit the use of MRI in the diagnosis of women's breast and prevent general availability of screening examination using this method.

New and better imaging methods are needed to improve the detection of breast cancer. Any method which improves the detection of breast cancer, both in the diagnosis and screening, will help to reduce mortality from this disease.

This paper shows the preliminary results of the research of biological structures and healthy breast tissue in vivo obtained with the ultrasonic tomography device, which for several years has been developed by the Polish manufacturer of ultrasound scanners—DRAMIŃSKI S.A. in cooperation with the Laboratory of Ultrasound Technology at the Chair of Acoustics and Multimedia of the Faculty of Electronics at Wroclaw University of Technology. Currently, DRAMIŃSKI S.A. aims at getting an ultrasonic tomography device to market. Advanced work towards building a prototype of the ultrasonic tomograph to examine the breasts of women is currently underway at few research centers in the world (Poland, USA, and Germany) [2, 5, 10, 17, 22]. Despite the use of similar methods, the device models developed differ significantly. The compact description of the existing advanced prototypes of devices is presented in [17] for comparison. The advantages of the ultrasonic tomograph developed by Polish company DRAMIŃSKI compared to the devices by research centres in the USA and Germany, are: the capability to obtain relatively high-quality images, short measurement time, fast data acquisition, signal processing and image reconstruction with still relatively simple design of the ultrasound transducer array and electronics. The ultrasonic tomograph by DRAMIŃSKI S.A. should be far cheaper in mass production compared to the designs of other research groups.

The aim of the study was mainly the analysis of results of in vivo ultrasonic tomography imaging of biological media and female healthy breast tissue with regard to the distributions of local values of relevant acoustic parameters, which allow recognizing and distinguishing the specific inner structures.

2 Ultrasonic Tomography Device

The model of the Polish ultrasonic tomography device currently consists of four main blocks: array block, signal block, water system, and the unit for control and visualization (Fig. 1).

Each of these blocks is still at the development and studies stage aimed at building a prototype ultrasonic tomography device for diagnostic screening of breast in women.

The array block is the most characteristic one. Circular array comprising 1024 omni-directionally arranged ultrasonic transducers [3, 17, 18], operating on a frequency of 2 MHz is its main element. Each transducer can be a transmitter and a receiver. Circular array is coupled to a mechanical system positioning it at an appropriate depth (Z-axis) and is responsible for its movement in this axis. The array surrounds the patient's breast hanging vertically down and laid horizontally on the device bed. The movement range is chosen to encompass the entire breast from its base to the nipple. The array does not touch the breast at any point on the Z axis position. Water provides a suitable acoustic coupling between the array and the breast. Therefore, the array-breast system is positioned in the array chamber, which is com-

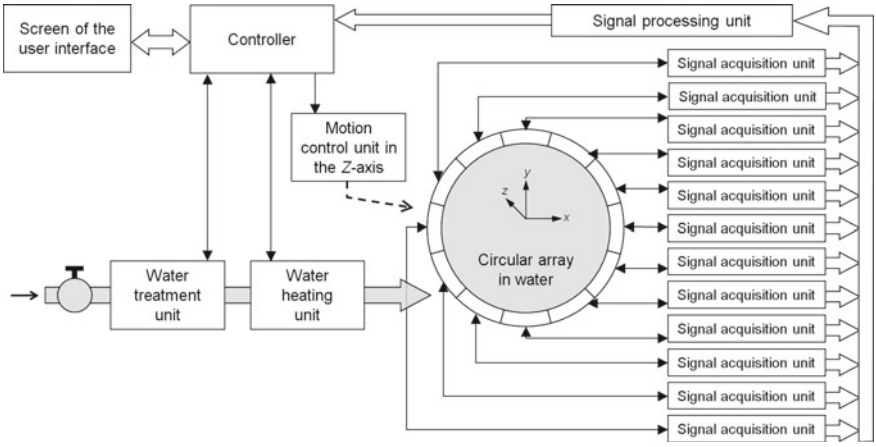


Fig. 1 The block diagram of the ultrasonic tomography device

pletely filled with purified water. The chamber is a waterproof container with a hole in top which enables to put the breast into the array ring.

The water system is a block containing elements for preparing and supplying water into the array chamber. It consists of the water conditioning block and water heating block. Conditioning is mainly aimed at water purification and is important for the protection of array conditioning from scaling. Water heating ensures an adequate temperature for establishing suitable conditions for ultrasound propagation. Water system is powered from the water mains through the appropriate valve.

The control and visualization block consists of a controller, array movement control block and user interface with the connected screen. The controller is designed to control the water system and the array movement control block. It also manages the final images obtained and the entire process of data collection and processing. In the water system, the controller mainly controls the process of filling and emptying the chamber with water, heating and stabilization of its temperature, purification and drawing of water from the mains. The array movement control unit is responsible for the initial setting of the array and a precise and stepping motion during the measurements. The controller also creates two-dimensional images of breast each cross-section in different device operation modes, i.e. ultrasound transmission tomography (UTT) and ultrasound reflection tomography (URT) [16–18]. The prototype device will also include the surround mode of standard ultrasonic echography (US) using several methods [14] (including *real-time* scanning for the selected 32-transducer circular array sector, the FASCI—full angle spatial compound imaging—image composed of sectoral images) and composing images of individual sections into the three-dimensional image. The controller will also communicate with the Internet and convert the resulting images into DICOM format. The screen user interface allows the operator to control the entire device, including the correct scan of the breast and preliminary evaluation of the resulting images.

The signal block has two components. It is the signal acquisition block and the data processing block. Each signal acquisition block supports a group of ultrasonic transducers. It consists of pulse generator excitating transmitters, switches for changing the operating mode of transducers between transmitting and receiving, the ADC system and FPGA circuit for the initial processing of received signals. FPGA determines the parameters of the collected signals and prepares them for the transmission to the data processing block. The data processing block collects data from all acquisition blocks and uses them to reconstruct two-dimensional cross-sectional images of the breast. The time of examination of one coronal cross-section of breast in UTT and URT modes concurrently is now 30 s. It will be reduced to approximately 1 s in the prototype, which will enable to examine the whole breast volume (over 200 coronal sections) including the reconstruction of images in tens of seconds.

3 Examined Structures and Imaging Method

Biological structures in the form of boiled egg without shell and in vivo breast of a healthy woman aged 52 were subjected to preliminary ultrasonic tomography examination. Water temperature in the tank was 26 °C during the egg, and 35.5 °C during breast examination. Due to the trade-off between the longitudinal resolution and the transmission range (the diameter of the circular array is 260 mm), the ultrasound wave with a frequency of 2 MHz (wavelength in soft tissue $\lambda = 0.77$ mm) was used. This paper presents the results of the examination of the same cross-section of an egg more or less at its maximum diameter and the results of the examination of the same coronal cross-section of breast just below the half height of the breast hanging freely in water (in the direction of the nipple). The structure of a raw egg and a female breast in vivo is shown in Fig. 2. Table 1 summarizes the average acoustic parameters determined for distilled water, hen's egg structures and tissues present in healthy

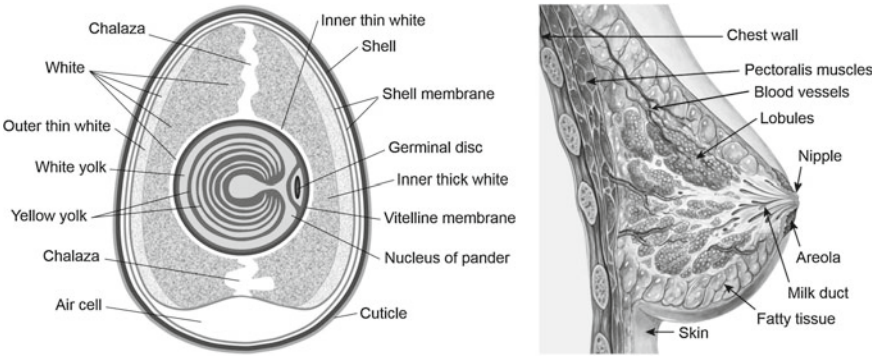


Fig. 2 The structure of: raw hen's egg on the *left*, and woman's breast on the *right* (Modified from P.J. Lynch, see http://patricklynch.net)

Table 1 Average values of acoustic parameters determined for distilled water, hen's egg structures, and tissues present in healthy breast [12, 13]

Structure			Ultrasound velocity c (m/s)	Ultrasound attenuation α (dB/m)	Acoustic impedance Z_a (MPa·s/m)
Distilled water	$T = 26\,^{\circ}\mathrm{C}$	$f = 2\,\mathrm{MHz}$	1499.36	0.74	1.49
Distilled water	$T = 35.5\,^{\circ}\mathrm{C}$	$f = 2\,\mathrm{MHz}$	1520.82	0.58	1.51
Raw egg's white	$T = 20\,^{\circ}\mathrm{C}$	$f = 2\,\mathrm{MHz}$	1530	46.04	1.58
Raw egg's yolk	$T = 20\,^{\circ}\mathrm{C}$	$f = 2\,\mathrm{MHz}$	1507	138.97	1.54
Boiled egg's white	$T = 20\,^{\circ}\mathrm{C}$	$f = 2\,\mathrm{MHz}$	1521–1536	30–70	1.60–1.61
Boiled egg's yolk	$T = 20\,^{\circ}\mathrm{C}$	$f = 2\,\mathrm{MHz}$	1504–1510	307.22	1.62
Artery wall	In vivo	$f = 2\,\mathrm{MHz}$	1600	610	1.69–1.72
Breast					
Ligaments	In vivo	$f = 1\,\mathrm{MHz}$	1750	470	1.94–2.14
Subcutaneous fat tissue	In vivo	$f = 2.5\,\mathrm{MHz}$	1470	171	1.40
Internal fat tissue	In vivo	$f = 2.5\mathrm{MHz}$	1470	180	1.40
Glandular tissue	In vivo	$f = 2.5\mathrm{MHz}$	1515	294	1.55

breast [12, 13]. The characteristic of the acoustic parameters of biological structures and tissues herein is a range of variation depending on the chemical composition, physical properties, individual differences, aging of biological material [13] and in the case of a boiled egg—its thermal history [12].

Data acquisition was carried out in the geometry of divergent projections by measuring the projection values of several different acoustic parameters of the selected section of the examined structure immersed in distilled water: the transit time and the amplitude of ultrasound pulses after passing (UTT mode), the envelope of scattered signals amplitude (URT mode) [15, 17]. UTT images were reconstructed using the filtered back projection (FBP) algorithm [6, 15] with the Hamming filter for the egg and the so-called stochastic filter for breast [4], and the URT images were reconstructed by using the synthetic focusing aperture algorithm for the system of transducers arranged on the inner ring (SAF) [17].

4 Results, Analysis and Conclusions

Figures 3, 4 and 5 show the cross-sectional images of the hen's egg reconstructed from measurements made using the developed ultrasonic tomography device, and the distributions of pixel values of these images along the x line for $y = -4$ mm. Individual images represent imaging capabilities for a fixed water temperature, the fixed frequency of the ultrasonic wave (Sect. 3), in 3 different modes: UTT with the visualization of the distribution of local values of ultrasound speed $c(x, y)$, the distri-

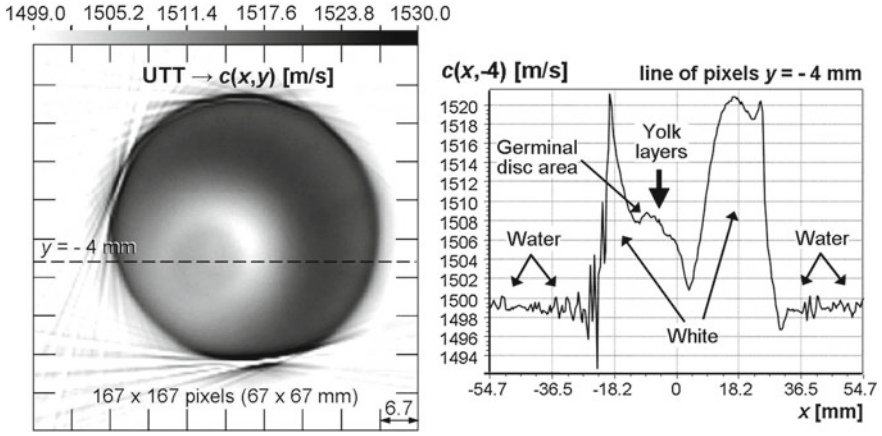


Fig. 3 UTT image of the distribution of local values of ultrasound speed in the hen's egg cross-section reconstructed from the measurements, and the distribution of pixel values of this image on the line $(x, -4)$

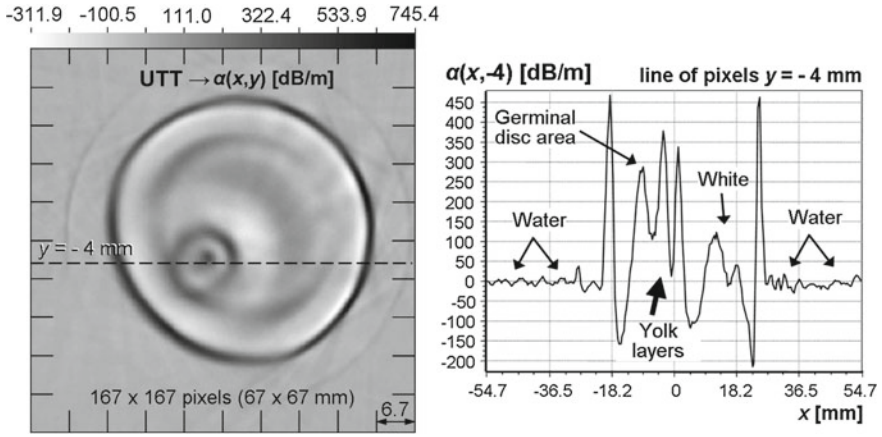


Fig. 4 UTT image of the distribution of local values of ultrasound attenuation in the hen's egg cross-section reconstructed from the measurements, and the distribution of pixel values of this image on the line $(x, -4)$

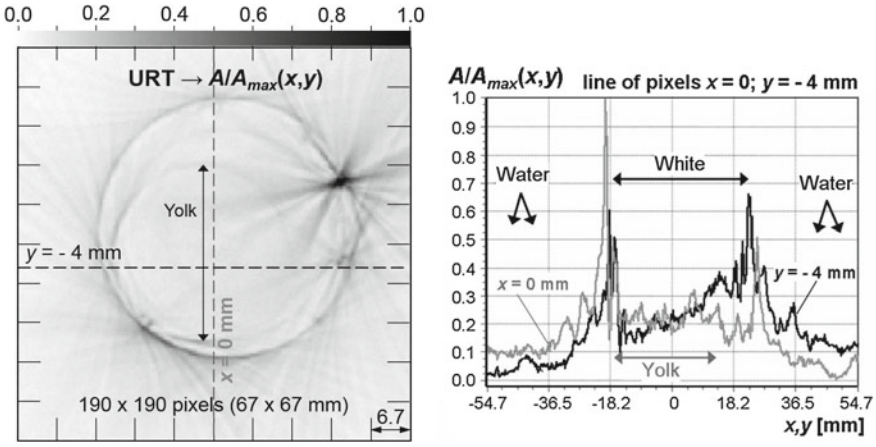


Fig. 5 URT image of the distribution of ultrasound scattering amplitude normalized values in the hen's egg cross-section reconstructed from the measurements, and the distribution of pixel values of this image on 2 lines perpendicular to each other: $(x, -4)$ *black* and $(0, y)$ *grey*

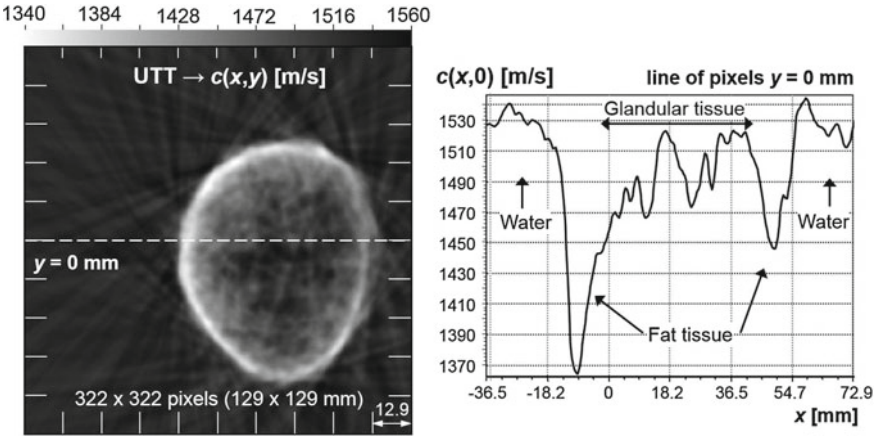


Fig. 6 UTT image of the distribution of local values of ultrasound speed in the breast coronal section reconstructed from the examinations in vivo, and the distribution of pixel values of this image on the line $(x, 0)$

bution of local values of ultrasound attenuation $\alpha(x, y)$, URT with the visualization of the distribution of normalized ultrasound scattering amplitude values $A/A_{max}(x, y)$.

Figures 6, 7 and 8 show the coronal breast cross-section images in a similar manner as in the case of the hen's egg.

When examining the conformity of acoustic parameters of the examined egg structure (Table 1; Fig. 3), we clearly see that the UTT image of the distribution of ultrasound speed can be regarded as a quantitative image: the resulting speed of ultrasound is about 1498–1500 m/s in water, 1518–1522 m/s in albumen and 1501–

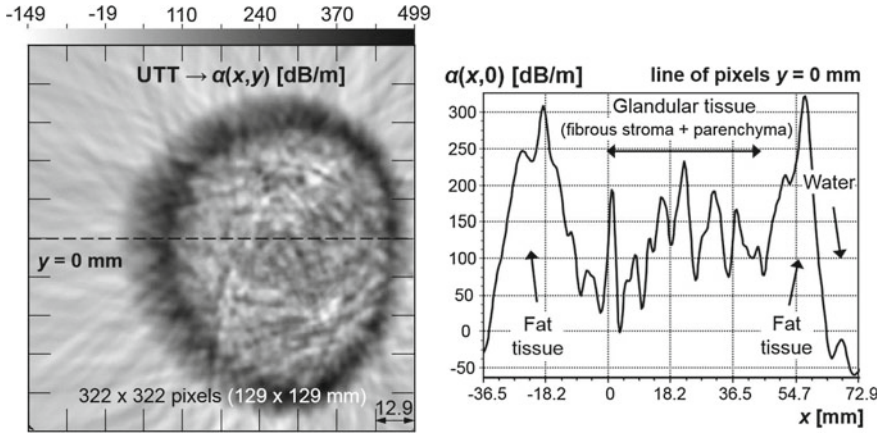


Fig. 7 UTT image of the distribution of local values of ultrasound attenuation in the breast coronal section reconstructed from the examinations in vivo, and the distribution of pixel values of this image on the line (x, 0)

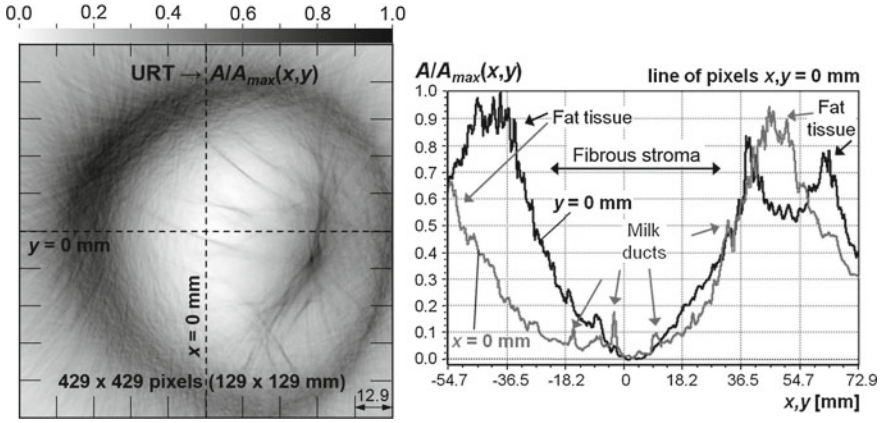


Fig. 8 URT image of the distribution of ultrasound scattering amplitude normalized values in the breast coronal section reconstructed from the in vivo examinations, and the distribution of pixel values of this image on 2 lines perpendicular to each other: (x, 0) *black* and (0, y) *grey*

1509 m/s in yolk. The water/albumen boundary is clearly visible, the albumen/yolk boundary is fuzzy and the structure of yolk is heterogeneous—the speed of sound decreases toward the edges of the yolk ball. There are also little ultrasound speed spikes in the yolk in the range of fractions of m/s, which suggests the presence of layers of white egg yolk and yellow egg yolk (Fig. 2). Thus it seems that quantitative characterization of lesions in breast tissue as compared to surrounding normal tissue and to distinguish e.g. malignant from benign lesions [23] is possible. UTT image of the distribution of ultrasound attenuation in the egg reflects the approxi-

mate real values of this parameter, but better exposes the boundaries of heterogeneous areas (Table 1; Fig. 4). Significant spikes and negative values of attenuation on the borders of such areas result from measurement errors, because also refraction, diffraction, and dispersion affects the measured amplitude of the ultrasonic wave pulses after passing, in addition to attenuation. However, this provides positive diagnostic information—the embryo area and yolk ball stratification are clearly visible in the egg structure (Figs. 2 and 4). URT image for egg (Fig. 5) carries the least information, but precisely exposes the water/albumen and albumen/yolk boundary as thin lines, whereas the yolk ball boundary is weak. This is understandable, because the reflection and scattering factor of ultrasound on the border of 2 media depends on the differences in their acoustic impedances ($R = (Z_2 - Z_1)/(Z_2 + Z_1)$), and this difference is negligible for albumen and yolk (Table 1).

UTT images of the cross-section of examined woman's breast in vivo (Figs. 6 and 7) enable to distinguish subcutaneous adipose tissue and glandular tissue, and heterogeneous areas can be observed in the structure of the latter that highlight demarcations of breast body lobes (Fig. 2). The size of the examined breast section in the UTT image of the ultrasound speed distribution is slightly smaller than the real one due to refraction of the ultrasonic wave at the transition of boundaries of water/fat/breast_tissue/fat/water, similar as in the case of standard ultrasound imaging [13]. The refraction which causes the multipath transitions of ultrasonic wave beam rays also affects the shape mapping errors in the UTT image of the ultrasound attenuation distribution—the ring of subcutaneous adipose tissue of breast is expanded and blurred. Due to moderate variations of the ultrasound speed in the breast structure (the highest for adipose tissue), these errors are not critical and can be further minimized by improving the algorithm of digital detection of acoustic parameters, the appropriate selection of water temperature and in the process of secondary calculation reconstructions [13]. You may also notice the conformity of acoustic parameters for the breast, but the spread of values is much higher than in the case of an egg due to its complicated structure. The reconstructed ultrasound speed is approximately 1515–1545 m/s in water, 1350–1450 m/s in the subcutaneous adipose tissue of breast (one shall expect the largest errors due to refraction here), 1450–1470 m/s in the internal adipose tissue of breast, 1510–1540 m/s in the glandular tissue of breast (Table 1; Fig. 6). UTT image of the distribution of ultrasound attenuation in the breast approximately reflects the real values of this parameter (see Table 1 and Fig. 7), however, it enables to distinguish the areas of glandular tissue, which significantly attenuate ultrasound waves compared to the internal adipose tissue. Subcutaneous adipose tissue around the breast has significantly excessive values of ultrasound attenuation due to refraction (Fig. 7).

Greater degree of distortion in the resulting cross-sectional images of the breast compared to the cross-sectional images of the egg result primarily from larger and more frequent refractions of ultrasonic wave beam rays caused by a complicated breast structure and a larger differences in acoustic parameters. In addition, the time of examination (currently 30 s) significantly affects the image distortion—the patient moves, the breast shakes as a result of heartbeat. In the case of egg sectional

examination, such errors result primarily from changes in measurement conditions due to temperature fluctuations and water movements.

When analyzing the qualitative URT image of breast, it should be noted that this method allows for 2-D visualization of the internal structures of the breast, which vary in echogenicity (and even 3-D when measuring many coronal sections), unlike in a standard ultrasonic imaging (Fig. 8). Shadows running radially from the center of the image towards its edges reflect connective tissue septa located between the lobes of the mammary gland and arranged radially around the nipple. These septa physiologically run to the skin and thoracic fascia (i.e., suspensory breast ligaments, Cooper ligaments), so they should be clearly visible on coronal sections of the breast, which is confirmed by the examination result (Fig. 8). The use of different modalities also allowed the visualization of different types of structures, with extended lumen converging around the nipple, whose courses correspond to the mammary ducts (Figs. 6, 7 and 8). Ducts leading out of breast follicular glands in a healthy breast form larger mammary ducts creating expansions-cisterns with a diameter of 5–9 mm, called bays sinuses, before the nipple aditus. Such structures are shown in Fig. 8. URT image of the breast as compared to the URT of the egg reveals many structural details, which result from large difference in acoustic impedance between adipose and glandular tissues in breast and ligaments, duct walls, and blood and lymph vessels (Table 1).

A visible mismatch in the sizes of UTT and URT breast images results currently from simplifying measurement procedures and the lack of appropriate corrections at the preliminary stage.

5 Summary

The innovative use of a combination of different imaging methods in the developed ultrasound tomography device (UTT, URT, and US in the prototype) allows the comprehensive characterization of breast tissue and provides complementary diagnostic information. It is also possible to obtain 3-D images allowing for presentation any cross-section of the breast by the intensity projection (MIP) or the multiplanar reconstruction (MPR) technique, analogous to CT. Compared to conventional breast ultrasonic imaging, UTT and URT methods enable to acquire a much larger amount of information whose clinical significance has not been well documented yet. Transmission images differentiate the heterogeneities of the examined structure with respect to their surroundings in a quantitative manner with good contrast resolution, but with blurred boundaries, which primarily allows their characterization. Reflective images visualize the examined structures more qualitatively, with a good spatial resolution, allowing in turn to estimate the sizes of heterogeneities due to sharpened boundaries. UTT images enable to visualize the adipose tissue and parenchyma in the breast structure along with lesions, and URT images visualize the fibrous stroma. Based on these results, it is expected that ultrasonic tomography with an appropriate fusion of UTT and URT image, as well as coronal US images can contribute

to the creation of a new standard for the diagnosis and treatment of breast cancer. The scheduled examinations of breasts of breastfeeding mothers, examinations using galactography, and the examinations of breasts with cancer will enable to refine the device and standardize the imaging results.

In contrast to X-ray mammography, the huge advantage of ultrasound tomography is not exposing the patient to ionizing radiation, therefore the study will be performed in almost all patients. It will also be possible to repeat the examination in order to test the efficiency of treatment. The examinations using the proposed device can be done also in patients with solid metal components in the body, which is contraindicated for magnetic resonance imaging. Another advantage of the proposed device is no need for mechanical compression of the breast, which is required in mammography and no need for an intravenous contrast agent, which is often required for the examination of breast using MRI. The limitation of the ultrasonic tomography are the contraindications in patient positioning on the abdomen (e.g. in the case of pregnancy) or too small size of the breast, preventing immersion in water under the operating circular array.

Acknowledgments Ultrasound breast imaging in vivo has been performed with the approval of the Bioethics Committee at the University of Warmia and Mazury (Resolution No. 22/2015), within the clinical research titled: "The evaluation of diagnostic value of breast examination using ultrasound transmission tomography and ultrasound reflection tomography".

References

1. Berrington de Gonzalez, A., Berg, C.D., Visvanathan, K., Robson, M.: Estimated risk of radiation-induced breast cancer from mammographic screening for young BRCA mutation carriers. J. Natl Cancer **I**(101), 205–209 (2009)
2. Duric, N., Littrup, P., Poulo, L., et al.: Detection of breast cancer with ultrasound tomography: first results with the Computed Ultrasound Risk Evaluation (CURE) prototype. Med. Phys. **34**(2), 773–785 (2007)
3. Gudra, T., Opielinski, K.J.: The ultrasonic probe for the investigating of internal object structure by ultrasound transmission tomography. Ultrasonics **44**(1–4), e679–e683 (2006)
4. Jain, A.K.: Fundamentals of Digital Image Processing. Prentice Hall, USA (1989)
5. Jirik, R., Peterlík, I., Ruiter, N., et al.: Sound-speed image reconstruction in sparse-aperture 3-D ultrasound transmission tomography. IEEE T. Ultrason. Ferr. **59**(2), 254–264 (2012)
6. Kak, A.C., Slaney, M.: Principles of Computerized Tomographic Imaging. IEEE Press, New York (1988)
7. Kerlikowske, K., Zhu, W., Hubbard, R.A., et al.: Outcomes of screening mammography by frequency, breast density, and postmenopausal hormone therapy. JAMA Intern. Med. **173**(9), 807–816 (2013)
8. Lambertz, C.K., Johnson, Ch.J., Mongomery, et al.: Premedication to reduce discomfort during screening mammography. Radiology **248**(3), 765–772 (2008)
9. Lehman, C.D., Isaacs, C., Schnall, M.D., et al.: Cancer yield of mammography, MR, and US in high-risk women: prospective multi-institution breast cancer screening study. Radiology **244**(2), 381–388 (2007)
10. Marmarelis, V.Z., Jeong, et al.: High-resolution 3-D imaging and tissue differentiation with transmission tomography. Acoust. Imag. **28**, 195–206 (2007)

11. O'Flynn, E.A.M., Ledger, A.E.W., deSouza, N.M.: Alternative screening for dense breasts: MRI. AJR **204**(2), W141–W149 (2015)
12. Opieliński, K.J.: Ultrasonic parameters of hen's egg. Molec. Quantum Acoust. **28**, 203–216 (2007)
13. Opieliński, K.J.: Application of Transmission of Ultrasonic Waves for Characterization and Imaging of Biological Media Structures. Printing House of Wroclaw University of Technology, Wroclaw (2011). (in Polish)
14. Opieliński, K.J.: Full angle ultrasound spatial compound imaging. In: Proceedings of 7th Forum Acusticum 2014, European Acoustics Association, Krakow (2014)
15. Opieliński, K.J., Gudra, T.: Ultrasonic transmission tomography. In: Sikora, J., Wojtowicz, S. (eds.) Industrial and Biological Tomography. Electrotechnical Institute, Warsaw (2010)
16. Opieliński, K.J., Pruchnicki, P., Gudra, T., et al.: Ultrasound transmission tomography imaging of structure of breast elastography phantom compared to US. CT MRI Arch. Acoust. **38**(3), 321–334 (2013)
17. Opieliński, K.J., Pruchnicki, P., Gudra, T., et al.: Imaging results of multi-modal ultrasound computerized tomography system designed for breast diagnosis. Comput. Med. Imag. Graph **46**, 83–94 (2015)
18. Opieliński, K.J., Pruchnicki, P., Gudra, T., Majewski, J.: Conclusions from a test of multi-modal ultrasound tomography research system designed for breast imaging. In: Proceedings of 7th Forum Acusticum 2014, European Acoustics Association, Krakow (2014)
19. Uematsu, T., Yuen, et al.: Comparison of magnetic resonance imaging, multidetector row computed tomography, ultrasonography, and mammography for tumor extension of breast cancer. Breast Cancer Res. Tr. **112**(3), 461–474 (2008)
20. Warren, R.M., Pointon, L., Thompson, D., et al.: Reading protocol for dynamic contrast-enhanced MR images of the breast: sensitivity and specificity analysis. Radiology **236**(3), 779–788 (2005)
21. Wiener, J.I., Schelling, K.J., Adami, C., Obuchowski, N.A.: Assessment of suspected breast cancer by MRI: a prospective clinical trial using a kinetic and morphologic analysis. AJR **184**(3), 878–886 (2005)
22. Wiskin, J., Borup, D., Johnson, S., et al.: Threedimensional nonlinear inverse scattering: Quantitative transmission algorithms, refraction corrected reflection, scanner design and clinical results. POMA **19**(075001) (2013)
23. Yang, J.N., Murphy, A.D., Madsen, E.L., et al.: A method for in vitro mapping of ultrasonic speed and density in breast tissue. Ultrasonic Imaging **13**, 91–109 (1991)

The Ultrasound Investigation of the Medial Head of Gastrocnemius Muscle

Ewelina Świątek-Najwer, Urszula Czajkowska and Ludomir J. Jankowski

Abstract The paper presents results of study of medial head of gastrocnemius muscle morphological parameters applying ultrasonography. The tests were performed on 6 probants on dominant limb. Basing on the ultrasound scans dataset the following parameters were measured: pennation angle, fiber bundles length, mean echo intensity inversely proportional to the muscle density, muscle thickness, anatomical and physiological cross-sectional area. The parameters were measured in 3 various ankle positions (95°, 110°, 130°), on 3 scans recorded in various locations (proximal, central and distal). The results of study were analysed in order to check the influence of ankle positions and scan position.

Keywords Biomechanics · Biomedical engineering · Sonography · Pennation angle · Anatomical cross-sectional area

1 Introduction

The human body contains 650 muscles, and their mass constitutes 30–50 % of total body mass. Biomechanical modelling of motor system requires data describing morphology, structure and functionality of muscles—actuators. The paper describes results of ultrasound imaging investigation of the medial head of gastrocnemius muscle (GCM) structure.

E. Świątek-Najwer (✉) · U. Czajkowska · L.J. Jankowski
Wroclaw University of Technology,
Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland
e-mail: ewelina.swiatek-najwer@pwr.wroc.pl
URL: http://biomech.pwr.wroc.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_17

1.1 Imaging of Muscle Structural and Functional Parameters

The magnetic resonance imaging is a gold standard to measure volume and cross section of muscles. It enables visualization of all muscles both superficial and deep, identification of congenital myopathies, dystrophies and edemas [9]. However the interests on more available and less expensive ultrasound scanning to measure muscles structural and functional parameters is constantly increasing.

Ultrasound imaging provides non-invasive, safe real-time imaging of soft-tissues shape and structure. After calibration of ultrasound probe and by tracking its position and orientation, that imaging might be also applied in three-dimensional measurement of bone anatomical landmarks, bone geometry and planning of bone deformity correction surgery [5, 12, 13, 16].

Ultrasound imaging of muscle enables measurement of pennation angle, muscle tear and hematoma diagnosis [9]. However analysing the recorded ultrasound scan, it is impossible to test the proximity outside the scan plane [14].

Bénard provided useful recommendations required to perform repeatible and objective ultrasound measurements. Crucial task is to register ultrasound scan in the fibers plane, because plane deviation influences measured angle and fiber length. The scans should be registered with high contrast fibers echo [3]. Barber et al. suggested that the scanning parameters such as focusing depth, power, gain should be adjusted to provide optimal visibility of colagen tissue [1].

Bénard et al. noticed high influence of ultrasound probe inclination on the result of scanning [3]. Also Whitaker et al. proved that during ultrasound examination change of probe orientation about 10° and pressing the probe on skin influences the recognized thickness value significantly [14]. Jung et al. performed measurement of gastrocnemius muscle at 25 % of tibia length on the knee side, while changing the orientation of ultrasound probe every 5° [8]. Chow et al. proved the diversity of bundle length depending on location of scans, measuring the parameter in 10 scannig areas [4]. Barber et al. also suggested to register scans while the probe is in longitudinal line of muscle, as the long axis of transducer is colinear with fibers acting line in the middle of muscle belly [2]. Barber et al. recommended to perform measurement on scan in the middle of it for optimally visible fiber [2].

In various works the researchers have measured following parameters of lower limb muscles: cross-sectional area, fiber length, pennation angle, muscle density, physiological cross-sectional area, thickness and width. These parameters are measured on scans registered in various locations and orientations and in various limb position. Often the scans are taken for patients lying and resting in prone position, or in sitting position with legs in complete extension and supported backs with possible access for probe [7].

In order to measure the cross-sectional area, thickness, width, muscle density and pennation angle of rectus femoris muscle both Worsley and Scanlon performed scan in two thirds distally from anterior superior iliac spine towards superior aspect of patella [10, 15]. Worsley registered the scans for various contraction conditions (10, 20, 30, 50, 75 % of Muscular Voluntary Contraction).In the case of vastus lateralis

the scan was registered in the half of most prominent point of trochanter major from the lateral condyle [15].

Scanlon proposed also a formula to calculate physiological cross-sectional area basing on density measured as reciprocal of mean echo intensity in the muscle area, cross-sectional area, fiber length and pennation angle [10]. Narici recommends to calculate physiological cross section basing on volume estimated by MRI [7].

Measurements of anatomical cross section areas using ultrasonography described Lima et al. [6]. Two scans for rectus femoris muscle (15 cm above the patella and in half of thigh) were registered crosswise with minimal pressing and the area of manually limited region was calulcated.

Bénard recommended measurement of pennation angle both to the upper and lower aponeurosis, by approximation of their directions to the second order polynomial and the direction of visible fiber to the line [3]. Barber suggested to calculate pennation angle to the upper aponeurosis to minimize error [2].

Narici et al. measured muscle in 3 regions (proximal, central and distal) applying MRI and ultrasound and in various conditions—in rest and in contraction until the MVC level [7]. For each region pennation angle, fiber length and muscle thickness were measured.

An important aspect while measuring the muscles is also the position of joint. Barber et al. performed measurements in three ankle positions ($-15°$, $0°$, $15°$), controlled by goniometer) applying 3D ultrasound system [2].

To summarize, the researchers are still looking for standards how to measure structure and morphology with ultrasound imaging. Magnetic Resonance Imaging is so far considered as the only objective technique and provides ability to analyse substantial data for biomechanical modelling [1].

2 Materials and Methods

The aim of work was to characterize medial head of gastrocnemius muscles morphologic parameters applying ultrasound imaging. The ultrasound measurements were performed on 6 healthy subjects characterized in Table 1. All probands were informed in detail about measurement procedure and signed the consent to participate in it.

During the tests the ultrasound EchoBlaster 128 system (Telemed, Lithuania) with linear ultrasound probe (width 80 mm, maximal frequency 10 MHz) was apllied. Measurements were performed in subject lying prone position, on dominating right lower limb, in three positions of ankle: 95°, 110° and 130° controlled by Noraxon mechanical goniometer. Proband's knee was stabilized in the lying prone position. The ultrasound imaging system was applied in order to register scans to evaluate following parameters: pennation angle, fiber bundle length, muscle thickness, anatomical cross-sectional area, mean echo intensity. In order to analyse anatomical cross-sectional area and mean echo intensity three transverse scans were recorded on the muscle belly in proximal, central and distal location. Afterwards three scans were

Table 1 Probands data

Proband	Height (m)	Weight (kg)	BMI	Sex	Daily activity
1	1.76	93	30.02	Man	Significant
2	1.58	45	18.03	Woman	Medium
3	1.67	75	26.89	Woman	Small
4	1.87	92	26.31	Man	Medium
5	1.87	86	24.59	Man	Significant
6	1.62	63	24.01	Woman	Significant

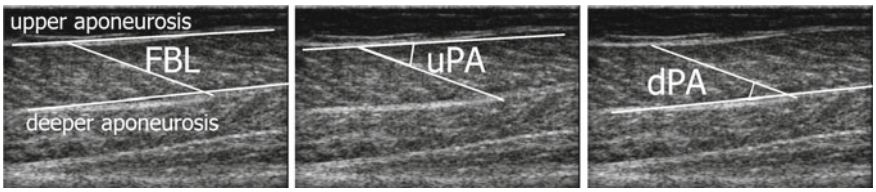


Fig. 1 Method of fiber bundle length, "upper" and "deeper aponeurosis" pennation angle evaluation

registered in the same position but after rotation about 90°. On these three scans muscle thickness, fiber bundle length, and pennation angle was evaluated. Measurements were performed for each ankle position: 95°, 110° and 130°.

All analyses of ultrasound scans were performed using ImageJ software [11]. Images were calibrated basing on probe width (80 mm), scanning depth (80 mm) and scan resolution (512 × 512 pixels). Fiber bundle length (FBL) was measured for fiber bundle with optimally visible direction as a distance between points where the fiber direction crosses deeper und upper aponeuroses. As the fiber bundle and aponeuroses directions are determined, the "upper aponeurosis" pennation angle (uPA) was measured between bundle direction and upper aponeurosis, and "deeper aponeurosis" pennation angle (dPA)—between bundle direction and lower aponeurosis (Fig. 1). The pennation angle (PA) was calculated as mean value from "upper" and "deeper aponeurosis" pennation angle.

The gastrocnemius muscle thickness (MT) was determined as the minor axis of ellipse approximated to the section of muscle on longitudinal scan. The anatomical cross-sectional area of gastrocnemius (ACSA) was determined as the area of figure limited manually in ImageJ software. The gastrocnemius muscle mean echo intesity (MEI) was measured basing on histogram of ACSA (Fig. 2).

The physiological cross-sectional area of gastrocnemius muscle (PCSA) was calculated basing on anatomical cross-sectional area, pennation angle, fiber bundle length and mean echo intensity. In the work of Scanlon et al. concerning rectus femoris muscle, the authors suggested to use in the formula anatomical cross-sectional area measured on scan in the half of thigh [10]. In this paper we compared results for all three scans to evaluate the influence of scan location on the obtained value.

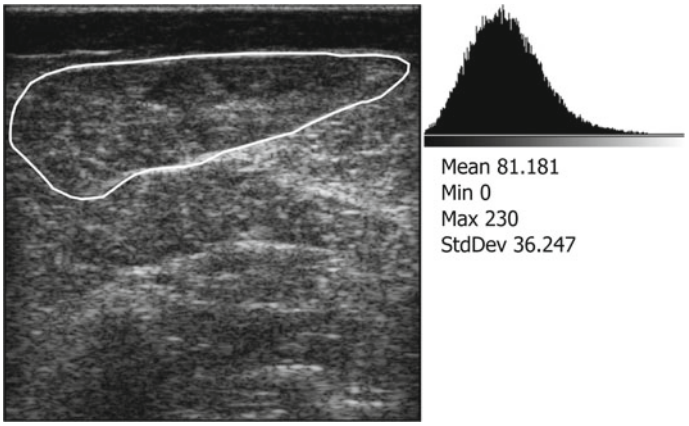


Fig. 2 Delineation of muscle contour and obtained echo intensity histogram

3 Results

The following pictures present the results of measurement of: fiber bundles length (FBL) (Fig. 3), muscle thickness (MT) (Fig. 4), "upper" (uPA) and "deeper" aponeurosis pennation angle (dPA) (Figs. 5, 6), pennation angle (PA) (Fig. 7), mean echo intensity (MEI) (Fig. 8), anatomical cross-sectional area—ACSA (Fig. 9) and physiological cross section area—PCSA (Fig. 10) on proximal, central and distal scan at 95°, 110° and 130° in ankle joint.

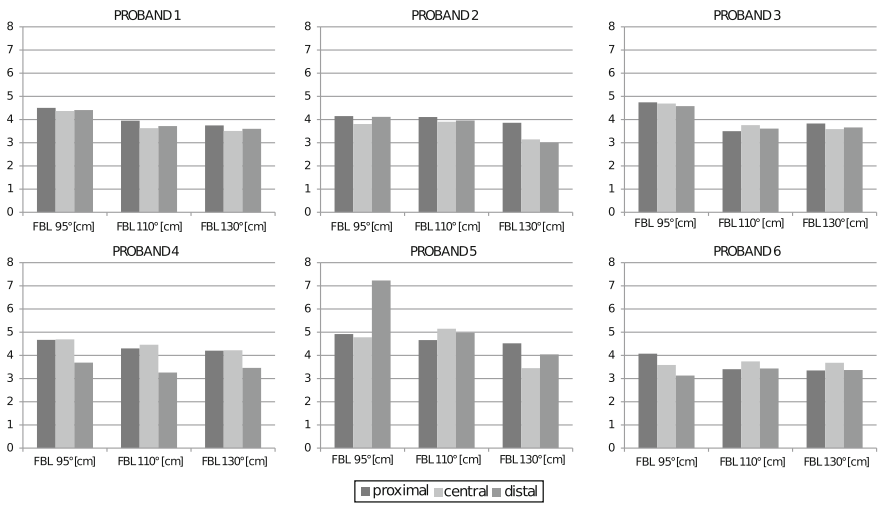


Fig. 3 Fiber bundles length (at 95°, 110°, 130°, proximal, central, distal scan)

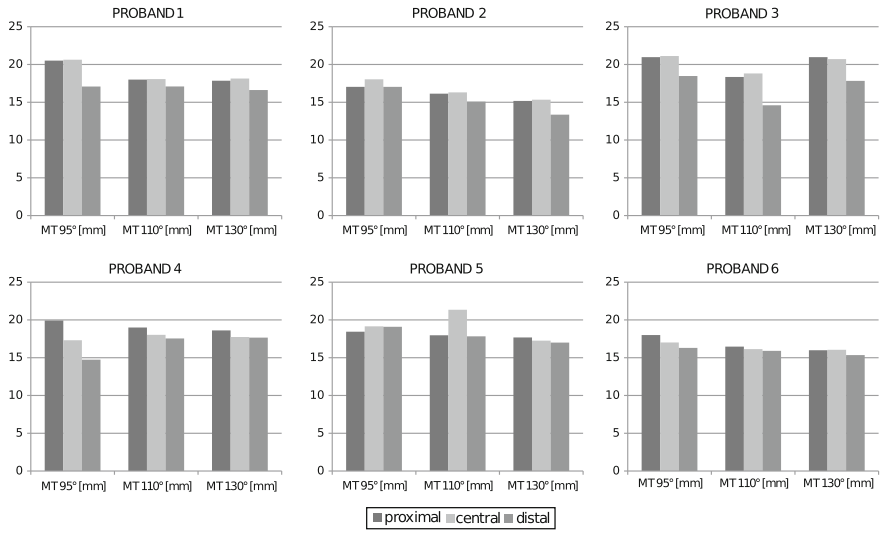


Fig. 4 Muscle thickness (at 95°, 110°, 130°, proximal, central, distal scan)

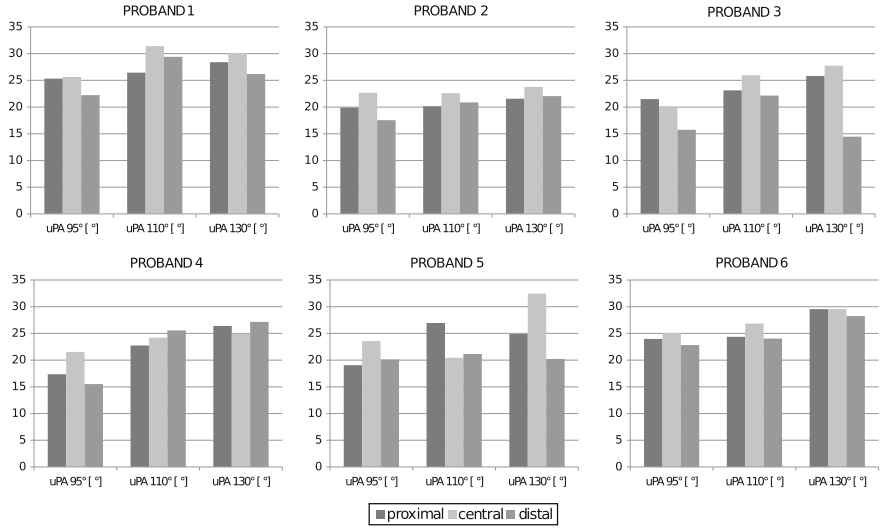


Fig. 5 "Upper aponeurosis" pennation angle (at 95°, 110°, 130°, proximal, central, distal scan)

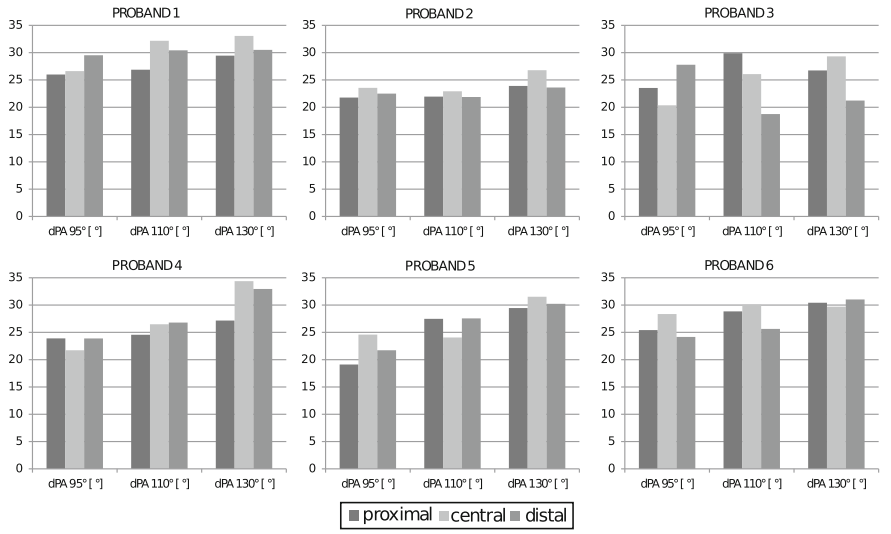


Fig. 6 "Deeper aponeurosis" pennation angle (at 95°, 110°, 130°, proximal, central, distal scan)

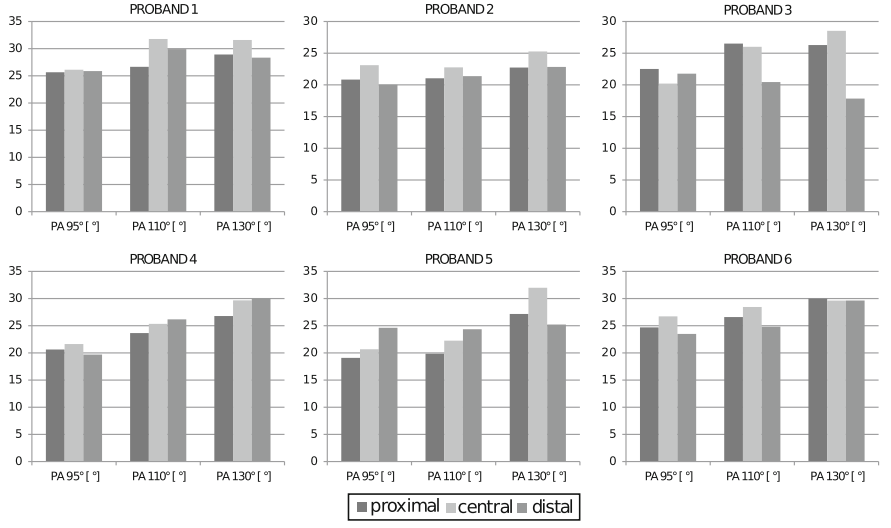


Fig. 7 Pennation angle (at 95°, 110°, 130°, proximal, central, distal scan)

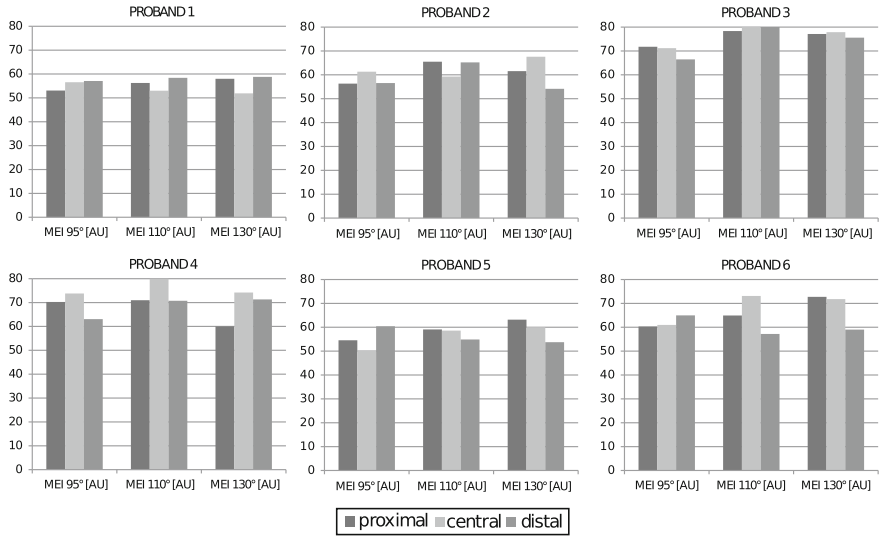


Fig. 8 Mean echo intensity measured at 95°, 110°, 130°, proximal, central, distal scan

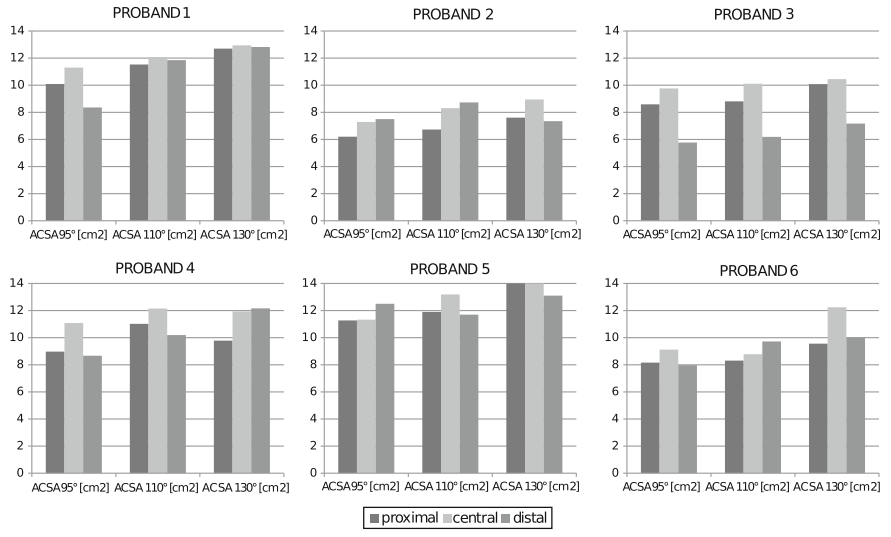


Fig. 9 ACSA measured at 95°, 110°, 130°, proximal, central, distal scan

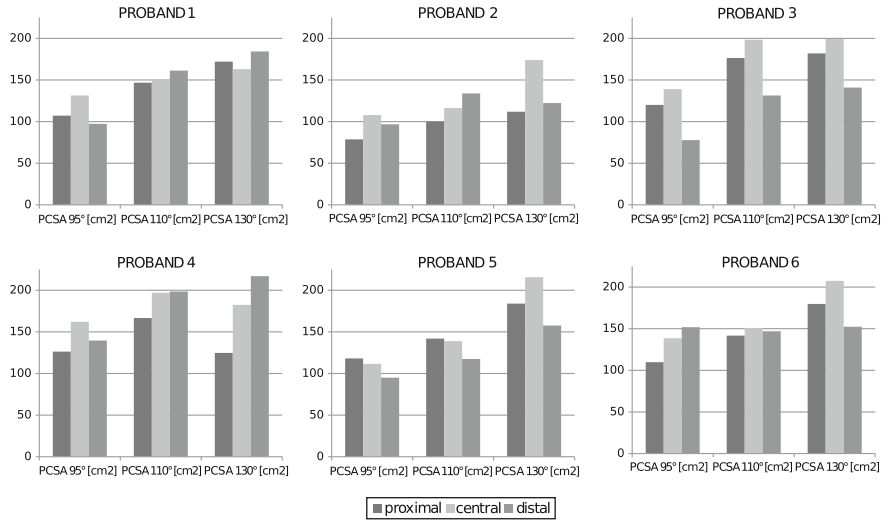


Fig. 10 PCSA measured at 95°, 110°, 130°, proximal, central, distal scan

Table 2 Mean values and standard deviations of GCM parameters

Parameter	Mean	SD
MEI (AU)	64.20	8.57
ACSA (cm^2)	10.06	2.20
FBL (mm)	40.35	6.89
MT (mm)	17.65	1.80
uPA (°)	23.73	4.00
dPA (°)	26.45	3.77
PA (°)	24.96	3.65
PCSA (cm^2)	146.17	35.23

Mean values and standard deviations for all measured parameters of the medial head of gastrocnemius muscle are presented in Table 2.

4 Discussion

The paper describes investigation of the medial head of gastrocnemius muscle morphology using ultrasound imaging. Obtained results of muscle morphological parameters measurements were analysed and referred to obtained by the other researchers.

The measured fiber bundle length equaled 40.3 mm ± 6.9 mm. Chow et al. reported the fiber bundle length in various scans locations as (44 . . . 50) (mm) (standard deviation for measurements performed in particular scan location equaled up to 10 mm) [4].

Chow et al. stated that the fiber bundle length should be lower for central scan position and higher for proximal and distal [4]. In our measurements that relationship was observed in at least one ankle position at four probands. Regarding the ankle angle, Narici confirmed that the fiber bundle length should decrease with increasing ankle angle [7]. In most of our results this relationship was also observed.

Both "upper aponeurosis" and "deeper aponeurosis" pennation angle demonstrated some variations. "Upper aponeurosis" pennation angle equaled 23.73° ± 3.99°. "Deeper aponeurosis" pennation angle equaled 26.45° ± 3.77°. The values are higher than results of Chow et al.: uPA (12 . . . 19) (°) and dPA (18 . . . 22) (°) [4]. Our results of averaged pennation angle measurements are close to obtained by Narici et al. (15 . . . 27) (°) while measuring the pennation angle at different ankle positions (90 . . . 150) (°) [7]. The mean pennation angle in our measurements equaled (22 . . . 27) (°) at three ankle positions 95°, 110°, 135°. Narici proved that the pennation angle should increase with increasing ankle angle [7]. In almost all our results that relationship was also observed. An important observation is also that in almost all cases the "deeper aponeurosis" pennation angle is higher than the "upper aponeurosis" pennation angle. It confirms the conclusion stated by Chow et al. [4]. The same author proved that the pennation angle measured on the middle scan should be higher by 3° than on proximal and distal scan, where the values should be similar [4]. In a part of our results this relationship was observed. All inconsistencies between our results and results of other researchers were likely caused by uprecise recognition of fiber direction or lower aponeurosis. The deviation of 2° was fairly possible. The topic is described in detail further in this section.

The measured muscle thickness value depended on scanning location and ankle joint position. The mean values of muscle thickness measured at particular probands (15 . . . 18) (mm) were slightly higher than obtained by Chow et al. [4] (12 . . . 15) (mm). The muscle thickness values obtained for proximal and distal scans positions according to Chow et al. [4] should be lower than the one obtained for the central scan position. That relationship was observed in a part of our results. In some cases the results were also similar to measured on scans in other locations. In most of our results the muscle thickness decreased with increasing ankle angle. Such relationship is logically justified, because as the angle increases the muscle is extending and its thickness gets reduced. The reason of any deviations in case of muscle thickness evaluation might have been strongly related with method of ellipse adjustement to the muscle area. That procedure can introduce substantial deviation up to 2 mm. It depends for example if the aponeuroses are classified as belonging to muscle region, and the echo of aponeuroses may be thick in certain probe inclination.

The results of these three described parameters (fiber bundle length, pennation angle and muscle thickness) measurements might have been burdened with unprecisely recognized deeper aponeurosis. The deeper aponeurosis, where the gastrocnemius muscle contacts with the soleus muscle, in some cases is complicated to delineate. That conclusion was also clarified at work of Bénard et al. [3]. The diversity is also caused by limited resolution of image. It is possible that the frequency was not enough adjusted to provide high quality visualization of deeper structures. That may be an explanation for discrepancies. A critical factor was also problematic

recognition of fiber bundle direction on certain scans. That may be caused by limited resolution, but also improper scan location while registering the fiber bundles direction and length. Bénard emphasized that the scanning plane should be parallel to fibers direction and measured direction and length of fiber bundles highly depends on plane of scanning [3]. Higher than $(-5 \ldots 5)$ (°) variation of scanning plane inclination introduce up to 25 % deviation in measured fiber bundle length [3].

The anatomical cross-sectional area measurement revealed strong dependence on angle in the ankle joint. The parameter was increasing with the angle in the ankle joint. The ACSA value was in 2/3 cases higher when measured on central scan than measured on proximal and distal scan. Unfortunately we could not compare the numerical results with other authors because these studies were performed on other muscles (for example Lima measured ACSA of the rectus femoris muscle [6]).

The mean echo intensity value is a parameter inversely proportional to the muscle density. The mean values of MEI at probands varied to a maximum of 15 AU (at Proband 3: MEI equaled 70 AU, at Proband 1: MEI equaled 55 AU). The difference between patients are rational and may be results of differences in hydratation and local concentration of collagen in tissue. The standard deviation of MEI measurement on various scans and in various ankle positions at particular probands varied $(2.5 \ldots 6.2)$ (AU). The mean echo intensity measured by Scanlon et al. for rectus femoris muscle equaled $(66.5 \ldots 93.4)$ (AU) [10]. In fact the mean echo intensity varies for particular muscles, however it seems that obtained values are likely justified.

The physiological cross-sectional area was increasing with the angle in the ankle joint. Narici proved that the PCSA value should increase by 34.8 % from relax to contraction [7]. In our study the increase varied between scans locations and probands significantly. The reason of that diversity is that the scans were not recorded in exactly the same locations for each ankle position. Indeed an useful development would be to regsiter the scans in exactly the same locations due to a tracking system application. Another issue is that measured values of PCSA were overestimated comparing to obtained by Narici et al. $(40 \ldots 60)$ (cm^2) [7]. The reason was probable misinterpretation of ACSA value, because remaining parameters, applied to calculate the PCSA, were similar to obtained by other researchers.

To summarize, the study enabled definition of morphologic parameters of medial head of gastrocnemius muscle. Most of parameters (except ACSA and PCSA) are similar to obtained by other researchers. Performed investigation also proved that the result of measurement highly depends on scan location and stabilized position of ankle. Moreover ultrasonography demands high experience while recording scan and its analysis. The researcher needs to find appropriate location of scan to be able to evaluate the morphological parameter and perform repeatable measurements. That conclusion does not diminish importance and advantages of ultrasound imaging in biomechanical study. That noninvasive, easy available, non-expensive imaging has a great potential, but the operator needs to be aware of diversity of results depending on conditions of measurement and system settings.

Acknowledgments 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. This work benefited from the use of the ImageJ public domain Java image processing program and available at www.imagej.nih.gov/ij/. This work was supported by Wroclaw University of Technology in frames of financial support of statutory activities in 2014 donated by Ministry of Science and Higher Education—inter alia: financing for conducting research and development work and related tasks, contributing to the development of young scientists and doctoral students. Grant no: B40051/K1008.

References

1. Barber, L., Barrett, R., Lichtwark, G.: Validation of a freehand 3D ultrasound system for morphological measures of the medial gastrocnemius muscle. J. Biomech. **42**(9), 1313–1319 (2009)
2. Barber, L., Hastings-Ison, T., Baker, R.: Medial gastrocnemius muscle volume and fascicle length in children aged 2 to 5 years with cerebral palsy. Dev. Med. Child Neurol. **53**, 543–548 (2011)
3. Bénard, M.R., Becher, J.G., Harlaar, J.: Anatomical information is needed in ultrasound imaging of muscle to avoid potentially substantial errors in measurement of muscle geometry. Muscle Nerve **39**(5), 652–665 (2009)
4. Chow, R.S., Medri, M.K., Martin, D.C.: Sonographic studies of human soleus and gastrocnemius muscle architecture: gender variability. Eur. J. Appl. Physiol. **82**(3), 236–244 (2000)
5. Keppler, P., Krysztoforski, K., Świątek-Najwer, E.: A new experimental measurement and planning tool for sonographic-assisted navigation. Orthopedics **30**(10 Supplement), 144–147 (2007)
6. Lima, K.M.M., Matta, T.T., Oliveira, L.F.: Reliability of the rectus femoris muscle cross-sectional area measurements by ultrasonography. Clin. Physiol. Funct. Imaging **32**, 221–226 (2012)
7. Narici, M.V., Binzoni, T., Hiltbrand, E.: In vivo human gastrocnemius architecture with changing joint angle at rest and during graded isometric contraction. J. Physiol. **496**(1), 287–297 (1996)
8. Park, E.S., Sim, E., Rha, D.-W., Jung, S.: Estimation of gastrocnemius muscle volume using ultrasonography in children with spastic cerebral palsy. Yonsei Med. J. **55**(4), 1115–1122 (2014)
9. Pillen, S.: Skeletal muscle ultrasound. Neurol. Res. **33**, 1016–1024 (2011)
10. Scanlon, T.C., Fragala, M.S., Stout, J.R.: Muscle architecture and strength: adaptations to short-term resistance training in older adults. Muscle Nerve **49**(4), 584–592 (2014)
11. Schneider, C.A., Rasband, W.S., Eliceiri, K.W.: NIH Image to ImageJ: 25 years of image analysis. Nature Methods **9**, 671–675 (2012)
12. Świątek-Najwer, E., Krysztoforski, K., Dragan, S.L.: The investigation of the lower limb geometry using 3D sonography and magnetic resonance. Measurement **45**(4), 702–710 (2012)
13. Świątek-Najwer, E., Otto, K., Krowicki, P.: 3D bone shape modelling basing on dataset recorded by ultrasound free-hand navigated probe information technologies in biomedicine 4. Book Series: Advances in Intelligent Systems and Computing, vol. 284, pp. 45–56 (2014)
14. Whittaker, J.L., Stokes, M.: Ultrasound imaging and muscle function. J. Orthop. Sports Phys. Ther. **41**(8), 572–580 (2011)
15. Worsley, P., Warner, M., Delaney, S.: The application of ultrasound imaging in the musculoskeletal modeling process. In: 55th Annual Meeting of the Orthopaedic Research Society, Las Vegas, USA, 22–25 Feb 2009, pp. 1501–1501 (2009)
16. Żuk M., Świątek-Najwer E., Pezowicz C.: Hip joint centre localization: evaluation of formal methods and effects on joint kinematics. Information technologies in biomedicine 4. Book Series: Advances in Intelligent Systems and Computing, vol. 284, pp. 57–67 (2014)

The Use of Doppler Effect for Tomographic Tissue Imaging with Omnidirectional Acoustic Data Acquisition

Tomasz Świetlik and Krzysztof J. Opieliński

Abstract The imaging of blood flow in blood vessels is the classic use of the Doppler effect in medicine. This method involves a transceiver ultrasonic probe which generates the acoustic signal in the form of a continuous wave which is reflected by moving biological components and returns with changed frequency. This effect can also be used to obtain the image of stationary tissue section, by moving the probe around the object. Such an imaging method is called Doppler Tomography (DT), or Continuous Wave Ultrasonic Tomography (CWUT). Currently, there is no comprehensive study of this method in the literature. This paper shows a simulation model which allows the reconstruction of the image consisting of infinitely small objects which evenly disperse the ultrasound wave. In order to assess the possibility of using DT in medicine, the obtained images were analyzed based on the parameters input to the reconstruction algorithm. Additionally, the exemplary waveform of a useful signal recorded of the real object of a small diameter is included.

Keywords Doppler tomography · Computational model · Image reconstruction · Continuous wave ultrasonic tomography

1 Introduction

A well-known use of Doppler effect in medicine is the measurement and imaging of blood flow velocity [11]. This method uses ultrasonic transceiver probe consists with a two piezoelectric transducers. The wave with a frequency of f_B generated by the transmitter, upon the penetration into the tissue is reflected from red blood cells moving in the blood vessel, and then returns with the change frequency of f_R. The difference between frequencies f_B and f_R is called the Doppler frequency f_D. The dependence between the Doppler frequency, transmitted frequency, received

T. Świetlik (✉) · K.J. Opieliński
Chair of Acoustics and Multimedia, Faculty of Electronics,
Wroclaw University of Technology, Wroclaw, Poland
e-mail: tomasz.swietlik@pwr.edu.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_18

frequency, and blood flow velocity v is shown in (1), where θ is the probe tilt angle with respect to the blood vessel, and c is the velocity of propagation of ultrasonic wave in the examined medium (flowing blood):

$$f_R - f_B = f_D = \frac{2 \cdot f_B \cdot v \cdot \cos(\theta)}{c} . \tag{1}$$

It is also possible to use the Doppler effect in an unconventional way for a different purpose, which is the tomographic imaging of tissue structure. The Doppler tomography (DT) also uses a double transducer ultrasonic transceiver probe placed in water, which generates a quasi-planar ultrasonic wave. However, in contrast to conventional methods, the probe is moved around or along the examined object, which is also placed in water. Due to the movement of the probe, the Doppler frequency will be added to or subtracted from the frequency of signals reflected from still inhomogeneities found in the object, according to the Doppler effect. This enables imaging the inclusions which effectively scatter the ultrasound wave inside tissue. There are two ways of DT data acquisition. In the first method, the probe moves along the examined stationary object. It is called linear geometry. Another method of DT data acquisition is circular geometry, in which the transducers move around the imaged stationary tissue section [5]. Due to the higher spatial resolution of the reconstructed image, the simplification of computer simulations and the simplicity of the measuring system, circular DT geometry is used in this paper. Currently, there is no comprehensive research on Doppler tomography in the literature, in particular, it lacks simulations, which would allow the analysis of the possibilities of using this method in medicine. To this end, this paper presents the method of DT measurement simulation and the results of image reconstruction with an example measurement of the actual signal required for imaging. The aim of the work is to test the visualization of compact inhomogenities submerged in a water-like medium using the modern, safe and non-invasive ultrasonic DT method. It allows to predict about applications in medical diagnostics. As a long-term result it may be possible to elaborate a new measurement device for *in vivo* imaging the interior of the human body (e.g. detection of pathological lesions in female breasts, examination of the bone of limbs).

2 Doppler Tomography for Circular Geometry Method

2.1 Measurement Data and Acquisition

Double transducer ultrasonic transceiver probe moves around the object for the circular geometry in Doppler tomography. It should be noted, however, that the same image will be obtained when the examined object rotates around its own axis, while the probe remains stationary, due to the nature of the Doppler effect. In the latter case, it is much easier to understand the concept of the DT method, and therefore it

has been used in this paper. Of course, we should keep in mind that in order to be able to use the device based on this method to study tissue *in vivo*, the probe should rotate around the examined object.

Figure 1a shows the acquisition of the data needed to reconstruct the object image in Doppler tomography for circular geometry. It shows the section of tissue in which there are two inclusions a and b which scatter the ultrasound wave well [5]. Note that in the rotation, the velocity components in direction of wave propagation for these two objects will be equal to, respectively, $\boldsymbol{v_a}$ and $\boldsymbol{v_b}$. Doppler frequency will be added to or subtracted from the frequency of the wave reflected from inclusions depending on the sense of said movement velocities according to the Doppler effect. In this case, at an angle of rotation of θ, circular speed of rotation ω_{turn}, transmission frequency of the ultrasonic wave f_B, and the propagation velocity of ultrasonic waves in tissue c, the formula for Doppler frequency which is generated from the inclusion located at a distance r from the center of rotation becomes:

$$f_D = 2 \cdot f_B \cdot \omega_{turn} \cdot r \cdot \frac{\cos(\theta)}{c} . \tag{2}$$

Formula (2) can be converted to a form which clearly shows the relationship between the Doppler frequency generated by the inclusion, and the velocity component of rotation speed in direction of wave propagation at a given angle of rotation. This relationship is shown in the formula (3):

$$v = \frac{f_D}{2 \cdot f_B} \cdot \frac{c}{\cos(\theta)} . \tag{3}$$

So it should be considered how the velocity component of movement changes in the area of image reconstruction. Two statements are true, which can describe the value of v depending on the position of the rotating scatter at a constant rotation angle. First, velocity components v have the same value (Fig. 1b) in the points on vertical lines (along the wave propagation). This also means that Doppler frequencies generated by so arranged inclusions will be the same (3). The second true statement is that as you move away from the center of rotation, the value $\boldsymbol{v}$ increases to the maximum $\boldsymbol{v}_{max}$ (Fig. 1b).

Knowing the speed of rotation and the diameter of the reconstruction area, we can easily determine the value $\boldsymbol{v}_{max}$, which in turn (using the formula (3)) allows to calculate of the maximum Doppler frequency $f_{D_{max}}$.

The second step involves determining the so-called sinogram [3]. It is an array with rows containing the applicable sums of amplitudes of Doppler frequencies, while the columns change according to the angle of rotation at which the Doppler signal is received. In order to fill such an array, we must first evenly separate subbands of a predetermined width Δf_D across the whole range $[-f_{D_{max}}, f_{D_{max}}]$. The parameter of Δf_D is called Doppler bandwidth. For a given rotation angle, the sum of amplitudes of the spectrum of Doppler frequencies derived from inhomogeneities occurring in the respective vertical line of the reconstructed area should appear in a

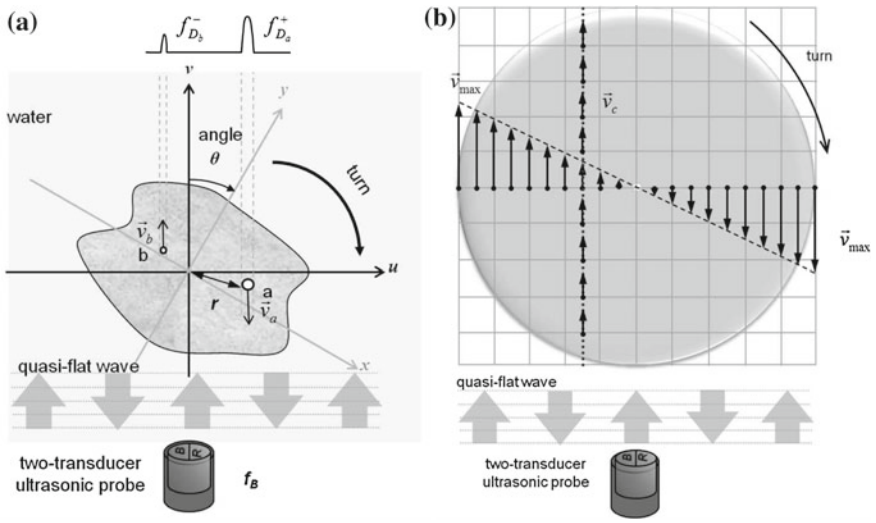


Fig. 1 An example of the acquisition of data necessary for image reconstruction in a circular geometry for DT (**a**) and the distribution of velocity components of the rotation speed in the direction of wave propagation in the image reconstruction area (**b**) [5]

given band. It is such information which is contained in Doppler signals. So, to get the sinogram line at a given angle of rotation (so-called projection [3]), said signal should be divided into Doppler bands in the range of $[-f_{D_{max}}, f_{D_{max}}]$ and write its amplitudes to the appropriate array cells. Thus prepared sinogram can be used to reconstruct an image in the similar way as in X-ray tomography.

2.2 *Example of a Sinogram Creation*

This section shows a simple example of how to create a sinogram in DT for the circular geometry. To make the example simpler and more understandable, we assume that we are dealing with imaging of objects which are infinitely small and scatter the ultrasound wave identically in each direction. They will be called pin objects in this paper. For each Doppler band in the sinogram, only the existence of a Doppler frequency derived from a given rotating object at a given angle of rotation has been recorded. This situation is shown in Fig. 2, which shows three pin objects placed on a rotating platform. The platform is rotated at a speed of 3.6 s per rotation of the corresponding $\omega_{turn} = 1.745$ rad/s, and its diameter is 10 cm. Objects a, b, c are placed at the distance of 1.5, 2.83, and 4 cm from the center of rotation. The transmitted wave frequency equal to 4.7 MHz has been used in the calculation, the same as in the measurement system shown in Fig. 5. The value between successive angles of rotation at which the Doppler signal was recorded is 7.2°. It can be assumed that the recording

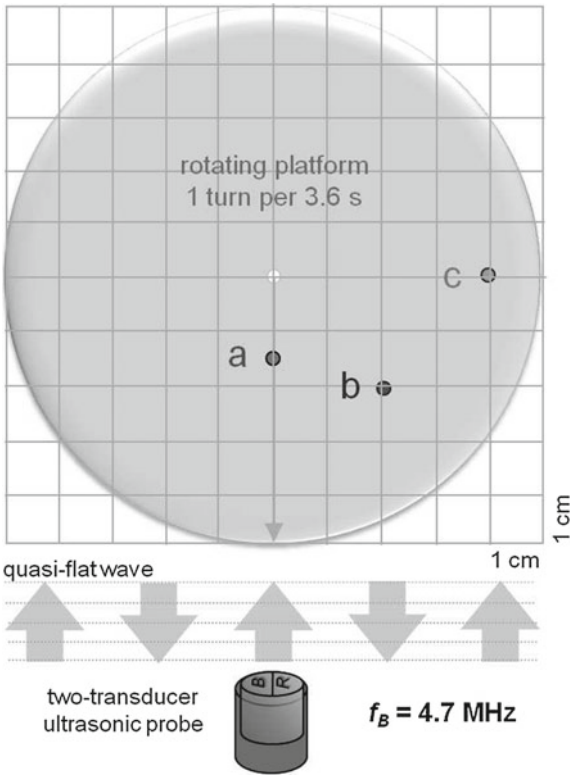


Fig. 2 The layout of three pin objects for which the sinogram has been simulated

time of the signal by the probe has a negligible impact on image reconstruction both in simulations and in actual measurements, because the velocity of ultrasound in water is about 1500 m/s, and the diameter of the imaged area is several centimeters. Figure 3 shows the result of sinogram calculation. Note that the maximum Doppler frequency is 510 Hz on the sinogram (Fig. 3). Doppler bandwidth has been set at 20 Hz, giving a total of 51 bands (so-called tomographic rays of projection [3]. The colors which determine the existence of a Doppler frequency correspond to the objects shown in Fig. 2.

2.3 Image Reconstruction Algorithm

There are two methods for image reconstruction in Doppler tomography [5, 7]. The first is a coherent method in which the Doppler signal is shown in the complex form and includes the amplitude and phase. In the second method, referred to as incoherent, the said signal is real. In the incoherent method, well-known and fast algorithms used in standard X-ray tomography (CT) may be adapted for image reconstruction. This was the reason for choosing this method of DT image reconstruction in this paper.

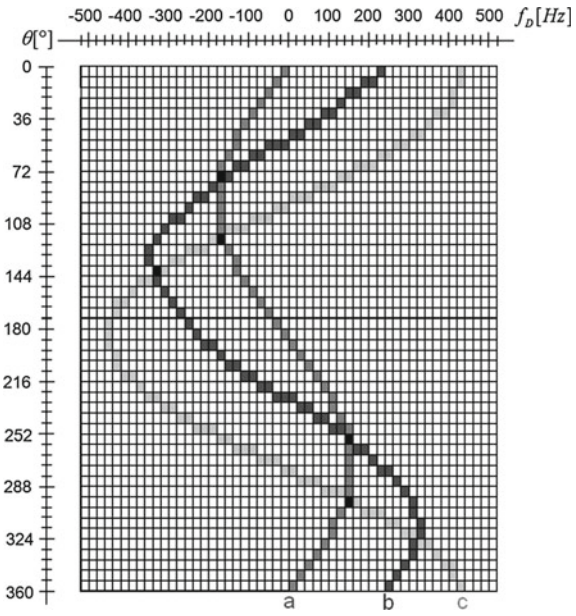


Fig. 3 Calculated example of the sinogram for three elements *a*, *b*, *c* shown in Fig. 2

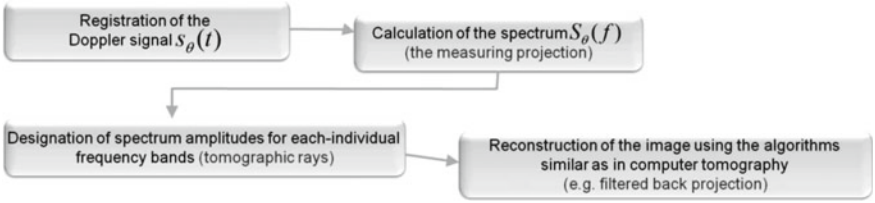


Fig. 4 Incoherent image reconstruction scheme for DT

Figure 4 shows the block scheme of incoherent image reconstruction for Doppler tomography in a circular geometry.

The first step of image reconstruction is to record the demodulated signals reflected, so those that contain only the Doppler frequencies $s_\theta(t)$ (Doppler signal). Then, the spectrum of such a signal $S_\theta(t)$ is determined for different angles of rotation θ and thus it creates tomographic projections. In the third step, the maximum range of Doppler frequency changes is calculated and divided evenly into sub-ranges (Doppler bands), in which Doppler spectrum amplitudes are recorded at a fixed angle. This process is also called the determination of tomographic measurement rays in the literature [3, 10]. The last step involves the reconstruction of the image using one of the algorithms used in CT [3]. Filtered back projection (FBP) algorithm with Ram-Lak filter [4, 10] is used for this purpose, being one of the fastest and most accurate in this family.

3 Recording of the Real Doppler Signal

3.1 Measurement System

In terms of obtaining real images of the objects it is most important to acquire a good quality Doppler signal. This signal is the basis for calculating the sinogram, which is then used to reconstruct an image. Pin objects can be used to investigate how small are the heterogeneities which still produce a measurable Doppler signal, and what are DT images of such objects. Measurement system shown in Fig. 5 has been developed for the generation and recording of the signal from the rotating pin. An ultrasonic blood flow detector is the most important device used for the construction of the system.

It is due to this matter that the reflected signal recorded by the probe is amplified and converted to a Doppler signal. Double transducer ultrasonic transceiver probe is placed in a tank filled with distilled water so as to receive only the signal reflected from the rotating pin. The rotating platform, on which the test object is placed is connected to a stepper motor. Power supply, controller motor, and function generator force the rotation at a predetermined speed. By sending the TTL signal of the appropriate frequency to the controller, we can determine the rate at which the platform rotates. In this case, the frequency of 200 Hz means one turn per second. The Doppler signal is recorded by an oscilloscope data acquisition card installed in a computer.

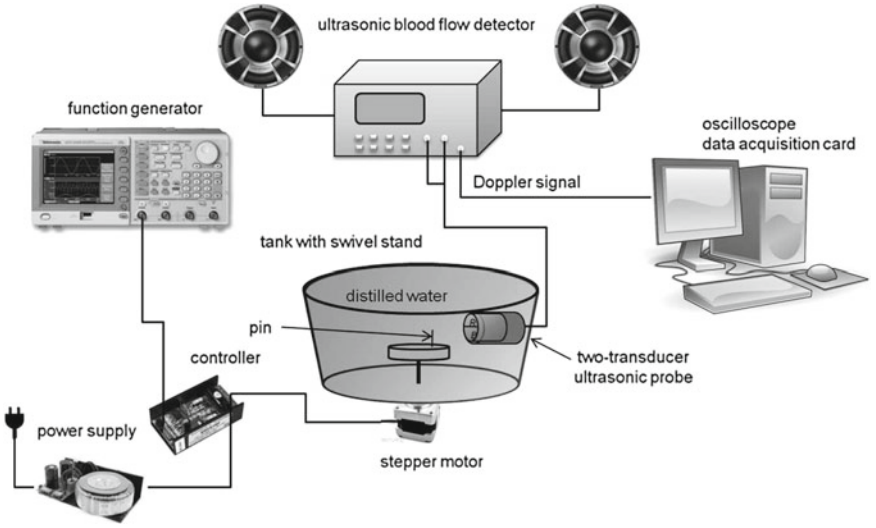


Fig. 5 Measuring system for the recording of the Doppler signal generated by the rotating pin

3.2 Example of Measurement

Doppler signal originating from three objects has been recorded. Two of them are pins with a diameter of 3 and 1 mm. The third test object was a filament with a very small diameter of 0.1 mm. Tests were performed at a speed of rotation of the platform equal to 2 rotations per second, which is a sufficient value so that Doppler frequency is at several hundred hertz for most angles of rotation. This enables to precisely determine this value. The diameter of the rotation platform was 5 cm, and the pins were placed 1.6 cm from the center of rotation. The frequency of the transmitted wave was 4.7 MHz. Figure 6a shows the system in which the measurements have been performed. Figure 6b shows the results of Doppler signal measurement for a pin with the diameter of 1 mm at an angle of rotation equal to 285°.

It shows clearly that the amplitude of the resulting signal varies in the vicinity of 8 V_{pp}. This is the case for all angles of rotation and for a pin with the diameter of 3 mm, as well. When measuring the filament with a diameter of 0.1 mm, the amplitude decreased and oscillated around 6 V_{pp}. Nevertheless, the quality of Doppler signal was still very good in terms of signal-to-noise ratio. It follows that it is theoretically possible to obtain an image of 3 and 1 mm pins and the filament with the diameter of 0.1 mm.

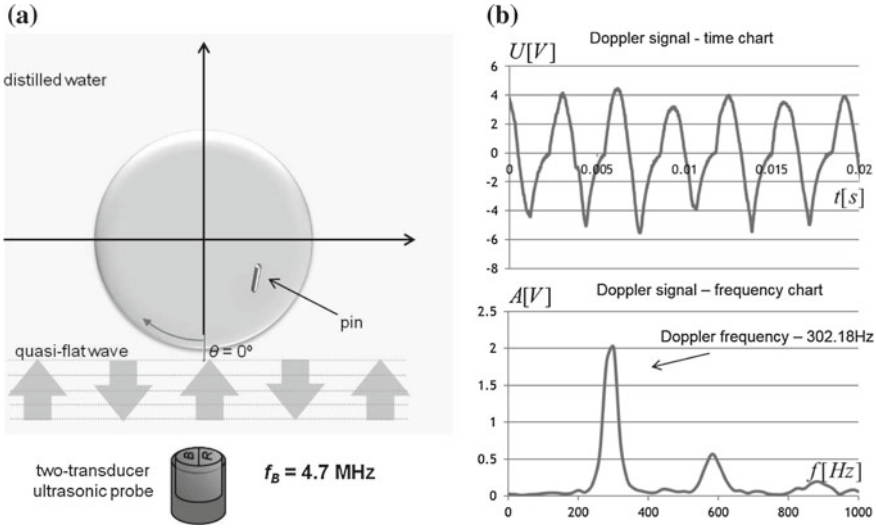


Fig. 6 Measuring system for Doppler signal from the rotating pin (**a**), and the results of measurements of the Doppler signal in time and frequency for the angle of rotation equal to 285° (**b**)

4 DT Imaging of Scattering Objects Using Simulated Data

4.1 Parameters and Assumptions

We demonstrated that it is theoretically possible to obtain a Doppler signal from small objects with a diameter of 3 and 1 mm, or even 0.1 mm. We shall now examine for which parameters of the reconstruction algorithm we will obtain the image for such objects. A computer simulation has been developed for this purpose. It assumed that the imaged object is infinitely small and scatters the ultrasound wave identically in each direction. This enables to analyze the blur of the point image caused only by the algorithm parameters. In addition, it is possible to optimize these parameters with the exception of errors related to the measurement process and wave propagation phenomena such as diffraction, refraction, and attenuation.

It is assumed in the simulation that the imaged object is placed on a rotating platform immersed in distilled water at 20 °C at coordinates $x = 0$, $y = -2$ cm. Additionally, the speed of rotation of the platform was set at 2 turns per second, the transmission frequency of the ultrasonic wave at 4.7 MHz, and the imaging area diameter at 10 cm. The Doppler frequency at a given angle of rotation for the pin has been determined from the formula (2), and the velocity of propagation in distilled water—from the empirical formula [6], which in this case gave the result of $c = 1482.38$ m/s. When creating sinograms for FFT operations in the FBP image reconstruction algorithm, the number of samples has been set at 512, and the quadrature size of the grid of pixels of the reconstructed image is equal to an odd number of measurement rays (Doppler bands) each time [10].

4.2 Effect of Number of Doppler Bands on DT Imaging

Based on the real measurements made at the measurement setup shown in Fig. 5, the number of Doppler bands have been determined, for which one should examine the effect on the quality of image reconstruction. This is necessary because too many bands can significantly decrease the precision of Doppler frequency determination and the number of angles is limited due to the recording time of the reflected signal. The determined range starts from 51 Doppler bands (so-called measurement rays [3]) and ends at 501. The values changed every 50. The number of signal recording angles during one turn was 400. Figure 7 shows the results of image reconstruction for ten numbers of Doppler bands in the range of 51–501. It may be noted that with the increase in the number of bands, the resolution and dynamics of the image increases, and the diameter of the pin cross-section image decreases.

The method of evaluating the quality of reconstructed images for the case of 301 Doppler bands is shown in Fig. 8. In Fig. 8a, the horizontal line shows the location where values of individual pixels were recorded. In turn, Fig. 8b shows the chart of normalized pixel values, where two key values are marked. The first of these is about

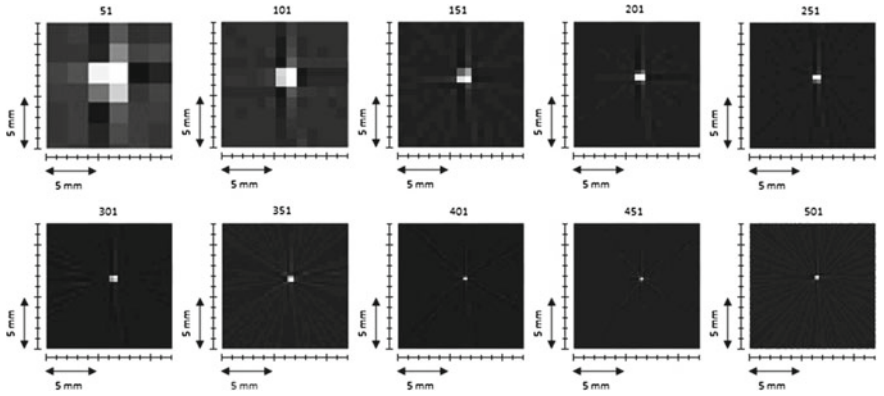


Fig. 7 Image reconstruction for the number of Doppler bands in the range of 51 to 501

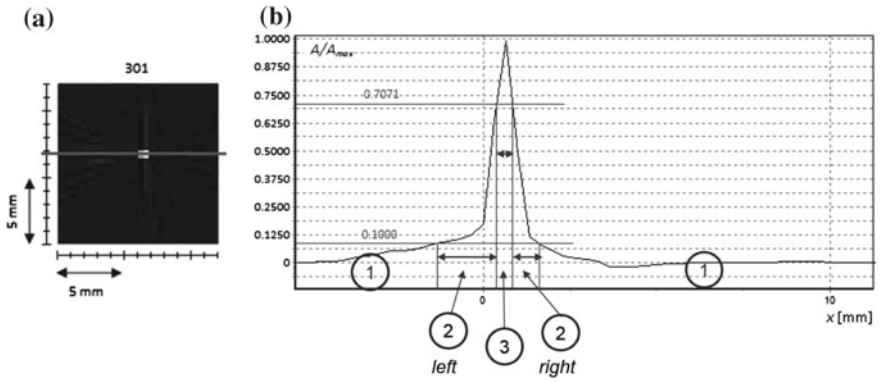


Fig. 8 The illustration of the recording method for pixel values in the image for the case of 301 bands (**a**), and the chart of pixel values in the reconstructed DT image for the case of 301 Doppler bands (**b**)

0.7 and corresponds to the 3 dB decrease, while the other is exactly 0.1, which is the decrease of 10 dB (noise cutoff level). This divides the chart into three types of areas. The area marked 3 (the pixel value above 3 dB) corresponds to the portion of the image, in which a pin can be seen very clearly, and the number 2 (between 3 and 10 dB) corresponds to the portion in which the image is blurred. The pin is almost not visible in the area number 1 (below 10 dB).

Figure 9 shows the charts of diameters of areas of the pin that is well visible in the reconstructed images (area 3) and the blur width of the pin image (area 2) for all examined numbers of Doppler bands. Its clearly shows that the width of areas decreases exponentially with the increasing number of bands. Values for the area where the pin image is very good (area 3) range from approximately 3 mm for 51 to 0.15 mm for 501 Doppler bands.

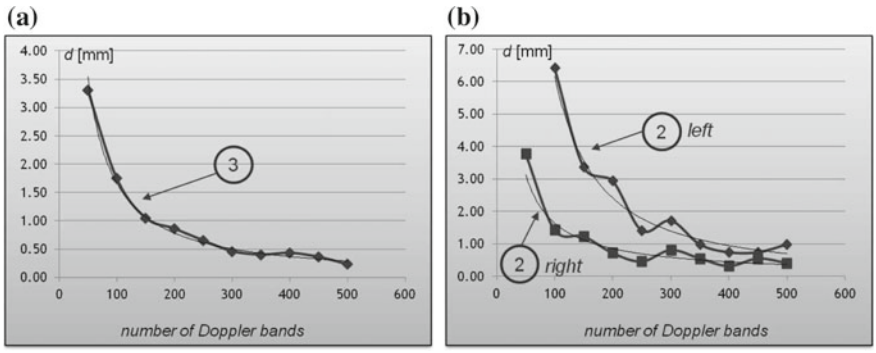


Fig. 9 The chart of widths of area 3 for the examined numbers of Doppler bands (**a**), and the chart of widths of 2 (*left* and *right*) areas for the examined number of Doppler bands (**b**)

Since the pin diameter is infinitely small, this provides a view of how theoretically small objects with what amount of blur can be imaged by DT in the circular geometry using the filtered back projection algorithm [4] for a given number of Doppler bands. The Ram-Lak filter was applied due to the linear characteristics of spatial frequencies. It allows to obtain the sharpest tomographic images in the case of a low noise level and with the exception of errors in the measurement data [1]. Therefore, this filter is perfect for simulation data. More smoothing filters will be tested in future studies with a real measurement data (e.g. Hamming or stochastic filter), due to the potential to significantly minimize the tomographic image distortions at the expense of blurring [2, 8, 9].

5 Conclusions

Both the results of measurements and DT image reconstructions from simulated data have shown that theoretically it is possible imaging the small inclusions contained in tissue of the order of 3, 1 or even 0.1 mm using the circular scanning. The maximum spatial resolution of less than 1 mm can be achieved already at 201 Doppler bands, even for relatively low frequency of the transmitted ultrasonic wave (4.7 MHz). This is a theoretical value, which obviously will deteriorate in real conditions due to the uncertainty of measurements, interference, noise, and phenomena accompanying the propagation of the ultrasound wave in inhomogeneous medium, that is, diffraction and refraction. However, it is always possible to reduce the object blur by using a higher frequency of the ultrasonic wave. Keep in mind, however, that the consequence will then reduce the wave penetration depth due to the increased attenuation.

Based on the preliminary research, we can predict that the Doppler tomography could be applied when examining woman's breasts tissue *in vivo* to detect cancer lesions which reflect the ultrasound waves well. Another way to use this method may

be the examination of the bone of limbs due to the circular geometry. It should also be noted that it is a very safe method for obtaining a cross-section and 3D image of tissue *in vivo* because of the use of ultrasonic waves. In addition, the device based on this method could be relatively inexpensive because only one rotating probe is used for the measurements.

The computer model of ultrasound Doppler tomography will developed with further research in such a way that it will be possible to simulate the measurement signals and taking into account the interference and the phenomena occurring in the ultrasound wave propagation. The DT examination setup in circular geometry will be developed in parallel, which will enable to examine the objects of increasingly complex structure including the female breast phantom along with the inclusions simulating neoplastic tissue.

References

1. Dobrucki, A.B., Opieliński, K.J.: Adaptation of image reconstruction algorithm for purposes of ultrasound transmission tomography (UTT). Arch. Acoust. **25**(4), 395–422 (2000)
2. Jain, A.K.: Fundamentals of Digital Image Processing. Prentice Hall, USA (1989)
3. Kak, A.C., Slaney, M.: Principles of Computerized Tomographic Imaging. IEEE Press, New York (1988)
4. Lewitt, R.M.: Reconstruction algorithms: transform methods. Proc. IEEE **71**(3), 390–408 (1983)
5. Liang, H.-D., Tsui, Ch.S.L., Halliwell, M., Wells, P.N.T.: Continuous wave ultrasonic Doppler tomography: interface focus **1**, 665–672 (2011)
6. Marczak, W.: Woda jako wzorzec w pomiarach prędkości propagacji ultradźwięków w cieczach (Water as a standard in measuring the speed of sound in liquids). Akustyka Molekularna i Kwantowa (Molecular and Quantum Acoustics) **17**, 191–197 (1996). [in Polish]
7. Nagain, K., Greenleaf, J.F.: Ultrasonic imaging using the Doppler effect caused by a moving transduser. Opt. Eng. **29**(10), 1249–1254 (1990)
8. Opielinski, K., Gudra, T.: Ultrasound transmission tomography image distortions caused by the refraction effect. Ultrasonics **38**(1–8), 424–429 (2000)
9. Opielinski, K., Gudra, T.: Multi-parameter ultrasound transmission tomography of biological media. Ultrasonics **44**(1–4), e295–e302 (2006)
10. Opieliński, K.J., Gudra, T.: Ultrasonic transmission tomography. In: Sikora, J., Wojtowicz, S. (eds.) Industrial and Biological Tomography. Electrotechnical Institute, Warsaw (2010)
11. Pozniak, M.A., Allan, P.L.: Clinical Doppler Ultrasound. Churchill Livingstone, Elsevier (2014)

Analytical Statistical Approach for Fan-Beam Scanners

Robert Cierniak

Abstract The main topic of this paper is strictly concerned with the 3D reconstruction algorithm for spiral cone-beam x-ray tomography. The approach proposed here is based on a fully analytical formulation of the reconstruction problem. This method can be classed as a statistical reconstruction method, which significantly improves the quality of the subsequently reconstructed images, so allowing a decrease in the x-ray dose absorbed by a patient during examination. In this paper, we proved that this statistical approach, originally formulated for parallel beam geometry, can be adapted for fan-beam geometry of scanner, with the direct use of projections, and consequently for helical cone-beam scanners. Computer simulations have shown that the reconstruction algorithm presented here outperforms conventional analytical methods with regard to the image quality obtained.

Keywords Image reconstruction from projections · X-ray computed tomography · Statistical reconstruction algorithm

1 Introduction

Currently, the most significant challenge in the field of medical computer tomography is the development of image reconstruction algorithms from projections which would enable the reduction of the impact of measurement noise on the quality of tomography images. This kind of approach is intended to improve high resolution image quality and, in consequence, reduce the dose of X-ray radiation while at the same time preserving an appropriate level of quality in the tomography images. The concept has found its application in the form of statistical reconstruction algorithms. In these algorithms, the probabilistic conditions of the measurements made in tomographic devices are taken into consideration in such a way as to reduce to a minimum

R. Cierniak (✉)
Czestochowa University of Technology, Armii Krajowej 36, Czestochowa, Poland
e-mail: cierniak@kik.pcz.czest.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_19

the impact of measurement noise on the reconstructed image obtained from those measurements.

So far, there are only a few commercial solutions to this problem. The most well-known are, for example: ASIR (Adaptive Statistical Iterative Reconstruction) introduced by General Electric, IRIS (Iterative Reconstruction in Image Space) and SAFIR (Sinogram Affirmed Iterative Reconstruction) constructed by Siemens, AIDR 3D (Adaptive Iterative Dose Reduction) promoted by Toshiba, and iDose developed by Philips. The basis of these developments is the statistical analysis within the image or the measurements obtained in tomographic scanners. Other producers, like Hitachi, Samsung, Nanjing Foinoe, Shimadzu, Neusoft, Shenzhen ANKE, Shangai and other Chinese companies, also have systems to reduce the radiation dose. However, there is no information in the world literature about their operating design. Noise and tomographic image distortions are reduced by iterative signal processing. The technical details of these systems are not revealed, but their practical application has been confirmed in many publications that can be found in radiological journals, e.g. [12].

A completely different approach, called MBIR (Model-Based Iterative Reconstruction) by its authors, is presented in publications like [1, 13], where a probabilistic model of the measurement signals is described analytically. An iterative reconstruction algorithm can be formulated on the basis of this. The reconstruction problem in this approach was reformulated into an optimization problem, using a method called ML (Maximum Likelihood). The objective in these solutions was devised according to an algebraic scheme for formulating the reconstruction problem [14]. An algebraic scheme has been selected in this case for one very obvious reason—the measurement noise can be modelled relatively easily, because each measurement is considered separately. Nevertheless, as is well known, such a scheme adds significant calculative complexity to the problem. The time for image reconstruction becomes unacceptable from the practical point of view. For instance, if the image resolution is assumed to be $I \times I$ pixels, the complexity of the algebraic problem is of the level of $I \times I \times number\_of\_measurements \times number\_of\_cross-sections$; a multiple of I to the power of four in total. This means that despite using a computer which is fast enough, one iteration may last up to one hour [9].

In order to reduce the calculation time to a minimum, a numerical approach has been developed for solving this type of reconstruction algorithm, as described in the article [15]. As a consequence, in 2013, this development had its debut under the name Veo—CT Model-Based Iterative Reconstruction. It was presented as a revolutionary solution. As a result of the conditions described above, the application of this approach made it necessary to purchase a computer with a huge calculating capacity. A potential user would therefore have to pay a lot of money. In addition, the system uses a specific reconstruction problem model. Such a model is defined by mathematicians as very ill-conditioned. There are many profound reasons for that, one of which is the necessity of applying so called regularization. Without this the image is distorted by noise. Although applying regularization prevents that, unfortunately, it is associated with a loss of optimality, which is inconsistent with

the objective in terms of the statistical conditions of the reconstruction problem. Image smoothing takes place instead, which may lead to the loss of many important elements from the images. These elements may be important for diagnostic purposes and therefore it is important that they be retained. It must be said that there are many similar solutions on the market, but they are still inadequate with respect to limiting the radiation dose. Thus there is still room for improvement in such systems, and there will be for a long time to come, because the ideal state, i.e. a zero dose absorbed by the patient during a scan, is unachievable.

The difficulties mentioned above connected with the use of an algebraic methodology can be limited by using an analytical strategy of reconstructed image processing. In previous papers we have shown how to formulate the analytical reconstruction problem consistent with the ML methodology for parallel scanner geometry [2, 3]. This strategy has been used for fan-beams [4], and finally for the spiral cone-beam scanner [5]. However, an approach to the reformulation of the reconstruction problem from parallel to real scanner geometries, called rebinning, was applied there. The rebinning approach in the 3D spiral version involves a nutating slice CT image reconstruction method described in the literature by its abbreviation ASSR (Advanced Single Slice Rebinning) [11]. Much more popular 3D reconstruction methods, which are implemented in practice, are FDK (Feldkamp)-type algorithms [10] that the use projections obtained from spiral cone-beam scanners directly (see e.g. [7, 8]). In this paper, we present a mathematical derivation of a method for the direct (i.e. without rebinning) adaptation of fan-beam projections to the statistical analytical reconstruction algorithm originally formulated by us. This solution is directly applicable to 3D spiral cone-beam scanner geometry.

2 Adoption of the 2D Analytical Approximate Reconstruction Problem to the Fan-Beam Projections

The 2D analytical approximate reconstruction problem was originally formulated for parallel scanner geometry [2, 3]. This iterative approach to the reconstruction problem can be considered as a model-based statistical method of image reconstruction [6].

$$\boldsymbol{\mu}_{\min} = \arg\min_{\boldsymbol{\mu}} \left(\frac{n_0}{2} \sum_{i=1}^{\mathrm{I}} \sum_{j=1}^{\mathrm{J}} \left(\sum_{\bar{i}} \sum_{\bar{j}} \mu^* \left(x_{\bar{i}}, y_{\bar{j}} \right) \cdot h_{\Delta i, \Delta j} - \tilde{\mu} \left(x_i, y_j \right) \right)^2 \right), \quad (1)$$

where coefficients $h_{\Delta i, \Delta j}$ are

$$h_{\Delta i, \Delta j} = \Delta_\alpha \sum_{\psi=0}^{\Psi-1} int \left(\Delta i \cos \psi \Delta_\alpha + \Delta j \sin \psi \Delta_\alpha \right), \quad (2)$$

and $\tilde{\mu}(i, j)$ is an image obtained by way of a back-projection operation; *int* (Δs) is an interpolation function used in the back-projection operation; every projection is carried out after a rotation by Δ_α.

This concept can also form the starting point for the design of a 3D reconstruction algorithm for spiral cone-beam scanner geometry. One of the principal reconstruction methods devised for the cone-beam spiral scanner is the generalized FDK algorithm. The FDK algorithm is a development of a conventional fan-beam reconstruction approach with direct use of measurements performed in a fan-beam scanner. In the traditional FDK approach, the fan-beam projections are filtered in two dimensions and then back-projected in three dimensions. This methodology is adapted to our original iterative model-based reconstruction concept.

However, before discussing the reconstruction in 3D, it is worth first returning to the implementation of the direct fan-beam reconstruction method in 2D. In our further discussions of the fan-beam system, we need to be able to express the trigonometric relationships in polar coordinates, as shown in Fig. 1.

We can use this diagram to derive the following trigonometric relationships between quantities in the parallel-beam system and those in the fan-beam system:

$$s = \mathrm{R}_f \cdot \sin\beta, \tag{3}$$

and

$$\alpha^p = \alpha^f + \beta, \tag{4}$$

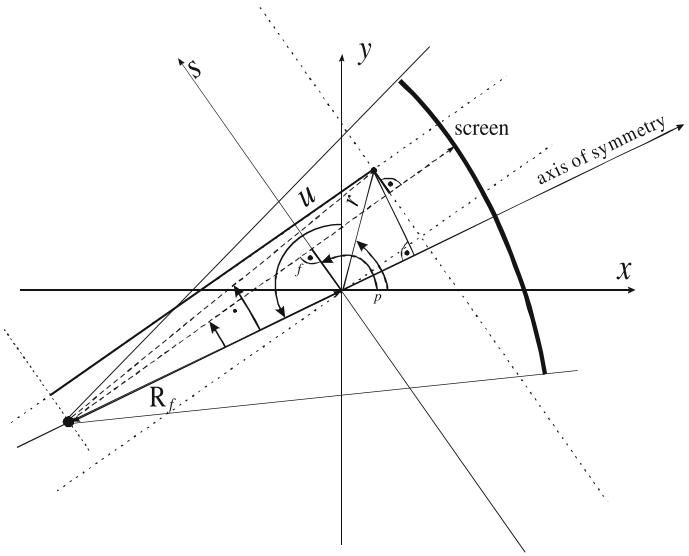


Fig. 1 Determination of the coordinates in the plane of a fan-beam projections geometry

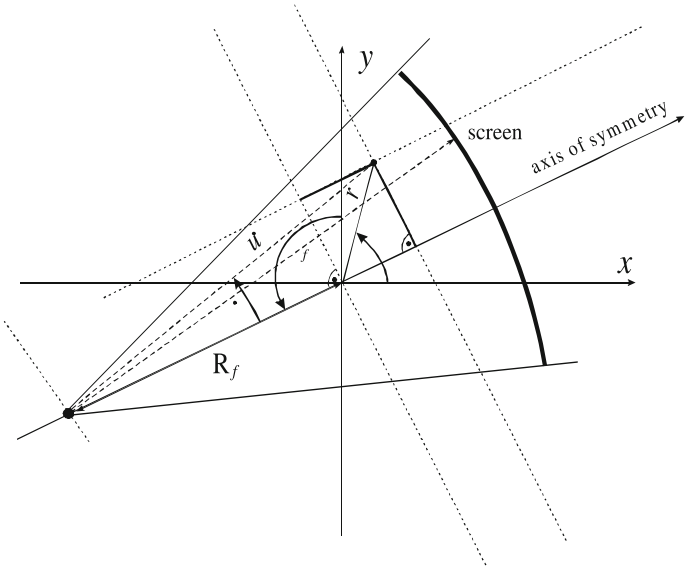


Fig. 2 Determination of the polar coordinates of a point in the plane of a fan–beam projection

where $p^p\,(s,\,\alpha^p)$ and $p^f\,\left(\beta,\,\alpha^f\right)$ are projections and their parameters, in parallel-beam and fan-beam geometrics, respectively.

The diagram in Fig. 2 shows how to determine the parameters of points inside the test cross-section during a fan-beam projection.

The following equations show relationships between selected parameters of a point during the projection $p^p\,(s,\,\alpha^p)$:

$$r\cos\left(\alpha^p-\phi\right)-s=\dot{u}\sin\left(\dot{\beta}-\beta\right), \tag{5}$$

$$\dot{\beta}=\arctan\left(\frac{\alpha^f-\phi}{\mathrm{R}_f+r\sin\left(\alpha^p-\phi\right)}\right), \tag{6}$$

$$\dot{u}^2=\left[r\cos\left(\alpha^p-\phi\right)\right]^2+\left[\mathrm{R}_f+r\sin\left(\alpha^p-\phi\right)\right]^2. \tag{7}$$

Taking into consideration the definition of the two-dimensional inverse Fourier transform, and the frequential form of the relation between the original image of a cross-section of an examined object represented by function $\mu\,(x,\,y)$ and the image obtained after the back-projection operation $\tilde{\mu}\,(x,\,y)$, we obtain:

$$\tilde{\mu}\,(x,y)=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\frac{1}{|f|}M\,(f_1,f_2)\,e^{j2\pi(f_1x+f_2y)}df_1df_2, \tag{8}$$

which, after converting to polar coordinates and using the projection slice theorem (taking into account a full revolution of the projection system), takes the form:

$$\tilde{\mu}(x,y) = \frac{1}{2}\int_{-\pi}^{\pi}\int_{-\infty}^{\infty} \bar{P}(f,\alpha^p)\, e^{j2\pi f(x\cos\alpha^p + y\sin\alpha^p)} df d\alpha^p. \tag{9}$$

Then, after transferring the projections into the spatial domain, we have the formula:

$$\tilde{\mu}(x,y) = \frac{1}{2}\int_{-\pi}^{\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \bar{p}^p(s,\alpha^p)\cdot e^{j2\pi f(x\cos\alpha^p + y\sin\alpha^p)} e^{-j2\pi fs} ds df d\alpha^p. \tag{10}$$

Arranging the right hand side of the above formula and changing the order of integration, we get:

$$\tilde{\mu}(x,y) = \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\pi}^{\pi} \bar{p}^p(s,\alpha^p)\, e^{j2\pi f(x\cos\alpha^p + y\sin\alpha^p - s)} d\alpha^p ds df. \tag{11}$$

Next, after converting the attenuation function into polar coordinates, we obtain:

$$\tilde{\mu}(r\cos\phi, r\sin\phi) = \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\pi}^{\pi} \bar{p}^p(s,\alpha^p)\, e^{j2\pi f[r\cos(\alpha^p-\phi)-s]} d\alpha^p ds df. \tag{12}$$

Note here that the substitution $x\cos\alpha^p + y\sin\alpha^p = r\cos(\alpha^p - \phi)$ refers to the point (r,ϕ), to which the reconstruction process applies, and the variable s specifies the location on the screen.

Of course, we should also take into account the application of the interpolation function used during the back-projection operation, which should be placed appropriately (a frequency representation of this function) in the formula above to finally obtain:

$$\breve{\mu}(r\cos\phi, r\sin\phi) = \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\pi}^{\pi} INT(f)\, p^p(s,\alpha^p)\, e^{j2\pi f[r\cos(\alpha^p-\phi)-s]} d\alpha^p ds df. \tag{13}$$

By using Eqs. (5), (6) and (7) in the above equation and at the same time changing the limits of integration we obtain a relationship, which is fundamental for the fan-beam image reconstruction method:

$$\breve{\mu}(x,y) = \frac{R_f}{2}\int_{-\infty}^{\infty}\int_{-\beta_m}^{\beta_m}\int_{-\beta}^{2\pi-\beta} INT(f)\, p^f\left(\beta,\alpha^f\right)\cos\beta\, e^{j2\pi f\dot{u}\sin(\dot{\beta}-\beta)} d\alpha^f d\beta df, \tag{14}$$

and further

$$\breve{\mu}(x, y) = \frac{R_f}{2} \int_{-\beta_m}^{2\pi-\beta_m} \int_{-\beta_m}^{\beta_m} p^f\left(\beta, \alpha^f\right) \cos\beta \int_{-\infty}^{\infty} INT(f)\, e^{j2\pi f \dot{u} \sin(\dot{\beta}-\beta)} df d\beta d\alpha^f. \quad (15)$$

Unfortunately, there is a serious drawback associated with the use of the fan-beam reconstruction method formulated like this. It stems from the dependence of Eq. (15) on the parameter $\dot{u}$, which poses certain practical problems when carrying out the calculations during the reconstruction process. It now becomes necessary to determine a different form of the interpolation function for every point of the object's cross-section. This is because $\dot{u}$ represents the distance of the point (r, ϕ) from the radiation source. Therefore, by changing the angle α^f, we also change $\dot{u}$ and consequently we need to adjust the value of the interpolation function. We will take action later to eliminate this undesirable situation. The appropriate adjustment is based on a term in Eq. (15), which is reproduced here in a suitably amended form:

$$int(s) = \int_{-\infty}^{\infty} INT(f)\, e^{j2\pi f \dot{u} \sin(\dot{\beta}-\beta)} df. \quad (16)$$

In this equation, the integration is carried out with respect to the frequency f. The next step will be to make a substitution for f, using the following expression:

$$f^f = \frac{f \cdot \dot{u} \cdot \sin\beta}{R_f \cdot \beta}. \quad (17)$$

If at the same time we change the limits of integration, the convolving function will be modified to:

$$int^f(\beta) = \frac{R_f \cdot \beta}{\dot{u} \cdot \sin\beta} \int_{-\infty}^{\infty} INT\left(\frac{f^f \cdot f_0}{f_0^f}\right) e^{j2\pi f^f R_f \beta} df^f, \quad (18)$$

where

$$f_0^f = \frac{f_0 \cdot \dot{u} \cdot \sin\beta}{R_f \cdot \beta}. \quad (19)$$

Unfortunately, even here we encounter problems caused by the dependence of the cut-off frequency f_0^f on the parameter $\dot{u}$. On the other hand, if we were to establish a constant value for f_0^f it would mean that the reconstruction process for the point (r, ϕ) would have a different resolution (determined by the value of the cut-off frequency f_0) for every angle α^f. However, if we put aside the assumption of uniform resolution

for the resulting reconstructed image, then, by manipulating the values $\dot{u}$ and f_0, the varying value of f_0^f can be fixed as:

$$f_0^f = \mathrm{f}_0^f = \frac{1}{\mathrm{R}_f \cdot \Delta_\beta}. \tag{20}$$

Let us assume that we apply a linear interpolation function in formula (13), i.e. the following form of this function:

$$int_L(s) = \begin{cases} \frac{1}{\Delta_s}\left(1 - \frac{|s|}{\Delta_s}\right) & \text{for } |s| \le \Delta_s \\ 0 & \text{for } |s| \ge \Delta_s \end{cases}, \tag{21}$$

which is shown in Fig. 3a.

The frequency form of the interpolation function shown in Eq. (21) is given by this formula:

$$INT_L(f) = \frac{\sin^2(\pi f \Delta_s)}{(\pi f \Delta_s)^2}, \tag{22}$$

and is shown in Fig. 3b.

Taking into account in the formula (18) the proposed interpolation function given by (22), we obtain the following relation

$$int_L^f(\beta) = \frac{\mathrm{R}_f \cdot \beta}{\dot{u} \cdot \sin\beta} \frac{1}{\Delta_s'} \int_{-\infty}^{\infty} \Delta_s' \frac{\sin^2(\pi f \Delta_s')}{(\pi f \Delta_s')^2} e^{j2\pi f \mathrm{R}_f \beta} df, \tag{23}$$

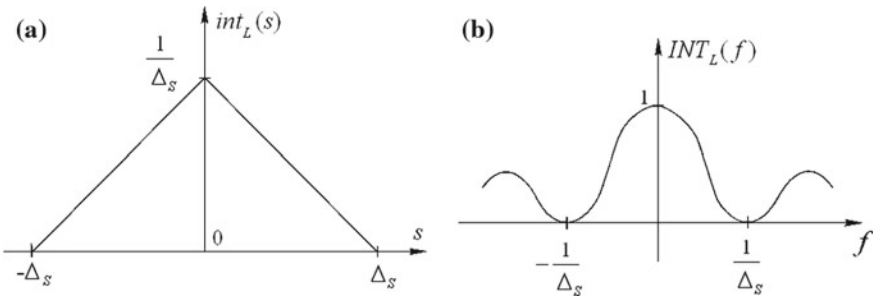


Fig. 3 The linear interpolation function: **a** a form of the function from Eq. (21); **b** a frequency spectrum of the interpolation function from Eq. (22)

where

$$\Delta'_s = \frac{f_0}{f_0^f}. \tag{24}$$

It is easy to show that it gives

$$int_L^f(\beta) = \frac{\mathrm{R}_f \cdot \beta}{\dot{u} \cdot \sin\beta} \begin{cases} \frac{1}{\Delta'_s}\left(1 - \frac{\mathrm{R}_f|\beta|}{\Delta'_s}\right) & \text{for } |\beta| \le \Delta'_s \\ 0 & \text{for } |\beta| \ge \Delta'_s \end{cases}, \tag{25}$$

and next, bearing in mind relations (20) and (24), it leads immediately to:

$$int_L^f(\beta) = \frac{\beta}{\dot{u} \cdot \sin\beta} \begin{cases} \frac{\Delta_s}{\Delta_\beta}\left(1 - \frac{\Delta_s|\beta|}{\Delta_\beta}\right) & \text{for } |\beta| \le \frac{\Delta_\beta}{\Delta_s} \\ 0 & \text{for } |\beta| \ge \frac{\Delta_\beta}{\Delta_s} \end{cases}. \tag{26}$$

Finally, if we assume that $\Delta_s = 1$, it gives

$$int_L^f(\beta) = \frac{\beta}{\dot{u} \cdot \sin\beta} int_L(\beta), \tag{27}$$

where

$$int_L(\beta) = \begin{cases} \frac{1}{\Delta_\beta}\left(1 - \frac{|\beta|}{\Delta_\beta}\right) & \text{for } |\beta| \le \Delta_\beta \\ 0 & \text{for } |\beta| \ge \Delta_\beta \end{cases}. \tag{28}$$

In consequence, returning to the formula (15), we obtain

$$\breve{\mu}(x,y) = \int_0^{2\pi} \int_{-\beta_m}^{\beta_m} p^f\left(\beta, \alpha^f\right) \frac{\mathrm{R}_f \cos\beta}{2\dot{u}} \frac{\Delta\beta}{\sin\Delta\beta} int_L(\Delta\beta)\, d\beta d\alpha^f. \tag{29}$$

Fortunately, we can linearize relation (29) by considering expressions inside the integration, namely $\frac{\Delta\beta}{\sin\Delta\beta}$. In the case of linear interpolation we use only a line of integrals from the neighborhood of a given pixel (x, y), then $\Delta\beta \le \Delta_\beta$, and $\sin\Delta\beta \simeq \Delta\beta$. Additionally, it is possible to omit the term $\frac{\mathrm{R}_f \cos\beta}{2\dot{u}}$ taking into account the fact that each projection value $p^f\left(\beta, \alpha^f\right)$ has its equivalent $p^f\left(-\beta, \alpha^f + \pi - 2\beta\right)$, as shown in Fig. 4.

Because of this we can notice that the sum of this pair of projections is proportional to $\frac{\dot{u}_1+\dot{u}_2}{4\dot{u}_1} + \frac{\dot{u}_1+\dot{u}_2}{4\dot{u}_2} = \frac{(\dot{u}_1+\dot{u}_2)^2}{4\dot{u}_1\dot{u}_2}$. This means that for $\dot{u}_1 \simeq \dot{u}_2$ this factor is equal to 1, and finally, we can write

$$\breve{\mu}(x,y) \simeq \int_0^{2\pi} \int_{-\beta_m}^{\beta_m} p^f\left(\beta, \alpha^f\right) int_L(\Delta\beta)\, d\beta d\alpha^f, \tag{30}$$

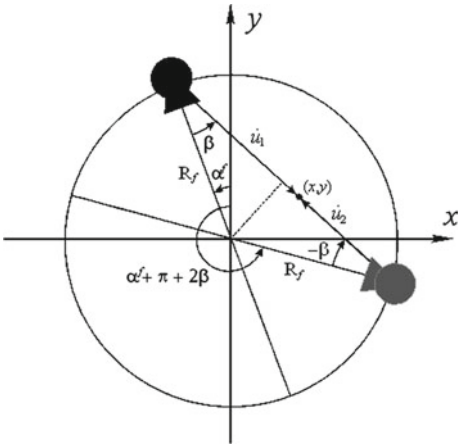


Fig. 4 Selecting complementary projection values

which is consistent with a form of the formula of the back-projection operation for parallel beams. Therefore, formula (30) can be used directly to obtain a reference image for the analytical statistical iterative reconstruction algorithm which was originally formulated for parallel beam scanner geometry (1) and (2).

3 Experimental Results

In our experiments, we used projections obtained from a C-arm scanner. Parameters: $SDD = 1025\,\text{mm}$ (Source-to-Detector Distance); $R_f = 825\,\text{mm}$ (SOD (Source-to-AOR Distance)); number of views per rotation $\Psi = 538$; number of pixels in detector panel 756; detector width 307.2 mm. During the simulations, the size of the processed image was fixed at $I \times J = 512 \times 512$ pixels. The coefficients $h_{\Delta i,\Delta j}$ were precomputed before we started the reconstruction process and were fixed for the subsequent processing. We started the actual reconstruction procedure and performed a back-projection operation to get a blurred image of the x-ray attenuation distribution in a given cross-section of the investigated object. The image obtained in this way was then subjected to a process of reconstruction (optimization) using an iterative statistically-tailored procedure. The starting point of this procedure can be chosen from any standard reconstruction method, for example an FBP reconstruction algorithm. It is worth noting that our reconstruction procedure was performed without any regularization regarding the objective function described by (1).

A view of the reconstructed image after 30000 iterations is presented in Fig. 5b. For comparison, the image reconstructed using a standard FBP reconstruction method is also presented (Fig. 5a).

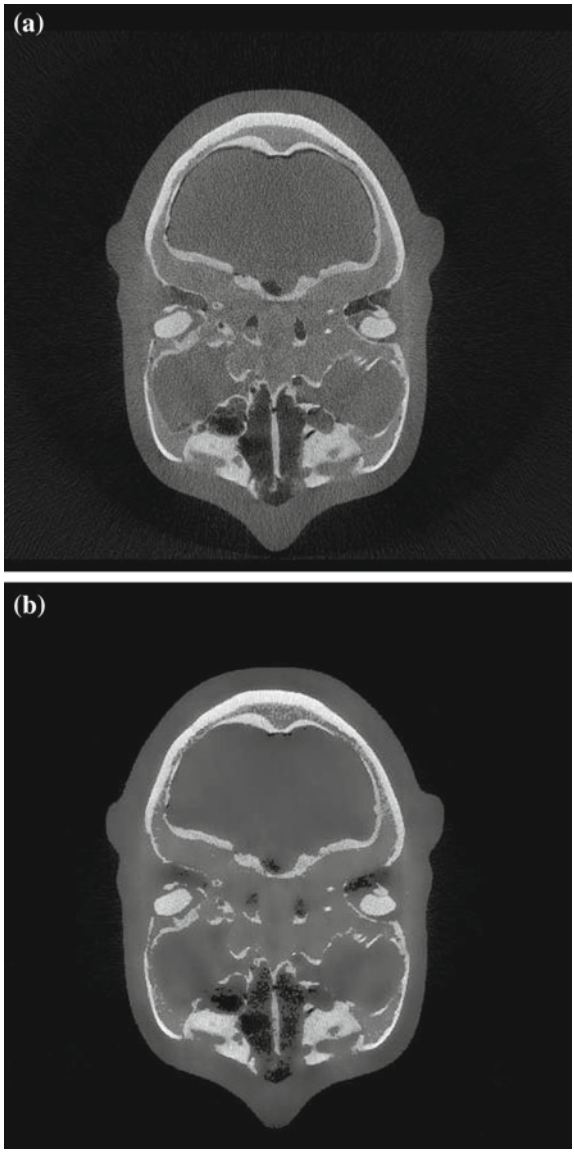


Fig. 5 View of the images reconstructed image using the standard FBP with Shepp-Logan kernel (**a**); reconstructed image using the method described in this paper after 30000 iterations (**b**)

4 Conclusion

In this paper we have presented a fully feasible statistical reconstruction algorithm for fan-beam projections. It has been proven that this statistical approach, originally formulated for parallel beam geometry, can be adapted for fan-beam scanner geometry, with the direct use of projections. Consequently, after only a few additional transformations, the algorithm can be used for helical cone-beam scanners. Simulations have been performed, which prove that our reconstruction method is very fast (thanks to the use of FFT algorithms) and gives satisfactory results with suppressed noise, without the introduction of any additional regularization term. It is worth underlining that in the algorithm shown here, no additional geometrical correction of the projection lines has been used. The coefficients $h_{\Delta i,\Delta j}$ are precalculated according to Eq. (4) before the start of the actual reconstruction procedure and are the same for all pixels in the reconstructed image.

References

1. Bouman, C., Sauer, K.: A unified approach to statistical tomography using coordinate descent optimization. IEEE Tran. Image Process. **5**, 480–492 (1996)
2. Cierniak, R.: A new approach to image reconstruction from projections problem using a recurrent neural network. Int. J. Appl. Math. Comput. Sci. **183**, 147–157 (2008)
3. Cierniak, R.: A new approach to tomographic image reconstruction using a Hopfield-type neural network. Int. J. Artif. Intell. Med. **43**, 113–125 (2008)
4. Cierniak, R.: New neural network algorithm for image reconstruction from fan-beam projections. Neurocomputing **72**, 3238–3244 (2009)
5. Cierniak, R.: A three-dimensional neural network based approach to the image reconstruction from projections problem. In: Rutkowski, L., Tadeusiewicz, R., Zadeh, L.A., Żurada, J. (eds.) LNCS 6113, pp. 505–514. Springer, Heidelberg (2010)
6. Cierniak, R.: An analytical iterative statistical algorithm for image reconstruction from projections. Appl. Math. Comput. Sci. **24**, 7–17 (2014)
7. Cierniak, R., Knas, M.: Ultrafast fully analytical iterative model-based statistical 3D reconstruction algorithm. In: Proceedings of the 12th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, Lake Tahoe, 16-21.VI.2013, pp. 521–524 (2013)
8. Cierniak, R., Knas, M.: Ultrafast iterative model-based statistical 3D reconstruction algorithm for x-ray computed tomography. In: 4th International Conference on Information Technologies in Biomedicine, Kamień Śląski, 2-4.VI.2014, Advances in Inteligent Systems and Computing 283, pp. 187–196 (2014)
9. DeMa,n B., Basu, S., Thibault, J.-B., Hsieh, J., Fessler, J., Bouman, C., Sauer, K.: A study of four minimization approaches for iterative reconstruction in X-ray CT. In: IEEE Nuclear Science Symposium and Medical Imaging Conference, San Juan (2005)
10. Feldkamp, L.A., Davis, L.C., Kress, J.W.: Practical cone-beam algorithm. J. Opti. Soc. Am. **1**(A) 9, 612–619 (1984)
11. Kachelrieß, M., Schaller, S., Kalender, W.A.: Advanced single-slice rebinning in cone-beam spiral CT. Med. Phys. **27**, 754–773 (2000)
12. Ren, Q., Dewan, S.K., Li, M., Li, J., Mao, D., Wang, Z., Hua, Y.: Comaprison of adaptive statistical iterative and filtered back projection reconstruction techniques in brain CT. Eur. J. Radiol. **81**, 2597–2601 (2012)

13. Sauer, K., Bouman, C.: A local update strategy for iterative reconstruction from projections. IEEE Tran. Sig. Process. **41**, 534–548 (1993)
14. Thibault, J.-B., Sauer, K.D., Bouman, C.A., Hsieh, J.: A three-dimensional statistical approach to improved image quality for multislice helical CT. Med. Phys. **34**, 4526–4544 (2007)
15. Zhou, Y., Thibault, J.-B., Bouman, C.A., Hsieh, J., Sauer, K.D.: Fast model-based x-ray CT reconstruction using spatially non-homogeneous ICD optimization. IEEE Trans. Image Process. **20**, 161–175 (2011)

3D Measurement of Geometrical Distortion of Synchrotron-Based Perforated Polymer with Matlab Algorithm

Michał Walczak

Abstract Nowadays, many fields of science, are basing on information provided in form of photographs, projections or cross-sections. Especially medical diagnosis became a process impossible to be conducted effectively and efficiently without use of modern tools of analysis. While analyzing data, researcher thinks of what he observes in terms of features of a sample, not artifacts that may be related with the method used to acquire the data. Microtomography is one of techniques that require specific experience and approach while analyzing data provided, because of the variety of artifacts present on the image. The aim of the article is to measure perforations of a sample, and compare the results to their actual shape.

Keywords Image processing · Algorithm · Microtomography

1 Introduction

X-ray Microcomputed Tomography is method basing on reconstruction of multiple images, so called projections (Fig. 1), of sample placed under changing angle in relation to the source of electromagnetic radiation and the detector of that radiation. Absorption of X-ray radiation can be calculated with Lambert-Beer law (1) [4, 6] and in traditional tomography this absorption process is a fundament of that technique.

$$l = l_0 \exp(-\mu x) \tag{1}$$

Besides of absorption, electromagnetic radiation observed on detector is function of numerous features of X-rays, such as different absorption based on different frequency of electromagnetic waves or deflection on edges of two different density factions of sample. Considering all the flaws of equipment, like fluctuation of current during scanning procedure, non-linear reaction of detectors, changing the source

M. Walczak (✉)
University of Dabrowa Gornicza, 41-300 Dabrowa Gornicza, Poland
e-mail: mwalczak@wsb.edu.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_20

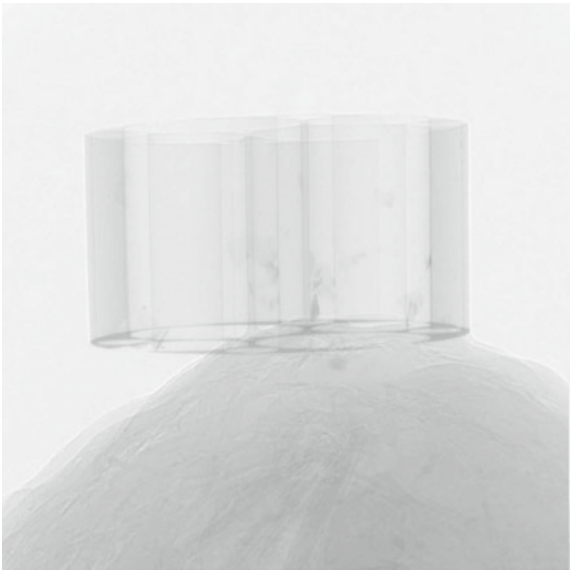


Fig. 1 One of the projections used during sample reconstruction

temperature during scanning procedure, or simple wear of that light source over the period of use, one may conclude, that images can be subject of many artifacts, objects or features present on image, related to the technique itself, not to the sample [3].

Some of artifacts, however, can be used as an advantage of this technique, for instance, phase contrast [1, 9], which is a technique of scanning basing on deflection of X-rays, and enhancing edges of two different density factions of sample. Some of artifacts can be suppressed or corrected, like so called Ring-Artifacts [5]. They can be suppressed with shifting detector or white reference correction, and using a monochromatic light source can exclude the beam hardening problem of absorption of heterochromatic X-rays.

One of the artifacts present on images of measured sample (Fig. 3E) is the artifact caused by the shift of the center of rotation, occurring in most of the cases, when sample moves slightly during scanning procedure. Most of reconstruction software should be capable of finding the center of rotation automatically, but if the sample moved from center of rotation, for example, due to thermal expansion of sample or mounting, it may be necessary to find the center of rotation manually. In this case, it was necessary to perform several reconstructions of singular slice, and basing on user experience choose the best value. After reconstructing several singular slices, for optimal value of Center Shift parameter this “mustache” like artifact is the least visible on the whole picture. For other values, artifacts present themselves more significant on some parts of the image.

Main motives for this research were to perform measurements of perforations dimensions changes on different depths of the perforations and evaluate deviations created during scanning procedure. Author’s previous study has proven that not only the perforation dimension shows distortion from the actual size, but also the very

shape of perforation appears to be distorted. Since all the perforations should be round (in y, z plane), algorithm previously designed and used for another study, as a tested and reliable source of measurements was proposed [7] to evaluate shape distortion.

2 Materials and Experiment

In this section the preparation of a sample, its scanning procedure and measurements algorithm will be presented.

2.1 Sample Preparation

The sample called SU8 (Fig. 2) is made with polymer with thin layer of gold on plane perpendicular to perforations. Perforations created with lithographic method based on X-ray radiation obtained in synchrotron are so precise, that we can call them ideal shape pattern [2], so all deformation visible should be related with imagining method.

Dimensions of perforations were measured with electron microscope to confirm precision of sample manufacturing. Sample consist of 9 perforations, but only the smallest one will be used for study since geometrical distortion is mostly visible within this perforation. Its actual size is 40 μm in diameter, which was confirmed by measurement performed with electron microscope.

2.2 Scanning Procedure

Sample was scanned with microCT XRadia scanner in Henry Moseley X-ray Imagining Facility, University of Manchester. Sample was mounted on a thin piece of

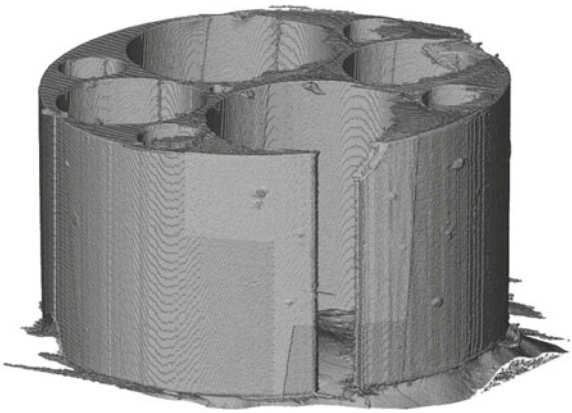


Fig. 2 Polymer perforated with synchrotron radiation (data visualisation)

Table 1 Scanning parameters using X-Radia

Voltage (kV)	Current (μA)	Optical magnification	Pixel size (μm)	Number of projections (N)	Angle of rotation (°)	Data type (B)
40	250	19.401×	0.61	1701	180	16

wood being fixed with double sided glue tape, so that both sample and mount have similar density in order to avoid significant attenuation differences and thus X-ray beam distortion. Sample is tilted by few degrees so that thin layer of gold would not be parallel to X-ray beam at all angles of rotation causing obstruction of X-rays. Scanning parameters are shown in Table 1. Projections were recorded over 180°.

2.3 Data Processing and Measuring Algorithm

Projections were reconstructed with standard filtered back projection based algorithm. White reference correction was applied. Center of rotation was found manually to ensure highest possible image quality and to minimalize potential artefacts.

Slices created that way were filtered with 3D median filter with $3 \times 3 \times 3$ structural element to decrease noise with lowest distortion of image. After filtering, the slices were cut to decrease volume of data in order to speed up calculations and decrease hardware requirements while post-processing of data. Segmentation of sample was based on region growing technique similar with global threshold. No additional segmentation process took place. At this point some impurities of sample are visible (Fig. 3), but some of perforations remain clear and suitable for measurements.

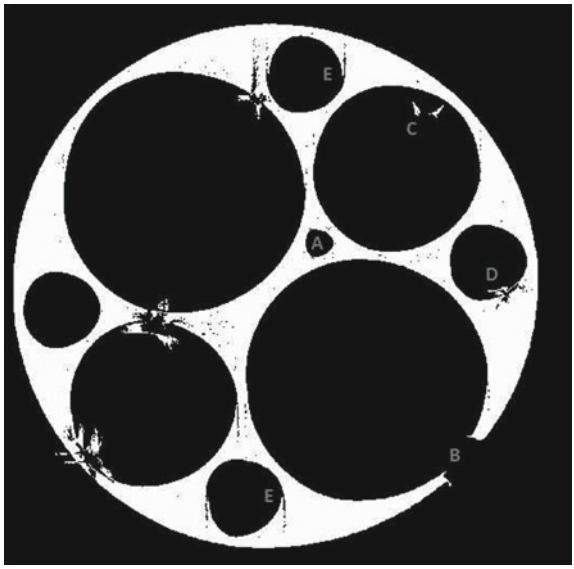


Fig. 3 SU8 Segmented slice. *A* perforation used for measurements. *B* sample damage. *C* sample impurities. *D* geometrical distortion due to scanning artifact. *E* artifact caused by shift of center of rotation

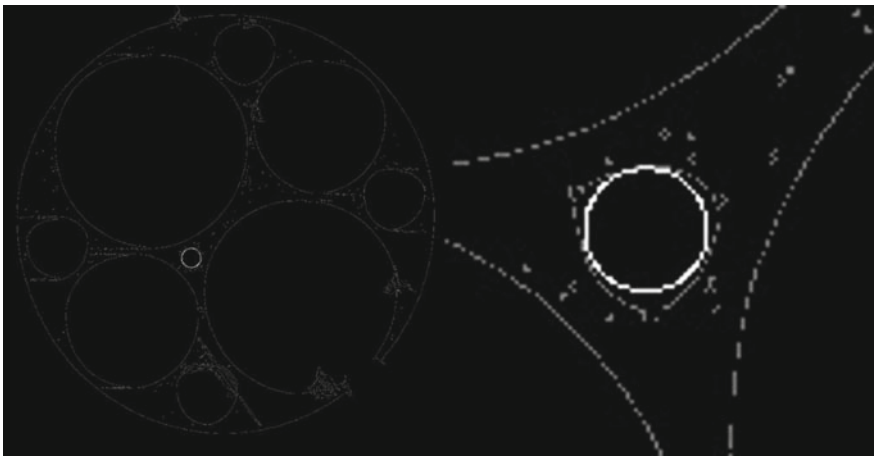


Fig. 4 Sphere (*white*) being placed in one of the perforation (*gray*)

Measuring Algorithm was written in Matlab® and proposed for studies on viviparous land snails [7]. Algorithm was designed to measure maximum sphere that can be placed inside of a progressing cavity in complex 3D environment. Basing on the assumption, that designed shape of perforation is round, measurements of radius of maximum sphere should result in measuring dimension of the perforation. Any deviation from this measurement should arise from algorithm accuracy and image distortion only (Fig. 4).

During the process of measurement algorithm searches immediate surrounding for points with maximum radius of sphere that could be situated in that place (Fig. 5). Algorithm calculates the distance from the nearest object (edge of perforation) basing on Euclidean distance. In a given number of iterations the algorithm moves back the center of sphere away from the nearest edge of the sample, and calculates the sphere radius again. This process is taking place till the center of sphere is placed the way that the sphere has a maximum radius. Next step is calculating place for next sphere basing on previous two calculated spheres. The algorithm was described in details in another authors paper [7, 8].

Because of the nature of the algorithm, it is possible to measure another structures for medical research purposes, such as blood vessels. Flow depends significantly on cross sectional area of the vessel, so its measurements accuracy would depend on 3D shape of that vessel. In the other hand, measuring of maximum sphere (or series of spheres) possible to be placed inside the vain could bring valuable data for angioplasty. Another application of the algorithm could be examination of patency of urethra or bronchi on CT images.

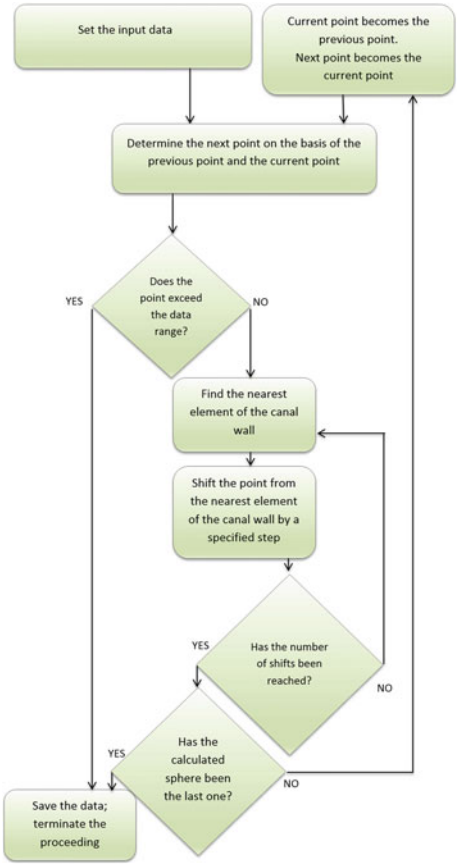


Fig. 5 Flow chart of the measuring algorithm

3 Results

Total execution time of measuring and placing the spheres on a personal computer took less than a minute. Implementation of measuring algorithm is presented in table containing three coordinates of center of the sphere and radius of that sphere. From that data, distance between spheres was calculated in order to measure progress of measurement, and simplify sphere position in perforation which can be considered as one dimensional structure.

Measured maximum sphere radius, its median and actual size are shown in Fig. 6. Visualization of spheres calculated by the algorithm are presented in Fig. 7. Due to small difference between spheres, visible structure can be seen as a canal-like structure. Also, cross section of one sphere being placed in the perforation is presented in Fig. 4.

Mean measured radius of sphere placed on the perforation is 18.15 μm. The length over which the measurement was performed is 420.13 μm.

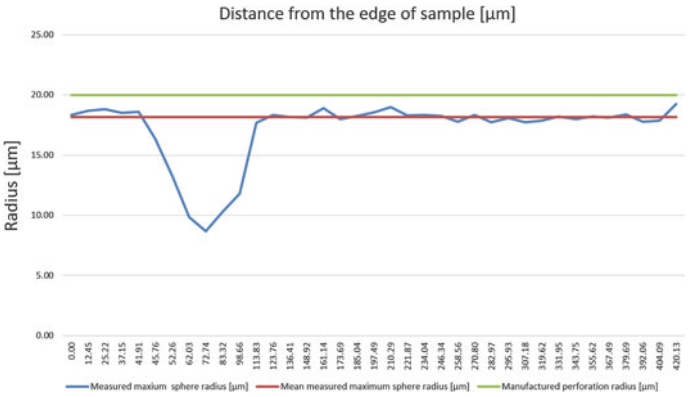


Fig. 6 Measured radius of maximum sphere placed in side of the perforation

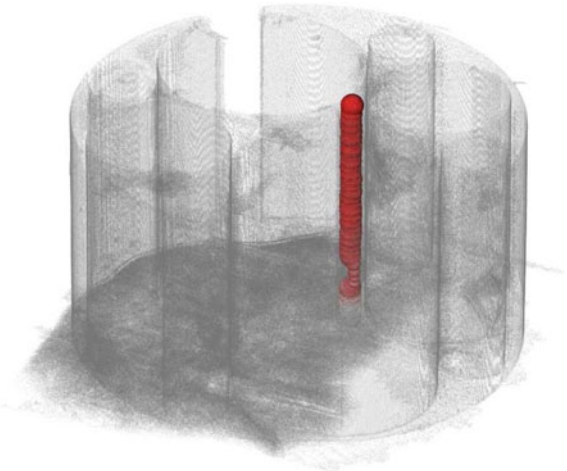


Fig. 7 Visualization of measured maximum spheres (*red*) placed inside of the perforation

4 Discussion

The measurements were focused only on the smallest perforation (40 µm diameter). Since scanning resolution is 0.6 µm we can expect representation of the structure with good resolution and distortion from actual size not greater than 1.5 %, but the measured radius equals 18.15 µm which is 90.75 % of actual size. Analyzing of the cross section of sphere placed in perforation (Fig. 4) may shed some light on cause of distortion of calculated diameters. Edge of the perforation visibly shows distortion from round shape of unknown origin. It is worth notifying that the pattern of sample perforations shows axial symmetry so the deflection of electromagnetic wave could cause visible shape distortion, but only if this distortion would share axial symmetry. In this case the distortion seems to have a triangle shape, that could not been caused by the deflection.

Results show greatest difference between manufactured and measured dimension of perforation at 72.74 μm of measurement path. This discrepancy is caused by the distortion in shape of perforation arisen during scanning procedure, visible on data, not in the real sample itself. The distortion of radius visible at 72.74 μm of perforation depth is cause by the impurity of the sample. Since the algorithm measures the diameter of maximum several times at various points of a perforation it is possible to measure diameter even if the canal is partially damaged or impure.

The results have proven that the distortion of the sample is greater than the scanning resolution itself. Besides of regular diameter of perforations the shape could be analyzed the way the shape distortion have been visible. Also, the measuring algorithm shows tolerance to canal impurity delivering the results despite partial perforation narrowing.

Acknowledgments The author of this article would like to express gratitude to PhD. Samuel MacDonald and Prof. Philip Whiters from Henry Moseley X-ray Imagining Facility at the University Of Manchester for their help during acquisition of data, and to PhD Agnieszka Banaś and PhD Krzysztof Banaś from Singapore Synchrotron Light Source from National University of Singapore for manufacturing the samples.

References

1. Arhatari, B.D.: Hight resolution phase contrast X-ray radiography. The University of Melbourne (thesis), pp. 1–169 (2011)
2. Moser, H.O., Chew, E.P., Kempson, V.C., et al: Singapore synchrotron light source-status, first results, program. Nuclear Instruments and Methods in Physics Research, pp. 536-540. Elsevier (2003)
3. Paganin, D., Barty, A., Mcmahon, P.J., Nugent, K.A.: Quantitative phase-amplitude microscopy III. The efect of noise. J. microsc. (2011)
4. Pawlicki, G., Pałko, T., Golnik, N., Gwiazdowska, B., Królicki, L.: Fizyka medyczna (Medical Physics). Akademicka Oficyna Wydawnicza Exit, Warszawa (2003). (in Polish)
5. Rashid, S., Lee, S.Y., Hasan, K.: An improved method for the removal of ring artifacts in high resolution CT imaging. EURASIP J. Adv. Sig. Process. **2012**, 93 (2012)
6. Stock, S.R.: MicroComputed Tomography. Methodology and Applications. CRC Press (2009)
7. Sulikowska-Drozd, A., Walczak, M., Binkowski, M.: Evolution of shell apertural barriers in viviparous land snails (Gastropoda: Pulmonata: Clausiliidae). Can. J. Zool. **92**, 205–213 (2014)
8. Walczak, M., Binkowski, M., Sulikowska-Drozd, A.: Maximum sphere method for shell patency measurements in viviparous land snails based on X-ray microcomputed tomography imaging computers in biology and medicine, vol. 64, pp. 187–196. Elsevier (2015)
9. Wilkins, S.W., Gureyev, T.E., Geo, D., Pogany, A., Stevenson, A.W.: Phase—contrast imaging using polychromatic hard X-rays. Nature (2011)

Improving Children Diagnostics by Efficient Multi-label Classification Method

Kinga Glinka, Agnieszka Wosiak and Danuta Zakrzewska

Abstract Using intelligent computational methods may support children diagnostics process. As in many cases patients are affected by multiple illnesses, multi-perspective view on patient data is necessary to improve medical decision making. In the paper, multi-label classification method—Labels Chain is considered. It performs well when the number of attributes significantly exceeds the number of instances. The effectiveness of the method is checked by experiments conducted on real data. The obtained results are evaluated by using two metrics: Classification Accuracy and Hamming Loss, and compared to the effects of the most popular techniques: Binary Relevance and Label Power-set.

Keywords Children diagnostics · Problem transformation methods · Labels chain · Multi-label classification

1 Introduction

Medical progress and the equipment development make possible collecting the increasing amount of patient data. Clinical records usually contain many parameters, which can be used for diagnostics purpose. However, in many cases, using intelligent computational methods is necessary to increase the process efficacy. A reliable data classification is one of the technique which can help medical staff in disease diagnosis, prevention and treatment.

Considering single-label classification problem may not be effective for medical data as remarkably often patients are affected by multiple illnesses. Dealing with the interaction of multiple comorbidities in a heterogeneous population of patients is mentioned as the big challenge to overcome [1]. In that case, multi-label classification technique may be helpful. In the medical informatics and data mining community

K. Glinka · A. Wosiak (✉) · D. Zakrzewska
Institute of Information Technology, Lodz University of Technology,
Wólczańska 215, Lodz, Poland
e-mail: agnieszka.wosiak@p.lodz.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_21

[14] there has already been discussed, that classifying patients according to their physiological values and laboratory tests may improve decision making process. We propose to use the multi-label classification for improving children diagnostics.

Medical classification very often has to deal with relatively small number of instances, which are usually described by many features and thus, specialized methods to deal with that kind of data are required. There exist universal techniques for multi-label classification that can be used for any dataset. However, they do not provide satisfactory accuracy in many cases, especially when sets of attributes are relatively large in comparison to the number of instances [3].

In the paper, we consider application of the problem transformation method, which performs well when the number of attributes significantly exceeds the number of instances. The technique was firstly introduced in [3], where its performance was examined by taking into account accuracy for two-label classification of datasets of images and music. In the current research, the technique is used for multi-perspective classification problem of children diagnostics, where cases are described by labels from separate sets. The method is validated by the experiments done on real data. The detailed statistical analysis was performed to compare the results with the ones obtained by using separate single-label classification tasks as well as application of the most commonly used problem transformation methods: Binary Relevance and Label Power-set.

The remainder of the paper is organized as follows. In the next section, relevant work is presented. Then, the proposed methodology is described. In the following section, the experiments and their results are discussed. Finally, some concluding remarks and future research are presented.

2 Relevant Research

There are two most commonly applied categories of multi-label classification. The first approach, based on adaptation methods, extends specific algorithms to obtain the classification results directly. Methods of the second category transform multi-label classification problems into single-label tasks. Such approach allows to apply well-known classification algorithms to multi-label classification problems.

There exist several transformation techniques [16]. The simplest approach consists in changing the problem into single label classifications. These techniques are very simple but the information on multi-label dataset is lost. As the most popular methods there should be mentioned Binary Relevance (BR) and Label Power-set (LP) techniques. The first method transforms multi-label problem into several binary classification tasks by using one-against-all strategy. The method produces the union of the labels, which are obtained by binary classifiers. Main disadvantage of BR technique consists in ignoring all the label correlations existing in a considered dataset [16].

LP creates new classes of all unique sets of labels which exist in the multi-label training dataset. Thus, every complex multi-label task can be changed into one single-label classification. LP method does not depend on a number or variety of labels assigned to the instances. The main disadvantage of creating new labels is that it may lead to datasets with a large number of classes and few instances which represent them [16].

In medicine area multi-label classification mostly has been applied to document classification. Zha et al. [20] consider free text clinical records categorization according to their specified disease labels. They proposed application of two algorithms: Sampled Classifier Chains and Ensemble of Sampled Classifier Chains, which intelligently select labels to be treated as extra input, what allows to exploit disease relation beneficial to classification. Experiments conducted on real datasets have shown the advantage of the proposed approaches. In the paper [12], the authors study the role of the feature selection, training data selection and probabilistic threshold optimization in improving different multi-label classification approaches. They have explored different transformation techniques to extract ICD-9-CM diagnosis codes from clinical narratives. The proposed method has been validated by using EMRs from two different datasets. Medical document records are very often taken into account to evaluate and compare different multi-label classification techniques (see example [9]).

Patient medical records were considered by Qu et al. [10], who introduced modifications of mult-label Bayesian classifier taking into account the dependency between two labels as well as a set of labels. The authors validated the proposed methods on real medical data of the pain medicine to predict effects of treatment plans.

3 Methodology

Cardiovascular and rheumatic diseases become a serious problem in developmental age population. Intra-uterine growth restriction (IUGR) is an important issue for paediatricians, as it occurs in about 3–10 % of newborns. Since the 1990s, it has been known that abnormalities during fetal growth may result in cardiovascular disease, including hypertension in adulthood [18]. Furthermore, primary arterial hypertension is diagnosed increasingly often and in patients at ever younger ages. Similarly, juvenile idiopathic arthritis (JIA) is the most common chronic rheumatological disease of childhood [8].

Early diagnostics of diseases enables prevention of morbidity and mortality caused by cardiovascular disorders in adulthood [19]. Therefore, finding an effective method for medical diagnosis in children is an important issue and many efforts are made to improve automated classification accuracy. Conventional approaches in medical diagnostics usually deal with single classification, which means that patient's data are gathered and analysed to asses one particular disease. Assigning more than one label may result in discovering also comorbidities and lead to improved and earlier diagnosis.

In the paper, two transformation methods based on separate single-label classification tasks are examined: Independent Labels and Labels Chain. The first technique executes all the classification tasks separately. The second one takes into account consequential labels in each following classification step. Let us consider the set of all the labels L and let K denotes a set of labels respective to the instance.

Independent Labels (IL) works similarly to Binary Relevance method, however, instead of $|L|$ binary classifiers $|K|$ multiclass classifiers should be learnt, assuming that the number of labels of each instance is known. IL is the approach where each label constitutes a separate single-label task. Thus, when labels belong to separate sets, single-class problems are considered. IL is competitive in time and computational complexity in the cases of the small number of labels per instance. However, the algorithm ignores existing label correlations during classification process.

Labels Chain (LC) is the improvement of IL method, that uses mapping of relationship between labels. LC algorithm also assumes the number of labels for instances to be known. It requires to learn $|K|$ multiclass classifiers, and consecutively uses result labels as new attributes in the following classification process. Thus, the classifications chain is created (the idea has been used so far only for binary classifications [11]). As classifications are not totally independent from themselves, such approach enables providing better predictive accuracy. Detailed descriptions of the both of the algorithms are presented in [3]. For the purpose of children diagnostics, IL is used as indirect method, improved by LC approach.

As the first and less restrictive metric used for evaluating classification results, *Hamming Loss* will be considered. It was proposed in [13] for evaluating the performance of multi-label classification. It calculates the fraction of incorrectly classified single labels to the total number of labels. As it is a loss function, the smaller value is connected with the better effectiveness of the algorithm. *Hamming Loss* is defined as:

$$HL = \frac{1}{N} \sum_{i=1}^{N} \frac{\text{xor}\,(Y_i, F(x_i))}{|L|}, \tag{1}$$

where: x_i are instances, $i = 1..N$, N is their total number in the test set, Y_i denotes the set of true labels and $F(x_i)$ is the set of predicted labels during classification process, operation $(Y_i, F(x_i))$ gives the difference between these sets.

Classification Accuracy (also known as *exact match*) is much more strict evaluation metric for multi-label classification, thus more suitable for medical diagnostics. Contrarily to the *Hamming Loss* measure, it ignores partially correct sets of labels by marking them as incorrect predictions, and requires all labels to be an exact match of the true set of labels. *Classification Accuracy* for multi-label classification is defined as [6]:

$$CA = \frac{1}{N} \sum_{i=1}^{N} I\,(Y_i = F(x_i))\ , \tag{2}$$

where: $I(true) = 1$ and $I(false) = 0$.

To verify experimental results, a detailed statistical analysis will be performed. The Friedman test will be used with the corresponding post-hoc analysis. It was proved to be suitable for comparison of many classifiers over multiple data sets [2].

4 Experiment Results and Discussion

To verify the proposed methodology, three different datasets have been considered, namely *JIA* for juvenile idiopathic arthritis classification, *IUGR* for medical problems concerning infants being born as small for gestational age, and *CTG* for cardiotocography examination. The *JIA* and *IUGR* datasets were collected from children hospitalized in the University Hospital No 4, Department of Cardiology and Rheumatology, Medical University of Lodz.

The *JIA* dataset is considered twice in the experiments. Firstly, *JIA* is denoted as *JIA#1* and consists of 50 instances with diagnosed juvenile idiopathic arthritis characterized by 16 attributes. Peripheral blood samples were obtained by laboratory tests, whereas parameters concerning synovial fluid were obtained during diagnostic punction. Detailed procedure of juvenile idiopathic arthritis treatment can be found in [15]. As a result, the dataset is defined using 16 attributes. Each case in *JIA#1* is described by two label sets: L_1 and L_2. First label set L_1 classifies JIA as one of three types: oligoarticular JIA, polyarticular JIA and systemic JIA. Another label set L_2 refers to disease activity, that can be low, medium or high. Second *JIA* dataset—denoted as *JIA#2*—is built by joining *JIA#1* with a group of healthy children. *JIA#2* consists of 67 cases and excludes one attribute, that has not been gathered for healthy children (1.25-dihydroxyvitamin D3). One more label in each label set denoted as '*none*' was included.

The *IUGR* dataset consists of instances represented by parameters of high blood pressure in children born with intrauterine growth restriction (IUGR). The characteristics includes perinatal history, risk factors for IUGR, family history of cardiovascular disease, nutritional status, echocardiography and blood pressure measurements. Each sample is defined by 120 attributes. The full medical description of data was presented in [18]. The whole dataset *IUGR* is considered twice building two separate subsets labeled by different label sets. Thus, data was fully described and has no missing values. First 50 instances are denoted as *IUGR#1* with two label sets: L_1 classifies children into *symmetrical*, *rather symmetrical* or *asymmetrical* types of IUGR, L_2 refers to nocturnal hypertension and determines whether the patient is a dipper or non-dipper. The second *IUGR#2*—is described by two label sets: the first label set L_1 pertains the same as for *IUGR#1*, while the second label set L_2 refers to high blood pressure diagnostics and classifies samples as *prehypertension* and *hypertension*.

The "CTG" dataset comes from UCI Machine Learning Repository [7]. It contains the Fetal Heart Rate, measurements from cardiotocography, and the diagnosis group classified by gynecologist. There are 21 attributes in the dataset, including 11 continuous, 9 discrete and 1 nominal scales. The details for the dataset were presented

Table 1 Datasets characteristics

Dataset	Instances	Attributes	Label set L_1	Label set L_2
JIA#1	50	16	3	3
JIA#2	67	15	4	4
IUGR#1	50	120	3	2
IUGR#2	24	120	3	2
CTG	650	21	3	10

in [5]. The number of observations covers 2126 cases in total. For the purpose of our investigations, subsamples of 650 instances are built. Each specimen is described by two labels derived from two separate classes. First label classifies cases as: N—normal, S—suspect, and P—pathologic. The second label refers to fetal heart rate (FHR) and takes values 1–10. The main characteristics of the considered datasets, such as the number of data instances, attributes or labels are presented in Table 1.

The aim of the experiments was to examine the performance of the proposed technique as supporting of children diagnostics. The experiments were conducted on five datasets: JIA#1, JIA#2, IUGR#1, IUGR#2 and CTG. During experiments *Classification Accuracy (CA)* and *Hamming Loss (HL)* were calculated, and compared for the considered IL and LC methods as well as distinct problem transformation techniques of multi-label classification: popular BR and LP techniques.

For all the datasets the experiments were conducted twice: for training (2/3 of all the instances) and test set (1/3 of instances), and also by using 10-fold cross-validation. Tables 2, 3, 4, 5, and 6 present average values of *CA* and *HL*, with standard deviations from ten repeated classification processes for randomized instances. In the case of LC method, two possible labels' orders were considered and the best values were taken into account from all single classifications.

All the examined techniques were conjunct with five single-label classifiers: *k*-nearest neighbours kNN, naive Bayes NB, support vector machine SVM, multilayer perceptron MLP and C4.5 decision tree [17]. The software implemented for experiments was based on WEKA Open Source [4] with its default parameters.

The last step of experiments concerned statistical analysis of considered classification techniques. All the classifiers over all the datasets were compared using the Friedman test with post-hoc analysis.

Results for *JIA#1* dataset are presented in Table 2 and Fig. 1. For all experiments and classifiers the best values of the both CA and HL were obtained for LC method. The higher result of CA, equal to $72 \pm 13.98\,\%$, and at the same time the smaller value $10 \pm 4.71\,\%$ of HL, were obtained by MLP classifier for divided dataset into 2/3 training and 1/3 test set. For 10-fold cross-validations the best value can be noticed also for considered LC method and MLP classifier with effectiveness of $56.71 \pm 3.71\,\%$. The worst effects were reported for BR method with C4.5 classifier ($29.66 \pm 18.63\,\%$ of CA) and SVM classifier ($27.49 \pm 0.50\,\%$ of HL), both for experiments with 10-fold cross-validation.

Table 2 JIA#1: Classification Accuracy (CA) (%) and Hamming Loss (HL) (%)

		Binary relevance	Label power-set	Independent labels	Lables chain
2/3 training set, 1/3 test set					
kNN	CA	43.75 ± 12.15	43.75 ± 12.15	43.75 ± 12.15	48.00 ± 16.87
	HL	25.63 ± 5.65	25.63 ± 5.65	25.63 ± 5.65	20.67 ± 7.98
NB	CA	33.13 ± 11.43	48.75 ± 11.71	50.00 ± 13.82	68.00 ± 16.87
	HL	19.38 ± 3.52	20.83 ± 7.15	19.38 ± 4.40	12.00 ± 6.13
SVM	CA	43.13 ± 16.26	50.00 ± 8.84	50.00 ± 8.84	66.00 ± 16.47
	HL	26.46 ± 3.11	25.62 ± 4.82	25.62 ± 4.82	16.67 ± 11.00
MLP	CA	46.88 ± 8.96	50.63 ± 15.15	61.25 ± 7.10	72.00 ± 13.98
	HL	17.71 ± 3.51	20.00 ± 6.07	15.42 ± 3.83	10.00 ± 4.71
C4.5	CA	33.75 ± 7.34	53.13 ± 7.93	48.13 ± 9.79	64.00 ± 18.38
	HL	18.34 ± 3.65	17.29 ± 4.06	19.79 ± 3.84	14.67 ± 8.20
10-fold cross validation					
kNN	CA	38.28 ± 2.29	38.28 ± 2.29	38.28 ± 2.29	39.89 ± 2.83
	HL	26.12 ± 0.78	26.12 ± 0.78	26.12 ± 0.78	25.92 ± 0.98
NB	CA	34.07 ± 2.62	44.29 ± 2.68	52.90 ± 4.07	53.90 ± 3.72
	HL	18.80 ± 1.02	21.24 ± 1.02	17.30 ± 1.36	16.90 ± 1.20
SVM	CA	43.48 ± 1.65	44.09 ± 0.28	44.09 ± 0.28	44.09 ± 0.28
	HL	27.49 ± 0.50	27.32 ± 0.04	27.32 ± 0.04	27.32 ± 0.04
MLP	CA	44.90 ± 5.09	47.49 ± 3.12	56.70 ± 5.29	56.71 ± 3.71
	HL	17.27 ± 1.83	20.71 ± 1.40	15.70 ± 2.28	15.37 ± 0.97
C4.5	CA	29.66 ± 4.51	50.69 ± 4.14	48.88 ± 5.54	48.88 ± 5.54
	HL	18.68 ± 1.40	18.31 ± 2.02	19.32 ± 2.67	19.25 ± 2.52

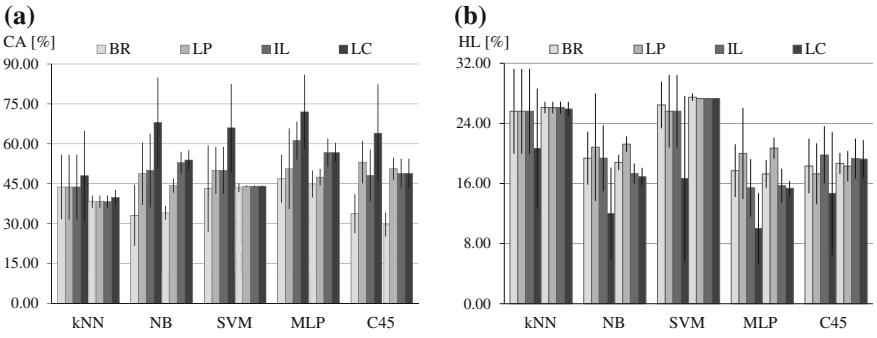


Fig. 1 JIA#1: two series—2/3 test and 1/3 training set, and 10-fold cross-validation. **a** Classification Accuracy (CA) (%). **b** Hamming Loss (HL) (%)

Table 3 and Fig. 2 present results of experiments for *JIA#2* dataset. Also this time, in all the cases the best effectiveness of the classification process was observed for LC technique without cross-validation, especially with C4.5 classifier—82.86 ±

Table 3 JIA#2: Classification Accuracy (CA) (%) and Hamming Loss (HL) (%)

		Binary relevance	Label power-set	Independent labels	Lables chain
2/3 training set, 1/3 test set					
kNN	CA	49.09 ± 6.36	49.09 ± 6.36	49.09 ± 6.36	54.29 ± 15.67
	HL	18.07 ± 2.70	18.07 ± 2.70	18.07 ± 2.70	15.36 ± 6.08
NB	CA	30.45 ± 12.31	57.73 ± 5.27	63.18 ± 6.23	68.57 ± 17.56
	HL	15.85 ± 2.47	13.30 ± 2.28	11.36 ± 2.34	9.64 ± 5.85
SVM	CA	0.00 ± 0.00	30.45 ± 7.74	33.64 ± 8.62	42.86 ± 17.82
	HL	26.02 ± 2.30	30.00 ± 4.61	28.07 ± 3.43	25.00 ± 4.12
MLP	CA	39.09 ± 9.63	50.45 ± 7.86	51.82 ± 8.08	71.43 ± 9.52
	HL	16.53 ± 2.75	18.07 ± 4.20	15.91 ± 4.18	8.93 ± 3.47
C4.5	CA	49.55 ± 14.60	68.18 ± 7.42	64.09 ± 9.69	82.86 ± 13.13
	HL	11.42 ± 3.02	9.09 ± 2.27	10.68 ± 2.95	5.00 ± 3.84
10-fold cross validation					
kNN	CA	50.45 ± 2.61	50.45 ± 2.61	50.45 ± 2.61	46.87 ± 2.56
	HL	18.36 ± 1.15	18.36 ± 1.15	18.36 ± 1.15	19.52 ± 1.00
NB	CA	41.49 ± 3.36	54.93 ± 2.09	61.19 ± 2.90	61.79 ± 2.46
	HL	13.71 ± 0.94	14.22 ± 0.83	12.05 ± 1.09	12.02 ± 0.84
SVM	CA	0.00 ± 0.00	32.84 ± 0.00	32.84 ± 0.00	32.84 ± 0.00
	HL	25.47 ± 0.49	27.99 ± 0.00	27.99 ± 0.00	27.99 ± 0.00
MLP	CA	41.34 ± 3.30	50.60 ± 2.28	49.10 ± 2.48	50.90 ± 1.92
	HL	16.72 ± 1.11	17.35 ± 1.37	15.93 ± 0.58	15.37 ± 0.93
C4.5	CA	47.46 ± 3.13	62.84 ± 2.94	60.00 ± 3.90	60.75 ± 3.52
	HL	12.76 ± 1.34	10.60 ± 0.94	11.49 ± 1.33	11.46 ± 1.19

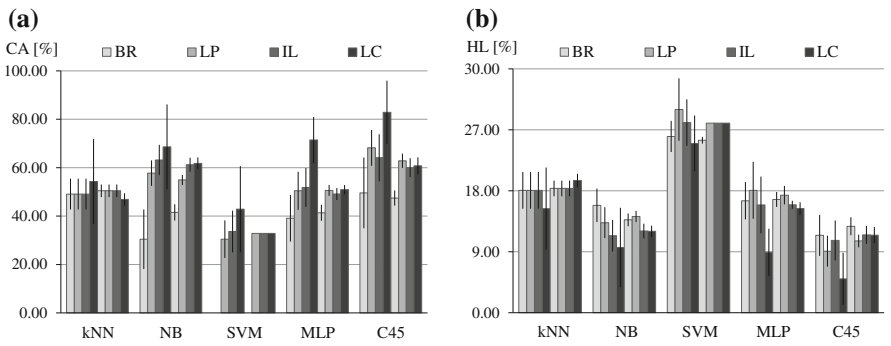


Fig. 2 JIA#2: two series—2/3 test and 1/3 training set, and 10-fold cross-validation. **a** Classification Accuracy (CA) (%). **b** Hamming Loss (HL) (%)

13.13 % of CA and 5.00 ± 3.84 % of HL values. For comparison, the lower value of CA was even equal to 0.00 % (for BR method) and the higher one of HL—30.00 ± 4.61 % (for LP technique), both in conjunction with SVM classifier. This time the worst results are comparable for two series of experiments, with and without 10-fold cross-validation.

The results for the third dataset *IUGR#1* are placed in Table 4 and Fig. 3. Similarly to the previous dataset, the best results of both metrics were obtained for the conjunction of LC method and C4.5 classifier, for divided dataset into training and test sets—62.00 ± 27.41 % of CA and 19.20 ± 15.64 % of HL. The worst results can be again noticed for Binary Relevance with SVM (CA equal to 17.50 ± 13.11 %) and with NB (HL equal to 39.75 ± 6.40 %), also this time without cross-validation. 10-fold cross-validation allowed to obtain results of CA from about 17–50 % and HL equals about 22–39 %.

Table 5 and Fig. 4 show results for the *IUGR#2* dataset. Also in this case, LC method gave the best results—70.00 ± 18.92 % of CA and 14.67 ± 9.84 % of HL, both for NB classifier and no cross-validation. With 10-fold cross-validation the best

Table 4 IUGR#1: Classification Accuracy (CA) (%) and Hamming Loss (HL) (%)

		Binary relevance	Label power-set	Independent labels	Lables chain
2/3 training set, 1/3 test set					
kNN	CA	32.50 ± 9.22	32.50 ± 9.22	32.50 ± 9.22	50.00 ± 25.39
	HL	32.75 ± 6.61	32.75 ± 6.61	32.75 ± 6.61	23.20 ± 12.19
NB	CA	18.13 ± 11.95	25.00 ± 12.84	28.75 ± 14.19	40.00 ± 18.86
	HL	39.75 ± 6.40	37.50 ± 7.73	39.00 ± 7.66	28.00 ± 10.83
SVM	CA	17.50 ± 13.11	30.63 ± 10.40	30.63 ± 10.40	46.00 ± 18.97
	HL	33.25 ± 2.06	31.25 ± 4.29	31.25 ± 4.29	24.00 ± 10.67
MLP	CA	31.88 ± 9.52	41.88 ± 10.64	39.38 ± 10.64	54.00 ± 18.97
	HL	30.00 ± 7.22	29.25 ± 6.78	30.50 ± 7.62	23.20 ± 10.96
C4.5	CA	28.13 ± 9.43	41.88 ± 16.42	49.38 ± 9.52	62.00 ± 27.41
	HL	29.63 ± 5.65	25.25 ± 6.17	22.75 ± 5.20	19.20 ± 15.64
10-fold cross validation					
kNN	CA	26.20 ± 3.19	26.20 ± 3.19	26.20 ± 3.19	28.60 ± 2.99
	HL	37.68 ± 1.43	37.68 ± 1.43	37.68 ± 1.43	36.72 ± 1.10
NB	CA	17.00 ± 2.36	27.60 ± 3.98	27.40 ± 2.67	27.40 ± 2.67
	HL	39.48 ± 1.92	34.56 ± 2.84	36.72 ± 2.11	36.56 ± 2.17
SVM	CA	30.60 ± 3.27	32.00 ± 0.00	32.00 ± 0.00	32.00 ± 0.00
	HL	31.80 ± 1.35	31.20 ± 0.00	31.20 ± 0.00	31.20 ± 0.00
MLP	CA	32.40 ± 5.06	42.80 ± 3.29	40.40 ± 5.32	42.80 ± 3.91
	HL	29.20 ± 2.10	29.36 ± 1.73	28.72 ± 2.98	27.36 ± 2.44
C4.5	CA	32.80 ± 3.79	42.60 ± 3.27	50.20 ± 3.58	50.20 ± 3.58
	HL	25.64 ± 2.02	25.20 ± 1.57	22.16 ± 2.17	22.16 ± 2.17

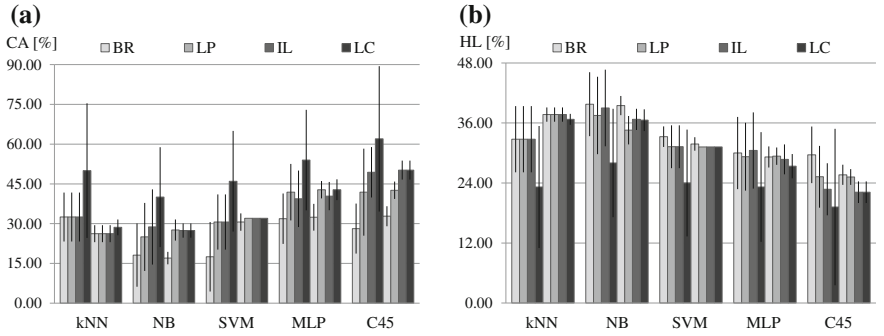


Fig. 3 IUGR#1: two series—2/3 test and 1/3 training set, and 10-fold cross-validation. **a** Classification Accuracy (CA) (%). **b** Hamming Loss (HL) (%)

result of effectiveness was much lower, equals 37.91 ± 6.65 %. The worst results of CA can be noticed for the conjunction of BR and SVM (15.00 ± 14.19 %), and the higher value of HL was obtained for three methods: BR, LP and IL with kNN classifier (equal to 46.00 ± 6.15 %), all with no cross-validation.

The last dataset *CTG* gave the most comparable results for the tested methods (Table 6, Fig. 5). However, also in this case, LC method allowed to obtain the highest value of CA—78.82 ± 1.62 % for C4.5 and experiments for divided dataset. Also this time BR (with SVM classifier) proved to give the worst CA result, while the higher value of HL can be noticed for the both LP and IL methods, for conjunction with SVM classifier and with cross-validation.

Summing up the investigations, conducted experiments on all the datasets, gave generally the best classification effectiveness for proposed LC technique. The difference between commonly used BR and LP algorithms, and the proposed method is from about 1 % to even 50 %. What is more, experiments conducted for divided datasets into training and test sets (proportion 2/3–1/3) and again with 10-fold cross-validation present the potential of the proposed method in different classification conditions.

Statistical analysis showed the advantage of LC method. The overall significance level of differences for all mentioned classification techniques using the Friedman test was below 0.01. The detailed post-hoc analysis was performed by comparing absolute differences between average ranks of classifiers over all the considered datasets. Values of differences between LC and the rest of classification approaches for all classifiers were above the threshold of 0.05 significance level, what proves that LC method outperformed the rest of the techniques.

Comparing IL and LC methods, one can conclude, that experiments showed the advantage of using multi-label techniques for multi-perspective classification problems of children diagnostics, what indicates on the importance of mapping dependencies between labels during classification process.

Table 5 IUGR#2: Classification Accuracy (CA) (%) and Hamming Loss (HL) (%)

		Binary relevance	Label power-set	Independent labels	Lables chain
2/3 training set, 1/3 test set					
kNN	CA	18.75 ± 10.62	18.75 ± 10.62	18.75 ± 10.62	56.67 ± 27.44
	HL	46.00 ± 6.15	46.00 ± 6.15	46.00 ± 6.15	22.67 ± 16.69
NB	CA	33.75 ± 13.24	45.00 ± 13.44	46.25 ± 15.65	70.00 ± 18.92
	HL	29.75 ± 8.45	30.00 ± 8.82	30.50 ± 9.56	14.67 ± 9.84
SVM	CA	15.00 ± 14.19	25.00 ± 10.21	28.75 ± 8.44	50.00 ± 17.57
	HL	37.50 ± 7.17	37.00 ± 10.59	34.00 ± 8.43	20.00 ± 7.03
MLP	CA	18.75 ± 8.84	26.25 ± 10.94	23.75 ± 10.94	43.33 ± 16.10
	HL	39.75 ± 6.61	40.50 ± 5.99	38.50 ± 6.69	29.33 ± 10.52
C4.5	CA	17.50 ± 12.08	35.00 ± 18.45	35.00 ± 11.49	66.67 ± 15.71
	HL	33.75 ± 8.92	36.50 ± 11.32	29.50 ± 7.25	18.67 ± 9.33
10-fold cross validation					
kNN	CA	20.00 ± 4.30	20.00 ± 4.30	20.00 ± 4.30	22.08 ± 5.22
	HL	45.83 ± 3.17	45.83 ± 3.17	45.83 ± 3.17	44.83 ± 2.15
NB	CA	24.17 ± 3.83	39.17 ± 3.51	35.42 ± 4.05	35.42 ± 4.05
	HL	34.25 ± 2.24	33.50 ± 2.00	37.00 ± 2.19	37.00 ± 2.19
SVM	CA	20.42 ± 6.35	17.50 ± 6.46	29.17 ± 0.00	29.17 ± 0.00
	HL	36.00 ± 2.11	39.50 ± 3.43	33.33 ± 0.00	33.33 ± 0.00
MLP	CA	19.17 ± 4.48	25.83 ± 8.52	20.83 ± 5.20	22.92 ± 4.50
	HL	37.33 ± 2.38	40.17 ± 5.90	38.50 ± 2.77	38.00 ± 2.33
C4.5	CA	27.50 ± 7.40	32.50 ± 8.52	37.92 ± 6.65	37.91 ± 6.65
	HL	29.33 ± 4.33	35.00 ± 5.09	28.17 ± 3.81	28.17 ± 3.81

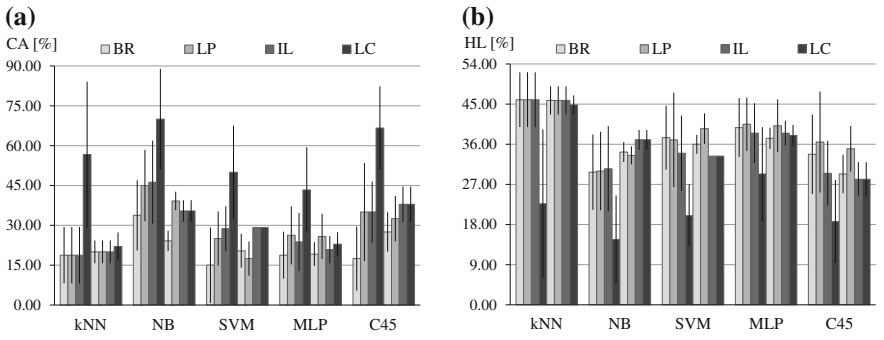


Fig. 4 IUGR#2: two series—2/3 test and 1/3 training set, and 10-fold cross-validation. **a** Classification Accuracy (CA) (%). **b** Hamming Loss (HL) (%)

Table 6 CTG: Classification Accuracy (CA) (%) and Hamming Loss (HL) (%)

		Binary relevance	Label power-set	Independent labels	Lables chain
2/3 training set, 1/3 test set					
kNN	CA	68.18 ± 2.90	68.18 ± 2.90	68.18 ± 2.90	69.02 ± 6.23
	HL	6.39 ± 0.75	7.97 ± 0.80	7.97 ± 0.80	7.30 ± 1.14
NB	CA	24.67 ± 3.42	69.35 ± 2.88	63.60 ± 3.58	67.47 ± 4.48
	HL	12.71 ± 0.83	8.35 ± 0.70	8.66 ± 0.69	7.84 ± 1.21
SVM	CA	3.55 ± 0.74	33.13 ± 2.87	33.04 ± 3.01	37.75 ± 3.92
	HL	10.64 ± 0.44	13.52 ± 0.68	13.52 ± 0.70	12.65 ± 0.93
MLP	CA	62.20 ± 4.15	76.26 ± 3.71	73.46 ± 3.14	76.76 ± 3.47
	HL	5.09 ± 0.35	6.71 ± 0.92	6.57 ± 0.62	5.75 ± 0.98
C4.5	CA	63.83 ± 2.50	77.15 ± 2.34	73.88 ± 3.43	78.03 ± 2.75
	HL	4.88 ± 0.58	6.60 ± 0.57	6.64 ± 0.66	5.32 ± 0.78
10-fold cross validation					
kNN	CA	70.12 ± 2.05	70.12 ± 2.05	70.12 ± 2.05	69.43 ± 2.15
	HL	6.08 ± 0.45	7.76 ± 0.53	7.76 ± 0.53	7.35 ± 0.49
NB	CA	23.66 ± 3.43	69.63 ± 2.57	63.08 ± 2.92	63.85 ± 2.75
	HL	13.13 ± 0.52	8.31 ± 0.60	8.80 ± 0.59	8.64 ± 0.66
SVM	CA	5.43 ± 0.91	33.23 ± 1.09	33.11 ± 0.95	33.52 ± 1.03
	HL	10.52 ± 0.26	13.58 ± 0.33	13.57 ± 0.33	13.48 ± 0.34
MLP	CA	64.49 ± 1.66	77.17 ± 1.53	74.38 ± 1.66	75.38 ± 1.59
	HL	4.84 ± 0.23	6.65 ± 0.32	6.52 ± 0.36	6.18 ± 0.36
C4.5	CA	66.71 ± 1.90	78.82 ± 1.62	76.45 ± 1.87	77.10 ± 1.78
	HL	4.42 ± 0.26	6.45 ± 0.40	6.33 ± 0.46	5.87 ± 0.45

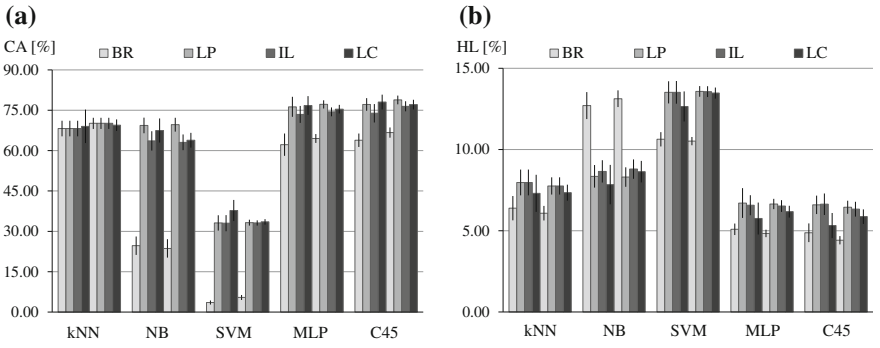


Fig. 5 CTG: two series—2/3 test and 1/3 training set, and 10-fold cross-validation. **a** Classification Accuracy (CA) (%). **b** Hamming Loss (HL) (%)

5 Concluding Remarks

In the paper, improving children diagnostics by using multi-label classification is considered. The application of the problem transformation method Labels Chain, which performs well when the number of attributes significantly exceeds the number of instances, is examined. The proposed methodology is evaluated by experiments done on real patient records. Investigations showed the good performance of the LC technique on the considered medical datasets taking into account two metrics: Classification Accuracy and Hamming Loss. Statistical analysis based on the Friedman test with post-hoc analysis also proved the advantage of the proposed method. What is more, experiment results showed that using multi-label classification can improve multi-perspective classification of children diagnostics, while dependencies between labels are mapped during classification process.

Future research will consist in further investigations of the problem transformation methods for medical diagnostics purpose taking into account different evaluation metrics as well as different kind of datasets. As the result building the tool which will support doctors in diagnostics process will be planned.

References

1. Bromuri, S., Zufferey, D., Hennebert, J., Schumacher, M.: Multi-label classification of chronically ill patients with bag of words and supervised dimensionality reduction algorithms. J. Biomed. Inform. **51**, 165–175 (2014). doi:10.1016/j.jbi.2014.05.010
2. Demsar, J.: Statistical comparisons of classifiers over multiple data sets. J. Mach. Learn. Res. **7**, 1–30 (2006)
3. Glinka, K., Zakrzewska, D.: Effective multi-label classification method for multidimensional datasets. In: Andreasen, T., et al. (eds.) Flexible Query Answering Systems 2015, Advances in Intelligent Systems and Computing, vol. 400, pp. 127–138. Springer International Publishing, Switzerland (2016)
4. http://www.cs.waikato.ac.nz/ml/weka/index.html
5. Huang, M.-L., Yung-Yan, H.: Fetal distress prediction using discriminant analysis, decision tree, and artificial neural network. J. Biomed. Sci. Eng. **5**, 526–533 (2012)
6. Kajdanowicz, T., Kazienko, P.: Multi-label classification using error correcting output codes. Appl. Math. Comput. Sci. **22**(4), 829–840 (2012)
7. Lichman, M.: UCI Machine Learning Repository http://archive.ics.uci.edu/ml. University of California, School of Information and Computer Science, Irvine, CA (2013)
8. Mosley, E.: Practical guide to treating children with juvenile idiopathic arthritis. Paediatr. Child Health **25**(12), 587–591 (2015). doi:10.1016/j.paed.2015.09.001
9. Nasierding, G., Kouzani, A.B.: Comparative evaluation of multi-label classification methods. In: 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2012), pp. 679–683 (2012)
10. Qu, G., Zhang, H., Hartrick, C.T.: Multi-label classification with Bayes' theorem. In: 2011 4th International Conference on Biomedical Engineering and Informatics (BMEI), pp. 2281–2285 (2011)
11. Read, J., Pfahringer, B., Holmes, G., Frank, E.: Classifier chains for multi-label classification. In: Buntine, W., Grobelnik, M., Mladenic, D., Shawe-Taylor, J. (eds.) Machine Learning and Knowledge Discovery in Databases, LNCS, vol. 5782, pp. 254–269. Springer, Heidelberg (2009)

12. Rios, A., Kavuluru, R.: Supervised extraction of diagnosis codes from EMRs: role of feature selection, data selection and probabilistic thresholding. In: 2013 IEEE International Conference on HealthCare Informatics, pp. 66–73 (2013)
13. Schapire, R.E., Singer, Y.: BoosTexter: a boosting-based system for text categorization. Mach. Learn. **39**(2/3), 135–168 (2000)
14. Sun, J., Sow, D., Hu, J., Ebadollahi, S.: Supervised patient similarity measure of heterogeneous patient records. SIGKDD Explor. **14**, 16–24 (2012)
15. Szymańska-Kałuża, J., Cebula-Obrzut, B., Smolewski, P., Stańczyk, J., Smolewska, E.: Imbalance of Th17 and T-regulatory cells in peripheral blood and synovial fluid in treatment naive children with juvenile idiopathic arthritis. Cent.-Eur. J. Immunol. **1**, 71–76 (2014). doi:10.5114/ceji.2014.42128
16. Tsoumakas, G., Katakis, I., Vlahavas, I.: Mining Multi-label data. In: Maimon, O., Rokach, L. (eds.) Data Mining and Knowledge Discovery Handbook, pp. 667–685. Springer, Boston (2010)
17. Witten, I.H., Frank, E., Hall, M.A.: Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, San Francisco (2011)
18. Zamecznik, A., Niewiadomska-Jarosik, K., Zamojska, J., Stanczyk, J., Wosiak, A., Moll, J.: Intra-uterine growth restriction as a risk factor for hypertension in children six to 10 years old. Cardiovasc. J. Afr. **25**(2), 73–77 (2014). doi:10.5830/CVJA-2014-009
19. Zamojska, J., Niewiadomska–Jarosik, K., Wosiak, A., Lipiec, P., Stanczyk, J.: Myocardial dysfunction measured by tissue Doppler echocardiography in children with primary arterial hypertension. Kardiologia Polska 2015 **73**(3), 194–200 (2015). doi:10.5603/KPa2014.0189
20. Zha, R.-W., Li, S.-Z., Lu, J.-M., Wang, X.: Clinical multi-label free text classification by exploiting disease label relation. In: 2013 IEEE International Conference on Bioinformatics and Biomedicine, pp. 311–315 (2013)

Building Intelligent Classifiers for Doctor-Independent Parkinson's Disease Treatments

Artur Szymański, Stanisław Szlufik, Dariusz M. Koziorowski, Piotr Habela and Andrzej W. Przybyszewski

Abstract In Parkinson's disease (PD) neurodegenerative processes related to nerve cell deaths have different time courses in individual patients, requiring individually adjusted treatments to be performed under the supervision of an experienced neurologists. In this project, we have compared results of treatments performed by a neurologist with predictions of the data mining systems: WEKA and RSES (Rough Sets). We have analyzed a PD patients database of 800 neurological records of 94 patients from three different groups: (1) BMT—best medical treatment—patients only on pharmacological treatment, (2) POP—postoperative patients that have received an implanted DBS electrode before our study, (3) DBS—patients with stimulating electrodes that were implanted during our study. We have divided the data into training and testing sets with help of the machine learning improved classifiers based on a Random Forest (WEKA) or on decomposition trees (RSES). These reached an accuracy of 87 % in LDOPA dosage estimation, 80 % for UPDRS III and over 90 % for UPDS IV estimations. The precision of our individual patients symptoms and medications dosage predictions should be sufficient to give neurologists objective advice on optimal treatment for particular patients.

Keywords Data mining · Parkinson's disease · WEKA · RSES · DBS

A. Szymański (✉) · P. Habela · A.W. Przybyszewski
Polish Japanese Academy of Information Technology, Warsaw, Poland
e-mail: artur.szymanski@pja.edu.pl

S. Szlufik · D.M. Koziorowski
Faculty of Health Science, Neurology, Medical University, Warsaw, Poland

A.W. Przybyszewski
Medical School, University of Massachusetts, Worcester, USA

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_22

1 Introduction

Neurodegenerative diseases (ND) as of today are not curable disorders. The second most common ND after the Alzheimer Disease (AD) is PD. The best medical treatment (BMT) is the most common PD treatment that effectively improves a majority of the motor symptoms. In the certain cases, when the pharmacological treatment has strong side effects (fluctuations related to on-off medications effects) patients undergo the Deep Brain Stimulation (DBS) surgery. In our study, the DBS treatment was achieved by the electric stimulation of the STN (subthalamic nucleus) that is a part of the basal ganglia system affected in PD.

Both treatments been used for many years (DBS was invented in 1987), however symptoms development processes are different between patients, as well patients react differently to such treatments as BMT or DBS. Adjusting individually for each patient different treatment is complicated, time consuming, needs expertise knowledge of the neurology doctor and frequent patient visits in the clinic. In order to improve individual PD treatment, we propose to create a model that can help to assess those treatments settings with aid of the AI system.

There are many examples of the intelligent methods applications in diagnosis and treatments for various diseases, among them for another ND: AD. In 2014 Padma and Sukanesh [5] have used wavelet statistical texture features to classify when AD starts. Zhang et al. [11] created kernel support machine with 80 % accuracy in classifying of AD based on Magnetic Resonance Imaging. Based on their AI research, Zhang and Wang [10] have published work on classification of AD affected regions of brain with the accuracy over 90 %.

Our previous models of the computer-aided treatments are demonstrated in several papers on machine learning applications in the various sets of measurements related to PD [3, 8, 9]. In our previous work we have proposed a model that was based on the Diffusion Tensor Imaging data combined with MRI in order to optimize the DBS parameters and their effectiveness to improve patient's symptoms (with help of Rough Set Theory [8]). In the following paper [9] we have extended our model with support of WEKA Random Forest implementation. We have also performed simulations of the neuronal population model that results were in agreement with the clinical DBS effects. Another approach was undertaken by Kubis et al. [3] to facilitate usage of the eye movement measurements with help of the Fuzzy Rough Set, which showed the importance of Eye Movement (EM) measurements, in the PD diagnosis.

In this work we have used neurological and EM data to create intelligent classifiers applied to the moderate size DB (database) of PD patients. It is an extension to our previous work on creation of the computer assist models in order to help and make more objective and doctor-independent diagnosis and treatment of PD.

2 Methods

We have created database with the neurological examination results of PD patients. All patients were divided into three groups with respect to their PD symptoms and treatment. Those groups were:

- BMT—best medical treatment—patients treated only pharmacologically,
- POP—postoperative patients that have received an implanted DBS electrode before our study,
- DBS—patients with stimulating electrodes that were implanted during our study.

Each record in our database represents results of neurological examination of POP and DBS patients in four sessions designed as S1: MedOffDBSOff, S2: MedOffDBSOn, S3: MedOnDBSOff, S4: MedOnDBSOn. During the first session (S1) the patient was off medications (L-Dopa) and DBS stimulators was OFF; in the second session (S2) the patient was off medication, but the stimulator was ON; in the third session (S3) the patient was after his/her doses of L-Dopa and the stimulator was OFF, and in the fourth session (S4) the patient was on medication with the stimulator ON. In BMT patients had only two sessions S1: MedOff, S2: MedOn. Changes in motor performance, behavioral dysfunction, cognitive impairment and functional disability were evaluated in each session according to the UPDRS:

- UPDRS I—mentation, behavior and mood,
- UPDRS II—activities of daily living,
- UPDRS III—motor examination,
- UPDRS IV—complications of therapy.

Each record contains results of the patient's neurological examination during his/her following visits every several months. We have ordered all visits as a number of months from the beginning of our project. In the case of the database (DB) of patients divided into groups, the first testing was performed before the DBS surgery (PREOP) and the following tests were performed postoperatively (POSTOP) after DBS electrode implementation.

Our DB consists of the standard personal (age, gender, etc.) and the neurological data with time of the examination relative to the DBS surgery.

In addition to the standard neurological data, we have also collected the eye movement measurements of PD patients (Fig. 1). The EM measurements were performed by means of the saccadometer (Ober Consulting, Poland). This infrared device has recorded only fast reflexive eye movements (saccades) and it calculated automatically: latency, amplitude, duration and velocity of the eye saccades. The saccadometer was attached to the patient's head that help to record the eye movements not disturbed by the patient's head movements. During measurement patient was placed in the front of solid object like a wall in proximity of around 60 cm. Procedure begins with display of the red fixation marker in the center of the patient's visual field. In the next step, the fixation point disappears and a light spot appears on the left or the right side in the patient's visual field. Patient's task is to fixate eyes on the red

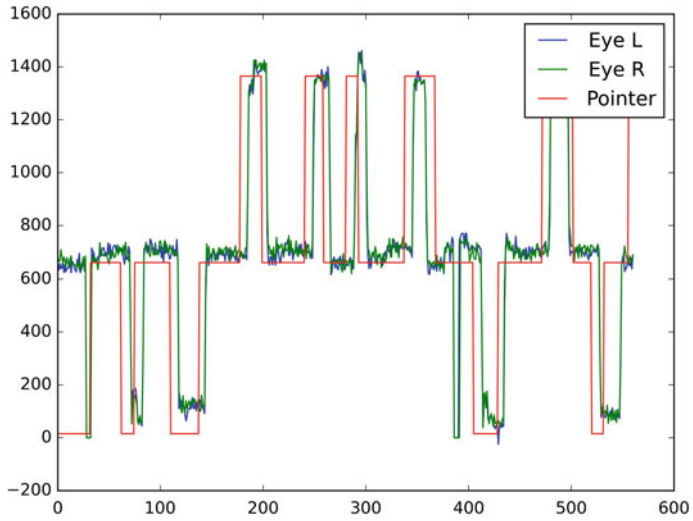


Fig. 1 Saccadic EM recording, *red line* representing pointer position, *green* and *blue lines* representing gaze position for *right* and *left* eye

marker, and follow light spot movements with a steady head. During each session, patient performs multiple reflexive saccades registered by the saccadometer.

All EM parameters were recorded for each eye separately. Since our other measurements were not related to any side of the body we have used attribute marked with ALL keyword, describing average for both eyes for each measurement.

As our database contains missing values for which we have calculated the missing value ratio that is represented by the first two digits of each attribute's name and gives percent of records with missing values for the whole data set. This ratio was calculated once and is relevant only for the whole data set.

We have analyzed our database with the following AI tools: Waikato Environment for Knowledge Analysis (WEKA) and Rough Set Exploration System (RSES).

Our database measurements were imported to WEKA through the flat file in a string format, meaning all attributes were initially defined as string fields. This approach allowed for easy data import into the attribute relation file format (ARFF) in WEKA. The advantage of this method was that we do not have to worry during import process about exact data format, also about writing validation tools and punctuations that can be system specific. Those issues could be dealt later by using the robust data manipulation filters included in the WEKA.

As a first step we have performed attribute conversion to the native types using attribute filters in WEKA, like 'StringToNominal' for attributes with limited defined subset of the possible values. These attributes were for example DBS, BMT or timing (visits dates) parameter (as explained above was strictly defined number of months after the surgery). Other attributes were converted to the numbers using 'StringToNumber' filter in WEKA.

After performing initial type conversion on all attributes, we have performed binning on certain attributes which we used further for classification using the discretize filter from WEKA. This procedure helped us create ranges of values thus boosting accuracy of the model. In this filter we have few options that can be set and influence further results, those are:

- Number of bins,
- Use equal frequency,
- Automatically find number of bins.

After dataset preparation, we have applied two different machine-learning algorithms: Random Forest from WEKA [6] implementation and Rough Set Theory classifier using the decomposition tree classifier from RSES [1]. After initial classification, data have been migrated to RSES format using native ARFF import function. All classifiers based on machine learning algorithms were assessed through application of the 4-fold cross validation method (Table 1).

Table 1 Excerpt from the decision table of our classifiers

00_PatientNumber	05/PD/BMT 2013	05/PD/POP/ 2009	49/PD/DBS/ 2013
00_YearOfBirth	1965.0	1956.0	1956.0
00_Gender	0.0	1.0	1.0
00_GroupName	BMT	POP	DBS
00_MonthsAfterDBS	18.0	0.0	0.0
LDOPA	'(560–1120]'	'(1120–1680]'	'(1120–1680]'
02_DBS	0.0	0.0	0.0
02_BMT	1.0	0.0	0.0
44_UPDRS_I	0.0	0.0	5.0
45_UPDRS_II	3.0	13.0	28.0
02_UPDRS_III	10.0	27.0	37.0
44_UPDRS_IV	2.0	6.0	14.0
47_HYscale	1.0	2.5	4.0
Schwab and England	90.0	70.0	50.0
58_SaccadesLatencyMeanALL	229.0		249.0
58_SaccadesDurationALL	54.0		61.0
58_SaccadesAmplitudeALL	26.6		9.2
58_SaccadesPeakVelocityALL	948.0		313.0

It presents data for above mentioned attributes as example in three patients, from groups: BMT, POP and DBS

3 Results

As a first, we have built a classifier able to predict UPDRS subsection for each patient based on the one-visit data. This approach is useful when we want to automatize classification of patient's symptoms based on the neurological tests and on the eye movement measurements.

Those tests were run on all instances from the database, and we had used discretized filter to bin classification attribute into the five groups. Attributes that were included in this model were:

- 00_PatientNumber
- 00_YearOfBirth
- 00_Gender
- 00_GroupName
- 00_MonthsAfterDBS
- 25_DiseaseDuration
- 28_LDopaDosage
- 02_DBS
- 02_BMT
- 58_SaccadesLatencyMeanALL
- 58_SaccadesDurationALL
- 58_SaccadesAmplitudeALL
- 58_SaccadesPeakVelocityALL

After extracting above attributes, we have divided our data into three groups and performed classification for different UPDRS subscales. Best results were acquired for UPDRS three and four:

UPDRS III:

- BMT: 80.83 %
- POP: 55.89 %
- DBS: 49.9 %

UPDRS IV:

- BMT: 80.48 %
- POP: 87.60 %
- DBS: 94.07 %

In the second approach, we have built classifiers for the medical treatment dosage (the most common PD medical treatment is LDOPA). This information is useful in in the systems for assisting with the medication dose adjustment for PD patients. We have used following attribute set:

- 00_PatientNumber
- 00_YearOfBirth
- 00_Gender

- 00_GroupName
- 00_MonthsAfterDBS
- 25_DiseaseDuration
- 02_DBS
- 02_BMT
- 44_UPDRS_I
- 45_UPDRS_II
- 02_UPDRS_III
- 44_UPDRS_IV
- 03_UPDRS_TOTAL
- 47_HYscale
- 47_SchwabEnglandScale
- 58_SaccadesLatencyMeanALL
- 58_SaccadesDurationALL
- 58_SaccadesAmplitudeALL
- 58_SaccadesPeakVelocityALL

In the third approach, we have dived patients into three groups, and LDOPA dosage into five groups. We have obtained the following results for LDOPA dosage classification:

- ALL: 87 %
- BMT: 78 %
- POP: 91 %
- DBS: 93.6 %

In Table 2 we can see detailed accuracy matrix for ALL group:

There are true positives rates at the level of above 90 % for all classes that gives a good result because, as one can notice, in the confusion matrix (Table 3) the number of elements in different classes is not even.

Table 2 Table of the accuracy for five different LDOPA dosage classes; True Positives (TP)—instances classified correctly as given class, False Positive (FP) rate—instances classified wrongly as given class, precision—defined as TP/(TP + FP), recall—defined as TP/(TP + FN), F-measure – 2 * Precision * Recall/(Precision + Recall), ROC Area—showing measure how optimal our classifier is (values closer to 1)

	TP rate	FP rate	Precision	Recall	F-measure	ROC Area	Class
Weighted Avg.	0.95	0.033	0.905	0.95	0.927	0.94	'(-inf–560]'
	0.958	0.045	0.902	0.958	0.929	0.971	'(560–1120]'
	0.9	0	1	0.9	0.947	0.993	'(1120–1680]'
	0.917	0	1	0.917	0.957	0.988	'(1680–2240]'
	0.95	0.007	0.95	0.95	0.95	0.998	'(2240–inf)'
	0.938	0.023	0.941	0.938	0.938	0.973	

Table 3 Confusion matrix presents number of the right (diagonal) and wrong (off-diagonal) LDOPA dosage classifications

a	b	c	d	e	<–	Classified
38	1	0	0	1	a	a = '(-inf–560]'
2	46	0	0	0	b	b = '(560–1120]'
0	4	36	0	0	c	c = '(1120–1680]'
1	0	0	11	0	d	d = '(1680–2240]'
1	0	0	0	19	e	e = '(2240–inf)'

For example: class a = (-inf–560) being correctly classified 38 times, and wrongly: two times classified as b and one time as d and one time as e class

We have rerun previous tests in RSES using four fold validation and decomposition tree classifier with results as follows:

UPDRS III classifier:

- POP group, 52 % accuracy, 25 % coverage
- BMT group, 91 % accuracy, 67.9 % coverage
- DBS group, 62 % accuracy, 36.9 % coverage

UPDRS IV classifier:

- POP group, 46.7 % accuracy, 23 % coverage
- BMT group, 57.7 % accuracy, 13.1 % coverage
- DBS group, 77 % accuracy, 25.4 % coverage

LDOPA dose classifier:

- ALL groups, 86.9 % accuracy, 30.2 % coverage
- POP group, 85.2 % accuracy, 39.6 % coverage
- BMT group, 96.9 % accuracy, 47.2 % coverage
- DBS group, 93.2 % accuracy, 39.4 % coverage

RSES classifiers were effective especially when predicting LDOPA dosage, but the coverage was still relatively low.

4 Discussion

There are many different approaches to computerize PD disease assessments. One of those attempts has proposed to estimate UPDRS by a multimodal interface called Virtual Touch Pad [4]. In his work, authors used camera recording of the standard finger tap and hand gestures (as in the UPDRS) tests and by classifying them into classes approximate the UPDRS scale. In our work, results of the automatic eye movement measurements, with help of other standard patients data like disease duration, medication dosage, etc., were sufficient to provide resulting classes of UPDRS for a given patient, with the accuracy reaching 90 % for some groups of patients.

Other studies [2, 7] have used the motion capture (MoCap) technology in order to obtain higher accuracy and quantitative assessment of the motor dysfunctions in PD patients. In [2] authors focused on predicting severity of motor symptoms in Parkinson's patients with DBS stimulation (on/off) and off-drugs. They have conducted study on these subjects and succeeded in building classifier by means of SVM for discriminating mild vs. severe symptoms with 90 % mean accuracy. In [7] authors have used the MoCap technology to measure walking parameters (changes in max/min hip angles and their velocities) and on this basis predicted the UPDRS. The UPDRS were divided into 6 classes and prediction was over 90 % for patients in different sessions: with DBS (off/on) and medication (off/on). However, predictions of the session number were only above 30 % because effects of medication/DBS in individual patients were strongly variable as registered by the gait trajectories changes [7]. The MoCap measurements give precise and objective data, but this technique is time consuming and requires specialized equipment. In our present study, we have used a standard neurological data and the eye movement (EM) measurements. The EM gave us additional objective information for classifiers that significantly increased our predictions. As the sample group of PD patients in our research was larger in comparison to other studies [2, 7] our actual results for a certain group of patients (different medical procedures) gave better predictions.

5 Conslusions

We have successfully created several classifiers for predicting values such as UPDRS III, IV or LDOPA dosage for the PD patients. It is important to highlight that this type of classifiers could be extremely useful in computer aided PD diagnosis and treatment.

As mentioned in the introduction, the treatment adjustment for a particular patient is not only very time consuming but also depend heavily on the doctor experience. Models that predict possible setup of PD treatment with accuracy over 90 % are very helpful and may determine future therapies. However, we did not get such accuracy for all our groups and for all parameters. One major reason is that there is a large variability of the disease duration, age, symptoms, etc. We probably need more data in order to get more consistent subgroups. We have obtained good predictions for the dosage of medication (LDOPA) in POP and DBS groups with accuracy level over 90 %. It was 78 % accuracy for the BMT (medication only) group, probably because medication doses for different patients are strongly variable.

Our further work will include additional EM measurements (antisaccades or pursuit EM) to increase accuracy for all patient groups and to extend precision of the classification e.g. stimulation current in DBS. Another approach is to use the machine learning algorithms in order to find correlation between different attributes that may demonstrate unknown functional connections opening up the possibilities of new treatments for PD patients.

Acknowledgments This work was partly supported by projects from Polish National Science Centre (Dec-2011/03/B/ST6/03816).

References

1. Bazan, J.G., Szczuka, M.: The rough set exploration system. In: Peters, J.F., Skowron, A. (eds.) Transactions on Rough Sets III. Number 3400 in Lecture Notes in Computer Science, pp. 37–56. Springer, Berlin, Heidelberg (2005)
2. Das, S., Trutoiu, L., Murai, A., Alcindor, D., Oh, M., De la Torre, F., Hodgins, J.: Quantitative measurement of motor symptoms in Parkinson's disease: a study with full-body motion capture data. In: Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, pp. 6789–6792 (2011)
3. Kubis, A., Szymański, A., Przybyszewski, A.W.: Fuzzy Rough sets theory applied to parameters of eye movements can help to predict effects of different treatments in Parkinson's patients. In: Kryszkiewicz, M., Bandyopadhyay, S., Rybinski, H., Pal, S.K. (eds.) Pattern Recognition and Machine Intelligence. Number 9124 in Lecture Notes in Computer Science, pp. 325–334. Springer International Publishing (2015). doi:10.1007/978-3-319-19941-2_31
4. Kupryjanow, A., Kunka, B., Kostek, B.: UPDRS tests for diagnosis of Parkinson's disease employing virtual-touchpad. In: 2010 Workshop on Database and Expert Systems Applications (DEXA), pp. 132–136 (2010)
5. Padma, A., Sukanesh, R.: Segmentation and classification of brain CT images using combined wavelet statistical texture features. Arab. J. Sci. Eng. **39**(2), 767–776 (2013)
6. Prinzie, A., Van den Poel, D.: Random forests for multiclass classification: random MultiNomial logit. Expert Syst. Appl. **34**(3), 1721–1732 (2008)
7. Przybyszewski, A.W., Boczarska, M., Kwiek, S., Wojciechowski, K.: Rough set based classifications of Parkinson's patients gaits. In: Nguyen, N.T., Attachoo, B., Trawiński, B., Somboonviwat, K. (eds.) Intelligent Information and Database Systems. Number 8398 in Lecture Notes in Computer Science, pp. 525–534. Springer International Publishing (2014). doi:10.1007/978-3-319-05458-2_54
8. Szymański, A., Przybyszewski, A.W.: rough set rules help to optimize parameters of deep brain stimulation in Parkinson's patients. In: Ślęzak, D., Tan, A.H., Peters, J.F., Schwabe, L. (eds.) Brain Informatics and Health. Number 8609 in Lecture Notes in Computer Science, pp. 345–356. Springer International Publishing (2014)
9. Szymański, A., Kubis, A., Przybyszewski, A.W.: Data mining and neural network simulations can help to improve deep brain stimulation effects in Parkinson's disease. Comput. Sci. **16**(2), 199 (2015)
10. Zhang, Y., Wang, S.: Classification of Alzheimer disease based on structural magnetic resonance imaging by kernel support vector machine decision tree. Prog. Electromagnet. Res. **144**, 185–191 (2014)
11. Zhang, Y., Dong, Z., Phillips, P., Wang, S., Ji, G., Yang, J., Yuan, T.F.: Detection of subjects and brain regions related to Alzheimer's disease using 3d MRI scans based on eigenbrain and machine learning. Front. Comput. Neurosci. **9**, 66 (2015)

Part II
Signal Processing

Towards Pulse Wave Velocity Estimation Using Specific ECG and PPG Algorithms

Stefan Borik and Branko Babusiak

Abstract The presented article details the design and evaluation of specific algorithms used for pulse wave velocity estimation. We chose to analyze both the ECG and PPG signals in order to estimate the time delay between the time when the blood is ejected into the aorta and the time when the pulse wave arrives to the point where the PPG probe is placed. Algorithms for the detection of R-waves in ECG signal and detection of slope beginnings in PPG signal were used.

Keywords Pulse wave velocity · Electrocardiography · Photoplethysmography · Algorithms · Arterial disease

1 Introduction

Arterial systems can be diagnosed using various methods. Especially useful are low-cost methods based on ECG or PPG signal acquisition and processing. Additionally, electrocardiography (ECG) is often the only method capable of detecting pathological heart processes. Generally speaking, cardiovascular diseases refer to any disease that affects the cardiovascular system, such as cardiac diseases, vascular diseases of the brain and kidneys and peripheral arterial diseases [3]. ECG is based on measurements of the electrical activity of the heart and plays an irreplaceable role in cardiology and internal medicine. Photoplethysmography (PPG) is based on the application of electromagnetic radiation from the visible spectrum to subcutaneous tissues and evaluation of reflection thereof. The detected signal corresponds to blood volume changes in the investigated area. PPG is widely used for evaluating the pulse wave which propagates from the aorta to peripheral circulation [1]. Combination of both

S. Borik (✉) · B. Babusiak
Faculty of Electrical Engineering, Department of Electromagnetic and Biomedical Engineering, University of Zilina, 010 26 Zilina, Slovakia
e-mail: stefan.borik@fel.uniza.sk
URL: http://fel.uniza.sk/ktebi

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_23

mentioned methods allows the evaluation of the pulse wave velocity, the so-called golden standard in vascular system diagnostics, which can be useful to diagnose possible pathological processes such as atherosclerosis [7].

2 Materials and Methods

2.1 Pulse Wave Velocity

Pulse wave velocity (PWV) is considered as the gold standard for the evaluation of arterial system properties. When the left ventricle ejects blood into aorta, the aorta diameter will change based on the heart activity and the pulse wave propagates through the arterial tree towards periphery. Different approaches are possible when detecting and measuring the pulse wave progress. By selecting two measurement points, we can use two photoplethysmography probes [2]. Another method uses a combination of impedance cardiography (ICG) [4] or impedance plethysmography with electrocardiograph. However, we chose a method based on the combination of ECG and PPG [5]. Combination of both mentioned signals enables the evaluation of time parameters pertaining to the blood ejected into the aorta. This corresponds to the R-peak of the QRS complex in the ECG record. Subsequently, PPG is used for the detection of the propagating pulse wave, corresponding with the end point of the arterial system whose parameters are evaluated. Calculation of the PWV is based on the following equation:

$$PWV = \frac{d}{t_{\text{delay}}}, \tag{1}$$

where d is distance between aorta beginning and PPG sensor placement, t_{delay} is pulse wave transition time (see Fig. 1).

2.2 Automatic R Wave Detection in ECG Signal

Detection of the R wave in the ECG signal is necessary for pulse wave velocity estimation. The R wave is part of the QRS complex (Fig. 1). The QRS complex

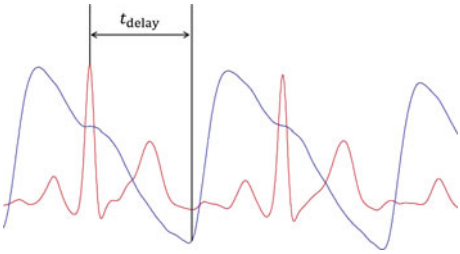


Fig. 1 Estimation of time delay between ECG and PPG

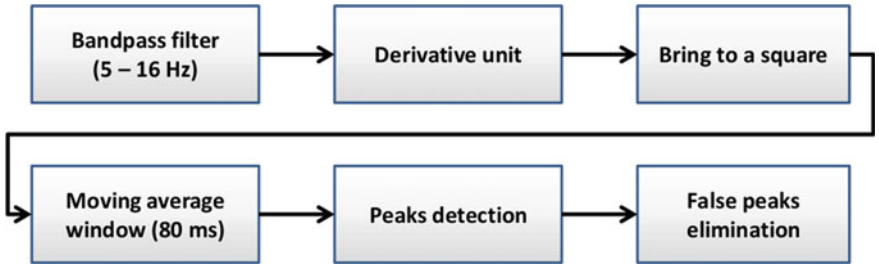


Fig. 2 Algorithm for automatic R wave detection

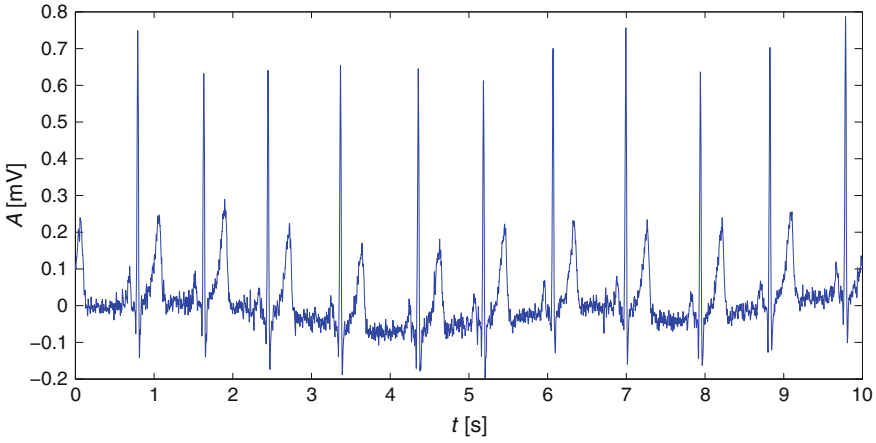


Fig. 3 Measured ECG

represents the rapid depolarization of the right and left ventricles. We employed an algorithm based on Pan-Tompkins method [6]—the actual algorithm for automatic R wave detection is shown in block diagram in Fig. 2.

A bandpass filter is used to emphasize the QRS complex. The main energy of the QRS complex lies in the frequency range from 5 to 16 Hz. A digital FIR filter was used to attenuate frequencies outside the scope of the mentioned frequency range. The transition band of the designed filter has a width of 2 Hz and attenuation in the stop band is equal to 80 dB. The filter order depends on the sampling rate of the ECG signal and can vary significantly because the detection algorithm is universal for various ECG recordings. The measured (original) ECG is displayed in Figs. 3 and 4 shows the resulting signal after filtration.

A derivative unit is applied to the filtered signal to highlight differences in the signal. The said unit calculates differences between adjacent elements of the signal vector. The resulting signal is shown in Fig. 5.

The next operation transforms the signal shown in Fig. 5 into positive values—however the resulting signal shown in Fig. 6 is raw and contains many peaks and local extrema (maxima and minima). This type of signal is complicated for further

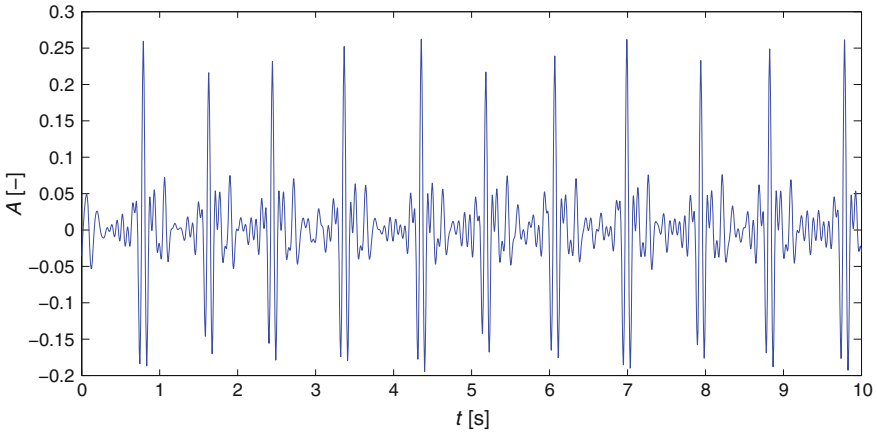


Fig. 4 ECG signal after digital filtration

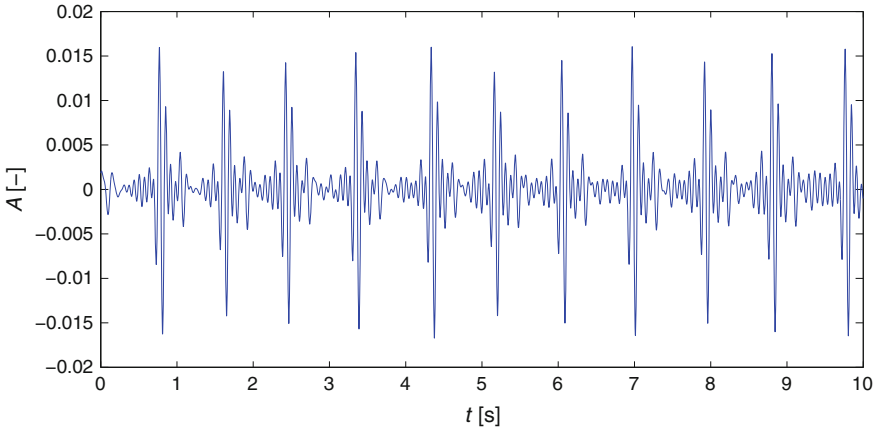


Fig. 5 First derivative of the signal

analysis and smoothing must be applied to enable better localization of maxima which correspond to the R wave of the ECG signal.

The signal is smoothed using a moving average filter and the said filter is defined as:

$$y[i] = \frac{1}{M} \sum_{j=1}^{M-1} x[i + j], \tag{2}$$

where $x[]$ is the input signal, $y[]$ is the output signal, M is the number of points (samples) used in the moving average.

The M parameter in Eq. (2) is also the length of the moving average window. The length of the window is set to 80 ms because this time interval corresponds with the duration of the QRS complex. The time of 80 ms will match with 80 samples

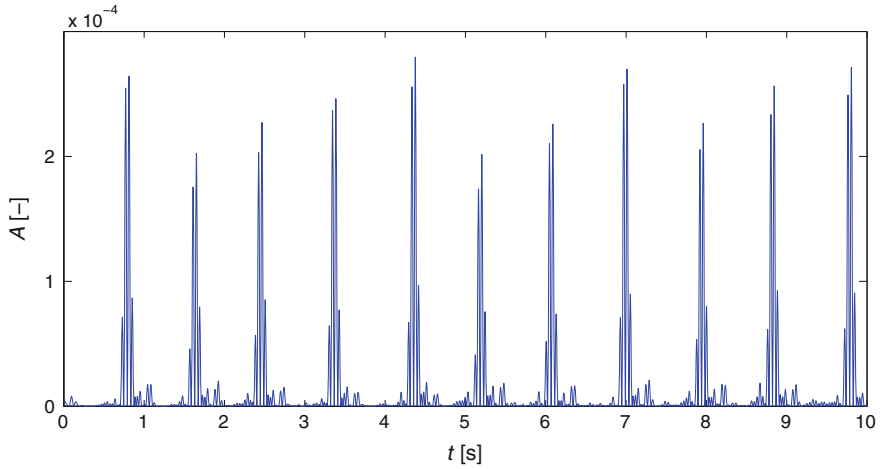


Fig. 6 Squared signal

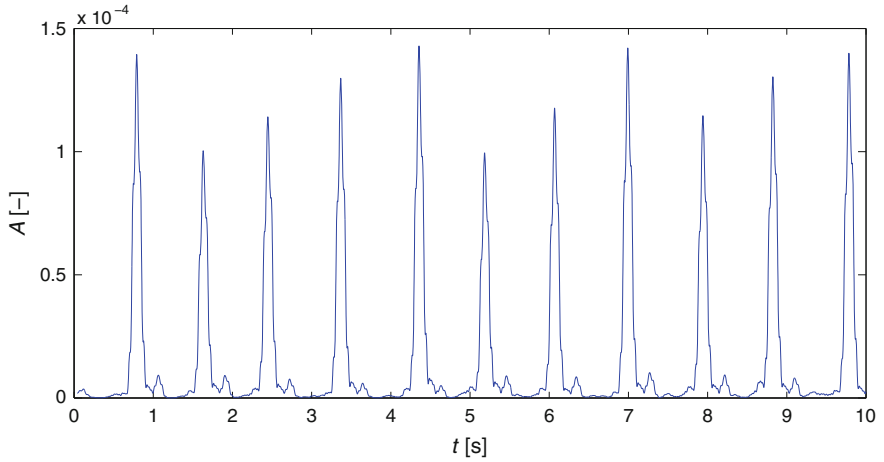


Fig. 7 Resulting averaged signal

($M = 80$) if the sampling rate is equal to 1 kHz. The span for the moving average windows is ($M - 1$). The signal generated by the moving average filter is displayed in Fig. 7.

The local maxima of the signal in Fig. 7 are detected in every point where the signal slope changes from ascending to descending. However, only certain local maxima represent R wave peaks. False peaks are discarded if:

- they are lower than the defined threshold. The said threshold value is equal to the average value of all local maxima;
- they precede or follow larger peaks by less than 200 ms (elimination of very high BPM);

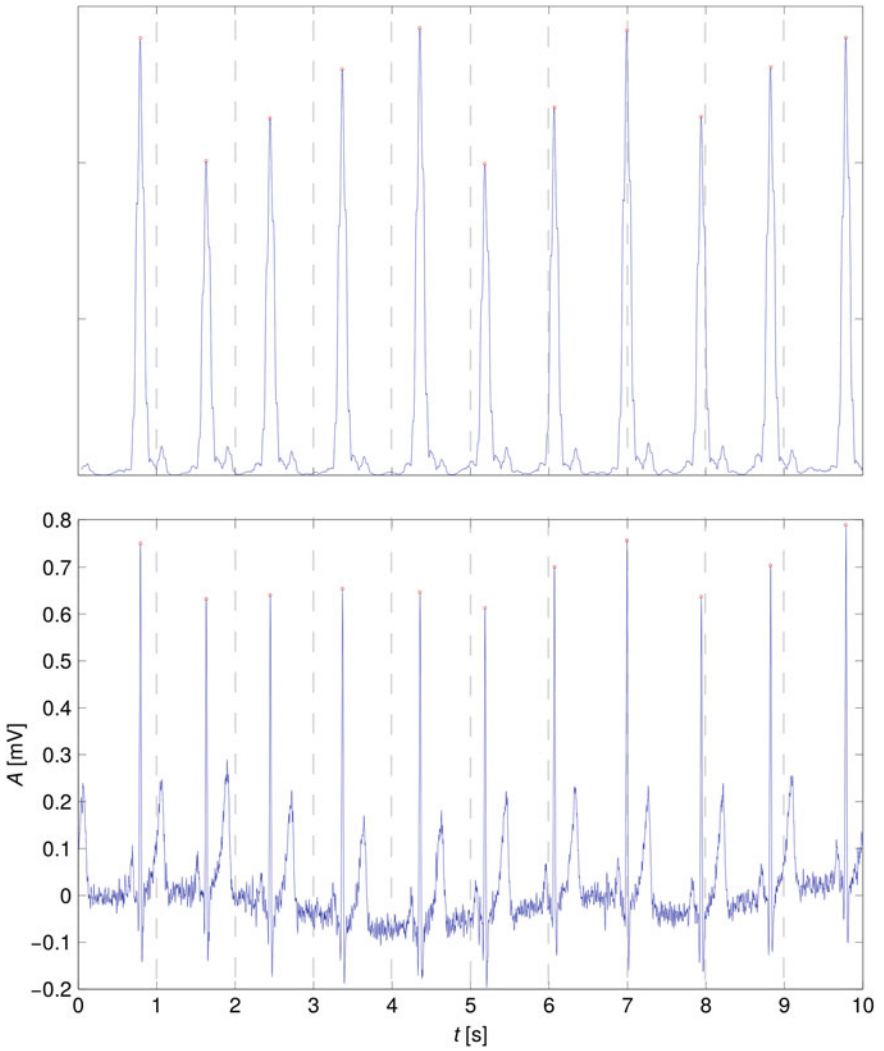


Fig. 8 Successful detection of R wave in ECG signal

- they are within 360 ms of a previous detection and if the peak value is not at least half the value of previous peak (peak is assumed to be a T-wave).

Undetected R waves are avoiding by applying the following rule:

- If no R wave has been detected within 1.5 R-to-R intervals then the largest peak is classified as R wave.

Detected peaks are indicated by red squares in the averaged signal and also in the original ECG signal (Fig. 8). All R waves were successfully detected by proposed algorithm and false R waves were not observed in the signal because the ECG was measured under calm conditions.

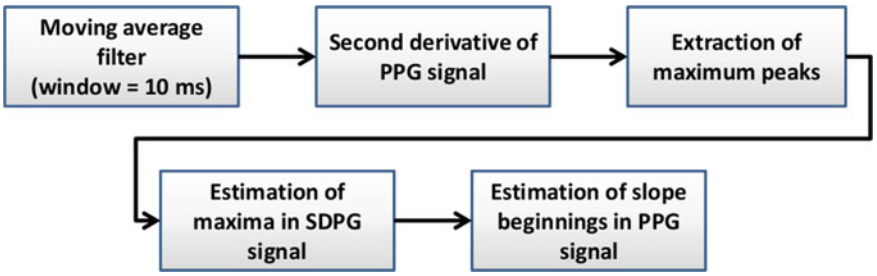


Fig. 9 Algorithm for automatic detection of slope beginnings in the PPG signal

Fig. 10 Measured PPG signal

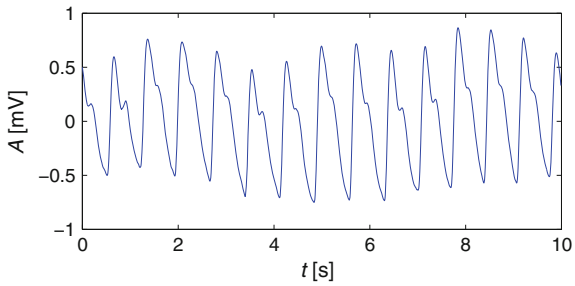
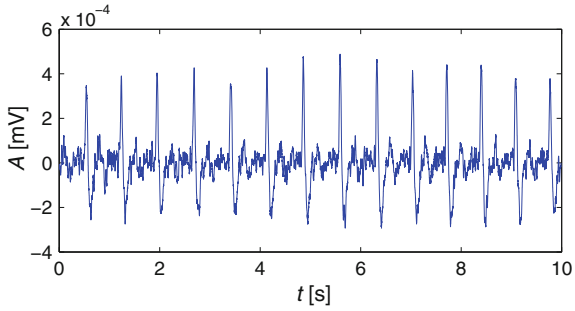


Fig. 11 Second derivative of the PPG signal



2.3 *Automatic Detection of PPG Slope Beginning*

We also designed an algorithm for automatic detection of slope beginnings in the PPG signal (Fig. 9).

The measured PPG signal (Fig. 10) can contain noise and therefore it should be preprocessed. The first step of the proposed algorithm consists in the application of a moving average filter with window length set to 10 ms.

After the PPG signal filtering, the second derivative of PPG curve (SDPG) is calculated (Fig. 11). From the physical point of view, the SDPG corresponds to curvature of the PPG signal. The first detected maximum peak position corresponds to a slope beginning in the PPG signal (Fig. 12).

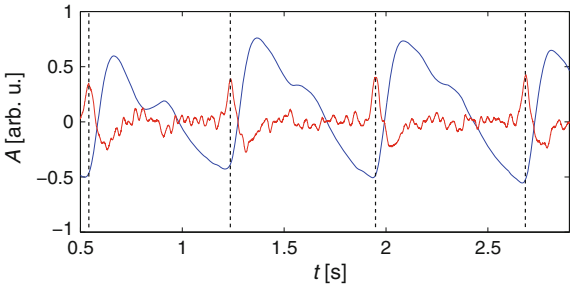


Fig. 12 Comparison of the PPG and the SDPG

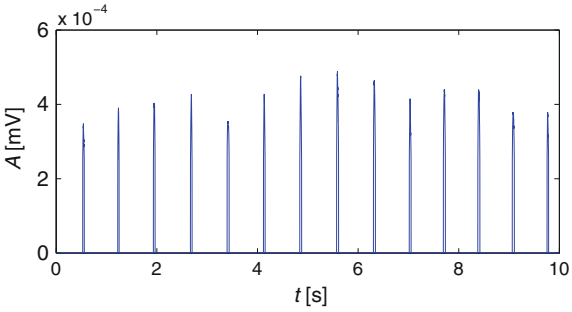


Fig. 13 Peaks extracted from SDPG signal

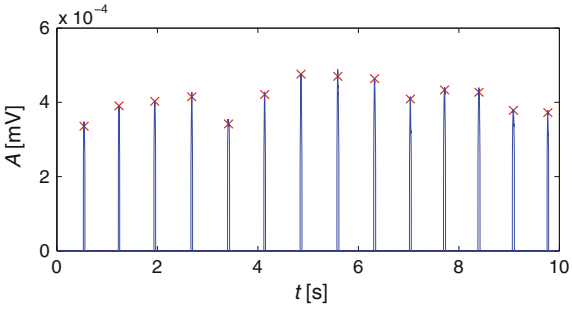


Fig. 14 Detected and marked SDPG peaks

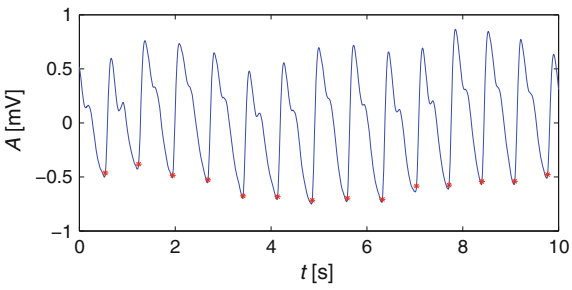


Fig. 15 Detected slope beginning in PPG signal

The SDPG peak extraction is performed in the next step—values lower than the threshold are set to zero and the resulting signal is shown in Fig. 13.

Once we obtain the SDPG peaks intervals it is possible to estimate and mark the peak positions (Fig. 14). Successfully detected SDPG peaks can be correlated with the PPG signal. An example of successful detection and estimation is shown in Fig. 15.

3 Results

The developed algorithms were evaluated on signals acquired from young and healthy subjects in both seated and relaxed positions. We used the BIOPAC MP36 to acquire the continuous 3-lead ECG and single point PPG using finger clip sensor. The previously described detection algorithms were implemented into a GUI application (Fig. 16) for better signal visualization and manipulation. Detection algorithms can be applied to ECG and PPG signals independently. The input data is loaded directly from MATLAB variable file generated by the BIOPAC MP36 device. It is possible to change the sampling rate and visible signal range by selecting the appropriate option from the program menu. Figure 16 shows sample results from the application—the first plot displays the ECG signal and detected R—waves and the second plot shows

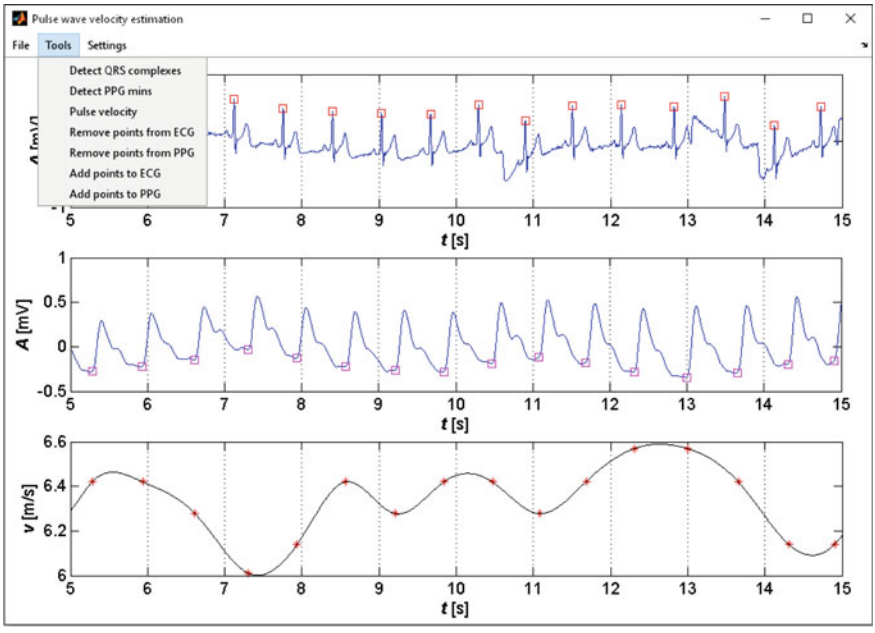


Fig. 16 GUI for automatic estimation of PWV

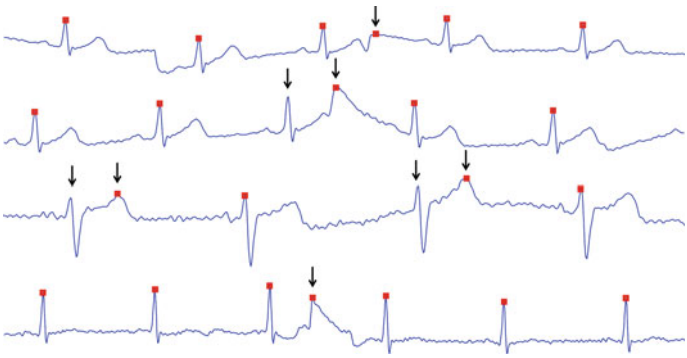


Fig. 17 Signal artifacts causing QRS false detection

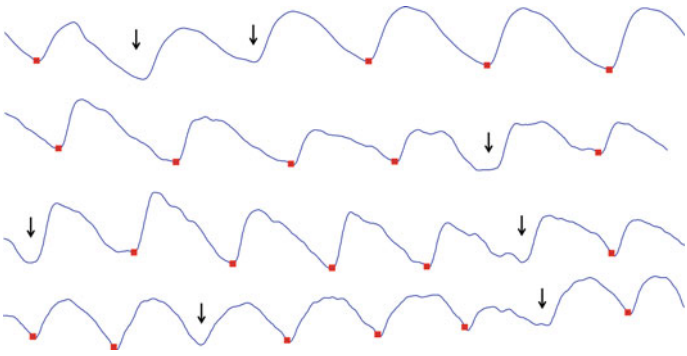


Fig. 18 Signal artifacts causing false detection of the PPG slope beginnings

the PPG signal with detected slope beginnings. The last graph contains information about the PWV, whose mean value was estimated to be around 6.2 m/s. PWV values change over time due to respiration which in turn influences the blood pressure.

Algorithm reliability was tested on 9 subjects whose ECG and PPG signals were acquired simultaneously. The QRS complex detection reliability varied from 35 to 100 % (Fig. 19). The beginnings of PPG slopes were detected with a reliability ranging from 65 to 100 % (Fig. 19). False detections were caused due to the presence of signal artifacts—this is especially obvious in signals acquired from subjects no. 4 and no. 9 (primarily motion artifacts). Figures 17 and 18 depict signal patterns which are problematic for our automatic detection algorithm. The automated ECG processing was also problematic for high amplitude T waves—these were erroneously detected as R waves (Fig. 17). On the other hand, the algorithm for PPG slope beginnings detection is sensitive to curvatures of the PPG signal. The occurrence of extra waves in the PPG signal also influences the detection reliability (Fig. 18). False detections are marked by black arrows in both figures. The mentioned errors can be eliminated by applying operations in the Tools program menu (Fig. 16)—erroneously detected R waves and PPG mins can be erased by selecting the Remove points operation and

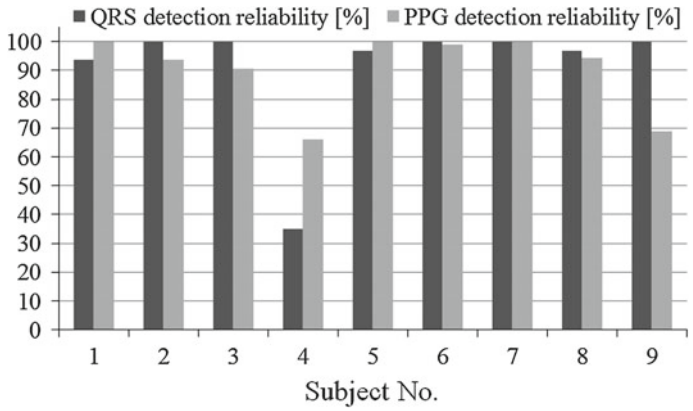


Fig. 19 Reliability of automatic algorithms

undetected R waves and PPG mins can be added via the Add points option. The points are added and removed manually in an interactive manner using single mouse click after selection of appropriate operation from the program menu.

4 Conclusion

This article focuses on the design and testing of algorithms for automatic detection of R waves in ECG signal and slope beginnings in PPG signal. We tested both algorithms on ECG and PPG signals acquired from young and healthy subjects. Based on our results we can conclude that both algorithms work properly, however further fine tuning is necessary for subjects which are not in a seated or relaxed position (during moving, exercise, etc.). Pulse wave velocity evaluation offers invaluable information about the state of the arterial system. Future work will thus focus on investigation of PWV from bigger patient group for the purpose of estimating arterial elasticity of the selected arterial system part. On the other hand, it is also possible to combine real-life measurements with electromechanical analogy model of arterial system to determine certain standard (expected) arterial system parameters (elastic modulus of arterial wall, blood viscosity or density) and confront these with data obtained from measurements. This combination could ultimately improve results obtained from conventional diagnostic methods such as ECG and PPG.

Acknowledgments This work was supported by the Slovak Research and Development Agency under the contract No. APVV-14-0519.

References

1. Allen, J.: Photoplethysmography and its application in clinical physiological measurement. Physiol. Measur. **28.3**, R1 (2007)
2. Borik, S., Cap, I.: Investigation of pulse wave velocity in arteries. In: 35th International Conference on Telecommunications and Signal Processing (TSP) (2012)
3. Kelly, B.B., Fuster, V. (eds.): Promoting Cardiovascular Health in the Developing World: A Critical Challenge to Achieve Global Health. National Academies Press (2010)
4. Koivistoinen, T., et al.: Pulse wave velocity reference values in healthy adults aged 26–75 years. Clin. Physiol. Funct. Imaging **27**(3), 191–196 (2007)
5. Loukogeorgakis, S., et al.: Validation of a device to measure arterial pulse wave velocity by a photoplethysmographic method. Physiol. Measur. **23**(3), 581 (2002)
6. Pan, J., Tompkins, W.J.: A real-time QRS detection algorithm. IEEE Trans. Biomed. Eng. **3**, 230–236 (1985)
7. Van Bortel, L.M., et al.: Expert consensus document on the measurement of aortic stiffness in daily practice using carotid-femoral pulse wave velocity. J. Hypertens. **30.3**, 445–448 (2012)

Averaging of Time-Warped ECG Signals for QT Interval Measurement

Tomasz Moroń

Abstract The paper investigates the method of averaging of time-warped signal cycles when applied to ECG noise suppression for noise immune measurement of the QT interval. First, a classical method is presented and then two modifications aimed to raise its immunity to noise are applied. Then a few versions of the method, obtained by application of the modifications proposed, are compared. Experiments on signals from the so called QT database are performed. These experiments show that while the original method is rather sensitive to noise, the modifications introduced make it much more robust. Applied to ECG enhancement prior to the measurements of the QT interval the modified method significantly improves the measurements accuracy.

Keywords ECG · Dynamic time warping · Signal averaging

1 Introduction

Dynamic time warping (DTW) is a technique that uses dynamic programming [2] to determine the best alignment of two time series and to provide a measure of their morphological similarity. This similarity measure was developed for pattern recognition problems, for example for recognition of words [17, 20], for ECG beats clustering, sound classification and more generally for different biomedical time series clustering. In this paper we focus on application of this method to noise suppression for QT interval measurement. Analysis of biomedical signals is often possible only when the quality of a processed signal is high. When the noise component frequency-band overlaps the desired component, suppression of noise is particularly difficult but when processed signals are of repetitive character, a relatively simple method of

T. Moroń (✉)
Faculty of Automatic Control, Electronics and Computer Science,
Institute of Electronics, Silesian University of Technology,
Akademicka 16, 44-100 Gliwice, Poland
e-mail: Tomasz.Moron@polsl.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_24

coherent averaging appears advantageous [19]. Since early years of computerized electrocardiography this family of methods has successively been applied to ECG signal filtering [18]. Very important is the use of this kind of methods in exercise electrocardiography where to get high-quality signal we first have to perform suppression of high level noise. The family of methods that use repetitive shape of the signal to achieve noise suppression has long been modified and improved [1, 14]. With the progress in the analysis and interpretation of ECG, the methods forming an average template of an ECG cycle but losing information on its variability from beat to beat appear rather unacceptable, at least in some applications. To overcome this limitation, different methods of ECG enhancement were proposed [4, 21] among others the method of nonlinear state-space projections [22]. Its modifications [7, 8] appeared successful in different applications to ECG processing [6, 9].

In [3] a modification of the averaging methods was proposed. It was developed for evoked potentials enhancement. In this method, the operation of averaging is preceded by nonlinear alignment (with the use of DTW) of the respective evoked potentials. This method was applied to ECG processing [11]. However, it appeared that dynamic time warping introduces unfavorable correlation among the noise components of the ECG cycles that later undergo averaging. This way a very crucial requirement of averaging (the lack of correlation among the noise components) is violated and, as a result, averaging of time warped ECG signals is not very effective. Therefore a modification was proposed [10, 11], aimed to limit these unfavorable effects. In this modification the definition of the alignment costs is changed. In this study a second modification is added to improve the method capabilities to suppress noise.

In Sect. 2 the original method of averaging of time-warped signals is described and its modifications. In Sect. 3 investigations of the method performance when applied to ECG noise suppression are performed, and the method influence on the precision of the QT interval measurements is studied. Finally, conclusions are drawn in Sect. 4.

2 Methods

2.1 Dynamic Time-Warping

Lets consider two signals of possibly different length: $v(n), n = 1, 2, \ldots, N_v$ and $u(n), n = 1, 2, \ldots, N_u$. To perform their nonlinear alignment, we first calculate the costs of aligning samples of the first with the samples of the second signal. These costs are classically defined as:

$$d_{i,j} = (u(i) - v(j))^2 . \tag{1}$$

Each $d_{i,j}$ corresponds to the alignment of $u(i)$ and $v(j)$. The warping path that minimizes the total cost of the alignment is given as:

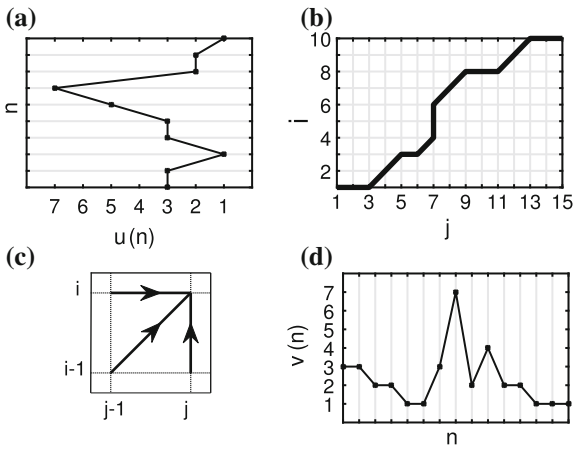


Fig. 1 Graphical representation (**c**) of the allowed step directions corresponding to usually applied boundary, continuity and monotonicity conditions, and an example of a warping path (**b**): (i, j) = {(1, 1), (1, 2), (1, 3), (2, 4), (3, 5), (3, 6), (4, 7), (5, 7), (6, 7), (7, 8), (8, 9), (8, 10), (8, 11), (9, 12), (10, 13), (10, 14), (10, 15)} relating two signals: $u(n)$ (**a**) and $v(n)$ (**d**)

$$Q = \sum_{k=1}^{K} d_{i_k, j_k} \tag{2}$$

preserving the specified constraints (such as boundary, continuity and monotonicity conditions [20]) is searched for. It can be found with the use of dynamic programming [2].

In Fig. 1 we present an exemplary warping path relating two nonlinearly aligned signals. We can notice that to achieve nonlinear alignment of both signals, some elements of $u(n)$ and $v(n)$ had to occur more that ones in the warping path.

2.2 Averaging of Time Warped ECG Segments

In the first stage of ECG signal processing, we perform QRS complexes detection. This way we obtain so-called fiducial points $r_l, l = 1, 2, \ldots, L+1$, corresponding to the central position within the respective detected $L+1$ complexes. Then the processed signal is cut into the following L time segments: $x_l(n) = x(r_l - \Delta + n), n = 1, 2, \ldots, N_l$ where segments lengths are given by $N_l = r_{l+1} - r_l$, $l = 1, 2, \ldots, L$. In the experiments parameter Δ is given a value corresponding to a time interval of 100 ms. The first cycle $x_1(n)$ is considered to be a first template $t_1(n)$ and this template is nonlinearly aligned and averaged with further cycles. Each time after the alignment operation has been completed, the template is updated. After aligning $l-1$th template with next lth cycle, the lth warping path is obtained consisting of the successive pairs: $(i_k^{(l)}, j_k^{(l)}) | k = 1, 2, \ldots, K_l$ where K_l is the length of the new template that will be obtained according to (3).

$$t_l(k) = \frac{(l-1)t_{l-1}(i_k^{(l)}) + x_l(j_k^{(l)})}{l}, k = 1, 2, \ldots, K_l. \tag{3}$$

In the course of averaging the successive cycles the template length increases. Therefore the warping paths relating this template with the respective cycles are changing. Thus each time the template is updated the paths must be updated as well [11]. After the final Lth template has been constructed, the individual cycles are reconstructed [11] and then they are used to replace the corresponding segments of the processed ECG signals. In [10] the method of averaging was called as NonLinear Alignment Averaging Filter (NLAAF). In this work it will be denoted as $NLAAF_{CCNF}$ where CC stands for Classical Costs, and NF for Non Filtered signal; the meaning of these terms will be explained in the next subsections.

2.3 *Modification #1: Time Warping on Band-Pass Filtered Signals*

In [16] we proposed a simple way of signals decomposition based on a Moving Average (MA) filter. The MA filter calculates the mean of the signal samples that appear in a moving time window of the assumed length:

$$MA_K(z) = \frac{1}{K}\sum_{k=0}^{K-1} z^{-k} = \frac{1}{K}\frac{1-z^{-K}}{1-z^{-1}} \quad . \tag{4}$$

As we can see, the filter can be applied in either non-recursive or recursive way. Particularly favorable is the latter, very fast solution. The filter can be applied to perform low-pass filtering of ECG signals. It is also often used to form a high-pass filter for baseline wander suppression (5).

$$HP_K(z) = z^{-(K-1)} - (MA_K)^2. \tag{5}$$

According to the power spectral analysis of the ECG signal components performed in [23], we select a proper value of parameter K to obtain an appropriate filter (for this purpose). The cut-off frequency of such a filter should not exceed about 0.8 Hz to avoid suppression of diagnostically important components of the ECG signal [25]. Unfortunately a filter with such a low cut-off frequency can not suppress the very troublesome low frequency noise caused by patients movements. This type of noise, called as electrode motion artifacts, can reach up to about 5 Hz and overlaps [23] the spectrum of the diagnostically important P and T waves. Therefore it is extremely difficult to be dealt with. It can spoil the results of dynamic time-warping. To prevent this type of noise, we apply band-pass filtering to suppress the low-frequency electrode motion and also higher frequency noise. Again, we apply MA to form a simple band-pass filter:

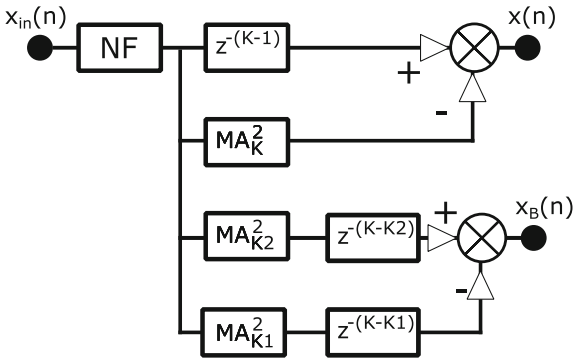


Fig. 2 Moving average based high-pass and band-pass filtering of the ECG signals. NF is a notch filter for powerline suppression

$$BP_{K,K1,K2}(z) = \left(z^{-(K-K2)} \cdot (MA_{K2}(z))^2\right) - \left(z^{-(K-K1)} \cdot (MA_{K1}(z))^2\right). \quad (6)$$

The filter parameters satisfy the condition K>K1>K2; K1 and K2 determine the left and the right cut-off frequencies of the filter, respectively. For larger K1 (and fixed K2) the left cut-off frequency is smaller. For larger K2 (and fixed K1) the right cut-off frequency is smaller. As we can see in Fig. 2, the filter can be applied parallel with the high-pass filter suppressing the baseline wander. This way we obtain signal $x(n)$ with powerline and baseline wander suppressed and a band-pass filtered signal $x_B(n)$ with electrode motion artifacts suppressed as well. In the proposed version of NLAAF, we perform dynamic time warping and averaging on a band-pass filtered signal $x_B(n)$, and each time a warping path is constructed, we paralelly use it align non-linearly and average the cycles of signal $x(n)$. This way we prevent the unfavorable/unfavourable influence of low frequency electrode motion artifacts on the results of time warping, and still we obtain the average signal which contains all relevant frequency components of the ECG. This version of the method is denoted as NLAAF_{CCBF}, where CC as before stands for Classical Costs, and BF for Band-pass Filtered signal (on which DTW is performed).

2.4 Modification #2: Of the Alignment Costs

In [10] the following definition of the alignment costs was proposed:

$$d'_{i,j} = \left\| \mathbf{x}^{(i)} - \mathbf{y}^{(j)} \right\|, \quad (7)$$

where $\|\cdot\|$ denotes the Euclidean norm, and vectors $\mathbf{x}^{(n)}$ and $\mathbf{y}^{(n)}$ are defined as follows:

$$\mathbf{x}^{(n)} = [x(n-v), x(n-v+1), \ldots, x(n), \ldots, x(n+v)]^T, \quad (8)$$

which means that they are composed of $2v + 1$ successive signal samples (with the nth sample being the central one). By such a redefinition of a cost matrix, we assure that longer signal intervals are matched to each other instead of single samples of warped signals. The method combining this modification with modification #1 is denoted as NLAAF_{VCBF} where VC stands for Vector norm based Costs and BF as before for Band-pass Filtered signal (that undergoes DTW).

3 Numerical Experiments and Results

To test the methods developed 6 ECG signals from the QT database [12] of the Physionet resources [17] were used. The signals were carefully inspected cycle by cycle and 80s intervals of very high quality were chosen for the experiments. Since the method proposed can be regarded as an extension of the very popular (in ECG processing) method of time averaging, it can be applied to ECG signals of different morphology, only if this morphology is repeatable. Therefore limitation of the experiments to six signals does not limit the possible applications of this method. The selected signals were used to simulate the desired ECG and they were contaminated with the EMG noise from the MIT-BIH database [15]. Since the signals from the QT database are sampled with the frequency of 250 Hz and the noise record with 360 Hz, the former was re-sampled with the frequency of 250 Hz (with the use of the *interp* method in Matlab environment [5]).

3.1 Selection of the Methods Parameters

To choose the proper parameters, we investigated the methods abilities to suppress noise. To this end to each signal simulating the desired ECG, we added muscle noise (from the MIT-BIH database [15]) of the appropriate level to obtain signals of a few assumed values of the signal-to-noise ratio (SNR). Since measurements of the QT interval are possible only for rather good quality signals, we selected the values of 5, 10 and 15 dB for the experiments. Each signal was processed with the use of the investigated methods and the results obtained were evaluated with the following noise reduction factor (NRF):

$$NRF = \sqrt{(\sum_{i=1}^{N} n(i)^2)/(\sum_{i=1}^{N} n_r(i)^2)}, \tag{9}$$

where $n(i)$ is the noise added to the desired signal, $n_r(i)$ is the residual noise remaining in the filtered signal (it is the difference between the filtered signal, obtained after averaging and reconstruction of individual ECG segments, and its desired component).

In Figs. 3 and 4 we presented the influence of the band-pass filter parameters $K1$ and $K2$ on the mean NRF obtained by the NLAAF$_{CCBF}$ method (which performs DTW on a band-pass filtered signals). As a criterion for selection of the best parameters, we used the mean NRF obtained for two levels of noise, corresponding to SNR of 10 and 15 dB. As we can see, $K1 = 32$ (Fig. 3) and $K2 = 8$ (Fig. 4). The parameters are presented in Table 1. Similarly we selected parameter v of the NLAAF$_{CCBF}$ method. For this parameter, we presented in Table 1 the value τ_v, which

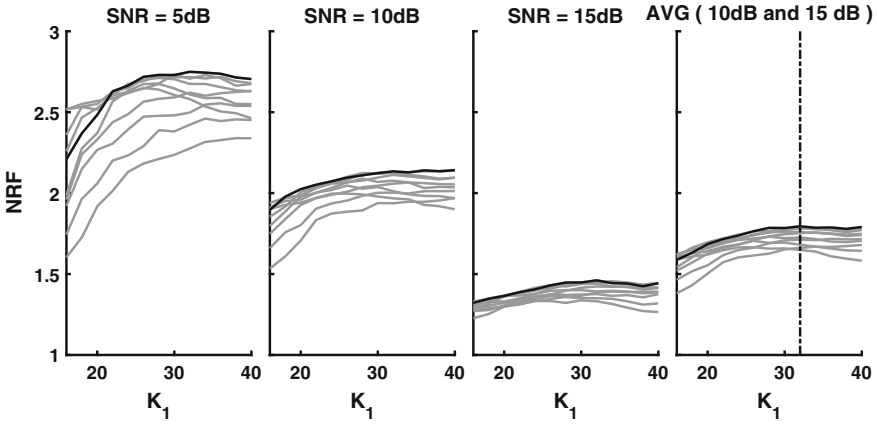


Fig. 3 The influence of parameter $K1$ (deciding on the left cut-off frequency of the band-pass filter) on the mean NRF obtained by the NLAAF$_{CCBF}$ method. In each *subplot*, different *curves* correspond to different values of parameter $K2$. *Vertical line* shows the chosen value of parameter $K1$ (for which the largest mean NRF was obtained in SNR of 10 and 15 dB)

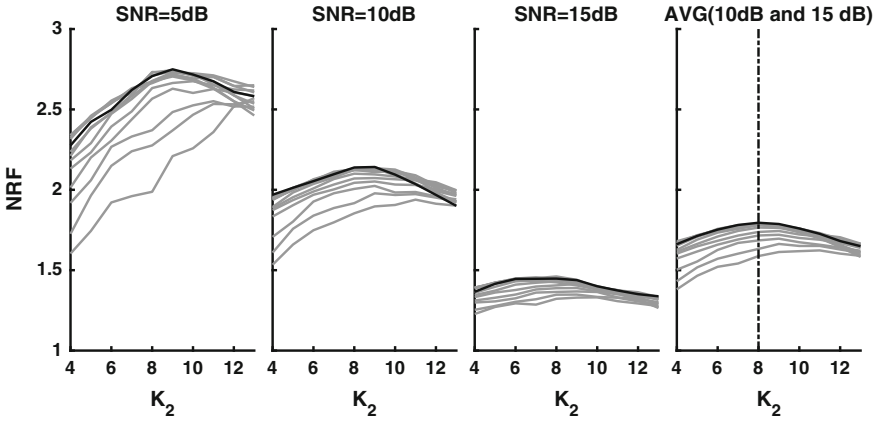


Fig. 4 The influence of parameter $K2$ (deciding on the right cut-off frequency of the band-pass filter) on the mean NRF obtained by the NLAAF$_{CCBF}$ method. In each *subplot*, different *curves* correspond to different values of parameter $K1$. *Vertical line* shows the chosen value of parameter $K2$ (for which the largest mean NRF was obtained in SNR of 10 and 15 dB)

Table 1 The mean NRF obtained by the compared methods in SNR of 10 and 15 dB

Method	K_1	K_2	V (ms)	NRF
NLAAF$_{CCNF}$				1.147
NLAAF$_{CCBF}$	32	8		1.795
NLAAF$_{VCNF}$			400	1.770
NLAAF$_{VCBF}$	40	10	400	2.027

corresponds to the time length of vectors defined by (8), used during calculations of the modified alignment costs. Analogically, we selected all these parameters of the NLAAF$_{VCBF}$ method, combining both modifications of the method. In the table, we have also presented the mean noise reduction factor obtained by these methods in SNR of 10 and 15 dB. As we can see, both modifications applied individually increased its value significantly and applied together allowed even better improvement of the NLAAF method ability to suppress noise. Thus the most effective version of the method (NLAAF$_{VCBF}$) will be compared in the next section to the original one (NLAAF$_{CCNF}$) with regards to their influence on the precision of the QT interval measurement.

3.2 Investigations of the Methods Influence on the Precision of the QT Interval Measurement

A QT interval is defined as a time elapsing between the beginning of a Q wave and the end of the T wave in the ECG signal. Since it is most often easier to locate precisely the peak of the T wave, this position is sometimes used as an offset of the QT interval. Thus the precision of the QT interval measurement depends on the precision of localization of these characteristic positions. In [24] thorough investigations of different factors that influence the precision of the QT interval measurements were performed. It was shown that it was advantageous to perform re-sampling with the frequency of 1000 Hz of the signals originally stored with the sampling frequency of 250 Hz. A similar approach was applied in [6] and it was also applied in this study. Therefore the signals used in the previous subsection were prepared in this way. Then they were processed with the use of the NLAAF$_{CCNF}$ and NLAAF$_{VCBF}$ methods and finally the measurements were performed on the desired signals, the noisy ones and on the signals enhanced by the two methods. The measurements were performed with the use of the method developed by Laguna et al. [13] and adapted in [6] to the sampling frequency of 1000 Hz (originally the Laguna's method was developed for 250 Hz). Parameters $K1$ and $K2$ selected in the previous subsection were multiplied by 4 to be adjusted to the increased sampling frequency (and similarly parameter K of the high-pass filter).

Exemplary results of the QT limits localization are presented in Fig. 5. As we can see, the measurements of the T wave peak and end in the noisy signal contained large errors. Application of the original NLAAF_{CCNF} method helped to preclude so large errors, still localization of the T wave end was not accurate. Application of the modified NLAAF_{VCBF} method allowed to improve the measurements much better.

On the basis of such measurements, we calculated the root mean squared error (RMSE) of the Q, T_{peak} and T_{end} determination. The results are presented in Figs. 6 and 7. The RMSE errors presented in Fig. 7 were obtained after trimming large errors (such as in Fig. 5b) to 30 ms for T_{peak} and T_{end}, and to 10 ms for Q. This operation was aimed to diminish the influence of large errors, which can relatively easily be detected and rejected during real measurements, on the RMSEs obtained.

As we can see both in Figs. 6 and 7 application of the original $NLAAF_{ccnf}$ method allowed to diminish errors of the T wave peak and T wave end determination but was less effective in detection of the Q wave beginning. But introduction of the two modifications proposed highly improved the results. All errors decreased significantly. If we compare Fig. 6 with Fig. 7, we can discern that for this method the RMSE errors are the same. It means that after application of this method no large measurement errors occurred. Thus we can conclude that this method makes the measurements not only more accurate but also more robust against high energy noise artifacts.

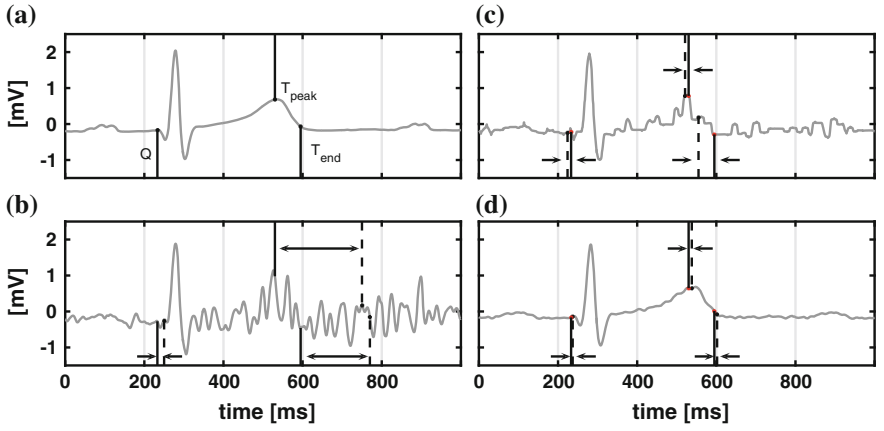


Fig. 5 Results of localization of: a Q wave onset (Q), a T wave peak (T_{peak}) and a T wave end (T_{end}) in the desired signal (*solid lines*) and in the noisy (SNR = 5 dB) and the filtered signals (*dashed lines*). *Arrows* indicate the measurement errors. **a** Desired component. **b** Noisy signal. **c** Filtered signal (for $\mathrm{NLAAF}_{\mathrm{ccnf}}$). **d** Filtered signal (for $\mathrm{NLAAF}_{\mathrm{vcbf}}$)

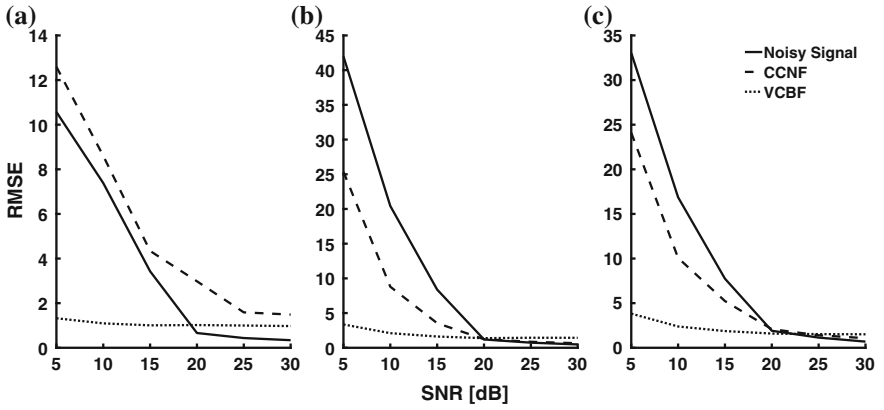


Fig. 6 The average deviation of the detected position of elements of ECG depending on the level of interference for noisy signal and signal after filtration by $NLAAF_{ccnf}$ ($K_1 = 32$, $K_2 = 8$) and $NLAAF_{vcbf}$ ($K_1 = 30$, $K_2 = 10$, $V = 400$). **a** Q. **b** T-peak. **c** T-end

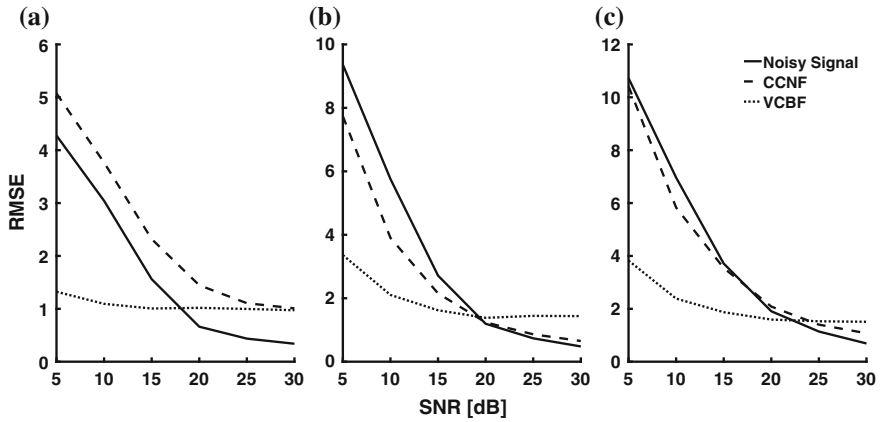


Fig. 7 The average deviation of the detected position of elements of ECG depending on the level of interference for noisy signal and signal after filtration by $NLAAF_{ccnf}$ and $NLAAF_{vcbf}$ taking into account the elimination of big errors. **a** Q. **b** T-peak. **c** T-end

4 Conclusion

In the paper we have shown how the method developed for evoked potentials processing can be adjusted to the enhancement of the ECG signals. Although the original version of this method can not suppress the ECG noise very effectively, introduction of modifications of the procedure of time warping significantly improves its capabilities. When applied to ECG processing prior to the measurements of the QT interval limits, the method significantly improved the results. It highly increased the measurements accuracy and made them much more robust against high energy noise.

Acknowledgments This research was supported by statutory funds of the Institute of Electronics (BKM-2015), Silesian University of Technology. The work was performed using the infrastructure supported by POIG.02.03.01-24-099/13 grant: GeCONiI–Upper Silesian Center for Computational Science and Engineering.

References

1. Bataillou, E., Thierry, E., Rix, H., Meste, O.: Weighted averaging using adaptive estimation of the weights. Signal Process. **44**(1), 51–66 (1995)
2. Bellman, R.E., Dreyfus, S.E.: Applied dynamic programming (1962)
3. Gupta, L., Molfese, D.L., Tammana, R., Simos, P.G.: Nonlinear alignment and averaging for estimating the evoked potential. IEEE Trans. Biomed. Eng. **43**(4), 348–356 (1996)
4. Hu, X., Nenov, V.: A single-lead ecg enhancement algorithm using a regularized data-driven filter. IEEE Trans. Biomed. Eng. **53**(2), 347–351 (2006)
5. IEEE Acoustics and Speech and Signal Processing Society. Digital Signal Processing Committee. Programs for digital signal processing. IEEE (1979)
6. Kotas, M.: Projective filtering of time-aligned ecg beats for repolarization duration measurement. Comput. Methods Programs Biomed. **85**(2), 115–123 (2007)
7. Kotas, M.: Projective filtering of time warped ECG beats. Comput. Biol. Med. **38**(1), 127–137 (2008)
8. Kotas, M.: Robust projective filtering of time-warped ECG beats. Comput. Methods Programs Biomed. **92**(2), 161–172 (2008)
9. Kotas, M.: Combined application of independent component analysis and projective filtering to fetal ecg extraction. Biocybern. Biomed. Eng. **28**(1), 75 (2008)
10. Kotas, M., Leski, J.M., Moroń, T.: Dynamic time warping based on modified alignment costs for evoked potentials averaging. In: Man–Machine Interactions, vol. 4, pp. 305–314. Springer (2016)
11. Kotas, M., Pander, T., Leski, J.M.: Averaging of nonlinearly aligned signal cycles for noise suppression. Biomed. Signal Process. Control **21**, 157–168 (2015)
12. Laguna, P., Mark, R.G., Goldberg, A., Moody, G.B.: A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG. In: Computers in Cardiology 1997, pp. 673–676. IEEE (1997)
13. Laguna, P., Thakor, N., Caminal, P., Jane, R., Yoon, H.R., Bayés de Luna, A., Marti, V., Guindo, J.: New algorithm for QT interval analysis in 24-hour holter ECG: performance and applications. Med. Biol. Eng. Comput. **28**(1), 67–73 (1990)
14. Łęski, J.M.: Robust weighted averaging [of biomedical signals]. IEEE Trans. Biomed. Eng. **49**(8), 796–804 (2002)
15. Luis, A., Ivanov, P.C., Glass, L., Mietus, J.E., Peng, C.K., Hausdorff, J.M., Moody, G.B., Mark, R.G., Stanley, H.E., et al.: Physionet mit-bih noise stress test database (2000)
16. Moroń, T., Kotas, M., Leski, J.M.: Principal component analysis and dynamic time-warping in subbands for ECG reconstruction. In: Man–Machine Interactions, vol. 4, pp. 315–325. Springer (2016)
17. Myers, C., Rabiner, L.R., Rosenberg, A.E.: Performance tradeoffs in dynamic time warping algorithms for isolated word recognition. IEEE Trans. Acoust. Speech Signal Process. **28**(6), 623–635 (1980)
18. Rautaharju, P., Blackburn, H.: The exercise electrocardiogram: experience in analysis of noisy cardiograms with a small computer. Am. Heart J. **69**(4), 515–520 (1965)
19. Rompelman, O., Ros, H.: Coherent averaging technique: a tutorial review part 1: noise reduction and the equivalent filter. J. Biomed. Eng. **8**(1), 24–29 (1986)
20. Sakoe, H., Chiba, S.: Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans. Acoust. Speech Signal Process. **26**(1), 43–49 (1978)

21. Schmidt, M., Baumert, M., Porta, A., Malberg, H., Zaunseder, S.: Two-dimensional warping for one-dimensional signals-conceptual framework and application to ECG processing. IEEE Trans. Signal Process. **62**(21), 5577–5588 (2014)
22. Schreiber, T., Kaplan, D.T.: Nonlinear noise reduction for electrocardiograms. Chaos: Interdisc. J. Nonlinear Sci. **6**(1), 87–92 (1996)
23. Thakor, N.V., Webster, J.G., Tompkins, W.J.: Estimation of QRS complex power spectra for design of a QRS filter. IEEE Trans. Biomed. Eng. **11**, 702–706 (1984)
24. Tikkanen, P., Sellin, L., Kinnunen, H., Huikuri, H.: Using simulated noise to define optimal QT intervals for computer analysis of ambulatory ECG. Med. Eng. Phys. **21**(1), 15–25 (1999)
25. Van Alste, J., Schilder, T.: Removal of base-line wander and power-line interference from the ECG by an efficient FIR filter with a reduced number of taps. IEEE Trans. Biomed. Eng. **12**, 1052–1060 (1985)

Localization of Epileptic Foci Using EEG Brain Mapping

Michaela Snajdarova and Branko Babusiak

Abstract The method for localization of epileptic foci is described in this paper. The source of epileptic seizure is localized from multi-channel electroencephalogram (EEG). There are many spectral methods for analysis of EEG signal and brain mapping is one of them. The brain mapping is used for visualisation of brain activity over the scalp. The mapping of time delay between channels can be applied for localization of epileptic foci. This type of mapping was implemented into developed application with intention to localize source of epileptic seizure. The software also provides detection of epileptic graphoelements and displays location of the epileptic focus in two or three-dimensional space. The method described in this paper is verified on real data with epilepsy.

Keywords Brain mapping · Electroencephalogram · EEG · Epileptic focus · Epilepsy

1 Introduction

Epilepsy is the most widespread disease of the brain diagnosed by EEG. Epilepsy is reflected in the form of repeatedly uncontrolled, sudden changes of brain activity called seizures. From a clinical point of view, epilepsy is defined as a fault of consciousness, behavior, motoric disorder or sensitive functions, or the different frequencies of these changes. Epileptic seizure is physiologically sudden and transient cerebral cortical activity. This disorder is uncontrollably accompanied by electric discharge in the gray matter of the brain. Electric discharge takes a few seconds, minutes or exceptionally hours. The patient can be absolutely without health problems in the time between seizures [8].

M. Snajdarova (✉) · B. Babusiak
Department of Electromagnetic and Biomedical Engineering, University of Zilina,
Univerzitna 1, 01026 Zilina, Slovakia
e-mail: michaela.snajdarova@fel.uniza.sk

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_25

The best and most accurate method of epileptic foci localization is using Functional Magnetic Resonance (fMRI). The MRI scan shows which part of brain is affected by epileptic seizures and what happens to brain activity during seizure. This method is quite expensive and the analysis of EEG is an appropriate alternative for localization of epilepsy. Electroencephalography is the neurophysiologic measurement of the electrical activity of the brain by recording from electrodes placed on the scalp. The resulting record is known as electroencephalogram (EEG) and represents an electrical signal from a large number of neurons. The electrode over the source of epilepsy captures epileptic signal earliest, the other electrodes captures seizure with some delay. It is very difficult to determine position of the origin of epilepsy from time-domain EEG because of fast spreading of epilepsy which results in very small time delays between EEG channels. The specific time delay brain mapping is used for localization of epileptic foci.

This paper deals with the new implementation of the time delay analysis between EEG leads using the topographic mapping. The proposed algorithm is able to detect the epileptic seizures and subsequently to determine area of the epileptic focus. The aim of our work was the origin localization of the seizures with standard EEG examination without support of additional methods e.g. fMRI.

2 Localization of Epileptic Focus in the EEG

2.1 EEG Data Format

Measured EEG data are stored in standard file format, the European Data Format (EDF). EDF stores multichannel data, allowing different sample rates for each signal. Developed software offers viewing nineteen channels of EEG. Records of EEG signal were digitalized at 256 samples/s with 12-bit resolution (Fig. 1). The international electrode placement system "10-20" was used in our records. A notch filter was used to remove artefact caused by electrical power line frequency.

2.2 Localization Method

There are multiple approaches to detect area of epileptic focus e.g. [5]. Our method which computes zone of epileptic focus was based on ideas by Brazier [1] and Gotman [2] further developed by e.g. [3]. The algorithm we used in our work is depicted in Fig. 2 and consists of following parts. The algorithm computes the phase of cross-spectrum between the two channels at all frequencies. The cross-spectrum is defined as Fourier transform of cross-correlation function R_{xy} between signal channels x and y:

$$G_{xy}(f) = FT\{R_{xy}(l)\}. \tag{1}$$

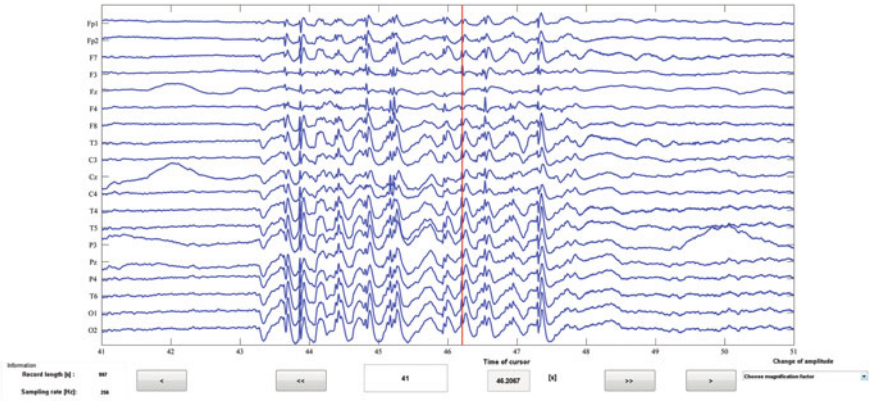


Fig. 1 EEG signal loaded in the application

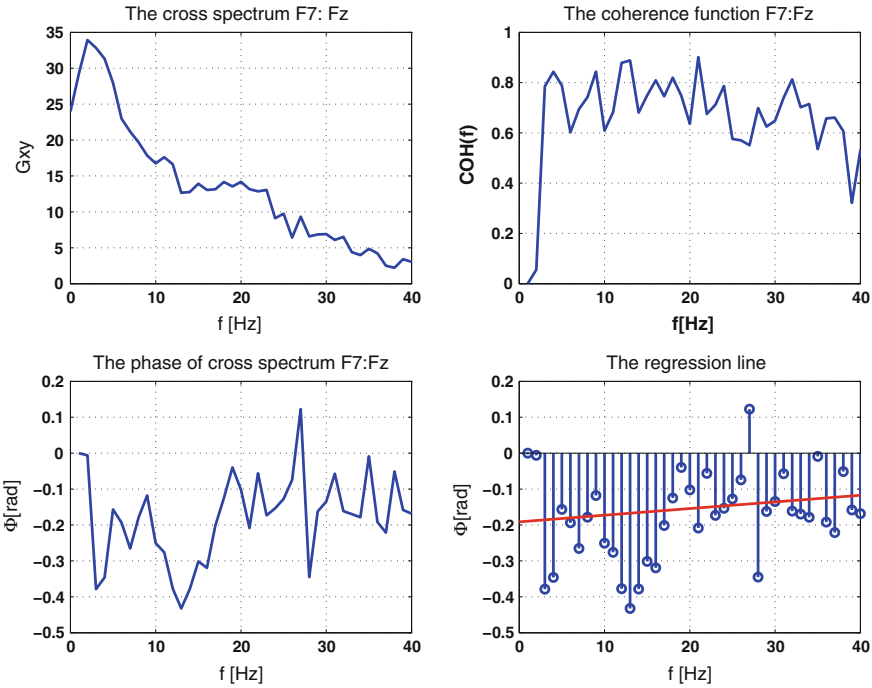


Fig. 2 Example of computations for channel F7:Fz

The spectrum is computed from signal in time interval with presence of epileptic seizure which is selected by detector. The automatic detector of epileptic seizures was described in previous co-author's work [8]. The values of electrodes in the middle line were considered as reference. Cross-spectrum is calculated between reference

electrode signal and the signal from the electrode in lateral line with reference electrode. By the help of the coherence function, it was discovered which frequencies were present in both channels. The values of the coherence function are between $\langle 0, 1 \rangle$. The zero value indicates no linear dependency between the two channels at specific frequency and the value of 1 indicates a perfect linear dependency. The limit of the coherence was estimated according to Otnes and Enochson. The degree of similarity of two signals has to be close to 1. In the next step the phase spectrum was modified because the phase shift can be estimated only at the existing frequencies in both signals. The values of the phase spectrum at frequencies which are not present in both channels according to low value of coherence function will be ignored. A regression line is calculated from modified phase spectrum by using method of the mean least-squares. An intercept point of the regression line and the value of phase spectrum is considered to be a time delay for each channel [4, 6, 7, 9].

A group delay $\tau(\omega)$ between two channels is defined as derivation of the phase spectrum τ according to the frequency f [6]:

$$\tau(\omega) = -\frac{d\phi}{d\omega}. \tag{2}$$

The relative delay Δt between two channels is determined by the following equation [6]:

$$\Delta t = \frac{\Delta\phi}{360 \cdot (\Delta f)}. \tag{3}$$

The high value of Δt represents a big phase shift between the correlated channels. The biggest positive value of time difference Δt of electrode is considered to be new reference electrode in next iteration and another electrodes are assumed to anticipate this electrode. The whole procedure is subsequently repeated. The Spectral Power Density (PSD) using Welch's method with a half-overlapping Hanning window type was calculated. There were used values of time differences at the most significant frequency in PSD. If the system shifts phase at certain frequency in negative sense the phase difference $\Delta\phi$ has the negative value what indicates signal delay compared to the reference electrode. Otherwise the positive value indicates phase overtaking. Then the biggest positive value of time difference indicates the electrode which is the closest to the epileptic focus because the phase difference on this position is most significant. This information is shown as graphical representation of the brain mapping in two and three-dimensional space. Described approach is illustrated in Fig. 3. The method of brain mapping (BM) shows the brain activity of individual parts of cortex which is measured by multiple electrodes on the scalp. The values in the areas without electrodes are computed by iterative linear interpolation. The numerical values of scalp points are encrypted into colour respecting defined colour range. The standard range passes from blue to red, where blue colour represents the lowest activity and red colour stands for highest activity. The BM can be conceived as a higher form of neuroimaging.

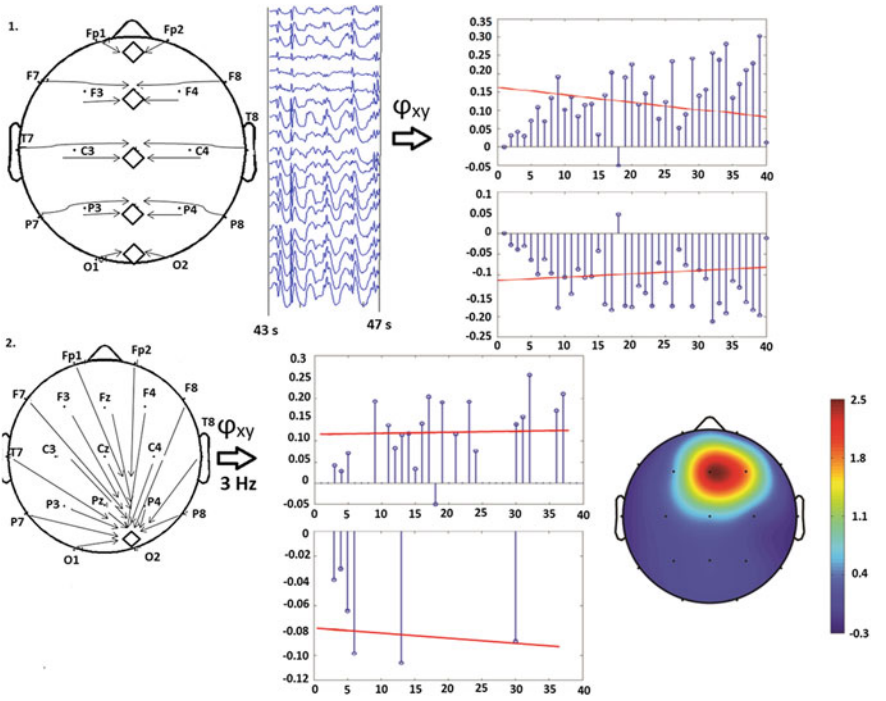


Fig. 3 The fundamentals of mapping brain activity using time delay in EEG records with epileptic seizure

3 Testing of the Proposed Algorithm

3.1 Verification of Method on Simulated Example

Our algorithm was verified by two different types of signals—simulated and real EEG. First of them consists of 19 harmonic signals y_1–y_{19}. These signals are mixture of harmonic waves with frequencies f_1–f_5 in range 1–5 Hz, amplitude A of 1 V and phase ψ_X which characterise a time shift of channel X. The signals are created according to the following equations:

$$\begin{aligned} y_1 &= A\,(2\pi f_1 t + \psi_1) + A\,(2\pi f_2 t + \psi_1) + \cdots + A\,(2\pi f_5 t + \psi_1) \\ y_2 &= A\,(2\pi f_1 t + \psi_2) + A\,(2\pi f_2 t + \psi_2) + \cdots + A\,(2\pi f_5 t + \psi_2) \\ &\cdots \\ &\cdots \\ y_{19} &= A\,(2\pi f_1 t + \psi_{19}) + A\,(2\pi f_2 t + \psi_{19}) + \cdots + A\,(2\pi f_5 t + \psi_{19}) \end{aligned} \tag{4}$$

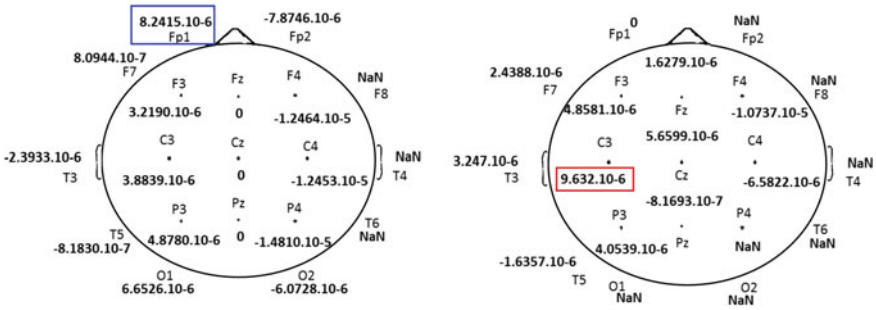


Fig. 4 The values of time delay in both steps for each channel

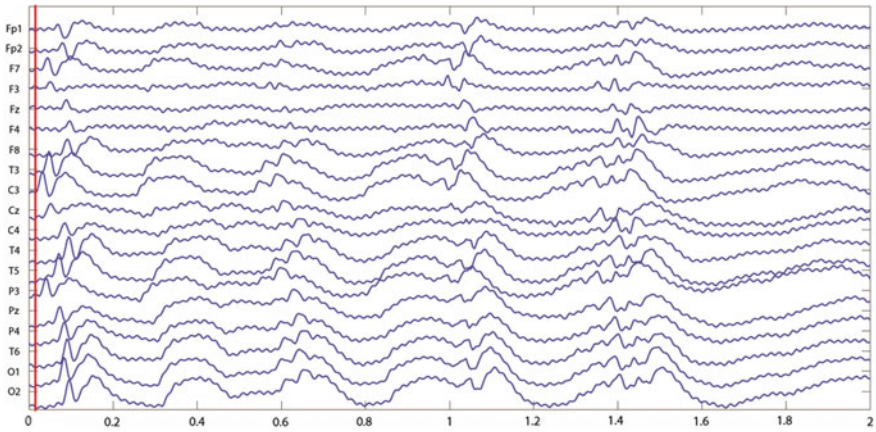


Fig. 5 Time differences observed at the beginning of epileptic seizure

In our experiment, we assumed the epileptic focus in position close to the electrode C3 (channel 9). In this position the phase of harmonic signal $\psi_9 = 0$. In Fig. 4 are noticed values of time delay for all electrodes.

At first, the middle electrodes are considered as a reference electrodes (left side of Fig. 4). The second step is shown on the right where the red rectangle (C3 electrode) characterizes the maximum time delay concerning new reference electrode (Fp1 electrode—blue rectangle from previous step). The position of signal source ($\psi = 0$) was successfully localized at C3 position.

3.2 Verification of Method on Real EEG Data

The second verification of proposed method was performed on the real data with epileptic seizure in EEG records of six patients with localization-related epilepsy. It is obvious from Fig. 5 that the electrode C3 captured the epileptic seizure first.

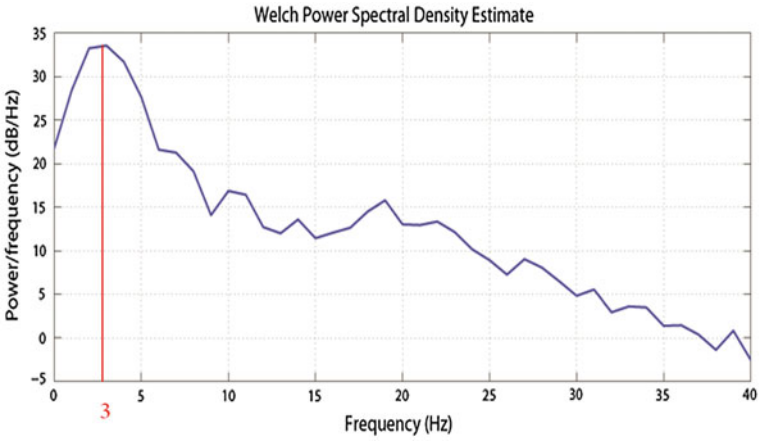
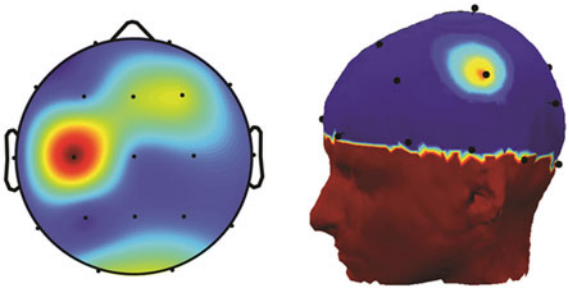


Fig. 6 Welch power spectral density of epileptic seizure

Fig. 7 Localization of epileptic focus



The most significant frequency was determined according to Welch Power Spectral Density at 3 Hz in segment with presence of seizure, as is illustrated in Fig. 6.

The result of our experiment is shown by two and three dimensional brain mapping visualisation technique. The result verified position of epileptic focus near to the electrode C3 (Fig. 7).

4 Conclusion

We created detector for automatic localization of epileptic seizures [8] and epileptic foci localization in MATLAB programming language. Both tools were implemented in application with graphical user interface. The application is designed for files in standard EDF format. Reliability of the algorithm depends on morphology and duration of epileptic seizure. The successful detection of the epileptic foci is possible only when epileptic seizure is observed in EEG record. Computing the phase of the cross-spectrum strongly depends on the determination of the start and the end of

epileptic seizures. Maximum resolution of the time shifts between channels is given by the sampling frequency, e.g. 3.9 ms for sampling rate of 256 Hz. Developed software is able to determine which electrode notices seizure earliest in record also when it is not possible by human eye. Localization of epileptic foci based on topographic mapping is powerful alternative method to another published methods e.g. method based on Low-resolution Electromagnetic Tomography (LORETA) published by [5]. Further BM techniques as Functional Magnetic Resonance (fMRI), a brain Positron Emission Tomography (PET) measure such properties as blood flow or metabolism. These functional brain imaging methods can be used as additional diagnosis of the epileptic zone localization.

Acknowledgments This work was supported by the Slovak Research and Development Agency under the contract No. APVV-14-0519.

References

1. Brazier, M.A.: Interactions of deep structures during seizures in man. In: Synchronization of EEG Activity in Epilepsies, pp. 409–427. Springer, Vienna (1972)
2. Gotman, J.: Measurement of small time differences between EEG channels: method and application to epileptic seizure propagations. Electroencephalogr. Clin. Neurophysiol. **56**(5), 501–514 (1983)
3. Gotman, J.: Noninvasive methods for evaluating localization and propagation of epileptic activity. Epilepsia **44**(12), 21–29 (2003)
4. Kobayashi, K., Ohtsuka, Y., Oka, E., Ohtahara, S.: Primary and secondary bilateral synchrony in epilepsy: differentiation by estimation of interhemispheric small time differences during short spike-wave activity. Electroencephalogr. Clin. Neurophysiol. **83**(2), 93–103 (1992)
5. Matysiak, A., Durka, P.J., Martinez-Montes, E., Barwiński, M., Zwoliński, P., Roszkowski, M., Blinowska, K. J.: Time-frequency-space localization of epileptic EEG oscillations. Acta Neurobiologiae Experimentalis **65**(4), 435–442 (2005)
6. Mohylova, J., Krajca, V.: Processing of biological signals. (Zpracovani biologickych signalu) VSB-TU Ostrava (2006)
7. Mormann, F., Lehnertz, K., David, P., Elger, C.E.: Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients. Phys. D: Nonlinear Phenom. **144**(3), 358–369 (2000)
8. Papezova, M., Babusiak, B., Faktorova, D.: Automatic detection of generalized epileptic seizures. Adv. Electr. Electron. Eng. **13**(2), 191–197 (2015)
9. Quiroga, R.Q., Kreuz, T., Grassberger, P.: Event synchronization: a simple and fast method to measure synchronicity and time delay patterns. Phys. Rev. E **66**(4), 041904 (2002)

The First Experience with the Use of Noise-Assisted Empirical Mode Decomposition Algorithm in the Processing of Multi-channel Electrogastrography Signals

Dariusz Komorowski

Abstract The electrogastrographic examination (EGG) is a non invasive method for the investigation of the myoelectrical stomach activities that is performed with the electrodes placed on the skin of the patient's abdomen. The main aim of EGG examination is the assessment of gastric slow wave propagation. Observation and measurement of the wave parameters changes (e.g. amplitude and frequency) facilitates diagnosing certain abnormalities and disorders associated with the impaired gastric activity. Such abnormalities are usually difficult to detect with the traditional diagnostic methods. One of the main tasks of this study is to extract the gastric slow wave, occurring naturally in the EGG signals, from the raw multi-channel EGG signal. In the presented work, in order to extract slow waves from multi-channel EGG, a modified method of the Multivariate Empirical Mode Decomposition (MEMD) called the Noise-Assisted Empirical Mode Decomposition (NA-MEMD) is proposed to apply. The use of NA-MEMD algorithm and Hilbert-Huang spectrum (HHS) seems to be the appropriate method for analysing non-linear and non-stationary signals and so is the multi-channel EGG.

Keywords Electrogastrography · Noise assisted empirical mode decomposition · Slow wave

1 Introduction

Like in a heart, the phenomenon of myoelectrical activity occurs in the stomach, and can be measured by cutaneous electrodes placed on the patient's abdominal skin. The method was first described by [1]; it concerned a single-channel EGG examination [1]. Until the 90s of the twentieth century the EGG study was not very popular due to the technical difficulties regarding the lack of appropriate recording

D. Komorowski (✉)
Faculty of Biomedical Engineering, Department of Biosensors and Biomedical Signals Processing, Silesian University of Technology, 40 Roosevelt Street, Zabrze, Poland
e-mail: dariusz.komorowski@polsl.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_26

equipment and computer systems for the EGG analysis. The development of this research method took place in 1993, when Chen and his colleagues presented the basic assumptions of one and multi-channel cutaneous EGG [3–5]. In 2003 EGG was considered as a method that helps distinguish healthy people from those with impaired gastric myoelectrical activity [21]. Due to its non-invasiveness and relatively small inconvenience, the method is increasingly used in paediatrics [26, 28]. Myocyte cells of the stomach generate three types of potential: membrane potential, slow waves and action potentials (spikes). Membrane potentials are the least interesting, as they exist in all living cells in the human body. Their function is to maintain the desired tonicity, and this is achieved by treatment with sodium-potassium pump. The basic bioelectrical activity of the stomach is determined by slow wave, which is generated by the gastric non-muscular cells called the Interstitial Cells of Cajal (ICC). Such cells are located between the circular and longitudinal muscle layers of corpus and antrum (the middle of the greater curvature) [7]. The rhythm of peristalsis generated by slow waves is referred as the basic electrical rhythm (BER) and has the frequency of about 0.05 Hz (3 cycle/min, cpm) for healthy humans. The BER itself does not cause contractions yet, but determines their frequency, as contractions occur when the depolarization associated with BER reaches the threshold and there appear spikes. Instead, BER is responsible for the frequency, speed and direction of peristaltic wave [15]. Single or multi-channel EGG signal is acquired from the disposable electrodes arranged on the surface of a patient's abdomen. It is assumed, that the frequency for the standard EGG signal is from 0.015 to 0.15 Hz. EGG signal amplitude is about 100–400 μV. The typical EGG examination takes about 2 h and consists of three stages: the first one generally lasts no longer than 30 min and is defined as the stage before a meal (pre-prandial), the second part takes about 5–15 min, including time when a person eats a standardized meal. The third part of the study, after the meal is digested, takes about 90–120 min and it is called a postprandial stage. The basic analysis of EGG involves determining the frequency of slow waves and classification of its rhythm. The determined frequency is called the dominant frequency (DF). DF values are calculated for some pre-prandial and postprandial parts. The part of the study during a meal is not analysed due to the very high level of disturbances in the EGG records. Normally, the frequency of the slow waves is determined by the short-time fast Fourier transform (STFFT) or using non-parametric methods (e.g. AR, ARMA). To calculate the spectra, according to the method used, the segments of 60–256 s length are used. When testing EGG, according to general medical standard administration, the frequency is given in cpm. Based on the DF values of slow waves, the classification of rhythm is performed for each individual pre-prandial and postprandial part. Additionally, the postprandial part is usually divided into 30 min parts. The rhythm is classified as: bradygastria DF (0–2 cpm), normogastria DF ⟨2–4 cpm⟩ and tachygastria DF ⟨4–9 cpm⟩ [21, 31]. On the basis of the rhythm classification, the percentage index of normogastria (NI) is determined. This index is expressed as the amount of DF values in the range of normal rhythm to the total amount of the DF values. The NI has adequate diagnostic significance.

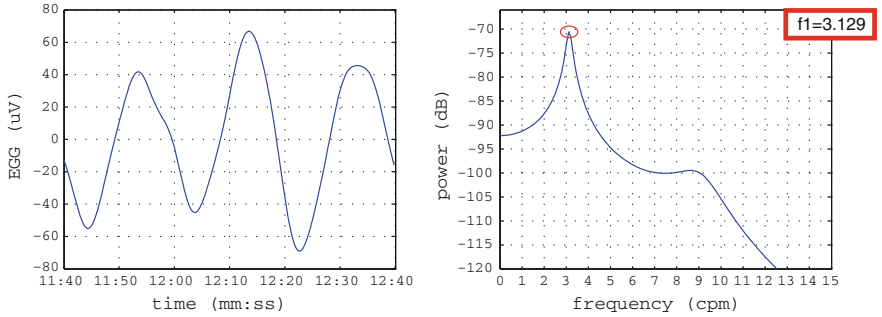


Fig. 1 The example of 1 min segment of EGG signal (*left*), its spectrum (*right*) with marked DF (*small red circle*)

The EGG examination also facilitates the determination of other parameters (the detailed description and application can be found in [19]). The example of 1 min length EGG signal, its spectrum and method of determining the value of DF are shown in Fig. 1.

In the presented work, in order to extract slow waves from multi-channel EGG, a modified method of the Multivariate Empirical Mode Decomposition is proposed to extract slow waves from multi-channel EGG. The use of NA-MEMD algorithm and Hilbert-Huang spectrum seems to be the appropriate method for analysing non-linear and non-stationary signals and so is the multi-channel EGG.

2 Method

The signals (time series), used in the presented study, were recorded using a prototype wireless four-channel biological amplifier of EGG signals [14]. The resolution of analog/digital (A/D) converters used in the amplifier was 24 bits, and sampling frequency was set to 250 Hz/channel. During the signal recording process, the C3 standard electrodes places were applied [19]. The time of examination was from 120 to 150 min. The recorded signals were filtered by means of anti-aliasing low-pass digital fourth-order Butterworth filter with the cutoff frequency set to 2.5 Hz. The filtering process was conducted in both the forward and reverse directions to ensure zero-phase [20]. Further signals were resampled with the new sampling frequency equal 5 Hz. Then, in order to extract EGG signals, the filtering process by using conventional zero-phase digital band-pass filter (0.015–0.15 Hz) was performed. The empirical mode decomposition (EMD) was introduced by Dr. Norden Huang in the mid 90s as an effective tool for adaptive local time-frequency decomposition. The EMD is used to obtain the intrinsic mode functions (IMFs) from the non-linear and non-stationary time series signal [11, 12]. Since then, the EMD has been successfully applied in various fields of engineering sciences, such as: the mechanics of fluid, geophysical surveys and also the study of biomedical signal. The EMD substantially

differs from the other methods of signal analysis e.g.: the Fourier transform or the wavelet transform that lacks the predetermined basis functions. In the case of the Fourier transform, it is a combination of sine and cosine functions, and in the case of the wavelet transform—various types of wavelets [16]. The existence of the known basis functions makes these traditional methods work well for linear and stationary signals, while the EMD method derives from the assumption that any signal consists of finite number of IMF, representing an embedded characteristic oscillation on separated time scale. The IMFs enables capturing certain physical characteristics of non-linear and non-stationary signals [32]. The considerations of the traditional EMD method are described in literature e.g. [6, 13, 29]. Because of the limitations of the standard EMD, in the scientific literature quite a lot of improvements of classical EMD are presented, e.g. Ensemble Empirical Mode Decomposition (EEMD). This method involves adding white Gaussian noise (WGN) to the original signal [30]. Some changes were also suggested in the operation of the EMD algorithm for analysing multi-channel signals. This new algorithm is called Multivariate Empirical Mode Decomposition (MEMD) and described in the [22–24]. In this study, to analyse multi-channel EGG signals, the latest version of MEMD algorithm (NA-MEMD) was used. This method involves creating additional signals containing different distributions of the white Gaussian noise, adding them as independent channels to the existing multi-channel signal and then performing the MEMD algorithm [18, 25]. The description of the applied method is presented below. The purpose of MEMD is the decomposition of the n-dimensional signal $x(t)$ into the sum of the n-dimensional intrinsic mode functions $x_m(t)$ (n-dimensional IMFs) and local median, so-called residuum $r(t)$ (also the n-dimensional),which may be either the average and constant trend. We can present it as an equation:

$$x(t) = \sum_{m=1}^{n} x_m(t) + r(t), \quad x, x_m \in R^n. \tag{1}$$

The main idea of the MEMD algorithm is to estimate the n-dimensional local mean. It is performed by creation the multiple real-valued projections of the original signal. Then, the extremes obtained from all such single-dimensional projections are interpolated componentwise to obtain multi-dimensional signal envelopes [25]. The detailed description of this algorithm is presented in [18, 22, 25]. Quote [18], the short explanation of the MEMD algorithm is presented below.

2.1 Instruction of the MEMD Algorithm

1. *Create V-point Hammersley sequence that will be used for uniformly sampling a (n-1) dimensional sphere.*
2. *Calculate the projections $q_{\theta v}(t)$ of the signal $x(t)$ along all directions of the vector $d_{\theta v}$ for $v = 1, \ldots, V$, this will allow to obtain a set of projections $\{q_{\theta v}(t)\}_{v=1}^{V}$.*

3. *Find the time instants* $\{t_{\theta v}^{i}\}_{v=1}^{V}$ *corresponding to the maxima of set of projections* $\{q_{\theta v}(t)\}_{v=1}^{V}$.
4. *Make interpolation* $\left[t_{\theta v}^{i}, x(t_{\theta v}^{i})\right]$ *to obtain the multivariate signal envelope* $\{e_{\theta v}(t)\}_{v=1}^{V}$.
5. *Calculate the mean of the V multidimensional space of envelopes.*

$$m(t) = \frac{1}{V}\sum_{v=1}^{V} e_{\theta v}(t). \tag{2}$$

6. *Extract the detail* $d(t) = x(t) - m(t)$.
7. *Check the stopping criterion for multivariate IMF. If* $d(t)$ *satisfies the criteria to stop it, use this procedure for* $x(t) - d(t)$, *otherwise use the* $d(t)$.

The stop condition (criterion) for multivariate IMF is similar to the condition proposed by Huang [13] for one-dimensional EMD method. The condition of equality between the number of extremes and the zero crossings is not prescribed here, because of the inability to define the extremes of the multi-dimensional signals properly [17]. In this study the NA-MEMD algorithm, which is a modified version of MEMD, was used for multi-channel EGG signal decomposition.

2.2 Instruction of NA-MEMD Algorithm [6, 24]

1. *Check if the input signal satisfies the criteria of an IMF if it does not, go to the next step, otherwise exit the algorithm.*
2. *Create an additional k-channel,* $(k \geq 1)$ *signal of the same length as the one of the input, containing uncorrelated white Gaussian noise.*
3. *Add the signal noise created in the previous step to the n-dimensional input signal* $(n \geq 1)$, *so as to obtain the new* $(n + k)$ *dimensional input signal.*
4. *Apply the MEMD algorithm to the signal created in the previous step.*
5. *Remove components for a channel containing noise from the obtained results; leave only those that match the original n-dimensional input signal.*

As a result of the processing signal by means of the NA-MEMD algorithm, the signal decomposition is transformed into individual IMF. The interpretation of the resulting components is a quite serious problem. Since our aim is to isolate the slow wave, the physical properties of this wave are known, i.e. the typical range of frequency. To choose an "appropriate" IMF, in this paper a frequency criterion is prepossessed. This criterion is based on the calculation for each taken IMF (for each EGG signal channel), defined for this purpose as the instantaneous normogastria index (INI). The value of this ratio determines the percentage of amount instantaneous frequency components between 2 and 4 cpm (0.033–0.66 Hz) to all amount values of the instantaneous frequency of the processed IMF. The value of this ratio is the highest for the chosen IMF. The values of instantaneous frequencies were calculated according to Hilbert-Huang transform (HHT) [12].

2.3 Hilbert-Huang Transform, Instantaneous Frequency

The HHT transform is created as a result of IMF and Hilbert transform (HT) [10, 12]. The Hilbert transform of any arbitrary time series signal x(t) of class L^P can be expressed by the following formula:

$$y(t) = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{x(\tau)}{t-\tau} d\tau, \tag{3}$$

where P indicates the Cauchy principal value of the integral. According to this definition, the $x(t)$ and $y(t)$ functions create the complex conjugate pair, thus the analytic signal $z(t)$ can be expressed as:

$$z(t) = x(t) + iy(t) = a(t)e^{i\theta(t)}, \tag{4}$$

where $i^2 = -1$, and

$$a(t) = (x^2(t) + y^2(t))^{1/2}, \quad \theta(t) = \arctan\left(\frac{y(t)}{x(t)}\right), \tag{5}$$

In the equations presented above a determinates the instantaneous amplitude, and θ represents the instantaneous phase. The instantaneous frequency can be derived from:

$$\omega = \frac{d\theta}{dt}. \tag{6}$$

In this case, both the amplitude and frequency are functions of time and can be expressed as a function of time and frequency $H(\omega, t)$. The Eq. (6) allows only calculating the main values of the phase. This causes a phase discontinuity, which period is equal to 2π, so that the frequency calculation is performed by using the following equation that derives directly from the Eq. (6).

$$\omega(t) = \frac{x(t)y'(t) - x'(t)y(t)}{x^2(t) + y^2(t)}, \tag{7}$$

where prime $'$ means the differentiation with respect to time. In practice, to avoid these differentiations and decrease time processing, the continuous frequency $\omega(t)$ can be approximated using one of the following formulae:

$$\omega(t) = \frac{1}{T} \arctan\left(\frac{x(t)y(t+T) - x(t+T)y(t)}{x(t)x(t+T) + y(t)y(t+T)}\right), \tag{8}$$

$$\omega(t) = \frac{1}{T} \arctan\left(\frac{x(t-T)y(t+T) - x(t+T)y(t-T)}{x(t-T)x(t+T) + y(t-T)y(t+T)}\right), \tag{9}$$

$$\omega(t) = \frac{1}{T} \arctan\left(\frac{x(t)y(t+T) - x(t+T)y(t)}{(x(t) + x(t+T))^2 + (y(t) + y(t+T))^2}\right), \tag{10}$$

where T is the sample period of time series signal [2].

3 Results

The results, obtained with the method described in this work are presented on the example of analysis of the real four-channel EGG signal. The EGG test was made for a light meal (400 ml of cold water). The presented results were obtained using the NA-AMD algorithm with one channel of WGN added. The NA-MEMD decomposition was made with V = 40 directions of projections. The Calculations were performed for two different values of noise to signal (N/S) ratio. The corresponding values were 0.1 and 0.2 respectively. In order to provide the required values of N/S, the EGG signal was standardized (carried out separately for each of the 4 channels) by dividing the values of the signal (samples) by the standard deviation (STD) of EGG signal. Because of the standard (conventional) method of analysis of the EGG, the examination is divided into about 30 min parts: pre-prandial (which usually takes about 30 min) and postprandial (2 or 3 parts of 30 min each). In the presented study the analysis of EGG signal was also made for separate 30 min stages. The calculations were carried out in Matlab using MEMD function implemented by Rehman and Mandic [22] and downloaded from [9]. The Hilbert-Huang spectra presented below were calculated using the function HHSPECTRUM, TOIMAGE and, DISPHHS written by Flandrin and Rilling [8, 27]. The example of decomposition of the one (A1) channel of 30-min time series 4-channel EGG signal (pre-prandial part) with NA-MEMD algorithm is shown in Fig. 2. The calculated values of the coefficients INI (%) for each IMF (Top 10) of 4-channel EGG (pre-prandial) for two values of N/S: 0.1 and 0.2 are shown respectively, in Tables 1 and 2.

The Hilbert-Huang spectra of the chosen IMFs (with the highest INI index) of four-channel EGG signal (pre-prandial part) are shown in Fig. 3. The spectra were obtained by NA-MEMD method with one added channel of WGN with N/S ratio equal 0.1.

Figure 4 presents the IMFs of each 4-channel EGG leads, selected on the basis of the criteria described above (max INI).

The classical spectra of "the best" IMFs (obtained by the Fourier transform) are presented in Fig. 5a–d. These spectra and results of HHT, presented in the Fig. 3, clearly show that the local energy of IMFs components is distributed mainly at the range of the characteristic slow wave frequency (0.05 Hz; 3 cpm).

The following figures show the selected (with the highest values of INI) IMFs for the postprandial stages of the four-channel EGG examination. The results were obtained for the same realization of WGN, but for the different N/S amounting to 0.1 (Fig. 6) and 0.2 (Fig. 7) respectively. These results show that decomposition EGG

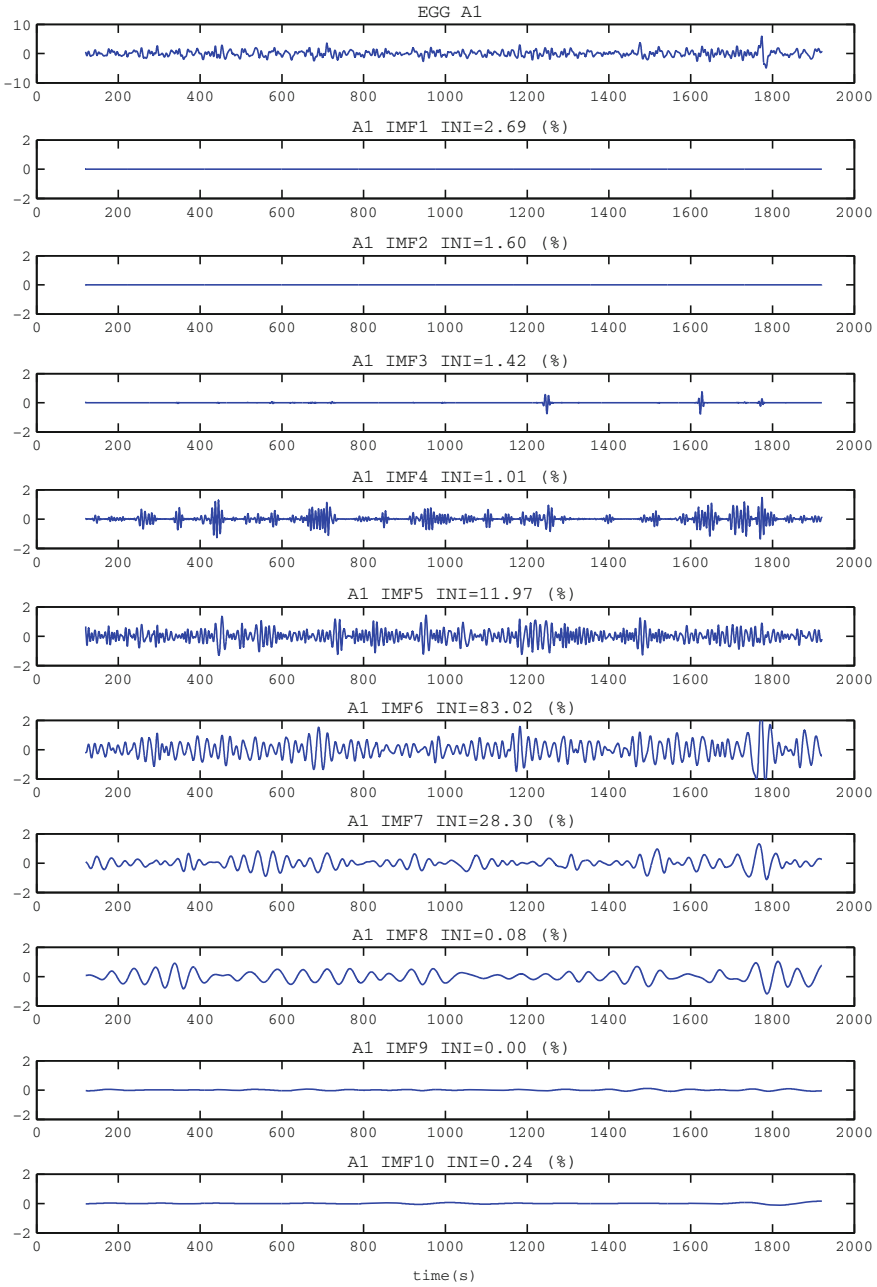


Fig. 2 The example of decomposition into IMFs of one (*A1*) channel 30-min time series of the 4-channel EGG signal (pre-prandial part), (with NA-MEMD method and N/S = 0.1)

Table 1 The values of INI (%) for the first 10 IMFs of four-channel EGG signal (pre-prandial), (N/S = 0.1)

	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10
A1	2.69	1.60	1.42	1.01	11.97	83.02	28.30	0.08	0.00	0.24
A2	2.33	1.31	1.43	0.80	11.19	85.93	17.01	0.64	0.00	0.03
A3	2.52	1.63	1.57	0.72	13.69	91.96	24.06	1.43	0.21	0.39
A4	2.63	1.85	1.73	0.98	11.43	90.72	19.54	2.48	0.02	0.20

Table 2 The values of INI (%) for the first 10 IMFs of four-channel EGG signal (pre-prandial), (N/S = 0.2)

	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10
A1	2.57	2.12	1.71	0.59	5.30	78.63	46.14	0.10	0.67	0.07
A2	2.69	1.13	1.50	0.71	4.86	82.79	44.78	1.10	1.13	0.02
A3	2.70	1.78	1.49	0.84	5.56	90.08	47.10	1.75	0.29	0.67
A4	3.08	1.27	1.89	0.57	4.93	86.74	42.97	1.88	0.14	0.26

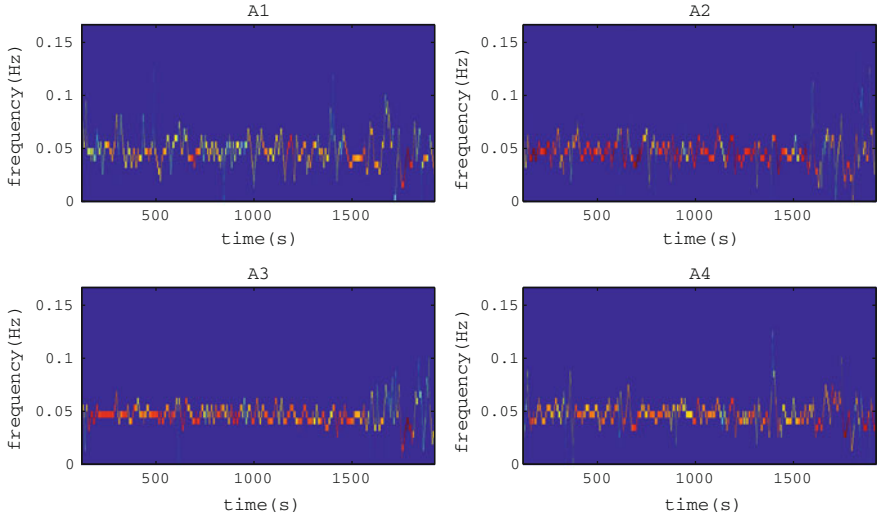


Fig. 3 The HHS of "the best" IMFs of the 4-channel EGG signals (pre-prandial part)

into IMFs by the means of NA-MEMD method is sensitive for N/S ratio and requires further investigation.

The Fig. 8 also presents the correlation between the IMFs of different EGG channels (leads), for example between A1 and A2, and A1 and A3. These diagrams of the correlation coefficients seem to be interesting material for further analysis of coupled and uncoupled slow waves, that may also occur in the EGG signal [4].

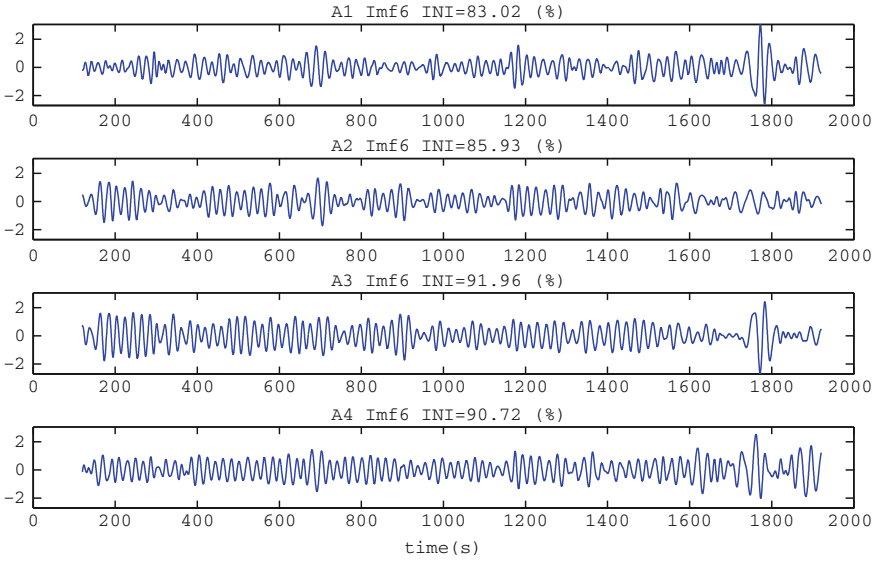


Fig. 4 "The best" (with the highest values of INI) IMFs of each channel (A1–A4) of the 4-channel EGG. The results were obtained for the pre-prandial part (N/S = 0.1)

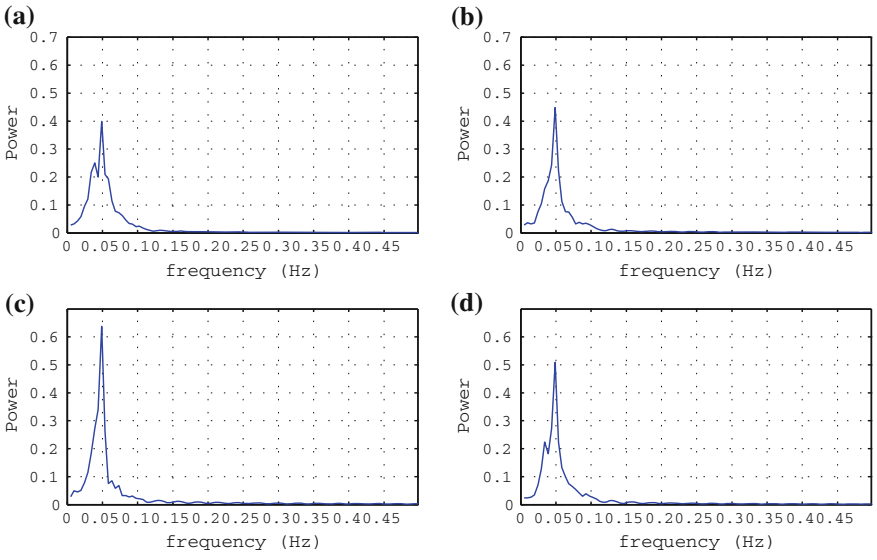


Fig. 5 The examples of Fourier spectra of "the best" IMFs of four-channel EGG (*A1–A4*)

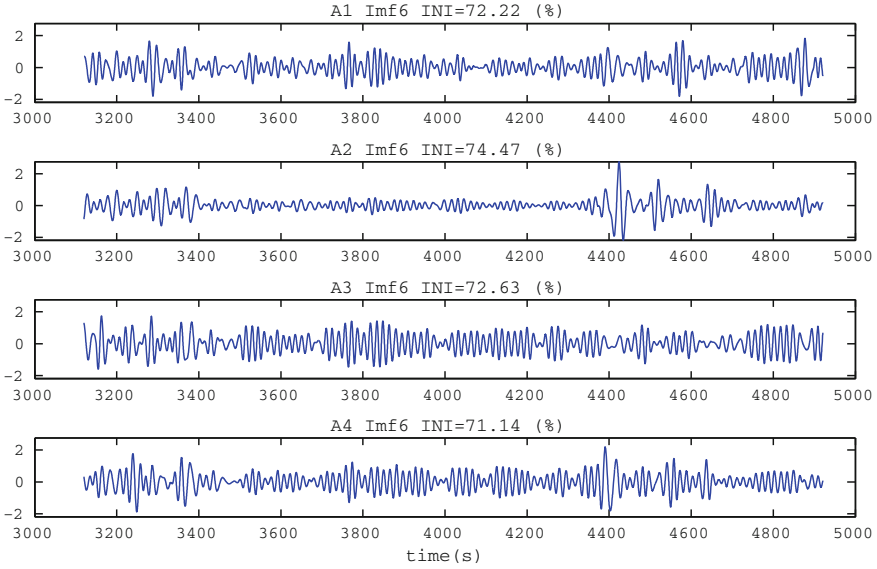


Fig. 6 "The best" (with the highest values of INI) IMFs of each channel (*A1–A4*) of the 4-channel EGG. The results were obtained for the postprandial part (N/S = 0.1)

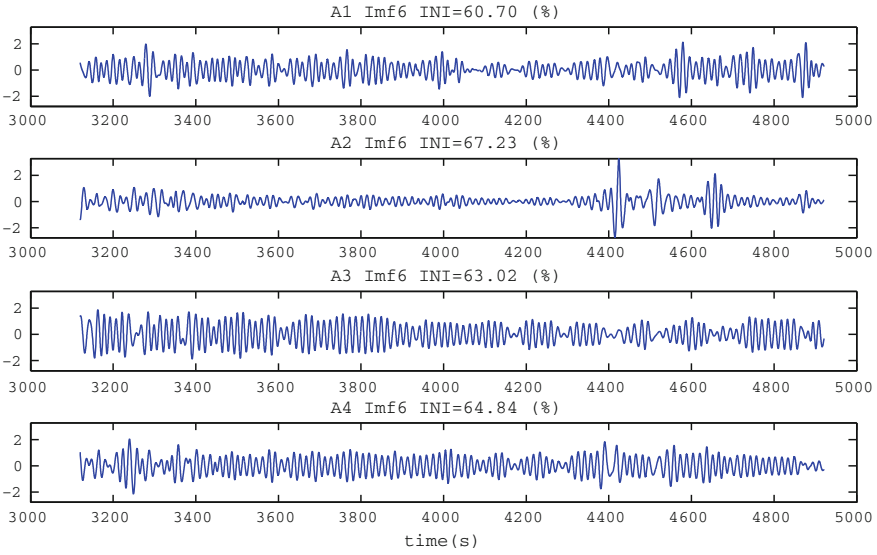


Fig. 7 "The best" (with the highest values of INI) IMFs of each channel (*A1–A4*) of the 4-channel EGG. The results were obtained for the postprandial part (N/S = 0.2)

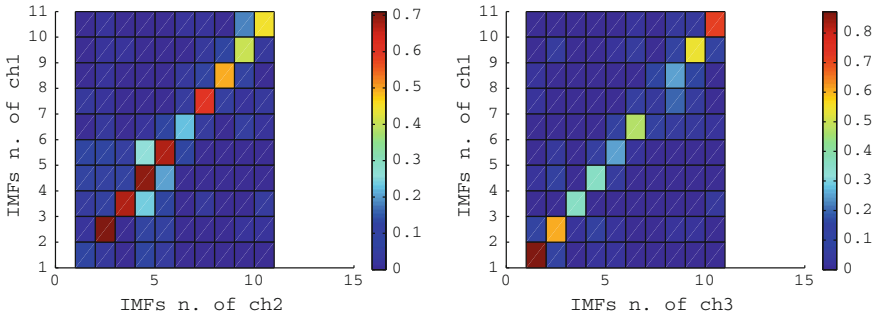


Fig. 8 The graphical representation of the correlation rations between the IMFs of selected channels, *A1–A2* (*left*) and *A1–A3* (*right*)

4 Conclusion

The preliminary results of analysis of four-channel EGG signals, obtained by both NA-MEMD and HHT methods, indicate relatively high potential for adaptively decomposing non-stationary and non-linear multi-channel EGG signals. After a series of experiments using the NA-MEMD algorithm for processing four-channel EGG signals, examples of which are presented in this work, it was found that the proposed algorithm works suitably for analysing EGG signals. The obtained results are promising (extraction of the slow wave component (0.05 Hz) works properly), however, they require further comparative analysis with the results of conventional methods used in the analysis of EGG signals. Such verification and comparison between healthy patients and those who suffer from stomach disorders may also lead to finding the criteria or parameters that would help diagnose certain diseases of the digestive system. The proposed method of analysis EGG is very interesting and perspective, which should be the basis for further research. There are several possible directions of such research, and they determine its effectiveness. These include, inter alia: evaluation of the influence of parameters changes in NA-MEMD algorithm (the number of projection directions, the number of additional channels of noise, the values of N/S ratio) on the obtained results. The study of the interrelationship between the IMFs, for example diagrams of correlation coefficients, seems to be interesting for further analyse of coupling and uncoupling between slow waves, which is crucial for gastric emptying. In the future, it would be advisable to confront the results of the analysis EGG by means of NA-MEMD method with clinical research and medical specialist's opinion. The applied algorithm should be also optimized in terms of its computational speed to reduce the processing time.

References

1. Alvarez, W.C.: The electrogastrogram and what it shows. JAMA **78**, 1116–1119 (1922)
2. Chen, J.D., Jegen-Kulcsar, M.: The empirical mode decomposition (EMD) method in MT data processing, vol. 33. Kolloquium Elektromagnetische Tiefenforschung, Hotel Maxicky, Decin, Czech Republic, 1–5 Oct 2007
3. Chen, J.D., McCallum, R.W.: Clinical application of electrogastrography. Am. J. Gastroenterol. **88**, 1324–1336 (1993)
4. Chen, J.D., McCallum, R.W.: Electrogastrographic parameters and their clinical significance. In: Chen, J.D., McCallum, R.W. (eds.) Electrogastrography: Principles and Applications, pp. 45–73. Raven, New York (1994)
5. Chen, J.D., Zou, X., Lin, X., Ouyang, S., Liang, J.: Detection of gastric slow wave propagation from the cutaneous electrogastrogram. Am. J. Physiol. **277**, G424–G430 (1999)
6. Flandrin, P., Rilling, G., Goncalves, P.: Empirical mode decomposition as a filter bank. IEEE Signal Process. Lett. **11**(2), 112–114 (2004)
7. Forys, S., Szadkowski, K., Chojnacki, C.: Elektrogastrografia w diagnostyce dyspepsji czynnościowej (Electrogastrography in the diagnosis of functional dyspepsia), Przegląd Gastroenterologiczny 2008 (In Polish)
8. http://perso.ens-lyon.fr/patrick.flandrin/emd.html
9. http://www.commsp.ee.ic.ac.uk/~mandic/research/emd.htm
10. Huang, N.E., Wu, Z.: A review on Hilbert-Huang transform: method and its applications to geophysical studies. Rev. Geophys. **46** (2008). doi:10.1029/2007RG000228A
11. Huang, N.E., Long, S.R., Shen, Z.: The Mechanism for Frequency Downshift in Nonlinear Wave Evolution. Academic Press Inc. Published by Elsevier Ltd. (1996)
12. Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C., Tung, C.C., Liu, H.H.: The empirical mode decomposition and the hilbert spectrum for nonlinear and nonstationary time series analysis. Proc. R. Soc. London, Ser. A **454**, 903–995 (1998)
13. Huang, N.E., Wu, M.-L.C., Long, S.R., Shen, S.S.P., Qu, W., Gloersen, P., Fan, K.L.: A confidence limit for the empirical mode decomposition and Hilbert spectral analysis. Proc. Roy. Soc. Lon. A **459**(2037), 2317–2345 (2003)
14. Komorowski, D., Pietraszek, S., Grzechca, D.: The wireless system for EGG signal acquisition. In: 19th IEEE International Conference on Electronics, Circuits and Systems. ICECS 2012, Seville, Spain, 9–12 Dec 2012
15. Konturek, S.J. (eds.): Fizjologia czlowieka: podrecznik dla studentow medycyny (Human Physiology: textbook), Elsevier Urban and Partner, Wroclaw wyd. 1, 2007 (In Polish)
16. Mallat, S.G.: A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans. Pattern Anal. Mach. Intell. **11**(7) (1989)
17. Mandic, D.P., Goh, V.S.L.: Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models. Wiley, Hoboken (2009)
18. Mandic, D.P., Rehman, N.U., Wu, Z., Huang, N.E.: Empirical Mode Decomposition-Based Time-Frequency Analysis of Multivariate Signals, Digital Object Identifier. doi:10.1109/MSP.2013.2267931. Date of publication: 15 Oct 2013
19. Medtronic A/S: Polygram NetTM Reference Manual. Skovlunde (2002) (Denmark)
20. Oppenheim, Alan V., Schafer, Ronald W., Buck, John R.: Discrete-Time Signal Processing, 2nd edn. Prentice Hall, Upper Saddle River (1999)
21. Parkman, H.P., Hasler, W.L., Barnett, J.L., Eaker, E.Y.: Electrogastrography: a document prepared by the gastric section of the american motility society clinical GI motility testing task force. Neurogastroenterol. Motil. **15**, 89–102 (2003)
22. Rehman, N.U., Mandic, D.P.: Multivariate empirical mode decomposition. Proc. Roy. Soc. Lon. A **466**(2117), 1291–1302 (2010)
23. Rehman, N.U., Mandic, D.P.: Empirical mode decomposition for trivariate signals. IEEE Trans. Signal Process. **59**(5), 2421–2426 (2011)
24. Rehman, N., Mandic, D.P.: Filterbank property of multivariate EMD. IEEE Trans. Signal Process. **59**(5), 2421–2426 (2011)

25. Rehman, N., Park, C., Huang, N.E., Mandic, D.P.: EMD via MEMD: multivariate noise-aided computation of standard EMD. Adv. Adapt. Data Anal. **5**, 1350007 (2013). doi:(10.1142/S1793536913500076)
26. Riezzo G., Russo F., Indrio F.: Electrogastrography in adults and children: the strength, pitfalls, and clinical significance of the cutaneous recording of the gastric electrical activity. BioMed Res. Int. **2013**, Article ID 282757, 14 pp (2013). doi:10.1155/2013/282757
27. Rilling, G., Flandrin, P., Goncalves, P.: On empirical mode decomposition and its algorithms. In: Proceedings of the 6th IEEE/EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP '03), Grado, Italy (2003)
28. Stroj, L., Krusiec-Swidergol, B., Kasicka-Jonderko, A., Jonderko, K., Blonska-Fajfrowska, B.: Application of electrogastrography in paediatrics I. Definition of normal ranges of parameters of an electrogastrogram in Polish children (in Polish). Wiad. Lek. **60**, 517–524 (2007)
29. Wang, G., Chen, X., Qiao, F.-L., Wu, Z., Huang, N.E.: On intrinsic mode function. Adv. Adapt. Data Anal. **2**(3), 277–293 (2010)
30. Wu, Z., Huang, N.E.: Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv. Adapt. Data Anal. **1**(1), 1–41 (2009)
31. Yin, J., Chen, J.D.Z.: Electrogastrography: Methodology. Validation and Applications. J. Neurogastroenterol. Motil. **19**, 5–17 (2013)
32. Zeiler, A., Faltermeier, R., Keck, I.R., Tome, A.M., Puntonet, C.G., Lang, E.W.: Empirical mode decomposition—an introduction. In: The 2010 International Joint Conference on Neural Networks (IJCNN), pp. 1,8, 18–23 July 2010

Assessment of Slow Wave Propagation in Different Phases of Food Stimulation in the Multichannel Electrogastrographic Signal (EGG)

Barbara T. Mika and Ewaryst J. Tkacz

Abstract The electrogastrogram (EGG), a cutaneous recording of electrical activity in the stomach is a mixture of 3 cycle per minute (3 cpm) gastric pacesetter potential known as slow wave, electrical activity from other organs and noise. Proper slow wave propagation is responsible for gastric peristaltic contractions, which are the basis for emptying of solids from stomach. Delay in the stomach emptying leads to some gastric disorders such as bloating, vomiting or nausea. To assess the slow wave propagation it is necessary to obtain slow wave in each channel of multichannel EGGs. In this paper combined methods: Independent Component Analysis (ICA) and adaptive filtering in the cosine transform domain was proposed to gain a purified EGG signal from each channel. Time shift between EGG signals from various channels was estimated by means of cross covariance analysis performed after adaptive filtering of each channel with reference signal obtained from blind sources separation by ICA algorithm. The effectiveness of that proposed methods was at first validated for the synthetic data and after was applied for human EGG, recorded before and after food stimulation.

Keywords Electrogastrography · Slow wave · Propagation

1 Introduction

Electrogastrography is a body surface, noninvasive measurement of electrical stomach activity. It can be acquired by attaching electrodes on the abdominal skin over the stomach. For the first time one channel EGG data was obtained in 1922 by Walter Alvarez [1] and during the long decades, analysis of the EGG signal was concentrated on the one channel recordings. With the introduction of multichannel electrogastrography in the second half of the twentieth century the EGG has become

B.T. Mika (✉) · E.J. Tkacz
Faculty of Biomedical Engineering, Department of Biosensors and Processing of Biomedical Signals, Silesian University of Technology, Zabrze, Poland
e-mail: b.mika@polsl.pl

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_27

an attractive tool for physiological and pathophysiological studies of the stomach. Computer simulation studies have suggested, that detection of slow wave propagation form multichannel EGGs is possible [3, 4, 11]. Gastric peristaltic wave that migrates from proximal to distal stomach are a basis for emptying stomach contents [9]. From the pacemaker area of stomach, located on the greater curvature between the fundus and corpus, spontaneous electrical depolarization and repolarization generates the myoelectrical excitation. The main component of gastric myoelectrical activity, called gastric slow wave has a frequency about 3 cpm (0.05 Hz). Slow waves originate in the pacemaker region and propagate circumferentially and distally toward the pylorus. For gastric peristalsis to spread as a mechanical wave, gastric slow wave must propagate through the gastric musculature.

The tunica muscularis of stomach is built of three layers of muscles: longitudinal, circular and oblique. However they are not able to awake electrical excitation. Particular network of cells named the Interstitial Cells of Cajal (ICC) are responsible for myoelectrical rhythmicity, recorded as the gastric pacesetter potentials [12, 15].

Proper slow wave propagation is a basic mechanism, which both controls and integrates gastric wall motility. The proximal corpus region generates the slow wave with the greatest frequency 3 cpm, so there is a time, for slow wave to propagate and initiate slow waves, in the other regions of stomach, before they generate their own event [9]. This phenomenon is called *coupling*. If the time required for slow wave propagation is too short for corpus pacemaker to entrain the distal pacemakers, there is the breakdown in the stomach frequency gradient. In these uncoupled cases the motility disorders occur and the spread of peristaltic contractions could be inhibited, causing delay in gastric emptying and such symptoms as nausea, bloating, vomiting or abdominal discomfort [9].

The aim of this paper is to examine the effectiveness of combining Independent Component Analysis (ICA), adaptive filtering and the cross-covariance for identifying the slow waves propagation. Parameters depict EGG properties mostly based on spectral analysis, where information about slow waves spread and coupling are lost, so this paper propose the new combination of known methods for assessing slow wave propagation in synthetic and human data.

2 Methods

2.1 Independent Component Analysis

Independent Component Analysis (ICA) is a powerful method for the blind source separation (BSS) [5–7, 14]. In this paper ICA algorithm was applied for multichannel EGG recording to obtain the reference signal for adaptive filtering. EGG data are a mixture consisting of gastric myoelectrical activity of stomach, the electrical activity of the adjacent organs and a various kind of noise or artifacts. All these signals may be considered to be independent because of their different biological origin

(as they are generated by different sources). Let's assume that n obtained by multichannel electrogastrography EGG signals $X_1(t), X_2(t), X_3(t), \ldots, X_n(t)$ are linear combinations of n unknown mutually, statistically, independent source components: $S_1(t), S_2(t), S_3(t), \ldots, S_n(t)$, such as electrical activity of stomach, heart, respiration or random noise.

If $\mathcal{X} = [X_1(t), X_2(t), \ldots, X_n(t)]^T$ and $\mathcal{S} = [S_1(t), S_2(t), \ldots, S_n(t)]^T$ then $\mathcal{X} = \mathcal{A} \cdot \mathcal{S}$, where $\mathcal{A}$ is unknown non-singular mixing matrix. The task of ICA algorithm is to extract the source signals $S_i(t)$ where $i = 1, 2 \ldots n$, only from their mixed measure $\mathcal{X}$, by estimating matrix $\mathcal{E} = \mathcal{A}^{-1}$ so as $\mathcal{S} = \mathcal{E} \cdot \mathcal{X}$. Each of vector $\mathcal{X}$ and $\mathcal{S}$ could be regarded as a random variable. The ICA method consists of two tasks: proper construction of so called *contrast function* and an optimization algorithm. The contrast function is a quantitative measure of stochastic independence of random variables S_i i.e. extracted source signals.

In this study the Matlab implementation of FastICA algorithm proposed by Hyvärinen and Oja [7], has been successfully applied for both simulated and human EGG data.

2.2 Adaptive Filtering

Adaptive noise cancellation from each channel of EGG data was performed, in the discrete cosine transform (DCT) domain, by the means of filter proposed by Liang [2, 10] with a reference signal derived by ICA algorithm from 4-channel EGG data. Let's vector $D = EGG_k = [d_1, d_2, \ldots d_T]$ denote EGG signals from k-channel of multichannel EGG recording of T samples. Adaptive filtering was based on the least mean square (LMS) algorithm to update the filter weights [3, 16]. If the mean square error $\frac{1}{T-N} \sum_{j=1}^{T-N} e_j^2$ for $e_j = d_j - y_j$, between y_j the filter output and d_j the primary input, reached the minimum, then the output signal $Y = [y_1, y_2, \ldots y_{T-N}]$ is the best, in terms of least square error, estimate of the primary input signal $D = EGG_k$ (N is the order of adaptive filter i.e. the number of filter weights).

Let's $F_j = [f_j, f_{j+1}, \ldots, f_{j+N-1}]$ is the reference signal obtained by ICA algorithm and $j = 1, 2, \ldots, T - N + 1$, where T denote the number of EGG_k samples and N is the order of adaptive filter (the number of filter weights).

Let's vector $Z_j = [z_j(1), z_j(2), \ldots, z_j(N)]$, will be the discrete cosine transform of vector F_j, $(DCT(F_j) = Z_j)$, and $W_j = [w_j(1), w_j(2), \ldots, w_j(N)]$, will be vector of filter weights. Output of adaptive filter y_j could be expressed by the formula:

$$y_j = Z_j \cdot W_j^T, \tag{1}$$

wherein the filter weights change according the formula:

$$w_{j+1}(n) = w_j(n) + \frac{\mu}{\frac{1}{N} \sum_{n=1}^{N} \left| z_j(n) \right|^2} e_j(n) z_j(n), \tag{2}$$

for $j = 1, \ldots, N-1$, $w_j(n)$ is n-th filter weight for the j moment and μ is the coefficient controlling the speed of convergence [2, 10].

The order of adaptive filter, used in this work has been set to $N = 8$ and the coefficient controlling the rate of convergence, has been experimentally fixed value to $\mu = 0.00375$. The adaptive filtering performed in the time domain, by means of LMS algorithm, is usually slow convergent. The convergence in this case, depends on the rate of $\frac{\lambda_{min}}{\lambda_{max}}$, where $\lambda_{min}, \lambda_{max}$ are the smallest and the largest eigenvalue of autocorrelation matrix of input signal. As the orthogonal transformation of input signal, decrease the range of eigenvalues of autocorrelation matrix, (which is proofed in the next section), it speeds up the convergence of LMS algorithm, at the same time [13]. In this paper the Discrete Cosine Transform (DCT) was chosen for transformation of the input signal [2, 10, 13].

Transform Domain Adaptive Filter Let's assume, that the Discrete Cosine Transform (DCT), chosen for the linear orthogonal transform is represented by it's orthogonal coefficients matrix P, so if $Y = DCT(X)$ then $Y = PX \wedge P-$ orthogonal i.e. $P^T = P^{-1}$. Let's denote $E\left\{YY^T\right\} = B$ and $E\left\{XX^T\right\} = A$.

Lemma 1 *For any vector $X = [x_1, x_2, \ldots x_n]^T$, of which norm is equal one $\|X\| = 1$ and any symmetric matrix A, there is the true relationship $\lambda_{min} \leq X^T AX \leq \lambda_{max}$, where λ_{min} and λ_{max} are respectively the largest and the smallest eigenvalue of matrix A [8].*

Theorem 1 *The spread of eigenvalues of matrix B is less than, or equal to, the spread of eigenvalues of matrix A.*

Proof If

$B = E\left\{YY^T\right\} = E\left\{PX(PX)^T\right\} = E\left\{PXX^TP^T\right\} = PE\left\{XX^T\right\}P^T = PAP^T$ then $B = PAP^T$. Let's w with $\|w\| = 1$ will be an eigenvector of matrix B for eigenvalue $\overline{\lambda_B}$. If w is an eigenvector of matrix B then $Bw = \overline{\lambda_B}w \Leftrightarrow w^TBw = w^T\overline{\lambda_B}w \Leftrightarrow w^TBw = \overline{\lambda_B}w^Tw \Leftrightarrow w^TBw = \overline{\lambda_B}\|w\|^2 \Leftrightarrow w^TBw = \overline{\lambda_B}$.

Let's take vector $P^Tw \in R^{nx1}$ and $P-$orthogonal ($P^T = P^{-1}$) under consideration so $\left\|P^Tw\right\|^2 = (P^Tw)^T(P^Tw) = w^TPP^Tw = w^Tw = \|w\|^2 = 1$.

If A is the symmetric matrix and $\left\|P^Tw\right\| = 1$ then according to the Lemma 1 $(P^Tw)^TA(P^Tw) \in [\lambda_{min}, \lambda_{max}]$ where $\lambda_{min}, \lambda_{max}$ are the smallest and the largest eigenvalue of symmetric matrix A ($\lambda_{min}, \lambda_{max} \in \mathbb{R}$) so $(P^Tw)^TA(P^Tw) \in [\lambda_{min}, \lambda_{max}] \Leftrightarrow (w^TP)A(P^Tw) \in [\lambda_{min}, \lambda_{max}] \Leftrightarrow w^T(PAP^T)w \in [\lambda_{min}, \lambda_{max}] \Leftrightarrow w^TBw \in [\lambda_{min}, \lambda_{max}] \Leftrightarrow \overline{\lambda_B} \in [\lambda_{min}, \lambda_{max}]$. □

2.3 Cross-Covariance

In order to determine the time shift, between signals from two various channels of EGG data, cross-covariance analysis was applied. In the first step, the signals from each channel were resampled and divided into 20 s parts, representing one cycle of

slow wave. In the second step, the following analysis was performed for each 20 s of EGG data. Iteratively, 20 scalar products were calculated, for two 20 s data from fixed two channels. After each iteration, first of two signals, was shifted one sample forward or backward, so each calculated dot product, was denoted by numbers of shifted samples. The largest value of scalar product means, the largest similarity of two 20 s signals and simultaneously, the number of shifted samples denotes, the time shift, of these investigated parts of signals. The same procedures was performed for all 20 s fragments of signals from two fixed channels. The number of samples, obtain for the largest scalar product, means the time lag, between each two 20 s time series, from individual channels [4]. Finally, average value for all of the 20 s pieces of tested signals, were found to be the time shift of these signals. Signal segmentation into 20 s pieces, gives the opportunity to observe gastric slow wave propagation.

3 Results

To test the efficiency of proposed method, two kinds of the synthetic data were designed for validation study and therefore the method was applied for 4-channel EGG human data. The first group of synthetic data was based on the fact, that slow wave could be simulated, by the sine function while the second group was constructed, using the real one channel EGG signal, recorded from a healthy person.

3.1 Simulation Results—Synthetic Data First Type

Let's assume, that the pure EGG signal, could be represented by the sine function with the frequency 3 cpm (0.05 Hz). To simulate the slow wave propagation, from corpus to pylorus, four source signals were taken into consideration: $S_1(t) = \sin(2\pi \cdot 0.05t)$ i.e. 3 cpm (0.05 Hz) slow wave, $S_2(t) = \sin(2\pi \cdot 0.05t + \frac{\pi}{2})$ 3 cpm slow wave with the phase shifted by vector $[\frac{-\pi}{2}, 0]$, it is about 5 s respectively to S_1, $S_3(t) = \sin(2\pi \cdot 0.05t + \pi)$ also 3 cpm slow wave shifted by vector $[-\pi, 0]$ i.e. about 10 s respectively to S_1 and $S_4(t)$ as random noise. The simulated 4-channel EGG data (denoted as a vector $\mathcal{X}$) was obtained by using the linear transform of vector $S = [S_1, S_2, S_3, S_4]$ with the mixing matrix $\mathcal{A}$.

$$\mathcal{X} = \mathcal{A}\mathcal{S} \Longleftrightarrow \underbrace{\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix}}_{EGG} = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 1.6 \\ 0.1 & 1 & 0 & 1.4 \\ 0.1 & 0.1 & 1 & 1.8 \\ 0 & 0 & 0.2 & 1.1 \end{pmatrix}}_{\mathcal{A}} \cdot \underbrace{\begin{bmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \end{bmatrix}}_{\mathcal{S}} \tag{3}$$

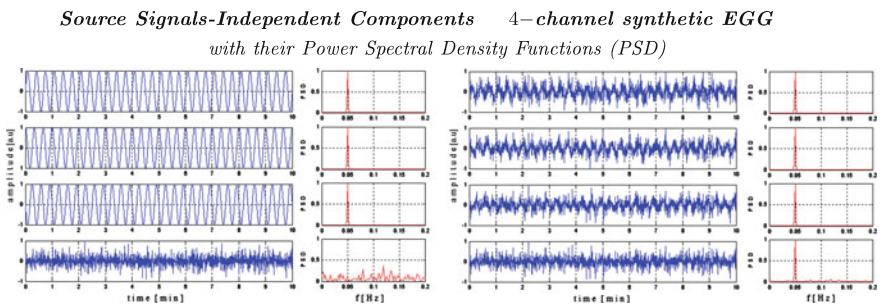


Fig. 1 On the *left*, there are the source signals S_1, S_2, S_3, S_4, on the *right*, simulated 4-channel EGG with their power spectral density function (PSD) next to

In Fig. 1, on the left hand, there are unknown source signals S_1, S_2, S_3, S_4 with encoded propagation of slow wave, which is the desired information, on the right hand, there are simulated 4-channel EGG. The power spectral density function (PSD) for each signal was calculated, as a product of squared absolute value of the Discrete Fourier Transform (DFT) of analyzed signal and the Hamming window function, divided by the energy of the window. The frequency of source signals according to the assumption are 3 cpm (0.05 Hz).

As the time shift could be changed or even lost in the EGG recording Fig. 2 (on right), it is difficult to detect slow wave propagation from raw, multichannel EGG. Some methods, which are able to recover the phase shift of source signals are needed. In order to assess the phase shift in the source signal from their recorded mixtures, in the first step the ICA algorithm was applied and two independent components ICA_1 and ICA_2 with frequency 3 cpm and phase shift about 5 s, were extracted from the EGG data Fig. 3.

They were next used as *reference signal 1* for adaptive filtering to enhance EGG signal in each channel. The filter order was 8 and the coefficient controlling the rate of convergence was experimentally established on $\mu = 0.00375$.

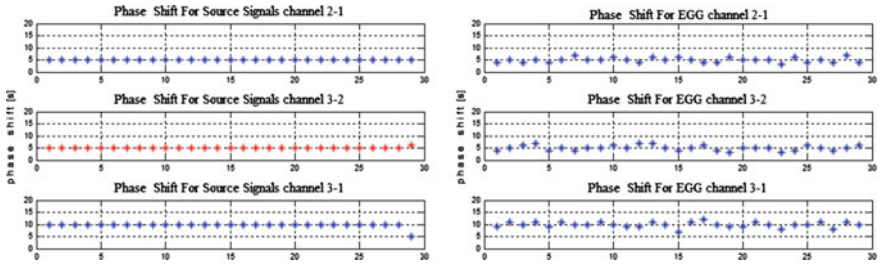


Fig. 2 The figure on the *left*, shows graphical representation of phase shift in the each 20 [s] of 10 [min] recording, between source signals: $S_2 - S_1, S_3 - S_2$ 5 s; $S_3 - S_1$ 10 s. The figure on the *right*, presents the time shift for EGG data

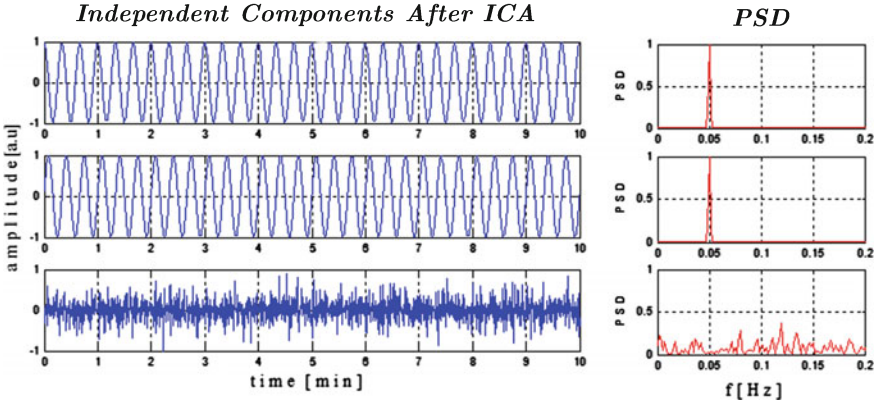


Fig. 3 Independent components: ICA_1, ICA_2, ICA_3 obtained after applying the ICA method with their PSD function next to

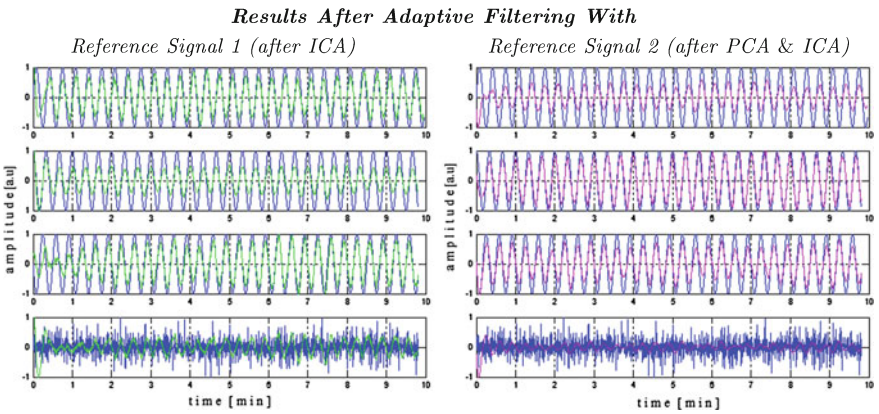


Fig. 4 On the *left*, the result of adaptive filtering for *reference signal 1*= ICA_1, on the *right*, the result of adaptive filtering for *reference signal 2* obtained after PCA & ICA

Both ICA_1 and ICA_2 give a good result in adaptive filtering (Fig. 4 result for ICA_1 is on the left side). In the next step the data dimension was reduced by the means of Principle Component Analysis (PCA) algorithm. The smallest eigenvalues of $(EGG)(EGG)^T$ matrix, indicating the noise amount in the multichannel EGG were omitted and a new reference signal denoted *reference signal 2* was obtained after ICA application. The result of adaptive filtering with the *reference signal 2* is shown in Fig. 4 on the right. Figure 5 presents the recovered phase shift after adaptive filtering with *reference signal 1 and 2* between channels 3-2 & 3-1. As the phase shift is recovered, the slow wave propagation could be identified.

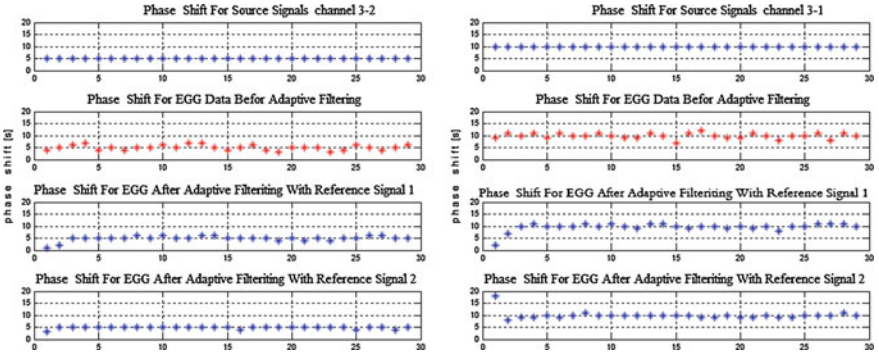


Fig. 5 The phase shift for: source signals; EGG data before adaptive filtering; EGG data after adaptive filtering with *reference signal 1*, regained by ICA and *reference signal 2*, obtained after PCA & ICA. On the *left*, the results for channels 3-2, on the *right*, for channel 3-1

3.2 *Simulation Results—Synthetic Data Second Type*

The second type of experimental data was created using a one channel real EGG signal. 10 min of one channel EGG signal with frequency 0.05 Hz (3 cpm), obtained from a healthy person was used, as the source signal S_1. In order to simulate the slow wave propagation, the S_2 signal of the same length was cut, from the same one channel EGG data, but shifted of 20 samples = 5 s respectively to signal S_1. In the same way, the third source signal S_3 was constructed, so the time lag, between the signal S_1 and S_3 was 10 s. As in the first type synthetic data, the S_4 source signal was a random noise, with the noise to signal ratio (the inverse of SNR), $NSR = std(S_4)/std(S_1)$ respectively 6 %, 17 % and 70 %. The presence of slow wave was concentrate in the first 3 channels and by means of mixing matrix $\mathcal{A}$ (the same as in the first type) the simulated 4 channel EGG (Fig. 6) was obtained for $NSR \in \{6\,\%, 17\,\%, 70\,\%\}$.

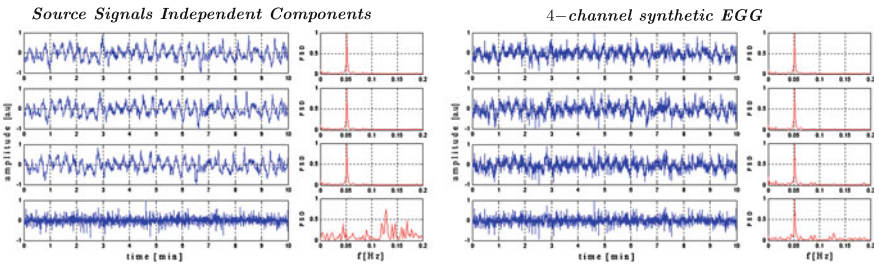


Fig. 6 On the *left*, there are the source signals S_1, S_2, S_3, S_4, on the *right*, simulated 4-channel EGG ($NSR = 70\,\%$) with their (PSD) function next to

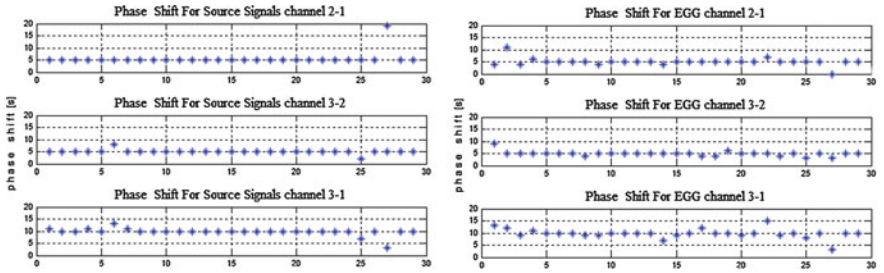


Fig. 7 The *left* figure, shows graphical representation of phase shift in the each 20 [s] of 10 [min] recording, between source signals: $S_2 - S_1$, $S_3 - S_2$ i.e. 5 s; $S_3 - S_1$ 10 s. The *right* figure, presents the time shift for EGG data

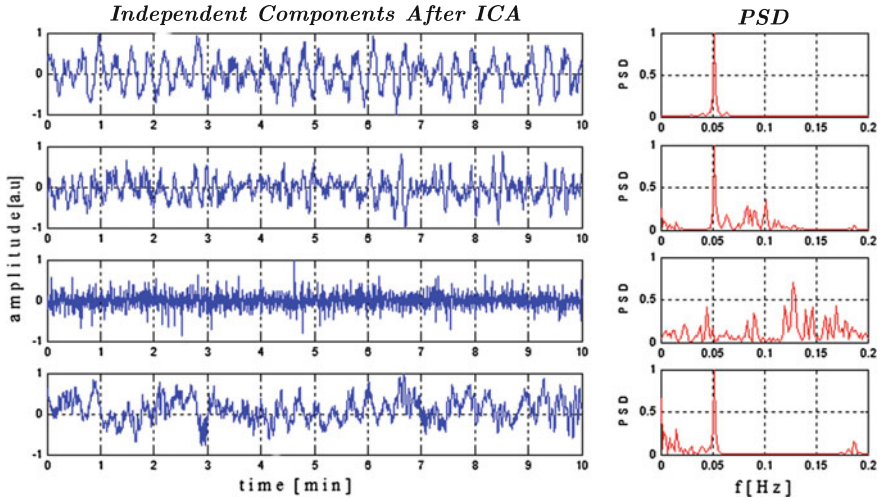


Fig. 8 Independent components: $ICA_1, ICA_2, ICA_3, ICA_4$ gained after applying ICA with their PSD function next to. ICA_1 served as *reference signal 1* for adaptive filtering

Figure 7, on the left, presents the graphical representation of phased shift for source signals: 5 s for $S_2 - S_1$, $S_3 - S_2$; 10 s for $S_3 - S_1$ and on the right, for created EGG data. The results gained for the synthetic data of second type were presented in the Figs. 8, 9 and 10.

4 Human EGG Data Analysis

To assess the slow wave propagation, in the human case, the analysis was performed, for 4 channel EGG data from healthy volunteers. Three phases of EGG registration has been taken under consideration, one before and two after, food stimulation. Each

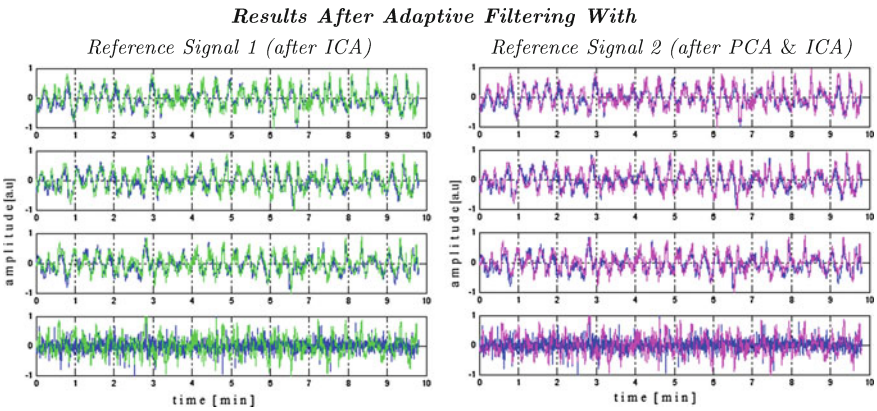


Fig. 9 The result of adaptive filtering, on the *left*, for *reference signal 1* = ICA_1, on the *right*, for *reference signal 2* gained after PCA & ICA application

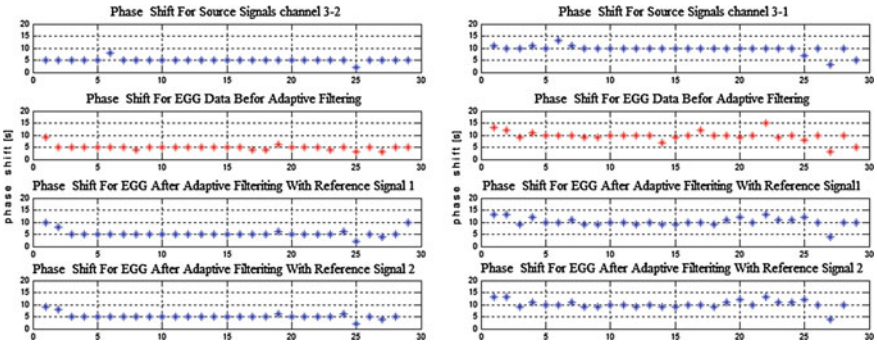


Fig. 10 The phase shift for source signals and EGG data, before and after adaptive filtering with *reference signal 1* obtained by ICA and *reference signal 2* gained after PCA & ICA, on the *left*, for channels 3-2, on the *right* for channels 3-1

phase, included 30 min registration of EGG signal with sampling frequency 4 Hz. Phase I, contained 30 min recording in a preprandial state, phase II, included 30 min registration of EGG signal directly after consuming 400 ml fruit yogurt containing 370 Kcal and phase III, concerned the following 30 min of registration, after 5 min break. EGG signal recording has been performed with standard four channel biosignal amplifier within the range of 0.9 ÷ 9 cpm (0.015 ÷ 0.15 Hz) with abdominal electrodes placement.

As it has been mentioned the EGG signal includes in addition to the basic normogastric rhythm 2–4 cpm (0.033–0.066 Hz), some pathological rhythms, covering the frequency range from 0.5 cpm up to 9 cpm (8–150 mHz): bradygastric rhythm [0.5,2) cpm ([10, 33) mHz), tachygastria rhythm (4, 9] cpm ([66, 150)mHz) and arrhythmia, without established range of frequency rhythm [12].

It is considered, [12] that for the healthy subjects the normogastric rhythm (2–4 cpm), covers up to 70 % of the whole periodicity of EGG signal. In order to, possess pure EGG components, for assessment of slow wave propagation, for each phase and each channel of EGG, the percentage of normogastric rhythm has been analyzed, before and after, adaptive filtering.

In the first step, of examination procedure, the application of ICA was applied, for 4-channel EGG signal to extract source signals (independent components). Afterwards, the integral of power spectral density function (PSD), in the limits for frequency of particular, earlier defined: brady, normo and tachygastric rhythm, was calculated for each obtained source signal. The independent component with the biggest normogastric rhythm contribution was used, as the reference signal for adaptive filtering. Therefore, adaptive filtering performed in the (DCT) transform domain was the basic preprocessing stage, improving quite complicated signal quality. Additionally the percentage of normogastric rhythm in the reference signal has been analyzed and the results before and after, adaptive filtering was compared.

For the phase I, two references signals were taken into consideration ICA_4 and $\overline{ICA_1}$. The second one was obtained after data dimension reduction by rejection of two the smallest eigenvalues of matrix $A = (EGG)^T(EGG)$ in the Principal Component Analysis (PCA) (Fig. 12). In the second case the quality of received EGG signal was better i.e. the amount of normogastric rhythm was higher. For such prepared 4-channel EGG signal the slow wave propagation between first and fourth channel was assessed.

4.1 Phase I—Before Food Stimulation

Tables 1, 2 and 3 contain the results of estimated level of normogastric rhythm in individual channels of EGG signal in the first phase i.e. before food stimulation. Figures 11, 12, 13 and 14 present obtained results.

Table 1 Values of area under the plot of PSD function for individual channel of 4-channel EGG (Fig. 11) before food stimulation

EGG	p_b	p_n	p_t	p_{b+n+t}	Brady	Normo	Tachy
1.	0.0006	0.0023	0.0005	0.0033	18.18 %	69.70 %	15.15 %
2.	0.0012	0.0021	0.0005	0.0038	31.58 %	55.26 %	13.16 %
3.	0.0016	0.0021	0.0005	0.0042	38.18 %	50.00 %	11.90 %
4.	0.0004	0.0014	0.0002	0.0020	20.00 %	70.00 %	10.00 %
$\overline{x}$	0.0010	0.0020	0.0004	0.0033	26.96 %	61.24 %	12.55 %
σ	0.0006	0.0004	0.0002	0.0010	9.50	10.17	2.16

p_b, p_n, p_t, denote respectively areas under the plot of PSD functions, for the range of frequency: brady, normo and tachygastric rhythm. **Brady**, **normo**, **tachy**, present percentage of brady, normo and tachygstric rhythm in each channel of EGG ($\overline{x}$-mean, σ-standard deviation)

Table 2 Values of area under the plot of PSD function for the range of frequency: brady, normo and tachygastric rhythm of reference signals ICA_4 and $\overline{ICA_1}$ (Fig. 12) obtained after rejection of two smallest eigenvalues of matrix $A = (EGG)^T(EGG)$ in the PCA analysis (Fig. 12)

ICA	p_b	p_n	p_t	p_{b+n+t}	Brady (%)	Normo (%)	Tachy (%)
ICA_4	0.0006	0.0023	0.0004	0.0034	17.65	67.65	11.76
$\overline{ICA_1}$	0.0004	0.0015	0.0002	0.0021	19.05	71.43	9.52

Table 3 Values of area under the plot of PSD function for individual channels (Fig. 13 on the right) of 4-channel EGG, after adaptive filtering (Fig. 13 on the left) with reference signal $\overline{ICA_1}$ (Fig. 12) obtained by ICA algorithm after rejection of two the smallest eigenvalues

EGG	p_b	p_n	p_t	p_{b+n+t}	Brady	Normo	Tachy
1.	0.0005	0.0016	0.0003	0.0024	20.83 %	66.67 %	12.50 %
2.	0.0004	0.0012	0.0004	0.0021	19.05 %	57.14 %	19.05 %
3.	0.0005	0.0017	0.0006	0.0028	17.86 %	60.71 %	21.43 %
4.	0.0004	0.0013	0.0002	0.0019	21.05 %	68.42 %	10.53 %
$\bar{x}$	0.0005	0.0015	0.0004	0.0023	19.70 %	63.24 %	15.88 %
σ	0.0001	0.0002	0.0002	0.0004	1.52	5.23	5.19

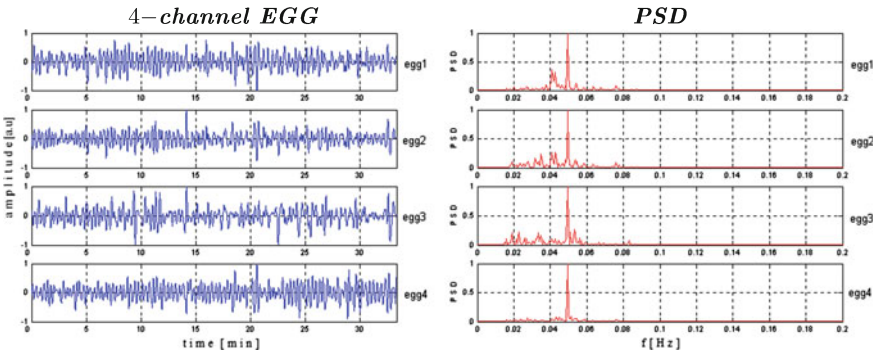


Fig. 11 On the *left* 4-channel EGG recording (phase before food stimulation). On the *right* plots of *power spectral density function* (PSD) for individual channels of 4-channel EGG signal (Table 1)

4.2 The Results for Two Postprandial Phases

The same procedure was applied for two next phases after food stimulation. Figure 15 presents obtained results.

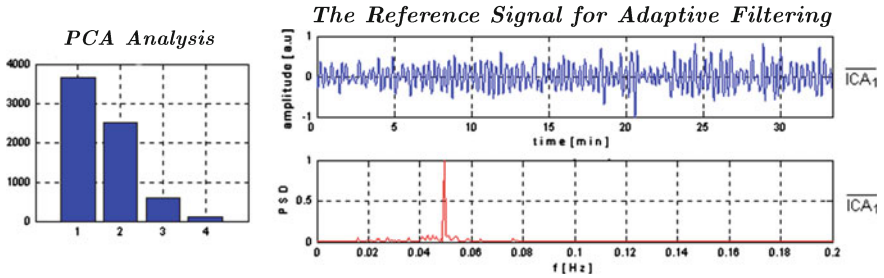


Fig. 12 On the *left*, the result of PCA analysis for EGG signal (from Fig. 11) presents the eigenvalues of matrix $\mathcal{A} = (EGG)(EGG)^T$ (phase before food stimulation). On the *right*, there is the reference signal $\overline{ICA_1}$ (Table 2) for adaptive filtering obtained by ICA algorithm after rejection of two smallest eigenvalues

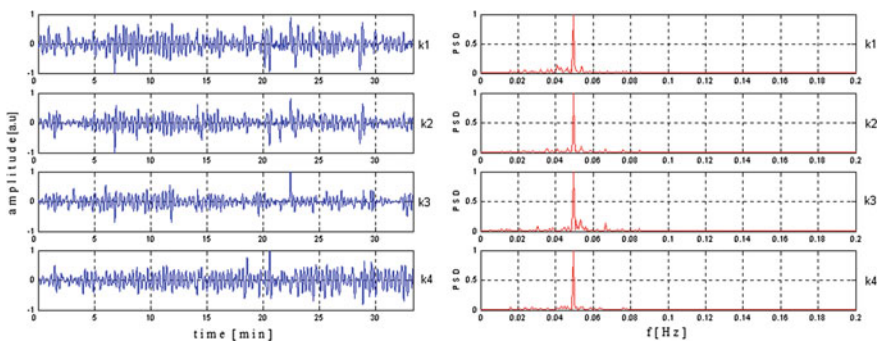


Fig. 13 The result of adaptive filtering with reference signal $\overline{ICA_1}$ (Fig. 12) for 4-channel EGG (Fig. 11) after rejection of two smallest eigenvalues of matrix $\mathcal{A} = (EGG)(EGG)^T$. On the *right*, plots of PSD function for individual channels of EGG data, regained after adaptive filtering with reference signal $\overline{ICA_1}$ (Fig. 12, Table 2)

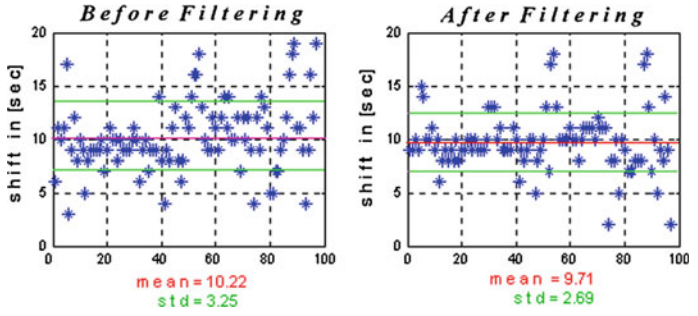


Fig. 14 Graphical representation of slow wave propagation for channel (1 & 4) of EGG data, before (Fig. 11) and after (Fig. 13), adaptive filtering with reference signal $\overline{ICA_1}$ (Fig. 12). Slow wave propagation time is approximately 10 s

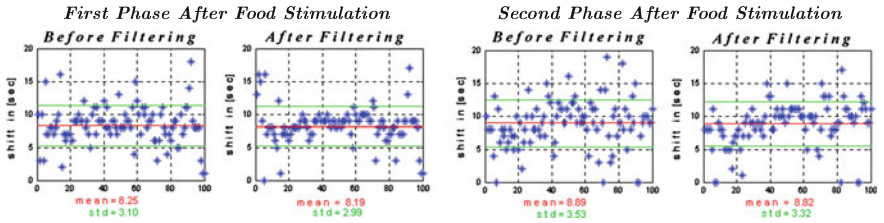


Fig. 15 Slow wave propagation, between channel (1&4) of EGG data, before and after adaptive filtering with reference signal obtained by ICA method, for the first and second phase after food stimulation. Slow wave propagation time is about 8 s, for the first phase and 9 s, for the final phase

4.3 Conclusions

Adaptive filtering with reference signal obtained by the ICA method applied for both simulated and real EGG data successfully reduced continuous noise in each channel of EGG. As ICA algorithm does not change the phase shift in the recovered signals the slow wave propagation could be examined using better quality of EGG signal. In this study the time for slow wave propagation (for healthy persons) between first and fourth channel was about 9.71 s in preprandial state and slightly decrease to 8.19 s in first postprandial state and increase 8.82 in the second postprandial state. The presented methods give opportunity to observe propagation of slow wave from antrum to pylorus present in the source signals, which could disappears in mixture raw EGG data. Some additional calculations and comparisons for real EGG multichannel recordings from diseased versus healthy person should be the subject of extensive study.

References

1. Alvarez, W.C.: The electrogastrogram and what it shows. J. Am. Med. Assoc. **78** (1922)
2. Chen, J.D.Z., Lin, Z.: Comparison of adaptive filtering in time-, transform-, and frequency-domain: an electrogastrograpihic study. Ann. Biomed. Eng. **22**, 423–431 (1994)
3. Chen, J.D.Z., McCallum, R.W.: Electrogastrogrpathy: measurement, analysis and prospective applications. Med. Biomed. Eng. Comput. **29**, 339–350 (1991)
4. Chen, J.D.Z., Zou, X., Lin, X., Ouyang, S., Liang, J.: Detection of gastric slow wave propagation from the cutaneous electrogastrogram. AJP Gastrointest. Liver Physiol. **277**, G424–G430 (1999)
5. Comon, P.: Independent Component Analysis, A new Concept? Elsevier, Special Issue on High-Order Statistics, vol. 36 (1994)
6. Hyvärinen, A.: Survey on independent component analysis. Neural Comput. Surv. **2**, 94–128 (1999)
7. Hyvärinen, A., Oja, E.: Independent component analysis: algorithms and applications. Neural Netw. **13**(4–5), 411–430 (2000)
8. Jurlewicz, T., Skoczylas, Z.,: Algebra Liniowa 2. (Linear algebra II) Oficyna Wydawnicza GiS, Wydanie 7, p. 94, Wrocław 2005 (In Polish)

9. Koch, K.L., Stern, R.M: Handbook of Electrogastrography. Oxford University Press (2004)
10. Liang, H.: Extraction of gastric slow waves from electrogastrograms: combining independent component analysis and adaptive signal enhancement. Med. Biomed. Eng. Comput. **43**, 245–251 (2005)
11. Liang, J., Chen, J.D.Z.: What can be measured from surface electrogastrography. Dig. Dis. Sci. **42**(7), 1331–1343 (1997)
12. Parkman, H.P., Hasler, W.L., Barnett, J.L., Eaker, E.Y: Electrogastrography: a document prepared by the gastric section of the American motility society clinical GI motility testing task force. Neurogastroenterol. Motil. **15**, 89–102 (2003)
13. Thakor, N.V., Zhu, Y.Z.: Application of adaptive filtering to ECG analysis: noise cancellation and arrhythmia detection. IEEE Trans. Biomed. Eng. BME **38**, 785–794 (1991)
14. Wang, Z.S., Cheung, J.Y., Chen, J.D.Z.: Blind separation of multichannel electrogastrograms using independent component analysis based on a neural network. Med. Biomed. Eng. Comput. **37**, 80–86 (1999)
15. Ward, S.M., Sanders, K.M.: Physiology and pathophysiology of interstitial cell of Cajal: from bench to bedside. Functional development and plasticity of interstitial cells of Cajal networks. AJP Gastrointest. Liver Physiol. **281**, G602–G611 (2001)
16. Widrow, B., Glover, J.R., McCool, J.M., Kaunitz, J., Williams, C.S., Hearn, R.H., Zeidler, J.R., Eugene Dong, J. Jr., Goodlin, R.C.: Adaptive noise canceling: principles and applications. In: Proceedings of the IEEE, vol. 63, pp. 1692–1716

Preprocessing Methods in Nuclear Magnetic Resonance Spectroscopy

Michal Staniszewski, Agnieszka Skorupa, Lukasz Boguszewicz, Maria Sokol and Andrzej Polanski

Abstract Magnetic resonance spectroscopy is currently used in chemistry and medicine as a diagnostic tool. Due to many imperfections that are present during measurement the signal has to be corrected by so called preprocessing methods or techniques. Some of them are performed by a scanner, but it is still necessary to improve the quality of the numerical signal. This paper presents a description of the most important preprocessing techniques which are applied by most current software and is an extension of the most currently reviews presented on this topic.

Keywords NMR · MRS · Nuclear magnetic resonance · Magnetic resonance spectroscopy · Preprocessing techniques

1 Introduction

Proton magnetic resonance spectroscopy (1H MRS) is a kind of analytic spectroscopy like 1H nuclear magnetic resonance (NMR), which provides information about chemical composition in vivo studies [4]. Due to the nature of measurement in a living organism, the technique of in vivo 1H MRS requires precise supporting programming, which is used at every stage of the analysis, in signal processing, and interpretation of results. In order to use the spectrum for the diagnosis, additional techniques have to be introduced to improve their quality, which are called preprocessing methods or techniques. The main preprocessing techniques may include filtering, phase shift correction, removal of unwanted components and correction of the signals with a dominant baseline.

M. Staniszewski (✉) · A. Polanski
Faculty of Automatic Control, Electronics and Computer Science, Institute of Informatics, Silesian University of Technology, Akademicka 16, Gliwice, Poland
e-mail: michal.staniszewski@polsl.pl

A. Skorupa · L. Boguszewicz · M. Sokol
Department of Medical Physics, Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Gliwice Branch, Wybrzeze Armii Krajowej st., Gliwice, Poland

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_28

The preprocessing techniques are partially applied on the side of a scanner, however the signal requires numerical quality enhancement. In terms of MRS, preprocessing can be performed either on the free induction decay (FID) side, in the time domain, or on the spectrum in the frequency domain. The methods of preprocessing were described accurately in many papers, books [15, 17], and were the subject of a few PhD theses [24, 27]. Particular preprocessing techniques and new methods were also presented in separate articles [7, 25, 26, 38]. In comparison to the works mentioned the authors of this work present a more comprehensive review which should give the reader information concerning what kind of preprocessing method should be applied depending on the reason of problem. Additionally, the methods for determination of chemical shift, relaxation correction and standardization and scaling, not included in the previous articles, are also enclosed. In the preprocessing stage the method of reiteration of HSVD was used [32, 33]. All resulting figures are presented in the form of absorption spectrum where the signal intensities are given in arbitrary units and the signal chemical shifts are expressed in ppm.

2 Preprocessing Techniques

2.1 Reading Data from Scanner

A common format for saving FID signals digitally has not been yet introduced and therefore the manufacturers of NMR scanners use their own file formats which usually vary. That is why the software being designed as a tool for analysis of NMR spectra has to deal with many different types of data. File reading always starts with the part containing the patient information and the parameters stored in the header. Then, the software has to read binary data representing the NMR signal in a complex form.

Philips Scanners In the Philips NMR scanners each MRS spectrum is stored in the form of two files: SPAR, in which spectrum acquisition parameters are stored as plain text and SDAT that contains a complex number FID signal in the binary form. Such a standard facilitates reading, because the text data which is to be processed and the binary data are read separately [23].

Siemens Scanners Standard data storage has been reduced to a single file with the extension .RDA. The file contains text with data acquisition parameters of the spectrum, as well as binary data, in which the FID signal is stored in the form of a complex number. The .RDA file will always start with the text section, which is marked as "Begin of header >>> <<<" at the beginning and "End of header >>> <<<", which is located at the end of the text section. The remaining part of the file is filled with binary data [13].

GE Scanners A scanner manufactured by GE Medical records the NMR spectrum in two files. File with the extension .7 stores the signal in a binary form, while the

file .7.shf contains text data, such as the parameters of the spectrum acquisition. Therefore, the reading procedure consists of two separate steps [12]. Along with the mentioned manufacturers there are several other companies, that produce NMR scanners. m. al. Toshiba Medical Systems [34], Bruker [2] and Varian [37].

2.2 Signal Filtration

A MR signal is affected by a random noise that results in artifacts, a low signal to noise ratio (SNR) coefficient, unwanted components and limited resolution. In order to minimize such undesirable effects the spectrum may be multiplied by a function of a precise shape, or its counterpart in the time domain may be filtered. This filtering process is also often called apodization or windowing. There are many filters and noise reduction methods available to be applied to the spectra. Filtration leads to a signal correction, but in most cases, one filter or method can be used depending on the degree of distortion. In such an approach there is also some risk that too much filtering might lead to the removal of the necessary information contained in the signal [9, 15, 17].

Apodization A measured NMR signal is loaded with a random function, which causes that every point of the FID to be distorted. It is assumed that the resulting NMR signal noise is a white Gaussian noise with a mean of zero, and a standard deviation of fixed value. Noise present in the FID is then transformed with a FT into the spectrum. In this case, the first method of improving signal quality chooses the appropriate acquisition time, or the total time of data sampling. The acquisition time is determined by multiplying the interval between the samples, called dwell time, and the number of samples N. The solution to the problem is to match an exact time of acquisition or to perform an appropriate cut-off of FID [15, 17].

Matching Filter The choice of filter depends mainly on the cause of the noise. The best fit for a filter is obtained by adopting a function where the attenuation coefficient is equal to the FID signal. This condition is met for a FID containing a single component. Unfortunately, due to NMR signal complexity, one factor may not be optimal for all components in the spectrum [15, 17].

Savitzky-Golay Filter One of the most popular filters used in digital analysis is a filter proposed in 1964 by Savitzky and Golay [28]. Its main advantage is the possibility of using the digital data, and smoothing the signals in such way that SNR increases. Filtration is based on local least-squares polynomial approximation, and involves fitting a polynomial to a set of input samples and then evaluating the resulting polynomial at a single point within the approximation interval. Such an approach is equivalent to discrete convolution with a fixed impulse response. The resulting signal is smoothed. The authors published the convolution coefficients for different subsets of polynomials and sizes [28]. The effects of the filter to the absorption spectrum are shown in Fig. 1.

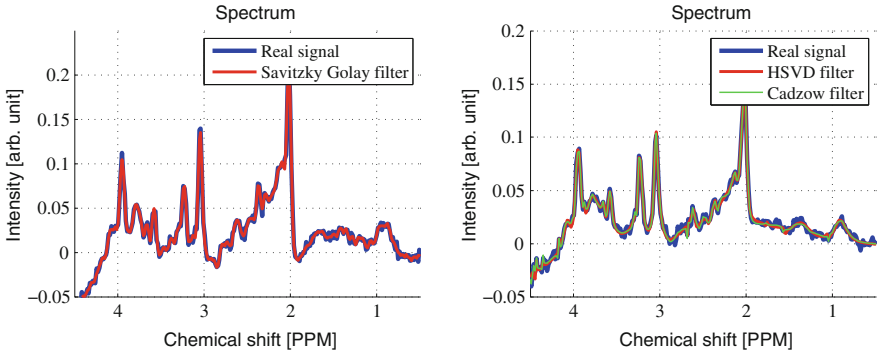


Fig. 1 Application of Savitzky Golay filter on the *left side*, and two other filters: HSVD and Cadzow on real spectrum, shown on *right side*

HSVD and Cadzow Filtering In these filtering techniques the HSVD method (or its modified version) may be used as introduced in [32, 33]. The main advantage of SVD is that the singular values of the diagonal matrix are arranged in descending order.

This means that the first singular values contain all the values corresponding to particular components and metabolites, and at the end the signal noise can be found. This means that SVD decomposition should first determine the correct number of FID components, and then remove the remaining values from the corresponding columns and rows to give a signal. It should be noticed that this method is not precise and is very sensitive to mistakes in the selection of a proper number of components—in case of a mistake a relevant part of the signal will be deleted. SVD decomposition was developed—basing upon Cadzow's work published in 1988—into a new method that removes noise from the signal [35].

2.3 Zero Filling

In the process of data acquisition, the FID signal is sampled by the converter and the number of samples is fixed. Hence if the score is not sufficient to reflect the signal, then further analysis might be unsuccessful. Before performing a Fourier transform, the MRS signal may be improved by adding values of zero, called zero filling, at the end of the FID [15]. Due to limitations of the fast Fourier transform, zero-filling must meet one basic assumption: after adding zeros to the FID signal a number of points must correspond to a power of 2. However, additional zeros in the signal only results in a visual improvement in spectral resolution. Thus, supplementing zeros only affects the improvement of resolution of the digital result, but does not add any new information to the spectrum.

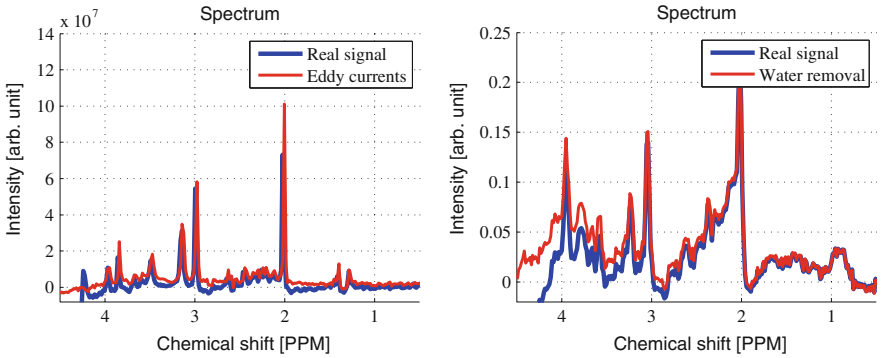


Fig. 2 Eddy currents correction (*left*) and influence of removal of water on real part of spectrum (*right*)

2.4 Eddy Currents Correction

Another artifact is caused by eddy currents, which are generated during the measurement process due to rapid and frequent changes of magnetic fields. The eddy currents introduce additional time-dependent magnetic fields, which influence the magnetization vector. For some scanners, a hardware correction for the eddy currents is implemented which improves signal quality, but in most cases software techniques are necessary to eliminate this problem on the application side. One of the best known approaches for such correction is to use a reference FID signal before (stated as ref) and after water suppression (stated as act). It is assumed that the FID signal acquired without water suppression is vitiated by eddy current effects. The correction involves taking the amplitude of FID (A^{act}), and subtracting from the phase of the signal acquired with water suppression (ϕ_{act}), the phase of the signal without water suppression (ϕ_{ref}) according to (1). Eddy currents mainly influence the phase shift and consequently the processing of MRS spectra [15]. The method allowing for the correction of eddy currents was proposed by Uwe Klose in 1990, Fig. 2 [19].

$$FID^{cor} = A^{act} e^{i(\phi_{act} - \phi_{ref})} \tag{1}$$

2.5 Removal of Unwanted Signals in FID

Removing unwanted components of the FID signal is sometimes necessary to improve data quality. It also leads to reduction of signal complexity, and consequently, improves the accuracy of signal parameters. A good example of such component that affects the spectral quality is water signal. The average content of water in the adult human body is about 60 %. Water is present in virtually every human tissue

and its concentration is much higher than that of other metabolites. For this reason, water suppression is an essential part of each NMR metabolic measurement. The data acquired with and without water suppression are stored in separate files or separate places in one file. However, due to scanner imperfection and the specific nature of water, even the spectra obtained with water suppression additionally require removal of the water signal. One way to remove a residual water signal is by constructing a software water filter, as low or high pass, using the resonance frequency of water protons [15].

Another approach is to model a residual water signal, using available methods, and to remove it from the NMR data. This can be done using the described HSVD method [3, 32, 33, 36]. First, the signal is modeled and the matrix product resulting from the decomposition SVD is obtained. Water, even in the residual signal, will be the dominant part—its component is first and its removal will consist of the rejection of the main singular values of the matrix from the beginning of the SVD. Alternatively, all components of the FID and their parameters are determined, and then the components existing in the expected range of the water signal are removed. A similar technique can be used to remove any metabolite signal available in the spectrum—this approach enables further quantitative analysis. The impact of the residual water signal removal from the absorption spectrum, and calibrated for an artifact caused by this removal, is shown in Fig. 2.

2.6 Phase Correction

In an ideal, theoretical case, acquisition of the FID signal is performed immediately upon removal of the external magnetic field, at a time $t = 0$. In a real situation there is a slight interval between switching off the external fields and the starting point of signal detection. This means that the starting point of the signal, described by a function of the cosine, of the real part of the spectrum MRS, does not coincide with the time $t = 0$—this shift introduces a phase shift. Zero order phase correction is independent of the frequency and is carried out by calculating a linear combination of the real- and imaginary parts of the FID. The main problem in this case is to find a single angle that will improve the spectrum. Correction operations can be performed manually by an operator, or automatically, which requires appropriate algorithms to be implemented. Figure 3 shows how the phase shift influences the absorption spectrum for different values of the angle [17, 41].

Manual Correction Manual correction requires direct operator intervention. In this case, the user manually adjusts phase shift basing upon his or her experience. Spectral phasing may be performed in the time domain or frequency domain; however, it is easier in the frequency domain. In order to change the phase shift in the FID signal, which is described by component RE^{FID} real and an imaginary IM^{FID}, at an angle that is ϕ_{cor}, one can use the Euler formulas and corrected signal written in the form (2).

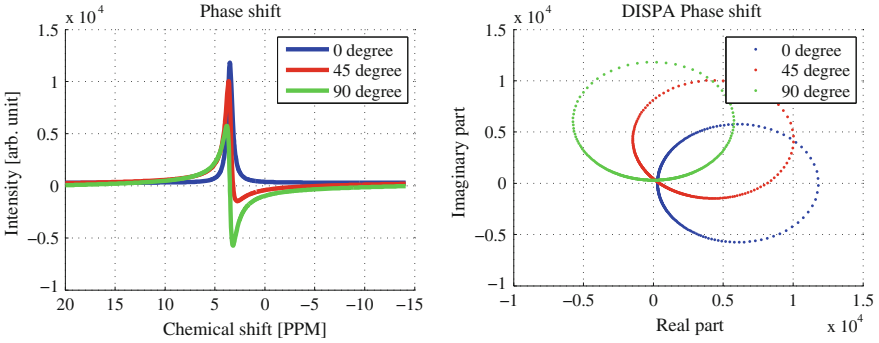


Fig. 3 Influence of phase shift on spectrum (*left*) and DISPA (*right*)

$$\begin{aligned} RE^{cor} &= \cos(\phi_{cor})RE^{FID} - \sin(\phi_{cor})IM^{FID} \\ IM^{cor} &= \sin(\phi_{cor})RE^{FID} + \cos(\phi_{cor})IM^{FID} \end{aligned} \tag{2}$$

Automatic Correction The earlier methodology is relatively simple, because the same tool only requires operator intervention and relies on their experience. When analyzing several spectra such an approach is acceptable, but in the case of analyzing dozens of spectra, human intervention will be time consuming. A possible solution is to apply automatic correction algorithms that estimate the phase error by using different optimization or minimization methods [1]. One of the first solutions consisted of an analytical dispersion versus absorption (DISPA) method for automatic phasing of the Fourier transform [14]. The proposed method uses a graph of relationships between parts of the dispersion, and the absorption of each point of the signal, in order to analyze the shape and phase of the spectrum lines (Fig. 3).

The classic DISPA method perfectly solves the problem of phase error for discrete spectral lines, however, in the case of in vivo spectroscopic data the lines are broad and overlap, which makes phasing more difficult. The modified method is based on the eDISPA algorithm [6]. This method creates the DISPA graphs for several signals in the spectrum and based on these graphs a two component function can be derived for the correction phase, which determines the maximum correct value of the phase adjustment. Another algorithm that solves the problem of automatic phase correction is associated with the concept of entropy [5].

2.7 Baseline Correction

The spectrum consists of the sharp peaks superimposed upon broad and slowly varying background called the baseline. The baseline often stems from various sources, such as the sample itself or from experimental or operator reasons. In the case of in vivo NMR, one of the main sources of distortion is short relaxation times of

the spins from macromolecules and lipids—much shorter than the relaxation times of low molecular weight metabolites. The signals due to macromolecules and lipids decay rapidly in the time domain and contribute signals with broad features in the frequency domain. Baseline distortions may also be caused by the residual water signal. Another important cause for baseline distortions in 1H NMR spectra is corruption of the first few data points in FID. These corrupted data points add low frequency modulations in the Fourier-transformed spectrum, and thus form the distorted baseline. The distorted baseline is a serious problem and needs to be corrected. The correction procedure involves background subtraction from the recorded spectrum, thus making the baseline flat [15].

Several correction methods exist that include removal of a portion of the broad FID components. An example of such an approach is modeling the signal using the HSVD method [32, 33]. In this case, the selected component with the highest value for the attenuation coefficient is removed from FID. Other methods use a suitable filter or simply remove the first few FID points, which could potentially be responsible for the baseline distortion. A common approach to baseline correction is to recognize several spectral points as potential baseline points, and then to perform a numerical combination of these points. The most common modification of this method is windowing. Figure 4 shows the effect of baseline correction to the absorption spectrum. There are also other baseline correction methods, which use different mathematical algorithms [21]. These may include the asymmetric least squares method (ALS), which uses a least squares algorithm, and adds a restriction to the second derivative used to smooth the line [8]. Modification of the windowing analysis may be a local median method [10], which is based on medians local search within a specified window and smoothing points using Gauss weights. Alternatively, an iterative adjustment of the polynomial [11] can be used.

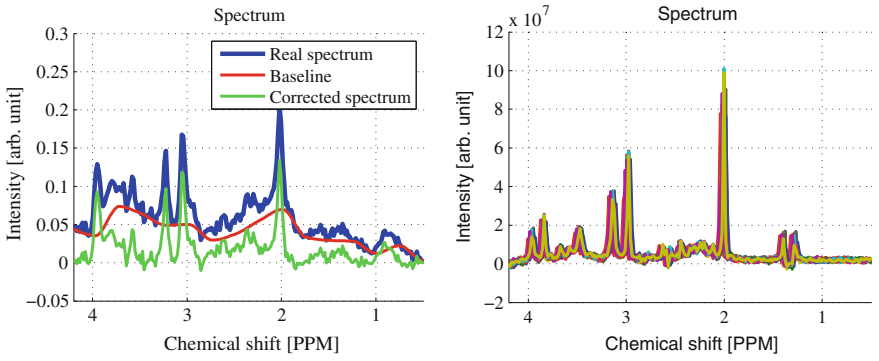


Fig. 4 Influence of baseline correction visible on real spectrum (*left*). 1H MRS spectra acquired in the same scanner from the phantom solution—27 plots are overlapped without any corrections in chemical shift showing some visible displacement (*right*)

2.8 Determination of Chemical Shift

In nuclear magnetic resonance spectroscopy, the chemical shift is the resonant frequency of a nucleus relative to a standard in a magnetic field. In proton spectroscopy the chemical shift is dictated by neighboring protons within the molecule. Therefore, metabolites can be characterized by their unique set of 1H chemical shifts. In MRS in vivo spectroscopy for the creatine signal—a singlet often plays the role of an internal standard. SpelAlign [39] matches peaks by moving portions of the spectrum. This is done by adding or removing points in the spectrum. Other algorithms for matching peaks in the NMR spectra, such as fuzzy warping [42], Bayesian alignment [18] and the FFT peak alignment [40] apply a predetermined pattern, which is translated into a set of spectra.

There has also been the development of the Progressive Consensus Alignment of NMR Spectra (PCANS) algorithm [31]. Through the progressive integration of many pairwise comparisons, this approach generates a single consensus spectrum as an output that is then used to adjust the chemical shift positions of the peaks from the original input spectra to their final aligned positions. This solution even allows for analysis of noisy spectra with peaks that vary in height and width. This approach starts from identification of the metabolite peaks. The result is a set of peak profiles that are described by peak positions, heights and widths. Figure 4 shows 27 absorption spectra acquired from the same phantom metabolite solution. Spectral referencing is absolutely necessary because inaccurate chemical shift may cause a mistake in metabolite identification.

2.9 Relaxation Correction

Relaxation affects the integral intensities of the NMR peaks if the spectra are not acquired under fully relaxed conditions. All signals of interest must have relaxed completely before each pulse, and if this condition is not met, the next pulse produces signals of lower intensities—thus their relative areas will be incorrect. In the case of 1H MRS in vivo, fulfilling the fully relaxed conditions is practically impossible. First, the T1 and T2 times of the metabolites are usually not known. Second, even if T1 and T2 are known, employing pulse sequence parameters—such as echo time TE and repetition time TR—which are appropriate for obtaining unsuppressed spectra, would markedly increase the time of medical examination. That is why there is a need to correct the spectral lines taking into account longitudinal and transverse relaxation effects [16, 22]. Currently used methods of relaxation correction of 1H MRS in vivo spectra are based on determination of the correction factor for relatively long, 135 ms, and short, 10 ms, TE. For shorter TEs the spectra are dominated by rapidly relaxing compounds, often incorporated into the baseline, whereas in the case of longer TEs the baseline is flat; however the calculation of the absolute concentrations metabolites is difficult [29]. In clinical use both types of sequences are often combined, mainly for spectral identification purposes. Calculation of metabolite concentrations is possible

with the employment of short TE and long TR values—however it is always a challenge. The best results of absolute concentrations are obtained for sequences with a short TE and the longest TR possible [20].

2.10 Methods for Standardization and Scaling

In order to perform the quantitative analyses of the spectra, the scaling and normalization procedures are required first. The scaling method is based on sensitivity correction factors for the transmitter and the receiver [30]. This is called an absolute quantification standard technique or simply a phantom method. Sensitivity factors are derived from measurements and saved as individual parameters for a particular spectrum. The (f_{cor}) factor is given in the formula:

$$f_{cor} = 10^{0.005(TG-65)} 2^{(6-\frac{R1}{2})+(30-R2)} f_{calib} \frac{1}{V}, \tag{3}$$

where TG is the sensitivity of the transmitter, R1 and R2 are the sensitivities of the digital and analog receivers, respectively, (f_{calib}) is the calibration coefficient, and V is a voxel volume. The other method normalizes the spectra using the water signal—calculated from the unsuppressed spectra. With knowledge of the water molar concentration in the phantom, the relative concentrations of different metabolites may be obtained [30].

3 Conclusions

This paper focused on preprocessing NMR data, steps in analysis and showing the possibilities of methods based on HSVD decomposition. These methods can be successfully used to remove unwanted signal components for NMR signal filtering or in performing phase correction. Processing workflow for 1H MRS in vivo spectra was presented with descriptions of use with various methods. Examples are provided, along with a literature review and comparisons to other works. This paper is extended to include knowledge about reading data, as well as including techniques that were not present in previous reviews, such as methods for determination of chemical shift, relaxation correction, and standardization and scaling. The purpose of our work is to stress the need for the use of appropriate correction procedures in preprocessing of NMR signals. The authors plan to uniform the presented methods and test them on real spectral data.

Acknowledgments This work has been supported by: projects for Young Scientist on Institute of Informatics BKM515/2014/9, BKM515/2015/9 (MS) and partly by infrastructure of POIG.02.03.01- 24-099/13 grant: GCONiI—Upper-Silesian Center for Scientific Computation (AP).

References

1. Binczyk, F., Tarnawski, R., Polanska, J.: Strategies for optimizing the phase correction algorithms in nuclear magnetic resonance spectroscopy. Biomed. Eng. Online **14**(Suppl 2), S5 (2015)
2. Bruker. http://www.bruker.com/products/mr/nmr.html
3. Cabanes, E., Confort-Gouny, S., Le Fur, Y., Simond, G., Cozzone, P. J.: Optimization of residual water signal removal by HLSVD on simulated short echo time proton MR spectra of the human brain. J. Magn. Reson. **150**(2), 116–125 (2001). San Diego Calif 1997
4. Cavanagh, J., Fairbrother, W.J., Skelton, A.G., Palmer III, N.J., Rance, M.: Protein NMR Spectroscopy: Principles and Practice. Academic Press (2010)
5. Chen, L., Weng, Z., Goh, L., Garland, M.: An efficient algorithm for automatic phase correction of NMR spectra based on entropy minimization. J. Magn. Reson. **158**(1–2), 164–168 (2002)
6. Cobas, C., Sykora, S.: Extended DISPersion-Absorption (eDISPA) Approach to Automatic Phasing of HR-NMR Spectra. Mestrelab Research
7. De Meyer, T., Sinnaeve, D., Van Gasse, B., Rietzschel, E.-R., De Buyzere, M.L., Langlois, M.R., Bekaert, S., Martins, J.C., Van Criekinge, W.: Evaluation of standard and advanced preprocessing methods for the univariate analysis of blood serum 1H-NMR spectra. Anal. Bioanal. Chem. **398**(4), 1781–1790 (2010)
8. Eilers, P.H.C.: Parametric time warping. Anal. Chem. **76**(2), 404–411 (2004)
9. Ernst, R.R.: Sensitivity enhancement in magnetic resonance. I. Analysis of the method of time averaging. Rev. Sci. Instrum. **36**(12), 1689–1695 (1965)
10. Friedrichs, M.S.: A model-free algorithm for the removal of baseline artifacts. J. Biomol. NMR **5**(2), 147–153 (1995)
11. Gan, F., Ruan, G., Mo, J.: Baseline correction by improved iterative polynomial fitting with automatic threshold. Chemom. Intell. Lab. Syst. **82**(1–2), 59–65 (2006)
12. GE Healthcare Worldwide. http://www.gehealthcare.com
13. Global Siemens Healthcare. http://www.medical.siemens.com
14. Herring, F.G., Phillips, P.S.: Automatic phase correction in magnetic resonance using DISPA. J. Magn. Reson. (1969) **59**(3), 489–496 (1984)
15. Jiru, F.: Introduction to post-processing techniques. Eur. J. Radiol. **67**(2), 202–217 (2008)
16. Jiru, F., et al.: The role of relaxation time corrections for the evaluation of long and short echo time 1H MR spectra of the hippocampus by NUMARIS and LCModel techniques. Magma N.Y. N **16**(3), 135–143 (2003)
17. Keeler, J.: Understanding NMR Spectroscopy, 2nd edn. Wiley-Blackwell, Chichester (2010)
18. Kim, S.B., Wang, Z., Hiremath, B.: A Bayesian approach for the alignment of high-resolution NMR spectra. Ann. Oper. Res. **174**(1), 19–32 (2008)
19. Klose, U.: In vivo proton spectroscopy in presence of eddy currents. Magn. Reson. Med. **14**(1), 26–30 (1990)
20. Kreis, R.: Quantitative localized 1H MR spectroscopy for clinical use. Prog. Nucl. Magn. Reson. Spectrosc. **31**(2–3), 155–195 (1997)
21. Liland, K.H., Almoy, T., Mevik, B.-H.: Optimal choice of baseline correction for multivariate calibration of spectra. Appl. Spectrosc. **64**(9), 1007–1016 (2010)
22. Mitchell, J., Chandrasekera, T.C., Gladden, L.F.: Obtaining true transverse relaxation time distributions in high-field NMR measurements of saturated porous media: Removing the influence of internal gradients. J. Chem. Phys. **132**(24), 244705 (2010)
23. Philips Healthcare. http://www.medical.philips.com
24. Poullet, J.-B.: Quantification and Classification of Magnetic Resonance Spectroscopic Data For Brain Tumor Diagnosis. Katholic University of Leuven (2008)
25. Poullet, J.-B., Sima, D.M., Van Huffel, S.: MRS signal quantitation: a review of time- and frequency-domain methods. J. Magn. Reson. **195**(2), 134–144 (2008)
26. Reynolds, G., Wilson, M., Peet, A., Arvanitis, T.N.: An algorithm for the automated quantitation of metabolites in in vitro NMR signals. Magn. Reson. Med. Off. J. Soc. Magn. Reson. Med. Soc. Magn. Reson. Med. **56**(6), 1211–1219 (2006)

27. Sava, A.R.C.: Signal Processing and Classification for Magnetic Resonance Spectroscopic Data with Clinical Applications. Katholic University of Leuven (2011)
28. Savitzky, A., Golay, M.J.E.: Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. **36**(8), 1627–1639 (1964)
29. Seeger, U., Mader, I., Nagele, T., Grodd, W., Lutz, O., Klose, U.: Reliable detection of macromolecules in single-volume 1H NMR spectra of the human brain. Magn. Reson. Med. Off. J. Soc. Magn. Reson. Med. Soc. Magn. Reson. Med. **45**(6), 948–954 (2001)
30. Skorupa, A., Wicher, M., Banasik, T., Jamroz, E., Paprocka, J., Kieltyka, A., Sokol, M., Konopka, M.: Four-and-one-half years experience in monitoring of reproducibility of an MR spectroscopy system-application of in vitro results to interpretation of in vivo data. J. Appl. Clin. Med. Phys. Am. Coll. Med. Phys. **15**(3), 4754 (2014)
31. Staab, J.M., O'Connell, T.M., Gomez, S.M.: Enhancing metabolomic data analysis with progressive consensus alignment of NMR spectra (PCANS). BMC Bioinform. **11**(1), 123 (2010)
32. Staniszewski, M., Binczyk, F., Skorupa, A., Boguszewicz, L., Sokol, M., Polanska, J., Polanski, A.: Comparison of black box implementations of two algorithms of processing of NMR spectra. In: Gaussian Mixture Model and Singular Value Decomposition. BIOSTEC, Lisbon (2015)
33. Staniszewski, M., Skorupa, A., Boguszewicz, L., Sokol, M., Polanski, A.: Reiteration of Hankel singular value decomposition for modeling of complex-valued signal. In: ICNAAM, Rhodes (2015)
34. Toshiba Medical System. http://www.toshiba-europe.com/medical/
35. Trickett, S.: F-xy Cadzow noise suppression. SEG Tech. Program Expanded Abs. **27**(1) (2008)
36. Vanhamme, L., Fierro, R.D., Van Huffel, S., de Beer, R.: Fast removal of residual water in proton spectra. J. Magn. Reson. **132**(2), 197–203 (1998). San Diego Calif 1997
37. Varian. http://www.agilent.com/home
38. Wilson, M., Reynolds, G., Kauppinen, R.A., Arvanitis, T.N., Peet, A.C.: A constrained least-squares approach to the automated quantitation of in vivo 1H magnetic resonance spectroscopy data. Magn. Reson. Med. Off. J. Soc. Magn. Reson. Med. Soc. Magn. Reson. Med. **65**(1), 1–12 (2011)
39. Wong, J.W.H., Cagney, G., Cartwright, H.M.: SpecAlign-processing and alignment of mass spectra datasets. Bioinform. Oxf. Engl. **21**(9), 2088–2090 (2005)
40. Wong, J.W.H., Durante, C., Cartwright, H.M.: Application of fast Fourier transform cross-correlation for the alignment of large chromatographic and spectral datasets. Anal. Chem. **77**(17), 5655–5661 (2005)
41. Worley, B., Powers, R.: Simultaneous phase and scatter correction for NMR datasets. Chemom. Intell. Lab. Syst. **131**, 1–6 (2014)
42. Wu, W., Daszykowski, M., Walczak, B., Sweatman, B.C., Connor, S.C., Haselden, J.N., Crowther, D.J., Gill, R.W., Lutz, M.W.: Peak alignment of urine NMR spectra using fuzzy warping. J. Chem. Inf. Model. **46**(2), 863–875 (2006)

Electrical Activity of Uterus as Reliable Information on Contractions During Pregnancy and Labour

Krzysztof Horoba, Janusz Jezewski, Tomasz Kupka, Adam Matonia, Robert Czabanski and Dawid Roj

Abstract In this paper an evaluation of analysis of uterine electrical signals as an alternative method to tocography for contractile activity monitoring is presented. A set of dedicated indices was defined to estimate the inconsistency of the number, location and other descriptive parameters of the corresponding contractions detected in simultaneously recorded mechanical and electrical activity signals. Research material comprised 57 recordings from three groups of pregnant women being: in the first uncomplicated pregnancy, with symptoms of the threatening preterm labour, and during the first period of the physiological labour. The highest consistency as for the number and location of contractions was noted for recordings acquired during labour. Obtained results show synchronization between the mechanical and electrical activity, which varies in different stages of pregnancy and labour, and which is stronger when the birth term approaches.

Keywords Uterine contractile activity · Electrohysterography · Tocography

1 Introduction

Monitoring of uterine contractile activity is an integral part of fetal monitoring which is a routine procedure performed during pregnancy and labour. It is aimed at assessment of fetal state which is accomplished through analysis of the fetal heart rate variability on the background of information on the contractile uterine activity [4, 15, 26]. Additionally, contractile activity monitoring enables to supervise a labour progress [3, 8, 25]. External tocography is the most common method being applied

K. Horoba (✉) · J. Jezewski · T. Kupka · A. Matonia · D. Roj
Institute of Medical Technology and Equipment ITAM, 118 Roosevelt Str.,
41-800 Zabrze, Poland
e-mail: krzysztof.horoba@itam.zabrze.pl

R. Czabanski
Faculty of Automatic Control, Electronics and Computer Science,
Institute of Electronics, Silesian University of Technology,
16 Akademicka Str., Gliwice, Poland

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_29

in fetal monitors [11]. Strain-gauge transducer with elastic strap is attached to maternal abdomen to measure a strength exerted by the uterine muscle through abdominal wall. Recorded signal is provided in a form of tocogram (Toco), which is analysed visually or automatically in computer-aided fetal monitoring systems in order to detect contractions and determine their amplitude and timing parameters [5]. Information on uterine contractile activity can be obtained also by recording the electrical activity of uterine muscle using electrodes attached to maternal abdomen [17, 29]. This approach called the electrohysterography (EHG) has been investigated to evaluate its usability in clinical practice [7, 23]. Additionally, a growing interest in pregnancy monitoring at home has led to a need for simple and user friendly measurement instrumentation [27]. Such instrumentation could be based on recording only the bioelectrical signals from abdominal electrodes: the fetal electrocardiogram to obtain information on the fetal heart rate, and electrohysterogram to provide information on the uterine activity [28]. However, as far as the tocography is the method widely used in clinical practice for many years, the electrohysterography needs to be evaluated as an alternative tool. In this paper we presented an original comparative methodology to accomplish this evaluation by defining a set of indices to compare the common time-domain properties of contractions detected in both mechanical and electrical signals. Original criteria to assume the two contractions—detected in tocogram and electrohysterogram—as the corresponding contractions were established. We also considered additional information on electrical excitation propagating through the uterine muscle that could be obtained exclusively from the electrical uterine activity signal.

2 Methodology

Simultaneous acquisition of mechanical (Toco) and electrical (EHG) uterine activity was carried out with a help of two fetal monitoring systems: MONAKO and KOMPOREL. The first one is a computer-aided system for fetal monitoring with automated analysis of fetal heart rate and uterine contractions. These signals are provided by a fetal monitor (Philips FM20), which is connected through an interface unit. The prime task of the second system is to record the fetal electrocardiogram from maternal abdomen, that enables determination of the fetal heart rate and analysis of fetal QRS morphology [13, 14]. However, the same measurement technique lets us to use this unit to record the EHG signal as well.

Each measurement channel of the bioelectrical recorder has the following settings: bandwidth –0.05 to 150 Hz, gain –2500 V/V, and the sampling frequency –500 Hz. During monitoring session, the Ag/AgCl electrodes were attached to the skin in the vertical median axis of the abdomen as it was shown in Fig. 1 distance between the electrodes constituting the differential channels was set at 5 cm. Three channels were used to record abdominal signals that enabled a selection of a good quality electrohysterogram. The first EHG1 signal was calculated as V1–V2, the second signal EHG2 as V2–V3, and the third EHG3 as V4–V0. The best EHG signal

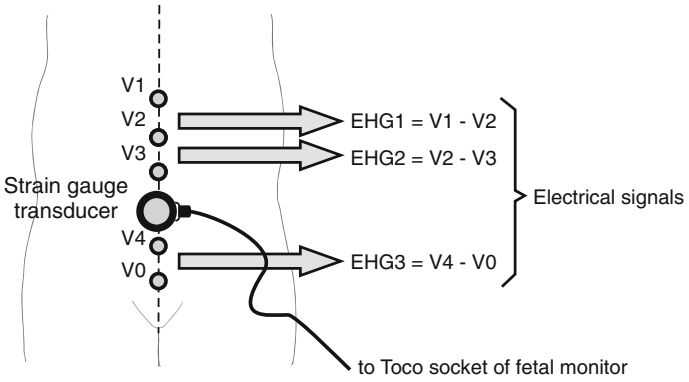


Fig. 1 Placement of electrodes and strain gauge transducer for simultaneous monitoring of the electrical and mechanical activity of uterine muscle through abdominal wall

was selected visually considering evident uterine activity to be compared with the mechanical activity signal. Since the frequency range of the acquired abdominal signals was limited by low-pass filtering to EHG frequency band equal to 5 Hz, the signals were downsampled to 20 Hz before further analysis.

The research material was collected in the Obstetrical Clinic of Silesian Medical University in Katowice, during routine monitoring sessions. It comprised the recordings (EHG and Toco signals simultaneously acquired) of total duration of 2575 min (average 45 min) acquired from 57 pregnant women. The patients were divided into three groups: Group I with 26 patients being in the first uncomplicated pregnancy, Group II with 18 women with symptoms of the threatening preterm labour, and Group III with 13 patients during the first period of physiological labour.

2.1 Contractions Detection

Due to a different nature of EHG and Toco signals the different methods have to be applied to obtain the contraction curve from these signals (Fig. 2a) [10]. Envelope of the EHG signal, which follows the change of electrical excitation amplitude during contractions (Fig. 2), is constituted from RMS values calculated in window of 1-min width and 3-s step. In turn, the Toco signal undergoes only low-pass filtering to remove interference coming from mother's body and breathing movements. Contractile episode is represented as an increase of the contraction curve above the so called basal tone. In the mechanical signal the basal tone represents some resting strength exerted by the uterine muscle on the strain gauge transducer when contractions do not occur, while in the electrical signal it refers to the resting bioelectrical activity of the uterus [16, 24]. Basal tone is constituted from the modal values of histogram obtained from the signal samples in each consecutive window of 4-min width and 1-min step.

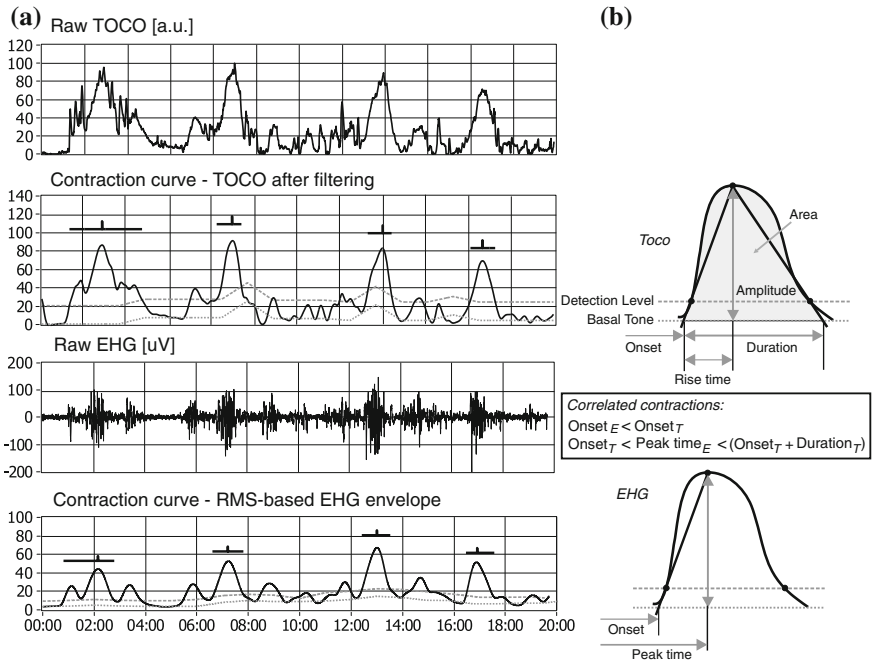


Fig. 2 Mechanical and electrical uterine activity signal with four contractions detected (**a**). Amplitude and timing contraction details together with criteria defining the corresponding episodes (**b**): the EHG onset before the Toco peak and the EHG peak within the contraction duration in Toco

Detection level is added to the basal tone that enables to recognize the candidate episodes whenever the contraction curve crosses the detection level. For the Toco signal, whose amplitude ranges from 0 to 100 arbitrary units, the detection level is set at 10 units above the basal tone. In case of EHG, whose amplitude range depends on the individual patient as well as measurement conditions, it is set dynamically by adding to the basal tone the value equal to 0.25 of the EHG envelope range within the consecutive 4-min window during the basal tone determination. Detected candidate episode is classified as contraction if it lasts longer then minimum duration and its amplitude exceeds the established minimum value (Fig. 2b). The minimum duration was established at 30 s, whereas the minimum amplitude, just like the threshold level, depends on the signal type. For the Toco signal the minimum amplitude has been set at 20 units above the basal tone, while for EHG it has been set as double distance between the basal tone and the detection level for each window being applied.

2.2 Inconsistency Estimation

In order to compare the contractions detected as a result of the analysis of both Toco and EHG signals, a set of indices to estimate the inconsistency of the number, location and the parameters describing these episodes were defined. Estimation of the

contractions inconsistency requires establishing the conditions to assume the two contractions—detected in mechanical and electrical signals—as the corresponding contractions. The conditions were determined basing on a phenomena being the background for electrical and mechanical uterine activity as well as on relation between both signals [2, 6]. Electrical excitation of myometrium cells is a source, while a mechanical contraction is a result, which starts after the cell reaches the depolarization phase [9, 18]. Hence, it should be expected that electrical activity precedes the mechanical contraction and the maximum of the electrical activity should occur in the rising phase of the contraction [20]. On the other hand, excitation propagation in the uterus is undetermined and the time needed for the excitation to reach the measurement area of electrodes may vary [19]. Finally, we assumed that two contractions—detected in mechanical and electrical signals—were considered as the same episode, when the EHG contraction reaches its maximum within the duration of the corresponding episode in tocogram and the onset of the EHG contraction occurs before the peak of Toco contraction (Fig. 2b). These conditions were verified empirically for a set of signals. Inconsistency indices were defined to take the minimum value of zero when the contraction parameters being compared are completely consistent. When estimating the inconsistency of one of the following parameter types: number, location, duration (relative measure), amplitude and area, the corresponding index reaches the maximum value of 1 for full inconsistency between the EHG and TOCO contractions. For the rest parameters describing the contraction in time: the onset time, rise time and duration, the corresponding inconsistency is quantified by the absolute differences between a given parameter determined for Toco contractions and corresponding episodes in EHG signal.

Number and Location When considering the numbers of contractions detected in both signals the proposed index *NI* takes a minimum value of zero if these numbers are equal, which means a full consistency as for the number of contractions. As for the contraction locations, the *LI* index quantifies what part of all detected contractions is constituted by the corresponding contractions. The full consistency ($LI = 0$) is reached if each Toco contraction corresponds to the single EHG contraction and vice versa.

$$NI = \frac{|N_T - N_E|}{\max(N_T, N_E)}, \tag{1}$$

$$LI = 1 - \frac{2}{N_T + N_E} N_C, \tag{2}$$

where: N_T and N_E are the number of contractions detected in Toco and EHG signal respectively, and N_C is the number of corresponding contractions.

Onset Time Taking into account that electrical excitation in EHG channel should precede the contraction recorded by Toco transducer, and due to a fixed position of Toco transducer in relation to the EHG electrodes during patient's monitoring, the Toco contraction is expected to start after some constant time since the corresponding electrical activity starts. This time shift Δ can be estimated as the mean value

calculated over the differences between the onset time of Toco contraction—O_T and the corresponding episode in EHG signal—O_E:

$$\Delta = \frac{1}{N_C} \sum_{i=1}^{Nc} (O_{Ti} - O_{Ei}) \quad [s]\,, \tag{3}$$

where: N_C is the number of corresponding contractions, O_{Ti} and O_{Ei} are the onset times of the i-th corresponding contraction in Toco and EHG signal, respectively.

Positive value of the shift Δ means that electrical excitation represented by contraction detected in EHG preceded the mechanical response in Toco signal. Since this shift originates from the physiological relationship between the electrical and mechanical activity, the onset time of Toco contractions is corrected with this shift. Thus, the inconsistency index referring to the onset time is defined as the root-mean-square error (RMSE):

$$OI = \sqrt{\frac{1}{N_C} \sum_{i=1}^{Nc} [O_{Ei} - (O_{Ti} - \Delta)]^2} \quad [s]\,, \tag{4}$$

where: N_C is the number of corresponding contractions, O_{Ti} and O_{Ei} are the onset times of the i-th corresponding contraction in Toco and EHG signal, respectively, and Δ is the constant shift between the mechanical contraction and electrical excitation.

Rise Time Inconsistency between the rise time of Toco contraction—R_T and the corresponding episode in the EHG signal—R_E is based on the absolute difference between these values, and for all corresponding contractions the inconsistency index is calculated as follows:

$$RI = \frac{1}{N_C} \sum_{i=1}^{Nc} |R_{Ti} - R_{Ei}| \quad [s]\,, \tag{5}$$

where: N_C is the number of corresponding contractions, R_{Ti} and R_{Ei} are the rise times of the i-th corresponding contraction in Toco and EHG signal, respectively.

Duration Inconsistency between duration of Toco contractions—D_T and the corresponding episode in EHG signal—D_E is estimated by two indices relying on the absolute difference between these values. For all corresponding contractions these two indices are calculated as follows:

$$DI = \frac{1}{N_C} \sum_{i=1}^{Nc} |D_{Ti} - D_{Ei}| \quad [s]\,, \tag{6}$$

$$DIR = \frac{1}{N_C} \sum_{i=1}^{Nc} \frac{|D_{Ti} - D_{Ei}|}{\max(D_{Ti}, D_{Ei})}, \tag{7}$$

where: N_C is the number of corresponding contractions, D_{Ti} and D_{Ei} are the durations of the i-th corresponding contraction in Toco and EHG signal, respectively.

The DIR index, being the relative measure, takes the following boundary values:
$DIR = DIR_{min} = 0$ *when* $D_{Ti} = D_{Ei}$ *for* $1 \leq i \leq N_C$,
$DIR \longrightarrow DIR_{max} = 1$ *when* $D_{Ti} \gg D_{Ei}$ *or* $D_{Ti} \ll D_{Ei}$ *for* $1 \leq i \leq N_C$.

Amplitude Since the amplitude of the corresponding contractions in Toco and EHG signals is represented in different scales and units, the inconsistency index AI has been based on linear correlation coefficient. In order to keep the same meaning as for other inconsistency indices (0—full consistency, 1—full inconsistency), and assuming that negative correlation and lack of correlation means full inconsistency, the index is defined as follows.

$$AI = \begin{cases} 1 - r_A & \text{for } r_A > 0, \\ 1 & \text{for } r_A \leq 0, \end{cases} \tag{8}$$

where: r_A is the correlation coefficient between two sets: $\{A_{T1}, A_{T2}, \ldots, A_{TNc}\}$ and $\{A_{E1}, A_{E2}, \ldots, A_{ENc}\}$, A_T and A_E are the amplitudes of corresponding contractions in Toco and EHG signal respectively, and N_C is the number of corresponding contractions.

Area The area under curve of the corresponding contractions in Toco and EHG signals is represented in different scales and units. Thus, the area inconsistency index UI has been defined on the same way as for the contraction amplitudes:

$$UI = \begin{cases} 1 - r_U & \text{for } r_U > 0, \\ 1 & \text{for } r_U \leq 0, \end{cases} \tag{9}$$

where: r_U is the correlation coefficient between two sets: $\{U_{T1}, U_{T2}, \ldots, U_{TNc}\}$ and $\{U_{E1}, U_{E2}, \ldots, U_{ENc}\}$, U_T and U_E are the areas of corresponding contractions in Toco and EHG signal respectively, and N_C is the number of corresponding contractions.

Contraction Shape Inconsistency between the shapes of corresponding contractions has been quantified basing on the cross correlation function R_{xy}. In our study the cross correlation function took the positive values only, thus the contraction shape inconsistency index SI, taking into account that value of one means full inconsistency, has been defined as:

$$SI = \frac{1}{N_C} \sum_{i=1}^{Nc} \left(1 - \max\left(R_{xy}\left(\tau_i\right)\right)\right), \quad \tau_i \in \langle -D_{Ti}, +D_{Ti} \rangle, \tag{10}$$

where: τ_i is the lag time for the i-th corresponding contraction in Toco and EHG signals (the lag time is changed according to the duration of Toco contraction D_{Ti}), and N_C is the number of corresponding contractions.

2.3 Uterine Electrical Activity Description

Besides the information on uterine contractions which corresponds to mechanical activity expressed by Toco signal, monitoring of electrical activity can provide information on electrical excitation propagating across the uterus [12, 22]. This type of potentially clinically useful information is manifested in the electrohysterogram by the action potential spikes within the bursts corresponding to contractions [1, 21]. It can be quantified by several spectral parameters as well as contraction intensity. The consecutive single burst segment is extracted from the EHG signal basing on information about the onset time and duration of contraction detected in the EHG envelope, and then the power spectrum density (PSD) is determined. The additional parameters, which can be obtained only from the signals of uterine electrical activity, underwent the analysis considering their correlation with the classical parameters of contractions in Toco signal: onset time O_T, rise time R_T, duration D_T, amplitude A_T and area U_T.

Contraction power P is calculated as a sum of amplitudes of frequency components of PSD multiplied by the frequency resolution:

$$P = \sum_{i=L}^{H} \left(p_i \cdot \frac{f_S}{N} \right), \tag{11}$$

where: p_i is the amplitude of i-th PSD component, H and L are the PSD component numbers corresponding to the upper and lower limit of the EHG frequency band respectively, f_S is the sampling frequency, and N is the number of samples within the contractile segment.

Maximum power frequency F_{max} indicates the component in PSD of the maximum power (amplitude):

$$F_{max} = k \cdot \frac{f_S}{N}, \quad p_k = \max\left(p_1 \ldots p_M\right), \tag{12}$$

where: k is the k-th PSD component of maximum amplitude, f_S is the sampling frequency, N is the number of samples, M is the number of frequency components in PSD (for the single-sided PSD the $M = N/2$), and p_k is the amplitude of the k-th PSD component.

Median frequency F_{med} represents the component frequency which splits the PSD into two parts of the same power:

$$F_{med} = j \cdot \frac{f_S}{N}, \quad \sum_{i=1}^{j} p_i = \sum_{i=j}^{M} p_i, \tag{13}$$

where: j is the j-th PSD component, f_S is the sampling frequency, N is the number of samples, M is the number of frequency components in PSD (for the single-sided PSD the $M = N/2$), and p_i is the amplitude of the i-th PSD component.

Contraction intensity refers to number of spikes within the burst in EHG corresponding to the uterine contractile activity. Action potential spike has a form of biphasic wave, which consists of two deflections in negative and positive direction. After smoothing by the moving average in the fifteen-sample window, the numbers of positive and negative peaks are determined within the burst. Only these peaks are considered as valid whose width is greater than 2 s, and amplitude exceeds 25 % of the contraction amplitude A_E. The average value of positive and negative peaks recalculated into 1-min interval is assumed as the intensity of the contraction I:

$$I = \frac{AP^+ + AP^-}{2} \frac{60}{D_E}, \tag{14}$$

where: AP^+ and AP^- are the number of positive and negative peaks respectively, detected in the contractile segment of duration D_E in seconds.

3 Result

In Group I the number of contractions detected in Toco was 224, while in EHG we recognized 372 contractions. Using the established criteria we found 122 corresponding contractions in that group. In other groups those numbers were as follows: Group II (178, 247, 134), Group III (164, 202, 108) respectively. In all groups the number of contractions detected in the EHG signals is higher than the number of contractions obtained from the Toco signals. This may prove the higher sensitivity of the bioelectrical method, though these numbers are also affected by a different approach to determination of the basal tone as well as the minimum values applied in the contraction detection conditions. For particular recordings the number and location inconsistency indices were calculated, and the inconsistencies of other parameters were estimated for the corresponding contractions. Mean values, standard deviations and the range of all indices determined for all the three groups of the signals are listed in Table 1.

The highest consistency of both the number of all contractions detected and the number of corresponding contractions was noted for Group III, which comprises the signals acquired during physiological deliveries. Both *NI* and *LI* index reached the lowest values of 0.20 and 0.35, respectively. It should be noted that the index value of zero means a full consistency, while one a full inconsistency. Such results can be considered as being expected, due to the fact that Group III comprises the signals with evident and strong contractile activity. In the signals collected in earlier weeks of pregnancy (Group I and Group II) the obtained inconsistency is higher. Inconsistency of the onset times is very similar in all groups, where OI takes a value of 21, 22 and 19 s, respectively. Slightly higher inconsistency was noted for the rise

Table 1 Statistical analysis of the indices describing the inconsistency of contraction parameters in all groups of recordings

Inconsist.	Group I			Group II			Group III		
Index	Mean ± SD	Min	Max	Mean ± SD	Min	Max	Mean ± SD	Min	Max
NI	0.39 ± 0.23	0.08	0.89	0.33 ± 0.28	0.00	0.79	0.20 ± 0.19	0.00	0.58
LI	0.55 ± 0.23	0.14	0.93	0.42 ± 0.26	0.00	0.88	0.35 ± 0.22	0.00	0.78
Δ [s]	7 ± 18	–38	32	10 ± 30	–36	91	17 ± 18	–5	54
OI [s]	21 ± 16	0	54	22 ± 18	0	59	19 ± 12	0	39
RI [s]	24 ± 12	7	52	28 ± 18	4	86	26 ± 13	10	56
DI [s]	37 ± 19	2	69	45 ± 24	12	97	34 ± 16	10	74
DIR	0.37 ± 0.15	0.06	0.64	0.46 ± 0.2	0.17	1.00	0.37 ± 0.16	0.07	0.76
AI	0.68 ± 0.27	0.00	1.00	0.63 ± 0.32	0.00	1.00	0.72 ± 0.24	0.39	1.00
UI	0.61 ± 0.32	0.00	1.00	0.72 ± 0.28	0.00	1.00	0.66 ± 0.33	0.01	1.00
SI	0.61 ± 0.08	0.39	0.74	0.65 ± 0.09	0.48	0.83	0.61 ± 0.07	0.49	0.73

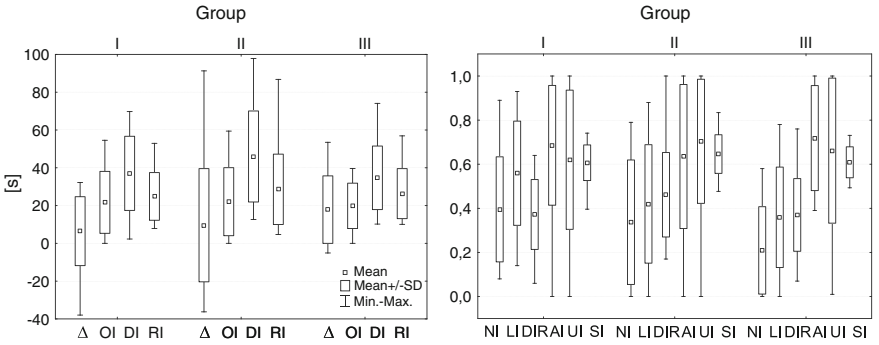


Fig. 3 Descriptive statistics of the inconsistency obtained for particular group of patients and concerning parameters of contractions detected in EHG and Toco signals: the time–domain parameters comprising the shift between the mechanical and electrical activity—Δ, onset time—O_I, duration—D_I, and rise time—R_I, as well as the number—*NI*, location—*LI*, relative duration—*DIR*, amplitude—*AI*, area—*UI* and shape—*SI* (B)

time of corresponding contractions that has been expressed by *RI* equal to 24, 28, and 26 s. Positive value of the shift Δ confirmed that electrical activity expressed as bursts in EHG precedes the mechanical response observed in Toco signal. Among the timing parameters of contraction the highest inconsistency was obtained for contraction duration, with the highest value of absolute difference $DI = 45$ s in the Group II. On average, the inconsistency of duration was equal to 0.37 for Group I and III, whereas for Group II comprising the records with symptoms of threatening labour, that inconsistency was higher –0.46. Taking into account the obtained values of *AI*, *UI* and *SI*, we can notice that the inconsistency of amplitude, area and shape is very high in all groups. In other words there is very weak linear relationship between these parameters of the corresponding contractions being detected in signal of mechanical and electrical uterine activity.

Table 2 Correlation coefficients determined between the parameters describing the bursts of action potential spikes (intensity I, power P, median frequency F_{med}, maximum power frequency F_{max}) and the parameters describing corresponding mechanical contractions (onset time O_T, duration D_T, rise time R_T, amplitude A_T and area U_T)

Group I					
	O_T	D_T	R_T	A_T	U_T
I	–0.056	–0.060	–0.054	–0.056	–0.060
P	–0.066	0.389	–0.065	–0.066	0.389
F_{med}	0.170	–0.133	0.167	0.170	–0.133
F_{max}	0.010	–0.181	0.006	0.010	–0.181
Group II					
	O_T	D_T	R_T	A_T	U_T
I	–0.074	0.052	–0.072	–0.074	0.052
P	–0.054	–0.005	–0.056	–0.054	–0.005
F_{med}	–0.005	–0.173	–0.010	–0.005	–0.173
F_{max}	–0.038	–0.178	–0.043	–0.038	–0.178
Group III					
	O_T	D_T	R_T	A_T	U_T
I	0.183	–0.035	0.181	0.183	–0.035
P	–0.085	0.010	–0.085	–0.085	0.010
F_{med}	0.121	0.055	0.118	0.121	0.055
F_{max}	–0.085	0.028	–0.087	–0.085	0.028

We noted a high dispersion for all inconsistency indices for the recordings in particular groups, that is expressed by high SD as well as the minimum and maximum values as presented in Fig. 3. The lowest dispersion was obtained for the recordings in Group III comprising the uterine activity signals acquired from the patients during the first period of physiological labour which showed an evident contractile activity. Among the inconsistency indices the lowest dispersion was observed for the contraction shape index *SI*. Correlation coefficients that have been determined between the parameters describing the burst of action potential spikes and the parameters of corresponding mechanical contractions are listed in Table 2. We can notice a low or very low relationship between the EHG action potentials and contraction in Toco signals. It was observed in all groups of recordings.

4 Conclusion

Both visual and automated analysis of EHG signals confirmed that it is possible to record a good quality signal of bioelectrical activity of the uterine muscle through the maternal abdominal wall. Higher number of contractions detected in the EHG signals

than in tocograms may indicate a higher sensitivity of the bioelectrical approach. However, we should take into account a significant impact of different algorithms for determining the basal tone as well as the limitation of automated threshold-based detection of contractions. The signal segment is classified as a contraction, if both the amplitude and duration of the transient increase above the basal tone exceed the fixed minimum values. Thus, there may be a situation when the contraction is not recognized because its parameters will be slightly smaller than the minimum values, while another one with very similar characteristics but exceeding these values is detected.

Obtained results concerning the number of contractions detected in both types of signals and a number of corresponding episodes show that synchronization between the mechanical and electrical activity varies in different stages of pregnancy and labour, and it becomes stronger as the birth term is approaching. Comparison of the parameters of corresponding contractions detected in Toco and EHG signals provided slightly worse results than we could have expected. Undoubtedly, limitations of the external tocography have affected the results obtained. Strain gauge transducer covers very small area above the uterine muscle, thus the real strength of contractions can be determined only approximately. It resulted in a very high inconsistency concerning both amplitudes and areas of corresponding contractions. Both methods measure the activity of the same uterine muscle, but of different type—electrical and mechanical, and what's more both methods use an indirect approach—the signals are acquired through the maternal abdominal wall. Without any doubt such approach affects the signals acquired and thus the inconsistency between them. Lack of correlation that have been noticed between the parameters describing the burst of action potential spikes and corresponding mechanical contractions may lead to conclusion that additional information can be obtained from the analysis of action potentials in electrohysterogram.

Taking into account the usability of the electrohysterography for classical fetal monitoring during pregnancy and for controlling the labour progress, this method may be considered as an alternative approach to tocography. Satisfying consistency of the number and location enables to associate the uterine contractions with decelerations recognized in fetal heart rate signal. Such situation when the fetus responds to contraction with a lowering its heart rate, is considered as a sign of fetal distress. Electrohysterogram provides not only the classical description of contractions based on amplitude and timing parameters, but it enables to determine a number of additional parameters associated with higher frequency components of the EHG signal that describe the physiological properties of the uterus. Evaluation of usefulness of this additional information in clinical practice, for example to early detection of symptoms of the threatening preterm labour, is the subject of many research studies being currently conducted.

Acknowledgments This scientific research work is supported by The National Centre for Research and Development of Poland.

References

1. Alamedine, D., Khalil, M., Marque, C.: Comparison of different ehg feature selection methods for the detection of preterm labor. Comput. Math. Methods Med. 9 (2013). http://dx.doi.org/10.1155/2013/485684
2. Alberola-Rubioa, J., Prats-Boludaa, G., Ye-Lina, Y., Valerob, J., Peralesb, A., Garcia-Casadoa, J.: Comparison of non-invasive electrohysterographic recording techniques for monitoring uterine dynamics. Med. Eng. Phys. **35**, 1736–1743 (2013)
3. Bajlekov, G., Rabotti, C., Oei, S., Mischi, M.: Electrohysterographic detection of uterine contractions in term pregnancy. In: 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 5851–5854 (2015)
4. Czabanski, R., Jezewski, J., Horoba, K., Jezewski, M.: Fetal state assessment using fuzzy analysis of the fetal heart rate signals—agreement with the neonatal outcome. Biocybern. Biomed. Eng. **33**(3), 145–155 (2013)
5. Czabanski, R., Wrobel, J., Jezewski, J., Leski, J., Jezewski, M.: Efficient evaluation of fetal wellbeing during pregnancy using methods based on statistical learning principles. J. Med. Imaging Health Inform. **5**(6), 1327–1336 (2015)
6. De Lau, H.D., Rabotti, C., Bijloo, R., Rooijakkers, M.J., Mischi, M., Oei, G.S.: Automated conduction velocity analysis in the electrohysterogram for prediction of imminent delivery: a preliminary study. Comput. Math. Method Med. 7 (2013)
7. Euliano, T.Y., Nguyen, M., Darmanjian, T.S., Mcgorray, S.P., Euliano, N., Onkala, A., Gregg, A.R.: Monitoring uterine activity during labor: a comparison of 3 methods. Am. J. Obstet. Gynecol. **208**(66), e1–66.e6 (2013)
8. Garfield, R.E., Lucovnik, M., Kuon, R.J.: Diagnosis and effective management of preterm labor. MGM J. Med. Sci. **1**, 22–27 (2014)
9. Horoba, K., Jezewski, J., Wrobel, J., Matonia, M., Czabanski, R., Jezewski, M.: Analysis of uterine contractile wave propagation in electrohysterogram for assessing the risk of preterm birth. J. Med. Imaging Health Inform. **5**(6), 1287–1294 (2015)
10. Jezewski, J., Horoba, K., Matonia, A., Wrobel, J.: Quantitative analysis of contraction patterns in electrical activity signal of pregnant uterus as an alternative to mechanical approach. Physiol. Meas. **26**, 753–767 (2005)
11. Jezewski, J., Roj, D., Wrobel, J., Horoba, K.: A novel technique for fetal heart rate estimation from doppler ultrasound signal. Biomed. Eng. Online **10**, 243–268 (2011). doi:10.1186/1475-925X-10-92/
12. Jezewski, J., Matonia, A., Czabanski, R., Horoba, K., Kupka, T.: Classification of uterine electrical activity patterns for early detection of preterm birth. In: Burduk, R., et al. (ed.) Computer Recognition Systems 8—CORES 2013, Advances in Intelligent Systems and Computing AISC. vol. 226, pp 559–568. Springer, Heidelberg (2013)
13. Kotas, M., Jezewski, J., Matonia, A., Kupka, T.: Towards noise immune detection of fetal qrs complexes. Comput. Methods Programs Biomed. **97**(3), 241–256 (2010)
14. Kotas, M., Jezewski, J., Horoba, K., Matonia, A.: Application of spatio-temporal filtering to fetal electrocardiogram enhancement. Comput. Methods Programs Biomed. **104**(1), 1–9 (2011)
15. Kupka, T., Jezewski, J., Matonia, A., Horoba, K., Wrobel, J.: Timing events in doppler ultrasound signal of fetal heart activity. In: Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 337–340. San Francisco (2004)
16. La Rosa, P.S., Nehorai, A., Eswaran, H., Lowery, C.L., Preissl, H.: Detection of uterine emg contractions using a multiple change point estimator and the k-means cluster algorithm. IEEE Trans. Biomed. Eng. **55**, 453–467 (2008)
17. Lucovnik, M., Manner, W.L., Chambliss, L.R., Blumrick, R., Balducci, J., Novak-Antolic, Z., Garfield, R.E.: Noninvasive uterine electromyography for prediction of preterm delivery. Am. J. Obstet. Gynecol. **204**, 228e1–10 (2011)
18. Moslem, B., Khalil, M., Marque, C., Diab, M.O.: Complexity analysis of the uterine electromyography. In: 32nd Annual International Conference of the IEEE EMBS, pp. 2802–2805. Buenos Aires, Argentina (2010)

19. Rabotti, C., Mischi, M., Laar, J.V., Oei, S.G., Bergmans, J.: Inter-electrode delay estimators for electrohysterographic propagation analysis. Physiol. Meas. **30**, 745–761 (2009)
20. Rabotti, C., Mischi, M., Oei, S.G., Bergmans, J.: Noninvasive estimation of the electrohysterographic action-potential conduction velocity. IEEE Trans. Biomed Eng. **57**, 2178–2187 (2010)
21. Radomski, D., Malkiewicz, A.: Identification of a nonlinear association between components of the electrohysterographical signal. Int. J. Electron. Telecom. **56**, 287–290 (2010)
22. Radomski, D., Grzanka, A., Graczyk, S., Przelaskowski, A.: Assessment of uterine contractile activity during a pregnancy based on a nonlinear analysis of the uterine electromyographic signal. In: Pietka, E., Kawa, K. (eds.) Advances in Soft Computing Information Technology in Biomedicine, pp. 325–334. Springer (2008)
23. Rooijakkers, M.J., Rabotti, C., Oei, S.G., Aarts, R., Mischi, M.: Low-complexity intrauterine pressure estimation using the teager energy operator on electrohysterographic recordings. Physiol. Meas. **35**, 1215–1228 (2014)
24. 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**(10), 1983–1989 (2006)
25. Vrhovec, J., Macek-Lebar, A.: An uterine electromyographic activity as a measure of labor progression. In: Steele, C. (ed.) Applications of EMG in Clinical and Sports Medicine, pp. 243–268. InTech (2012)
26. Wrobel, J., Horoba, K., Pander, P., Jezewski, J., Czabanski, R.: Improving the fetal heart rate signal interpretation by application of myriad filtering. Biocybern. Biomed. Eng. **33**, 211–221 (2013)
27. Wrobel, J., Jezewski, J., Horoba, K., Pawlak, A., Czabanski, R., Jezewski, M., Porwik, P.: Medical cyber-physical system for home telecare of high-risk pregnancy—design challenges and requirements. J. Med. Imaging Health Inform. **5**(6), 1295–1301 (2015)
28. Wrobel, J., Matonia, A., Horoba, K., Jezewski, J., Czabanski, R., Pawlak, A., Porwik, P.: Pregnancy telemonitoring with smart control of algorithms for signal analysis. J. Med. Imaging Health Inform. **5**(6), 1302–1310 (2015)
29. Ye-Lin, Y., Alberola-Rubio, J., Prats-Boluda, G., Perales, A., Desantes, D., Garcia-Casado, J.: Feasibility and analysis of bipolar concentric recording of electrohysterogram with flexible active electrode. Ann. Biomed. Eng. **43**(4), 968–976 (2015)

Computer-Aided Evaluation of Sibilants in Preschool Children Sigmatism Diagnosis

Zuzanna Miodońska, Michał Kręcichwost and Agata Szymańska

Abstract In this study an approach to the analysis of sibilant sounds is presented. The method includes the use of the signal spectral measures: amplitudes and frequencies of the fricative formants and the normalized spectral moments. Mel-Frequency Cepstral Coefficients are included in the feature vector as well. The feature sets have been classified using the Support Vector Machine with linear kernel. The experiments have been performed on the speech corpus containing 5 selected words pronounced by 60 preschool children with proper pronunciation or with natural pathologies. The best obtained accuracy of the classification is 94.48 %. The method could be used in a computer-assisted sigmatism diagnosis or therapy system.

Keywords Computer-aided pronunciation evaluation · Sibilants · Sigmatism diagnosis

1 Introduction

Computer-aided speech therapy and diagnosis systems become nowadays more and more appreciated by therapists and patients. However, computerized (automatic or semi-automatic) diagnosis of speech disorders, based on the voice recordings, is still not popular due to a small number of efficient tools for pathological speech signal processing.

Sigmatism, also known as lisping, is a speech disorder in which sibilant consonants (*s, z, c, dz; sz, ż, cz, dż; ś, ź, ć, dź*) are misarticulated. Depending on the articulators

Z. Miodońska (✉)
Faculty of Automatic Control, Electronics and Computer Science,
Silesian University of Technology, Gliwice, Poland
e-mail: zuzanna.miodonska@polsl.pl

M. Kręcichwost · A. Szymańska
Faculty of Biomedical Engineering,
Silesian University of Technology, Zabrze, Poland

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_30

position, several types of sigmatism can be distinguished. Each of them results in misarticulation of all or only a group of beforementioned phones.

Sigmatism is a very common disorder in children. Specific phoneme series occur consecutively with maturation, e.g. rustling series consonants (sz, ż, cz, dż) appear usually in fourth or fifth year of age [5] and are among the phonemes which appear last. Their pronunciation is difficult and often requires a speech therapy. A computer system able to detect and classify non-normative pronunciation of these sounds could help the speech therapists in the diagnosis and may become a useful home therapy tool.

1.1 Computer-Aided Sigmatism Diagnosis in Literature

Systems dedicated to pronunciation error detection are designed by many researchers. However, they are usually dedicated to people learning a foreign language [13, 20, 23, 24, 29, 32, 34], not speech therapy patients. Such solutions are called Computer-Assisted Language Learning (CALL) systems. Most of these systems are designed to detect a small number of selected systematic errors, usually the most common mispronunciations specific for second language learners. Some methods use only acoustic data (e.g. Goodness of Pronunciation algorithm [23, 31, 34] or Likelihood Ratio method [8, 16, 34]). Other employ additional phonological information [29, 32].

Pronunciation error detection methods often involve algorithms based on Hidden Markov Models (HMM) [13, 16, 21, 29, 34] or other probabilistic models. Most commonly used signal features are the Mel-Frequency Cepstral Coefficients (MFCC) [2, 8, 12, 14–16, 18, 21, 22] and their first and second-order time derivatives (Delta and Delta-Delta MFCC) [12, 14, 16]. At the same time, many mispronunciation detection studies focus on testing and development of new feature types [9, 28, 33].

The use of speech signal analysis in computer-assisted logopaedic tools is still not very common. Moreover, these solutions are usually aimed at adults. The processing of children speech signal requires different approaches, as their characteristics do not match feature values typical for adults. There are techniques intended to adapt children speech signal for typical analysis methods [10]. However, according to our knowledge the efficiency of such transformations has not been proven yet.

There are only few studies concerning systems dedicated to complex sigmatism diagnosis. Benselam et al. [1] proposed a computer aided pathological speech therapy application aimed at sigmatism in the Arabic language. The described method uses MFCC for modeling and HMM combined with Gausian Mixture Models (GMM) for speech segmentation. The phoneme evaluation is performed by an Artificial Neural Network (ANN). Another approach employs MFCC to construct a Gausian Mixture Model (GMM) [26]. The GMM's features (weights, mean vectors and diagonal elements of the covariance matrices of Gausian densities) are subsequently used as the speech signal models and classified using a Support Vector Machine (SVM) with polynomial kernel. Both studies focus on evaluation of only 2 or 3 selected sibilants

and both methods have been tested on speech corpora containing the recordings of adults or teenagers. Works concerning the evaluation of sibilants in children can hardly be found.

More studies dealing with sibilant acoustic features are limited to their recognition (without correctness evaluation). Extracted features are associated mostly with the noise, which is observable in all sibilants. The most popular measures are amplitudes and frequencies of the fricative formants (formants appearing above 3 kHz) [7, 17, 25], the minimal frequency of the noise [17], differences in the energy levels in different frequency bands [7] and spectral moments [11, 19] as well as measures related to them—kurtosis and skewness [7]. Many of these features have not been proposed in the field of sigmatism detection yet.

The aim of the project is to develop a method for automatic evaluation of sibilants to be employed in the sigmatism diagnosis. This study focuses on the correct and incorrect realizations of a fricative phoneme [Z] (*ż*), which is one of the most commonly mispronounced sounds among Polish children with sigmatism.

Our main contribution is the implementation of a classifier using spectral measures specific for sibilant sounds. To the best of our knowledge, these measures have not been used in the sigmatism diagnosis. The study is based on the speech corpus containing solely speech signal of children with proper pronunciation or with natural pathologies.

The paper is organized as follows. Section 2 provides a presentation of the proposed methods and the speech corpus. Section 3 presents the conducted experiments and the evaluation of the obtained results. The conclusions are discussed in Sect. 4.

2 Materials and Methods

2.1 Data Collection

In the study 5 words containing a fricative consonant [Z] have been chosen by a speech therapist (Table 1). In three words this phoneme is a first sound, whereas in the other two [Z] is in the middle, surrounded by vowels. The transcription of the dictionary is provided according to the extended SAMPA standard (Speech Assessment Methods Phonetic Alphabet) [6].

A group of 60 children at the age of 5 and 6 years has been selected. Each child has repeated the words after the speech therapist. Totally, 300 word instances have been recorded. Due to the background noise or careless pronunciation some utterances had to be excluded. The expert's classification has discarded another set of data that reflected the transitional pronunciation habits in children who were undergoing a speech therapy before. Finally, 50 recordings with correct realization of [Z] and 95 sigmatism occurrences were included in the speech corpus.

Audio recordings were made at the sampling rate of 44.1 kHz and the resolution of 16 bit.

Table 1 The words included in the speech database with the SAMPA phonetic transcription

Word	Pronunciation
strażak	['straZak]
róża	['ruZa]
żona	['Zona]
żarty	['ZartI]
żyrafa	['ZIrafa]

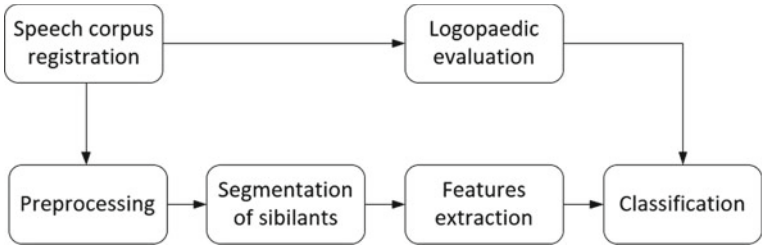


Fig. 1 The general workflow for the proposed method

2.2 General Workflow

The general workflow for the proposed method of evaluation of a phoneme is presented in Fig. 1.

First, the registered waveforms are preprocessed. Then, the sibilant part of each utterance is segmented. Next, the features are extracted. Finally, the classification of segments is conducted and verified with their logopaedic evaluation. The consecutive steps are described below.

2.3 Signal Preprocessing

At the beginning of the preprocessing, the region of interest (the part of the signal containing the analyzed fricative) is manually segmented from each recording. Then, it is annotated as correct or incorrect based on the speech therapists' opinions. Only this piece of the recording is used in further analysis.

Next, the signal is normalized according to the formula:

$$x'(n) = \frac{x(n) - \bar{x}}{\max |x(n) - \bar{x}|}, \tag{1}$$

where $x(n)$ is the nth sample of the signal, $\bar{x}$ is the mean of the signal, and $x'(n)$ is the nth sample of the resulting signal.

Subsequently, preemphasis is performed according to the formula:

$$x''(n) = x'(n) - 0.975 \cdot x'(n-1). \tag{2}$$

Finally, the signal is divided into 10 ms frames with 5 ms overlap.

2.4 Feature Extraction

Selected frames are subjected to a feature extraction phase, which contains 4 steps.

First, frequencies and amplitudes of 3 fricative formants (FF) are calculated. Three highest peaks of the spectrum envelope in the band above 3 kHz are chosen. A minimal distance between the peaks is set to 850 Hz to prevent from finding false peaks. Frequencies and amplitudes of the formants are median-filtered within a segment in order to remove outliers.

Then, the first four normalized spectral moments (SM) are extracted. Calculation of moments is performed using the power spectral density (PSD) [27]. A spectral moment of mth order is defined as:

$$M(m) = \sum_{k=0}^{\infty} |G(k)| * [f_k]^m, \tag{3}$$

where $G(k)$ is the power spectral density for the kth frequency band and f_k is the middle frequency of this band. A normalised mth spectral moment can be calculated as the quotient of the mth moment and the 0th moment:

$$M_{norm}(m) = \frac{M(m)}{M(0)}. \tag{4}$$

Next, 13 MFCC (including energy as the 0th MFCC) are calculated. The procedure of their extraction is widely described in the literature [2, 8, 12, 14–16, 18, 21, 22].

Finally, the evaluation of the proposed features is performed. Fisher linear discriminant analysis [30] (FLDA) is used to reduce the dimensionality of the features and maintain those which separate the two classes (pathology/normative) best.

2.5 Classification

The phoneme evaluation can be regarded as a binary classification problem. According to some of the beforementioned studies, we use SVM as a classification tool.

SVM constructs a hyperplane separating examples of two different classes using the training data. The separation is achieved with the largest margin between the

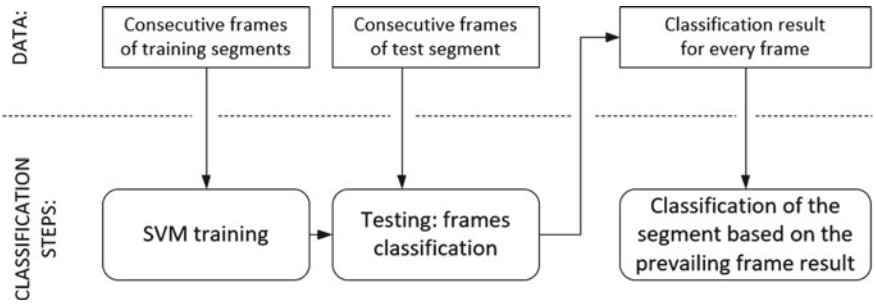


Fig. 2 The classification procedure

hyperplane and the nearest point of the training data [3]. In this study the SVM with linear kernel is employed. The Sequential Minimal Optimization method is used for finding the hyperplane.

A general classification procedure is presented in Fig. 2. An evaluated phoneme segment is modeled by a matrix of feature sets extracted for every 10 ms frame. As SVM cannot classify matrices, every frame needs to be classified independently. A classification of the segment is performed on the basis of the partial results.

During the first step, the feature vectors for all training segments are used for SVM training. Next, frames of the test segment are classified. The segment is then annotated according to the majority rule:

- if most of the frames are classified as pathology, the segment is marked as pathology,
- if most of the frames are classified as normative, the segment is marked as normative,
- if the number of the frames classified as normative and pathology is the same, the segment is assumed to be normative.

3 Experiments and Results

3.1 Testing Procedure

Leave-one-out cross-validation [4] has been used to define the accuracy of the classification. The scheme *training on 144 utterances, testing on 1 utterance* has been performed 145 times (for each segment once). Type I and type II errors have been counted, considering pathology as a positive result of the classification. Accuracy, sensitivity and specificity of the method have been calculated. This testing procedure has been repeated 10 times and the obtained efficiency measures have been averaged.

The classification of individual frames (not segments) has been performed as well. It has been used for evaluation of the relations between partial classification results and final segment classification.

3.2 Results and Discussion

The Full Feature Set (FFS) described in Sect. 2.4 contains 23 measures. As a result of FLDA, a feature vector containing 18 measures has been selected. Four MFCC and the frequency of the 3rd fricative formant have been excluded at this stage. The remaining parameter set is presented in Table 2 and is denoted as the Reduced Feature Set (RFS).

Table 3 presents the results obtained for individual frames classification. Table 4 shows the results for segments classification.

Apart from FFS and RFS, the feature sets including single types of parameters (only FF, only SM or only MFCC) have been tested. The purpose was to verify the usefulness of these features in the sigmatism detection. The accuracy exceeds 84.83 % and 91.03 % for FF and MFCC respectively. The efficiency and the sensitivity of the classification based on SM is significantly lower. However, the obtained accuracy appears promising and SM has been included in the further experiments.

The results obtained for the complex feature sets (FFS and RFS) are more satisfactory than those for single feature types. Therefore, it can be presumed that the

Table 2 The features included in Full Feature Set (FFS) and Reduced Feature Set (RFS)

	Features							
FFS	FF1	FF2	FF3	FFL1	FFL2	FFL3	SM1	SM2
RFS	FF1	FF2		FFL1	FFL2	FFL3	SM1	SM2
FFS	SM3	SM4	MFCC0	MFCC1	MFCC2	MFCC3	MFCC4	MFCC5
RFS	SM3	SM4	MFCC0	MFCC1	MFCC2	MFCC3	MFCC4	MFCC5
FFS	MFCC6	MFCC7	MFCC8	MFCC9	MFCC10	MFCC11	MFCC12	
RFS				MFCC9	MFCC10		MFCC12	

FFn—the frequency of the *n*th fricative formant, *FFLn*—the level (amplitude) of the *n*th fricative formant, *SMn*—the *n*th normalized spectral moment, *MFCCn*—the *n*th MFCC

Table 3 The results obtained for different feature sets in frames classification

Feature set	Accuracy (%)	Sensitivity	Specificity
MFCC	88.97	0.88	0.90
FF	84.83	0.82	0.90
SM	72.41	0.61	0.96
FFS	94.48	0.93	0.96
RFS	95.17	0.91	0.98

Table 4 The results obtained for different feature sets in segments classification

Feature set	Accuracy (%)	Sensitivity	Specificity
MFCC	91.03	0.89	0.94
FF	89.65	0.84	0.99
SM	77.93	0.67	0.98
FFS	92.41	0.91	0.96
RFS	94.48	0.91	0.98

Table 5 The comparison of the obtained results and the results reported in literature

Study	Feature set	Classifier	Accuracy (%)
Here	**RFS**	**SVM**	**94.48**
Benselam and Bencherif [1]	MFCC	ANN	89.00
Valentini-Botinhao et al. [26]	MFCC/GMM	SVM	86.00

proposed feature types are complementary. For both feature sets the frames classification has been more successful than the segments classification. This may result from the assumption of the phoneme correctness in the case of the same number of frames classified as correct and incorrect. Still, the sensitivity above 0.9 may be considered sufficient at this stage of the method development.

Eventually, the best results for segment classification have been obtained for RFS. This outcome has been compared with the results reported in literature (Table 5). Both cited studies have been based on MFCC and their modifications. In our tests the classification accuracy when using only MFCC has reached the level of 88.97 % for frames and 91.03 % for segments, which is close to the reported values. The experiments conducted within this study have shown, that the efficiency gain may be obtained by extending the feature vector by adding the fricative formants and the spectral moments.

4 Conclusion

In this paper an approach for the analysis of sibilant sounds has been presented. The study has been concentrated on the evaluation of the phoneme [Z] (*ż*) realizations. The proposed method uses the signal spectral measures that, to the best of our knowledge, have not been suggested for sigmatism diagnosis: amplitudes and frequencies of the fricative formants and the normalized spectral moments of the 1st–4th order. MFCC, which model the envelope of the spectrum, have been included in the feature vector as well. Fisher linear discriminant analysis has been performed on the proposed parameter set. Subsequently, the full and reduced feature sets have been classified using SVM with linear kernel. The best results have been obtained for the feature

set after the reduction. The accuracy reaches 94.48 %, with sensitivity 0.91 and specificity 0.98.

The experiments have been performed on the corpus of speech of 60 preschool children with proper pronunciation or with natural pathologies. The recorded dictionary have included 5 words containing the analyzed sibilant. The corpus has been verified and annotated by two speech therapy experts. The obtained results of classification suggest, that the proposed method could be used in a computer-assisted sigmatism diagnosis or a therapy system. As we have not used data with simulated disorders nor adults recordings, we have proven that the method is appropriate for children speech.

It can be assumed, that other spectral features associated with noise may also be relevant in the process of sibilants evaluation. These features, i.a. kurtosis or spectral energy in different frequency bands are planned to be tested within the next stages of the project development. The experiments considering other sibilant phones will also be conducted in the near future.

Acknowledgments The work has been partially financed by Polish Ministry of Science and Silesian University of Technology statutory financial support for young researchers BKM-508/RAu-3/2016.

References

1. Benselam, Z.A., Bencherif, M.G.: Arabic speech pathology therapy computer aidedsystem. J. Comput. Sci. **3**(9), 685–692 (2007)
2. Bugdol, M., Segiet, Z., Kręcichwost, M.: Pronunciation error detection using dynamic time warping algorithm. In: Piętka, E., Kawa, J., Więcławek, W. (eds.) Information Technologies in Biomedicine. Volume 284 of Advances in Intelligent Systems and Computing, vol. 4, pp. 345–354. Springer International Publishing Switzerland, Gliwice (2014)
3. Burges, C.J.C.: A tutorial on support vector machines for pattern recognition. Data Min. Knowl. Discov. **2**(2), 121–167 (1998)
4. Clarke, B., Fokoue, E., Zhang, H.H.: Principles and Theory for Data Mining and Machine Learning. Springer (2008). ISBN: 9780387981352
5. Demelowa, G.: Elementy logopedii (Elements of Speech Therapy). Wydawnictwa Szkolne i Pedagogiczne (1979) (in Polish)
6. Demenko, G., Wypych, M., Baranowska, E.: Implementation of grapheme-to-phoneme rules and extended SAMPA alphabet in polish text-to-speech synthesis. Poznań **7**(17) (2003)
7. Evers, V., Reetz, H., Lahiri, A.: Crosslinguistic acoustic categorization of sibilants independent of phonological status. J. Phonetics **26**(4), 345–370 (1998)
8. Franco, H., Ferrer, L., Bratt, H.: Adaptive and discriminative modeling for improved mispronunciation detection. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 7709–7713 (2014)
9. Ge, Z., Sharma, S.R., Smith, M.J.T.: Improving mispronunciation detection using adaptive frequency scale. Comput. Electr. Eng. **39**(5), 1464–1472 (2013)
10. Hagen, A., Pellom, B., Cole, R.: Highly accurate children's speech recognition for interactive reading tutors using subword units. Speech Commun. **49**(12), 861–873 (2007)
11. Haley, K.L., Seelinger, E., Mandulak, K.C., Zajac, D.J.: Evaluating the spectral distinction between sibilant fricatives through a speaker-centered approach. J. Phonetics **38**(4), 548–554 (2010)

12. Hu, W., Qian, Y., Soong, F.: A new neural network based logistic regression classifier for improving mispronunciation detection of L2 language learners. In: 2014 9th International Symposium on Chinese Spoken Language Processing (ISCSLP), pp. 245–249 (2014)
13. Hu, W., Qian, Y., Song, F.K., Wang, Y.: Improved mispronunciation detection with deep neural network trained acoustic models and transfer learning based logistic regression classifiers (2015)
14. Koniaris, C., Salvi, G., Engwall, O.: On mispronunciation analysis of individual foreign speakers using auditory periphery models. Speech Commun. **55**(5), 691–706 (2013)
15. Li, K., Meng, H.: Mispronunciation detection and diagnosis in l2 english speech using multidistribution deep neural networks. In: 2014 9th International Symposium on Chinese Spoken Language Processing (ISCSLP), pp. 255–259 (2014)
16. Liao, H.C., Guan, Y.H., Tu, J.J., Chen, J.C.: A prototype of an adaptive chinese pronunciation training system. System **45**, 52–66 (2014)
17. Lobacz, P., Dobrzanska, K.: Opis akustyczny glosek sybilantnych w wymowie dzieci przedszkolnych. Audiofonologia **15**, 7–26 (1999)
18. Miodonska, Z., Bugdol, M.D., Krecichwost, M.: Dynamic time warping in phoneme modeling for fast pronunciation error detection. Comput. Biol. Med. **69**, 277–285 (2016)
19. Nowak, P.M.: The role of vowel transitions and frication noise in the perception of polish sibilants. J. Phonetics **34**(2), 139–152 (2006)
20. Qin, Y., Wang, G.: A computer-aided Chinese pronunciation training program for English-speaking learners. In: 2014 International Conference on Asian Language Processing (IALP), pp. 154–157 (2014)
21. Rabiner, L., Juang, B.H.: Fundamentals of Speech Recognition. Prentice-Hall Inc, NJ (1993)
22. Sahidullah, M., Saha, G.: Design, analysis and experimental evaluation of block based transformation in MFCC computation for speaker recognition. Speech Commun. **54**(4), 543–565 (2012)
23. Strik, H., Truong, K.P., de Wet, F., Cucchiarini, C.: Comparing classifiers for pronunciation error detection. In: INTERSPEECH, ISCA, pp. 1837–1840 (2007)
24. Su, P.H., Wu, C.H., Lee, L.S.: A recursive dialogue game for personalized computer-aided pronunciation training. IEEE/ACM Trans. Audio Speech Lang. Process. **23**(1), 127–141 (2015)
25. Toda, M., Maeda, S., Honda, K.: Turbulent Sounds. Interface Explorations [IE] 21. De Gruyter Mouton, Berlin (2010)
26. Valentini-Botinhao, C., Degenkolb-Weyers, S., Maier, A., Noeth, E., Eysholdt, U., Bocklet, T.: Automatic detection of sigmatism in children. 1–4 (2012)
27. Vogel, F., Holm, et al.: Spectral moments and time domain representation of photoacoustic signals used for detection of crude oil in produced water (2001)
28. Wang, Y.B., Lee, L.S.: Toward unsupervised discovery of pronunciation error patterns using universal phoneme posteriorgram for computer-assisted language learning. In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 8232–8236 (2013)
29. Wang, H., Qian, X., Meng, H.: Phonological modeling of mispronunciation gradations in L2 english speech of L1 chinese learners. In: IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2014, Florence, Italy, May 4–9, 2014, pp. 7714–7718 (2014)
30. Welling, M.: Fisher linear discriminant analysis (2000)
31. Witt, S.M., Young, S.J.: Phone-level pronunciation scoring and assessment for interactive language learning. Speech Commun. **30**(2–3), 95–108 (2000)
32. Xu, S., Jiang, J., Chen, Z., Xu, B.: Automatic pronunciation error detection based on linguistic knowledge and pronunciation space. In: IEEE International Conference on Acoustics, Speech and Signal Processing, 2009. ICASSP 2009, Taipei, Taiwan, pp. 4841–4844. IEEE (2009)
33. Yuan, H., Xu, J., Zhao, J., Liu, J.: Improve low-resource non-native mispronunciation detection with native speech by articulatory-based tandem feature. In: 2013 IEEE China Summit International Conference on Signal and Information Processing (ChinaSIP), pp. 127–131 (2013)
34. Zhao, T., Hoshino, A., Suzuki, M., Minematsu, N., Hirose, K.: Automatic Chinese pronunciation error detection using SVM trained with structural features. In: Spoken Language Technology Workshop (SLT), 2012 IEEE, Kyoto, Japan, pp. 473–478. IEEE (2012)

Acoustic Mask for Air Flow Distribution Analysis in Speech Therapy

Michał Kręcichwost, Zuzanna Miodońska, Joanna Trzaskalik, Julia Pyttel and Dominik Spinczyk

Abstract Quantitative analysis of lateral sigmatism is a challenge for contemporary speech therapy. In the paper a prototype of an acquisition device for voice and lateral air flow registration is described. The proposed acoustic mask enables for registration of signal from 7 microphones placed in different positions in the area around the mouth. Tests considering air flows directed centrally and laterally have been conducted. Artificially generated air streams and lateral speech utterances have been considered. Energy distribution analysis has been performed to distinguish the direction of the interferences. The results confirm that the acoustic mask may be used in diagnosis and speech therapy of lateral sigmatism.

Keywords Computer-aided pronunciation evaluation · Sibilants · Sigmatism diagnosis

1 Introduction

Contemporary speech therapy procedures are often assisted with software or electronic tools. Computers provide interactive exercises, multimedia stimulus and remote therapy opportunities. However, they can be also used in logopaedic diagnosis.

Lateral sigmatism (also known as lisping) is a pronunciation disorder, which is found among children. It occurs when sibilant sounds (*s, z, c, dz; sz, ż, cz, dż; ś, ź, ć, dź*) are misarticulated because of inappropriate tongue position [1]. In normative realizations of these sounds the air flows from the mouth as a concentrated beam. In

M. Kręcichwost (✉)
Faculty of Automatic Control, Electronics and Computer Science,
Silesian University of Technology, Gliwice, Poland
e-mail: michal.krecichwost@polsl.pl

Z. Miodońska · D. Spinczyk
Faculty of Biomedical Engineering, Silesian University of Technology, Zabrze, Poland

J. Trzaskalik
Cardinal August Hlond University of Education, Myslowice, Poland

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_31

the described type of sigmatism the tip of the tongue touches gums or teeth. As a result, partial occlusion and distraction of the air stream occur. Such phenomenon is called dexter (right-side), sinister (left-side) or bilateral laterality, depending on the flow direction.

Lateral lisp diagnosis is a demanding task, as the position of articulators is difficult to be visually assessed [15]. The problem is even more complex because of many different (sometimes conflicting) linguistic descriptions of this disorder [1, 16–18]. An exact information about the direction of the air stream and the place of tongue-teeth contact is a key factor for an efficient therapy. Therefore, tools for speech diagnosis assistance may be appreciated by therapists and patients.

In recent years, the acoustic analysis of the speech signal has been employed for a variety of speech therapy tasks. Most of the studies are dedicated to mispronunciations detection [5, 13]. Projects concerning the evaluation of specific sounds are designed as well [10]. Among them, solutions for sigmatism diagnosis can be found [3, 4, 21]. There are many works concerning the acoustic features distinctive for different sibilant sounds [8, 12, 14, 19]. However, according to our knowledge no works focused specifically on lateral sigmatism have been conducted.

In the above-mentioned studies the speech registration has been conducted using single microphones. There are also projects which focus on the development of methods based on more complex measurement equipment, like microphone matrices [11]. They allow for multi-channel voice registration and generation of the map of acoustic pressures. However, they require that the speaker stays still during the recording, which makes the measurements hardly reproducible and is especially inconvenient for children. Devices for acquisition of signals other than speech, like electromyographs [2, 9] or electropalatographs [20], are also employed for articulators movement analysis. However, they require the use of electrodes. Therefore, they are more invasive than voice recorders.

The general aim of the overall project is (1) to design and implement a multi-channel speech acquisition device, (2) to develop a tool for computer-assisted diagnosis of lateral sigmatism, (3) to provide an acoustic description of articulation conditions in lateral sigmatism. In this paper we describe the first stage of the project, which is the development of a prototype of the acquisition equipment.

Our main contribution is the design and implementation of a prototype of an acoustic mask dedicated to spatial voice registration. The construction employs the idea of microphone matrices, yet eliminates their main drawbacks. As the mask is mounted on speaker's head, it allows for bigger freedom of movement and reproducible measurements. Moreover, the matrix is semi-cylindrical (not plane as usually). Therefore, the registration of the lateral air flow may be conducted more precisely.

The paper is organized as follows. Section 2 presents the preliminary study measurement station. Conducted experiments and the evaluation of the obtained results are described in Sect. 3. The conclusions are presented in Sect. 4.

2 Acoustic Mask

2.1 Methods

A measurement equipment dedicated to lateral sigmatism diagnosis should satisfy several assumptions. First, it should allow for the analysis of the distribution of the air flow coming from the mouth. Then, it should provide repeatable measurement conditions. Next, apart from air detection it should allow for the speech analysis. Therefore, the tool should employ high-quality microphones.

The core of the described prototype is the acoustic mask. It is composed of 3 parts (Fig. 1). The middle part contains a grid, which allows for installation of microphones in different geometric configurations. The electronic elements responsible for signal amplifying and transmission are placed in the upper part of the device. There are also 7 shielded outputs prepared for connection of sensors. These sensors are microphones by default, but they can be replaced with pressure sensors or others. The lower part of the mask provides stability and maintain the shape of the net. The adjustable fastener stabilizes the mask on the head of a subject. Repeatability of the positioning of the mask on the patient's head is provided by a suitable grid marker, which should coincide with the patient's philtrum.

A set of seven microphones placed on a semi-cylindrical grid allows for registration of the phenomenon of lateral air flow. The configuration of microphones may be easily changed. This feature is particularly important during the prototyping phase. Figure 2 presents several microphone arrangements which have been considered. In each of them the central microphone in front of the mouth has been used in order to provide a reference signal. The configuration presented in Fig. 3 has been chosen for tests.

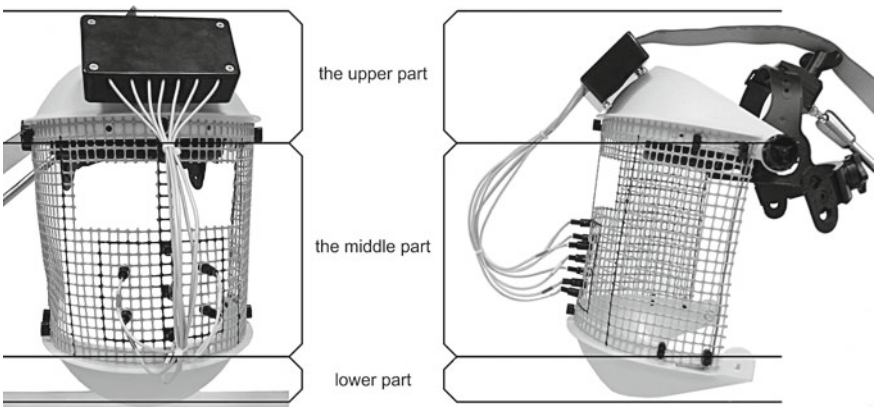


Fig. 1 The acoustic mask

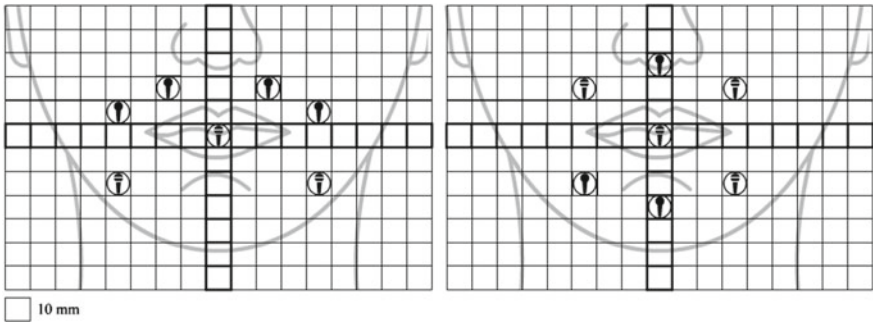


Fig. 2 Different microphone configurations

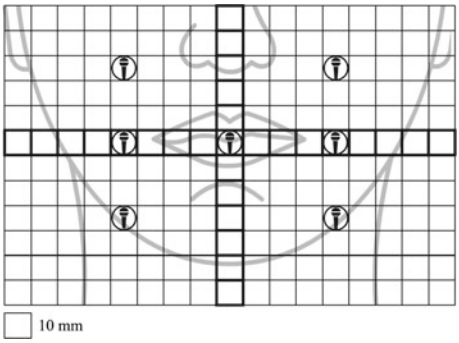


Fig. 3 The microphone configuration employed for the tests

The prototype consists of seven microphones POM-3044P-R [7]. This type of microphones does not require high voltage polarity as standard condenser microphones. They have an omnidirectional characteristic. Its frequency range covers the entire range of audible frequencies (20 Hz–20 kHz), and is completely stable from 300 to 8000 Hz. Because of a very weak output signal, seven-channel audio amplifier has been designed. Analog-to-digital signal is provided by a 16-bit 8-channel A/D converter USB-1608FS [6] equipped with a USB interface. It allows for 8 channel simultaneous recordings of the audio signal with a maximum sampling rate of 96 kHz while maintaining the timing, good quality and low latency in the transmitted signal.

Figure 4 presents a concept of application of the mask in speech therapy. A patient wears the mask and pronounces diagnostic word sequences. The speech signal is being registered in 7 channels and sent to the computer workstation. After the processing, the spatial energy of the patient's speech is visualized and presented to the speech therapist.

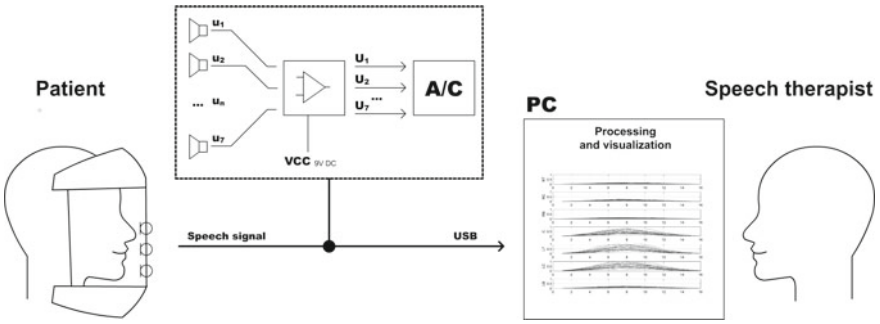


Fig. 4 Application of the acoustic mask in speech therapy

2.2 Evaluation Methods

The prototype has been tested considering the possibility of determining the direction of the air flow. The synthetic source of acoustic pressure has been used in order to provide repeatable conditions. Each registration has been performed using all 7 channels.

The signal analysis phase contains 3 steps. First, 10 ms frames are extracted and Hamming window is employed. Then, a quantitative analysis of the recorded signal is performed based on the frame energy defined as:

$$E_m = \sum_{k=0}^{n} {x_k}^2, \tag{1}$$

where E_m is the energy of the mth frame, x_k is the amplitude of a single sample and n is the number of samples in a processed frame.

Finally, the obtained energy values are normalized in order to allow comparison of signal energy in different channels:

$$E_{norm_m} = \frac{E_m - \min\limits_m(E_m)}{\max\limits_m(E_m) - \min\limits_m(E_m)}, \tag{2}$$

where E_m is the energy of the mth frame and E_{norm_m} is the normalized energy of the mth frame.

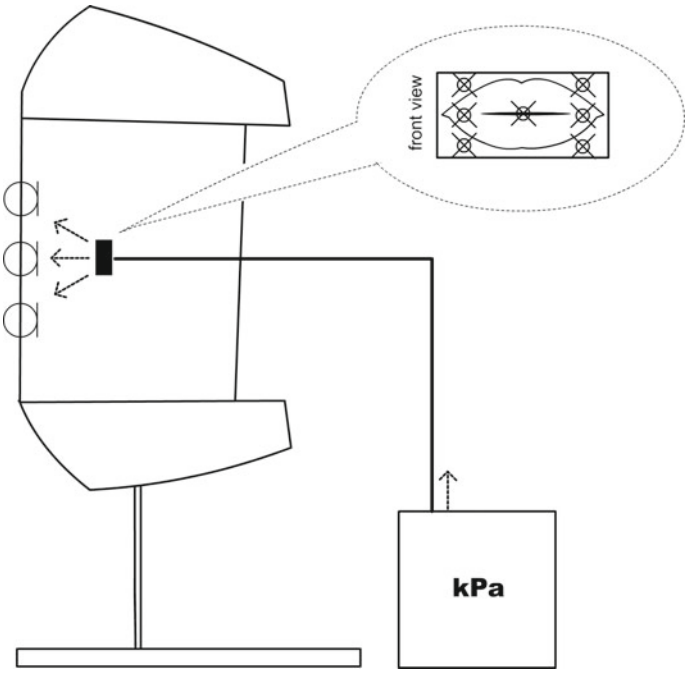


Fig. 5 Side profile of a measuring station

3 Experiments and Results

3.1 Experiment Setup

Two types of experiments have been performed. First, an acoustic pressure signal directed centrally and laterally has been measured. In the second experiment a speech signal analysis, that distinguishes a normal and three types of lateral pronunciation (bilateral, dexter and sinister), has been carried out. In both experiments the acoustic data has been collected by all 7 microphones.

The first experiment has been performed using the measuring station (Fig. 5). The station consists of the acoustic mask and a point source of the sound pressure. The point source of acoustic pressure has been placed in different positions in relation to the central microphone. This was to simulate the improper flow of the air during the incorrect articulation of sounds. The acoustic signal have been registered 20 times at each configuration. The acquired data has been analysed and evaluation measures have been calculated.

The second experiment has been carried out on the speech data representing the normative pronunciation and three types of lateral sigmatism. In each group 18 utterances have been recorded using the acoustic mask.

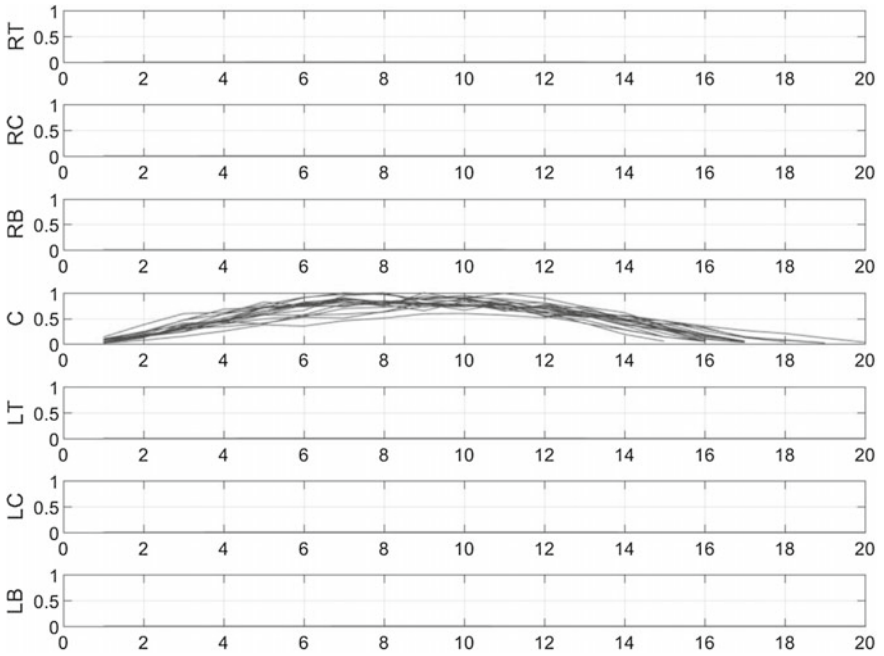


Fig. 6 Normalized energy plots with the air flow directed centrally for consecutive frames (microphones positions: L—left, C—center, R—right, T—top, B—bottom). The *vertical axis* represent normalized energy values. The *horizontal axis* represent the consecutive time frames

3.2 Results

The time plots (Figs. 6 and 7) present normalized energies calculated for consecutive frames in specific channels in the first experiment. Energy values for each recording are represented by individual lines. For centrally directed interference, the signal can hardly be noticed on channels other than the central one (Fig. 6).

Lateral air flows can be noticed as high levels of energy registered by left (Fig. 7a–c) or right (Fig. 7d–f) central microphones, respectively. Top (Fig. 7a, d) or bottom (Fig. 7c, f) interferences are observed on corresponding channels, but also on central ones. High signal energy can be observed for the central microphone, regardless of the direction of the interference. This suggests, that in every case the significant part of the signal energy is accumulated in the central pathway, yet laterality causes a visible division of energy between the channels.

The mean energies registered for each channel during different interferences have been presented in Table 1. The average energy over all the channels has been calculated (0.2004). The values above this threshold have been bolded, to mark the channels, where significant parts of the air flow have been registered.

The distribution of energy for speech dataset (the second experiment) between the central (C), left-central (LC) and right-central (RC) microphones is presented

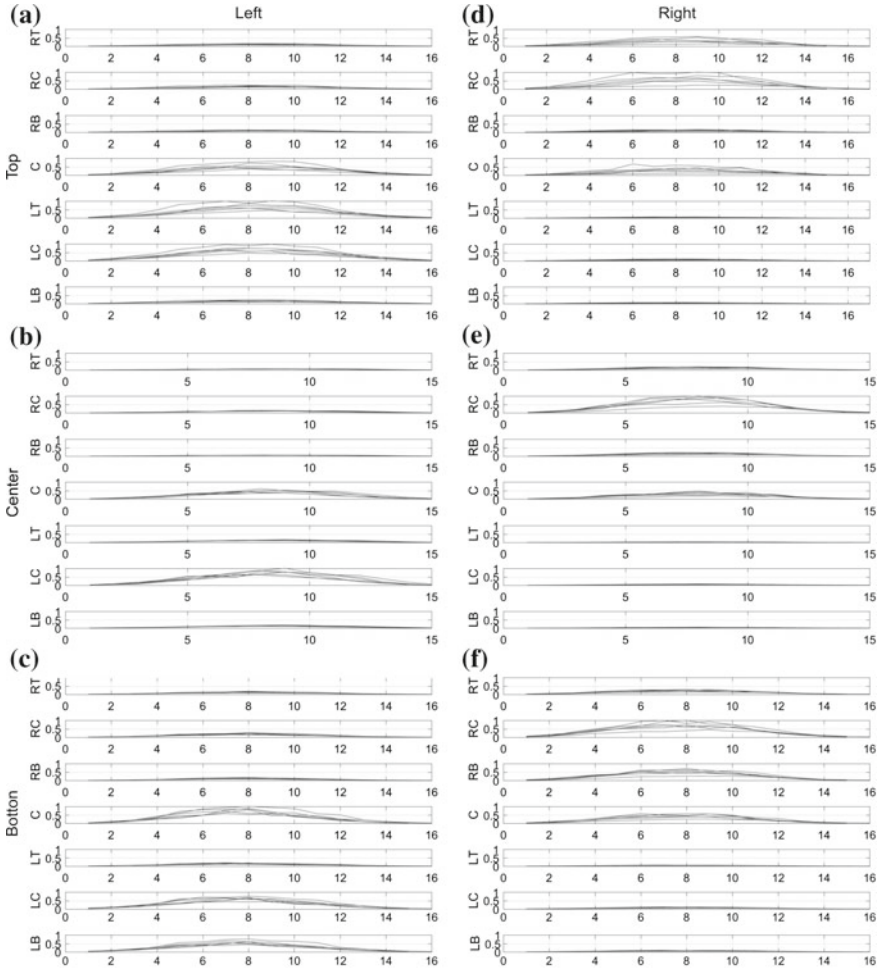


Fig. 7 Normalized energy plots with the air flow directed laterally for consecutive frames (microphones positions: L—left, C—center, R—right, T—top, B—bottom). The consecutive plots represent interferences directed to different parts of the microphone grid: **a** to the left-top part, **b** to the left-center part, **c** to the left-bottom part, **d** to the right-top part, **e** to the right-center part, **f** to the right-bottom part. The *vertical axes* represent normalized energy values. The *horizontal axes* represent the consecutive time frames

Table 1 Means and standard deviations for normalized energy for individual channels in the case of pressure disturbances given in different directions (microphones positions: L—left, C—center, R—right, T—top, B—bottom)

Flow direction							
Figures	Figure 7a	Figure 7b	Figure 7c	Figure 6	Figure 7d	Figure 7e	Figure 7f
Micr.	L. Top	L. Center	L. Bottom	Center	R. Top	R. Center	R. Bottom
LT	**0.4354** ± 0.1584	**0.4599** ± 0.1542	0.1125 ± 0.0364	**0.3454** ± 0.1264	0.0715 ± 0.0244	0.1073 ± 0.0364	0.0652 ± 0.0211
LC	0.0968 ± 0.0178	**0.5165** ± 0.1041	0.1048 ± 0.0190	**0.3246** ± 0.0748	0.0505 ± 0.0097	0.0884 ± 0.0184	0.0412 ± 0.0078
LB	0.1885 ± 0.0196	**0.4197** ± 0.0789	**0.3726** ± 0.0773	**0.5282** ± 0.1100	0.0946 ± 0.0172	0.1400 ± 0.0269	0.0866 ± 0.0176
C	0.0008 ± 0.0001	0.0011 ± 0.0002	0.0006 ± 0.0001	**0.8706** ± 0.0793	0.0007 ± 0.0001	0.0011 ± 0.0002	0.0006 ± 0.0001
RT	0.0370 ± 0.0181	0.0671 ± 0.0411	0.0469 ± 0.0271	**0.2932** ± 0.2008	**0.3056** ± 0.1046	**0.5367** ± 0.2753	0.0961 ± 0.1114
RC	0.0328 ± 0.0113	0.0730 ± 0.0195	0.0485 ± 0.0132	**0.3579** ± 0.0970	0.1850 ± 0.0948	**0.3494** ± 0.1756	0.1967 ± 0.0266
RB	0.0406 ± 0.0064	0.0754 ± 0.0123	0.0556 ± 0.0097	**0.3569** ± 0.0667	0.1724 ± 0.0260	**0.5622** ± 0.0933	**0.4029** ± 0.0751

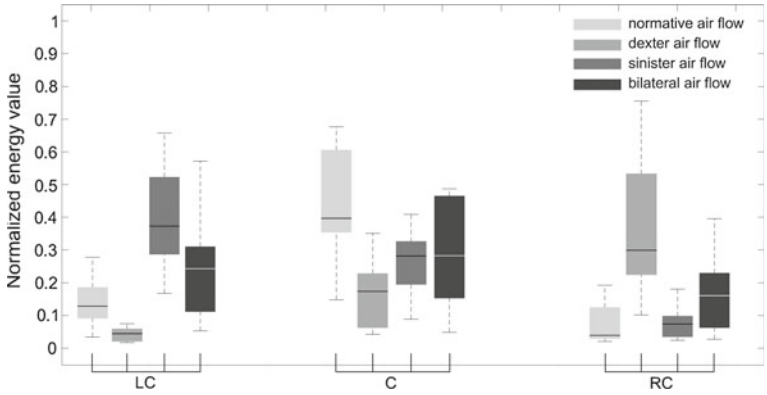


Fig. 8 Distribution of normalized energies on 3 central microphones depending on the direction of air flow during normative and pathological speech

in Fig. 8. For normative pronunciation case the median of normalized energy of the center microphone is several times greater than that of the side microphones. In the case of bilateral sigmatism the median values are at the same level for central and lateral microphones. For simulation of sinister and dexter sigmatism the median value for left and right microphone channels increase to a level comparable to normative signal level for central microphone channel. At the same time, a decrease of energy values for the center channel is observed.

The proposed design of the mask allows for analysis of the distribution of air flow energy on different channels. Thereby, the direction of the lateral interference may be distinguished.

4 Conclusion

In the paper a design and implementation of the acoustic mask has been presented. The concept of measuring station has been described as well. The obtained results confirm that the proposed design of the mask allows the spatial acquisition of speech signals. It is also possible to observe the changes of signal energy in cases of lateral and central air flow. Therefore, this mask may be employed for registration of pronunciation with improper flow of air e.g., lateral sigmatism.

The next stage of the project development will engage actual speech therapy patients, as the construction of the mask has to be tested on non-simulated speech data. These experiments are expected to provide acoustic criteria relevant for lateral sigmatism diagnosis.

Acknowledgments The work has been partially financed by Polish Ministry of Science and Silesian University of Technology statutory financial support for young researchers BKM-508/RAu-3/2016.

References

1. Antos, D., Demel, G., Styczek, I.: Jak usuwac seplenienie i inne wady wymowy (How to remove sigmatism and other speech disorders). Panstwowe Zaklady Wydawnictw Szkolnych, Warszawa (1978). (in Polish)
2. Bell-Berti, F.: An electromyographic study of velopharyngeal function in speech. J. Speech Hear. Res. (1976)
3. Benselam, Z.A., Guerti, M., Bencherif, M.: Arabic speech pathology therapy computer-aided system. J. Comput. Sci. **3**(9) (2007)
4. Botinhao, C., Noeth, E., Hornegger, J., Maier, A.: Speech classification for sigmatism in children. **1** (2009)
5. Bugdol, M.D., Segiet, Z., Kręcichwost, M.: Pronunciation error detection using dynamic time warping algorithm. In: Piętka, E., Kawa, J., Więcławek, W. (eds.) Information Technologies in Biomedicine. Volume 284 of Advances in Intelligent Systems and Computing, vol. 4. Springer, Switzerland (2014)
6. Data sheet of the employed analog-to-digital converter. http://www.mccdaq.com/usb-data-acquisition/USB-1608FS.aspx
7. Data sheet of the employed microphone. http://eu.mouser.com/ProductDetail/PUI-Audio/POM-3044P-R
8. Evers, V., Reetz, H., Lahiri, A.: Crosslinguistic acoustic categorization of sibilants independent of phonological status. J. Phonetics **26**(4) (1998)
9. Freitas, J., Teixeira, A.J.S., Silva, S.S., Oliveira, C., Dias, M.S.: Velum movement detection based on surface electromyography for speech interface. In: BIOSIGNALS 2014—Proceedings of the International Conference on Bio-inspired Systems and Signal Processing, ESEO, Angers, Loire Valley, France (2014)
10. Haley, K.L., Seelinger, E., Mandulak, K.C., Zajac, D.J.: Evaluating the spectral distinction between sibilant fricatives through a speaker-centered approach. J. Phonetics **38**(4) (2010)
11. Krol, D., Lorenc, A., Swiecinski, R.: Detecting laterality and nasality in speech with the use of a multi-channel recorder. In: 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2015)
12. Lobacz, P., Dobrzanska, K.: Opis akustyczny glosek sybilantnych w wymowie dzieci przedszkolnych (The acoustic description of sibilants in preschool children pronunciation). Audiofonologia **15** (1999)
13. Miodonska, Z., Bugdol, M.D., Krecichwost, M.: Dynamic time warping in phoneme modeling for fast pronunciation error detection. Comput. Biol. Med. **69** (2016)
14. Nowak, P.M.: The role of vowel transitions and frication noise in the perception of polish sibilants. J. Phonetics **34**(2) (2006)
15. Ostapiuk, B.: Dyslalia. O badaniu jakosci wymowy w logopedii (Dyslalia. About the evaluation of pronunciation in speech therapy). Uniwersytet Szczecinski, Szczecin (2013) (in Polish)
16. Skorek, E.: Oblicza wad wymowy (The aspects of speech disorders). Warszawa (2001) (in Polish)
17. Soltys-Chmielowicz, A.: Zaburzenia artykulacji. Teoria i praktyka (Articulation disorders. Theory and practice). Oficyna Wydawnicza Impuls (2011) (in Polish)
18. Styczek, I.: Logopedia (Speech therapy). Warszawa (1980) (in Polish)
19. Toda, M., Maeda, S., Honda, K.: Turbulent Sounds. Interface Explorations [IE] 21. De Gruyter Mouton, Berlin (2010)
20. Trochymiuk, A., Swiecinski, R.: Artykulograficzne badanie wymowy grzbietowej—studium przypadku (Articulography of dorsal pronunciation—a case study). Logopedia (2009) (in Polish)
21. Valentini-Botinhao, C., Degenkolb-Weyers, S., Maier, A., Noeth, E., Eysholdt, U., Bocklet, T.: Automatic detection of sigmatism in children. 1–4 (2012)

Longitudinal Voice Study (LoVoiS) Methodology and Preliminary Research Results

Marcin D. Bugdol, Monika N. Bugdol, Anna M. Lipowicz, Andrzej W. Mitas, Maria J. Bieńkowska, Agata M. Wijata and Dariusz Danel

Abstract The paper describes an approach to the kids and youth pubertal evaluation using voice signal. The results of preliminary study conducted on a group of 109 children (58 boys and 51 girls aged 10–18 years) has been presented. The analysis of the voice fundamental frequency proves that this parameter strongly depends on the age (for boys) and on the time of the first menarche (for girls). Such a method for girls and boys maturation assessment is very important for, among others, anthropologist in their studies on social inequalities and observing secular trends.

Keywords Pubertal · Voice analysis · Statistical analysis

1 Introduction

Despite decades of research conducted by anthropologists and paediatricians, puberty still requires further explanations. Due to its high sensitivity to environmental factors, the timing and tempo of maturation are useful tools to assess the living conditions in the society and to study the social inequality and intergenerational changes. The primary measure of the timing and rate of sexual maturation, which is used in epidemiological studies, is the age of menarche. Parallel research on male pubertal

M.D. Bugdol (✉) · M.N. Bugdol · A.W. Mitas · M.J. Bieńkowska · A.M. Wijata
Faculty of Biomedical Engineering, Department of Informatics
and Medical Equipment, Silesian University of Technology, Gliwice, Poland
e-mail: marcin.bugdol@polsl.pl

A.M. Lipowicz
Department of Anthropology, Wroclaw University of Environmental
and Life Sciences, Wrocław, Poland

D. Danel
Polish Academy of Science, Unit of Anthropology, Wrocław, Poland

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_32

timing is much more difficult because of the lack of equivalent marker of masculine pubertal development. Because the course of puberty varies significantly in both sexes (compared with females, males reach puberty later and develop divergent sexual characteristics) another markers of male development are needed.

The traditional methods used for research on boys include the age of peak height velocity, testicular volume, Tanner stages of genital and pubic hair development [5, 8] or endocrine biomarkers [4]. Unfortunately, none of these methods satisfy the user friendly criterion. Isolated measurements of body height do not allow to detect the age of peak height velocity, using the Tanner scale is troublesome (and nowadays unpractical to carry out in large group of people). Measuring the hormonal biomarkers, on the one hand, is expensive and on the other hand it requires the entire logistically complicated procedure (extraction of material, transport to the laboratory and conducting analyses), which in the case of field research faces many difficulties.

Therefore, there is a need to establish a method that will be noninvasive, but precise. It should easy and fast assess, without discomfort of the tested person and researcher, the pubertal stage of male. It is assumed that the analysis of the voice, which in boys change more markedly than in girls, could become such a tool. This need comes, among others, from the necessity of continuous monitoring of the entire population of children and adolescents (including both sexes), to study the intergenerational changes and to detect the underlying causes of social inequalities. This issue becomes important, especially while the results of epidemiological studies indicates that the rapid growth of obesity among children and youth is one of the hypothetical causes of an acceleration in maturation among girls [3].

Acoustic data which characterize the human voice and the voice break (mutation) have the advantage that their collecting do not involve highly qualified staff, and measuring equipment is widely available. Measurement techniques, depending on the objective function may be grouped into:

1. techniques oriented to determine the frequency bands for a high volume of different acoustic data generated for a long period of time,
2. techniques oriented to study individual sounds generated under specified measuring conditions.

In the first case, because of the possibility of spontaneous disorder measurement conditions, allowing potential errors in measurement, the appropriate length of the signal should be provided, which results in averaging these errors. In the second case, when provided repeatable measurement conditions, data should be extracted from the signal through the designation (e.g. indication) in the measuring window a single period (or several periods) of the analysed sound. In the case of continuous sample (case 1) the Fourier transform is a basic mathematical apparatus. In the case of precise sample for a periodic waveform analysis is carried out with Fourier series. In both cases, the analysis of the properties is conducted in the frequency domain; the first refers to the frequency range and accordingly formants ranges and in the second the fundamental frequency and its harmonics. The main difficulty of measuring implies

method error, related in case 1 to specifying the minimum recording time in different life situations, and in case 2 to the calculation of recording window [10, 12]. Both measurement approaches are exposed in the cited literature on the subject.

In order to examine the changes in the voice during puberty, it is necessary to register the sound samples with stable parameters at the free articulation of sounds. Adolescents participating in the study, especially aware of its purpose, may try to modulate their voice in order to achieve "better" results. Moreover, the awareness of being recorded may result in subject's anxiety (stress) and therefore undesirable changes in the speech signal. It is important to perform the sound recording in a way that prevents any voice modulation of the examined person and eliminating the stress of an unusual situation. The description of the various studies aimed at assessing voice changes during puberty that takes into account the described limitations, has been presented in the literature. The researchers recorded the voice of children and adolescents during introduction [8], reading famous passages [8, 14] singing scale and the song "Happy Birthday" [8], counting backwards [13, 14], naming the presented images [9], chanting sentences [7] or singing certain sounds (survey of boys singing in the choir) [13]. Defining the test sound set is basically impossible; unless the length of registration is assumed which makes the self-auto-correction imperfections of individual audio samples possible.

The most often parameter calculated in the studies concerning maturity of voice is its fundamental frequency. This value was used in works [2, 7, 8, 10, 13, 14]. In their studies Hodges-Simeon et al. [9] in addition to fundamental frequency they also induced the frequency and energy distribution of the first four formants. Further appointed parameters are: the frequency range of voice adopted by the speaker and the sound pressure level, which was applied in [7].

Common use of fundamental frequency is caused by its directly connection with the voice maturing and puberty. Study of Danish choir boys on the basis of data from 1994 to 2003 indicate median age at voice break of 14.0 years. The voice break assessed by conductor was a pubertal marker [11]. In [8] comparison between boy's Tanner stages (clinical classification of the stages of puberty) and voice parameters was described. There were abrupt changes in speaking and singing fundamental frequencies between Tanner stages G3 and G4. Sex differences in fundamental frequencies were described by Hodges-Simeon et al. [9]. Their study shows that female F_0 decreases linearly throughout development, whilst male F_0 changes are characterized by sharp drop around age 12–13.

Another problem in such type of research is the proper definition of the studies repetition period and selection of the subject group. The studies about voice changes at puberty require a collection of data from individuals of different age. Such a set may be obtained by examination of the large group of children including people before and after the mutation [7, 9]. Another approach is to register the voice in the same individuals multiple times with a certain interval. This approach has been used by Harries et al. [8]. Their study involved 26 boys aged 13–14 who were examined five times every three months. Thurman examined a larger group [13]. His study

involved 86 boys aged 12–15 years, who sang in the choir. They were examined 27 times at monthly intervals. Studies conducted for over 10 years have been described by Juul et al. [11]. 463 boys singing in the choir were evaluated by the conductor in terms of mutations every week during this period.

The aim of this paper is to present results of preliminary study concerning the possibility for assessing the pubertal stages based on voice changes. To the best of our knowledge, such studies are rather scarce and are not yet conducted in Poland.

2 Materials and Methods

2.1 Voice Registration Procedure

The following measurement procedure has been realized: prior to the test the idea of the experiment was presented to the participant. Upon entering the room the subject sat in a comfortable position in the front of the computer station for sound recording. The task of the person was to proceed according to instructions, which were presented on the computer screen: introduction, articulation for 3 s at a constant level the vowels (a, e, i, o, u) viewed on the computer screen (this stage was repeated three times), performing a computational task and reading a long text. All sentences and calculations were spoken aloud while maintaining the natural tone of the speech. The voice registration procedure has been presented in Fig. 1.

For each person 18 separate recordings have been obtained: three recordings of each vowel (15 in total), first name and last name, differentiated longer text and computational problem (3 recordings). The data have been recorded in 16 bits quality with 44100 Hz sampling rate and saved in WAVE audio format.

This study included 109 children (58 boys and 51 girls aged 10–18 years) and almost all of them will enrolled in next sessions. The parents' written authorization for performing the measurements cycle of children and youth with recordings of voice samples has been obtained prior to the experiments. The parents have completed a questionnaire on socio-economic status of the family and child's lifestyle considering his/her diet and physical activity and, for the girls, time of the first menarche.

2.2 Voice Analysis

In this research it has been focused on analysing only the vowels recordings as they contain the most stable voice signal (in terms of frequency and amplitude). The pre-processing step of speech analysis consisted only of removing the silent parts at the beginning and at the end of the signal. Next, the fundamental frequency has

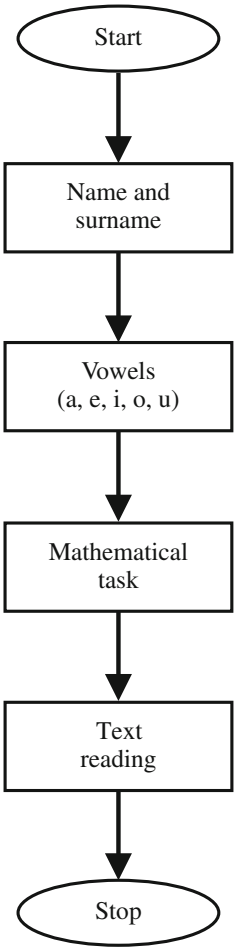


Fig. 1 The algorithm of voice registration

been calculated for every vowel using autocorrelation function [1]. Because three recordings of each phone were available, it has been decided to calculated median value as the parameter that will be the most representative for this vowel for a person.

3 Results

Due to the fact, that only one measurement of the three year cycle registration has been carried out so far, it is impossible to evaluate the boys' pubertal stage and therefore to evaluate the voice parameters that will describe this process. Such information

was available for the girls—it is the time of the first menarche. So the analysis of the voice fundamental frequency for females can be performed with regard to this parameter.

The fundamental frequencies of the girls' voices before and after menarche are statistically different for the groups at the level of significance $\alpha = 0.05$, compared using the Student's t-test for independent samples. In both groups the analysed variables were normally distributed and had equal variances. As expected, voice parameters, represented by the mean fundamental frequency, are lower after menarche (p-value between 0.01 and 0.05).

Figures 2, 3 and 4 show the voice fundamental frequency of the examined children and their age. Data is presented separately for girls and boys for the vowels A, E, I approximated by a polynomial of degree 2 [6]. The dependency exhibits a decreasing tendency (it is true also for vowels O and U).

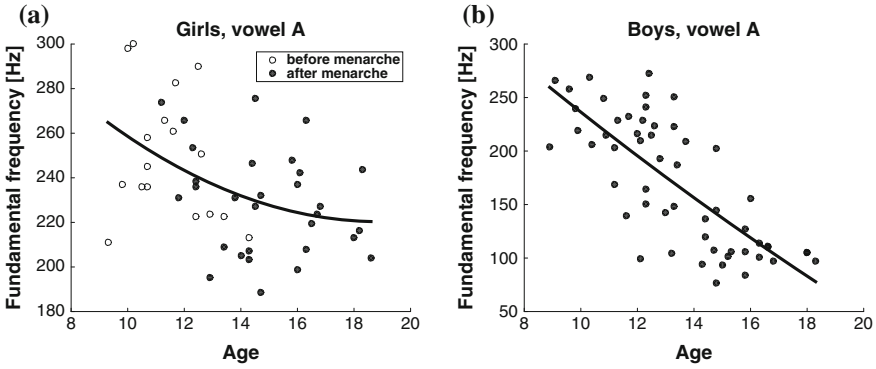


Fig. 2 Vowel A, **a** girls, **b** boys

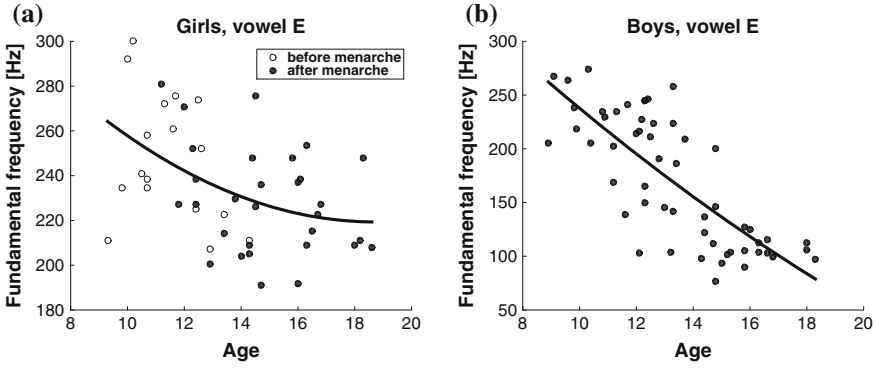


Fig. 3 Vowel E, **a** girls, **b** boys

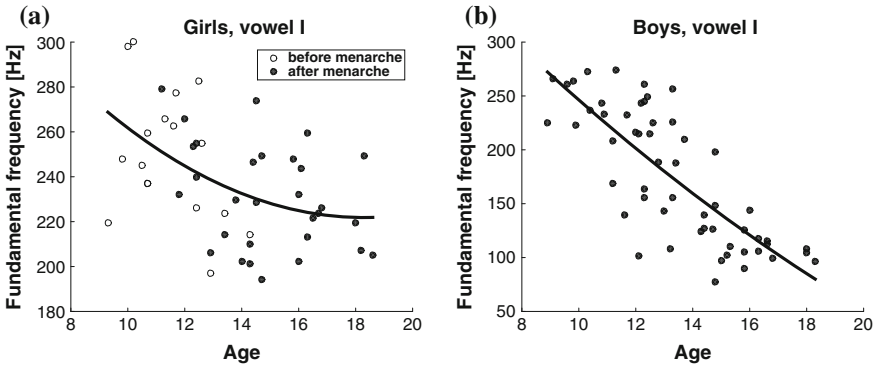


Fig. 4 Vowel I, **a** girls, **b** boys

In Fig. 5 means and their 0.95 confidence intervals for the fundamental frequencies for vowels A, E, I, O and U are presented. Each subfigure contains a comparison of data obtained for girls before and after menarche. It can be noticed, that the mean fundamental frequency is significantly higher for girls before menarche, which has been proven using the Student's t-test (Table 1).

4 Conclusions

In the paper the methodology and preliminary results of girls and boys pubertal evaluation using voice parameters has been described. The fundamental frequency has been extracted for each recorded vowel and an analysis has been carried out that proves that this parameter is changing in the function of youth's age, for both males and females. Also the comparison of mean values of fundamental frequencies for pre-menarcheal and post-menarcheal girls had shown that this method is fitted for assessing the pubertal stage of girls. So it could be assumed that similar adaptation will be obtained for boys. The proposed method has a significant advantages compared to the current approaches; it is non-invasive, low-cost and very fast. The next planned measurement sessions, which are going to last at least three years (in each half-year), will provide more data that will allow the phase of puberty in boys to be determined. So it will possible to create a model that determine the boys' maturity degree just on a single voice recordings what is the ultimate goal of this research.

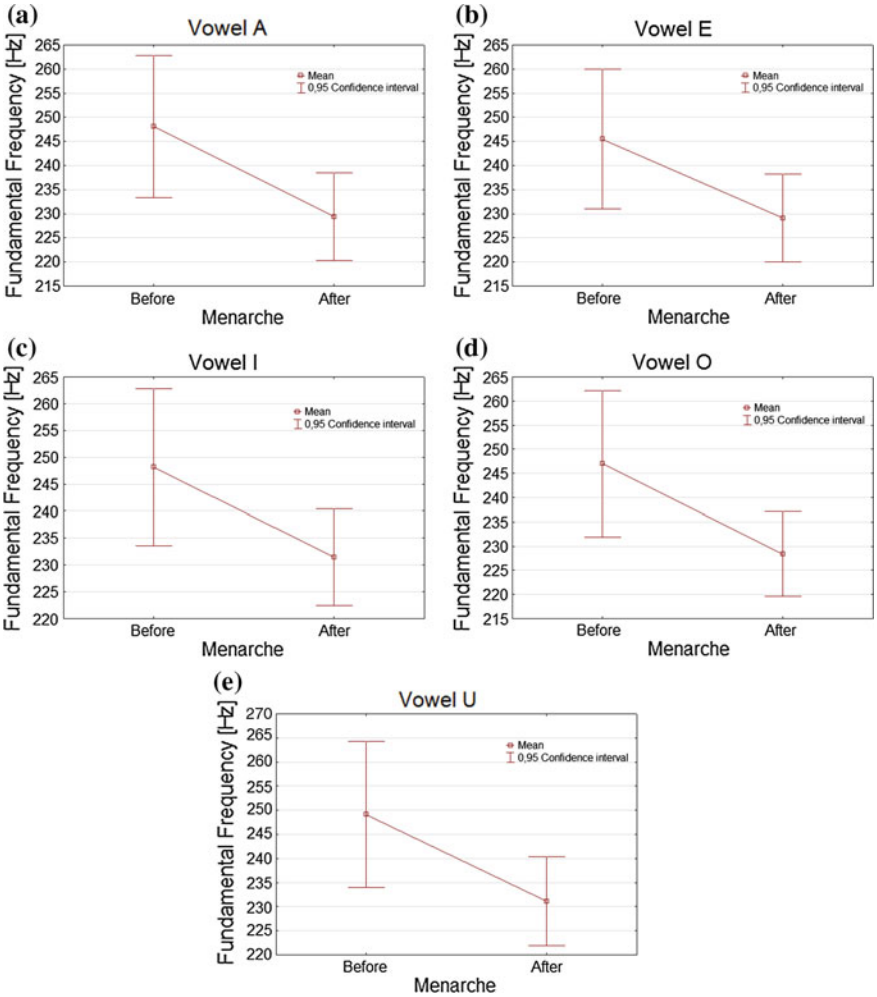


Fig. 5 Comparison of means and 0.95 confidence intervals of fundamental frequency for girls before and after menarche for the vowel **a** A, **b** E, **c** I, **d** O, **e** U

Table 1 The average values of girls' voices parameters

Parameter	Mean		Parameter change (%)
	Before menarche	After menarche	
Vowel A frequency	248.1127	229.3955	−7.5
Vowel E frequency	245.4972	229.1243	−6.7
Vowel I frequency	248.1793	231.4417	−6.7
Vowel O frequency	247.0075	228.3837	−7.5
Vowel U frequency	249.1335	231.1036	−7.2

References

1. Bugdol, M.D., Mitas, A.W.: Multimodal biometric system combining ECG and sound signals. Pattern Recogn. Lett. **38**, 107–122 (2014)
2. Chumlea, W.C., Schubert, C.M., Roche, A.F., Kulin, K.E., Lee, P.A., Himes, J.H., Sun, S.S.: Age at menarche and racial comparisons in US girls. Pediatrics **111**, 110–113 (2003)
3. Colon, I., Caro, D., Bourdony, C.J., Rosario, O.: Identification of phthalate esters in the serum of young Puerto Rican girls with premature breast development. Environ. Health Perspect. **108**, 895–900 (2000)
4. Ellis, B.J., Essex, M.J.: Family environments, adrenarche, and sexual maturation: a longitudinal test of a life history model. Child Dev. **78**, 1799–1817 (2007)
5. Eveleth, P.B., Tanner, J.M.: Worldwide Variation in Human Growth. Cambridge University Press, Cambridge (1990)
6. Fortuna, Z., Macukow, B., Wąsowski, J.: Metody numeryczne. Wydawnictwa Naukowo-Techniczne, Warszawa (1993)
7. Hacki, T., Heitmüller, S.: Development of the child's voice: premutation, mutation. Int. J. Pediatr. Otorhinolaryngol. **49**(1), 141–144 (1999)
8. Harries, M.L.L., Walker, J.M., Williams, D.M., Hawkins, S., Hughes, I.A.: Changes in the male voice at puberty. Arch. Dis. Child. **77**, 445–447 (1997)
9. Hodges-Simeon, R., Gurven, M., Cardenas, R.A., Gaulin, S.J.C.: Voice change as a new measure of male pubertal timing: a study among Bolivian adolescents. Ann. Hum. Biol. **40**(3), 209–219 (2013)
10. Hughes, A., Kumanan, M.: A wider perspective on puberty. Mol. Cell. Endocrinol. **254–255**, 1–7 (2006)
11. Juul, A., Magnusdottir, S., Scheike, T., Prytz, S., Skakkebaek, N.E.: Age at voice break in Danish boys: effects of pre-pubertal body mass index and secular trend. I. J. Androl. **30**, 537–542 (2007)
12. Mitas, A.W., Bugdol, M.D., Konior, W., Ryguła, A.: Phase angles of sound as a biometric feature. Inf. Technol. Biomed. 256–265 (2012)
13. Thurman, L.: Boys' changing voices: what do we know now? Choral J. **52**(9), 8–21 (2012)
14. Willis, E.C., Kenny, D.T.: Variability in speaking fundamental frequency in the adolescent voice. Proc. ICoMCS 172–175 (2007)

Pre and Post Menarche—Girls' Bodies Parameters Comparison

Anna M. Lipowicz, Andrzej W. Mitas, Marcin D. Bugdol, Monika N. Bugdol, Maria J. Bieńkowska, Agata M. Wijata and Dariusz Danel

Abstract The article presents changes in body composition in adolescent girls. A review of available measurement techniques is preceded by an analysis of the physiology of puberty, whose social significance is very high, especially in the present stage of development of civilization. Measurement techniques are classified depending on the used equipment, with attention to the bioelectrical impedance. Girls aged 9.3 to 18.6 y. from schools in Cieszyn were analysed. On the basis of their status of maturation they were divided into two groups: before menarche and after menarche. The comparison of body composition of pre– and post–menarcheal girls had shown the change in all body components. However, increased weight associated with the development and maturation is mainly the result of the growth of fatness, which was reflected in both the anthropometric parameters (waist, hip circumferences, skinfolds) and BIA analysis.

Keywords Girls · Bioelectric impedance analysis · BIA · Body parameters · Body composition · Maturation · Menarche

1 Puberty and Body Composition

Due do the increasing problem of a growing number of overweight and obese people there is an increased interest in the topic of changes in the human body composition. This applies particularly to children and young people, whom the frequency of people

A.M. Lipowicz (✉)
Institute of Anthropology, Wroclaw University of Environmental and Life Sciences, Wrocław, Poland
e-mail: Anna.Lipowicz@up.wroc.pl

A.W. Mitas · M.D. Bugdol · M.N. Bugdol · M.J. Bieńkowska · A.M. Wijata
Faculty of Biomedical Engineering, Department of Informatics and Medical Equipment, Silesian University of Technology, Zabrze, Poland

D. Danel
Polish Academy of Sciences, Unit of Anthropology, Wrocław, Poland

E. Piętka et al. (eds.), *Information Technologies in Medicine*, Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_33

with excess body weight increases (in some European countries) every year up to 2 % [24].

One of the critical periods relevant to shape the adult level of body adiposity and health status is a puberty. On the one hand, those children who have a higher body weight, mature earlier than slim children; on the other hand, obese child has a greater risk of being obese adults in the future. The timing of maturation may have a long-term effect on fatness, differences in fatness between early and late maturers were still evident at 30 years of age, with 26 % of early maturers being obese compared to 15 % of late maturers [10].

It is proposed, that menarche occurs when the teenager's body reaches the critical amount of body fat, which is necessary for pregnancy and lactation. Girls who gather the necessary amount of fat earlier, menstruate faster and faster are ready for reproduction [9]. It cannot be excluded that not only the total amount of fat tissue but also its distribution, especially in the lower parts of the body, plays the role in triggering the first menstruation. This is confirmed by the changes in distribution indicators during girls' adolescence, when the amount of body fat in the hips and buttocks is correlated with the age of menarche [16]. In addition, girls experiencing early menarche had greater abdominal subcutaneous fat than those experiencing menarche later, regardless of total body fatness.

2 Methods of Body Composition Measurements

Determining the body composition (including the body fat) is carried out using a variety of measurement techniques: anthropometric measurements, bioelectric impedance analysis (BIA), dual-energy X-ray absorptiometry (DEXA), magnetic resonance imaging (MRI), computed tomography (CT), ultrasonography (USG) and isotopic methods [5, 8, 12, 13, 21]. The most commonly used technique is bioelectrical impedance analysis, which is characterized by good availability and ease of measurement. BIA is noninvasive method, which allows to gain reliable information about human body composition [17] on the basis of difference in electrical conductivity of a determined frequency and low intensity ($\leq$1 mA) in water and fat compartments [12]. During examination, measurement of overall electrical impedance of the body is performed with a set of surface electrodes combined with data analysis system. This impedance is a derivative of resistance (active electrical resistance, R) and reactance (capacitive resistance X_c) [17, 22].

Analysis of registered data and results are based on electrical properties of the human body and basic physics [17]. Resistance (R) of unified object is proportional to its length and resistivity and inversely proportional to its cross sectional area. Impedance (Z) is a function of resistance (R) and reactance (X_c), which is inversely proportional to current frequency and electrical capacity of the

system [17]. Impedance formula has two different forms for serial (1) and parallel connection (2) [17]:

$$Z = \sqrt{R^2 + X_c^2}, \tag{1}$$

$$Z = \left(\sqrt{\frac{1}{R^2} + \frac{1}{X_c^2}} \right)^{-1}, \tag{2}$$

For a known resistance, height, body weight, age and sex parameters describing the composition of the human body may be determined [6, 8, 17]: Total Body Water (TBW), Intracellular Body Water (ICW), Extracellular Body Water (ECW), Body Cell Mass (BCM), Body Fat Mass (BFM), Fat Free Mass (FFM), Body Fat Distribution (BFD), Percent Body Fat (%BF), Skeletal Muscle Mass (SMM), Protein, Minerals, Body Mass Index (BMI), Target Weight, Basal Metabolic Rate (BMR), Waist-Hip Ratio (WHR), Visceral Fat Level (VFL). These parameters allow the assessment of body composition. It may be the basis for diagnosing irregularities and determining nutrition plan.

Numerous studies have reported changes in body composition during growth from pre–pubertal to post-pubertal periods. These changes concern not only the most frequently used FM (Fat Mass) and Percent Body Fat (%BF), but also Total Body Water (TBW), Minerals, Protein, Fat Free Mass (FFM), Skeletal Muscle Mass (SMM). The aim of the cross-sectional study was to describe differences in body composition between girls before and after menarche, with particular reference to body segments.

3 Materials and Methods

The study involved 47 girls aged from 9.3 to 18.6 years living in Cieszyn, Upper Silesia, Poland. On the basis of their status of maturation they had been divided into two groups: before menarche and after menarche. For all girls the body height (to the nearest 1 mm; anthropometer) and weight (to the nearest 0.1 kg; standard weight) had been measured. Next measurements of body circumferences (waist (WC), hips (HC), with an accuracy of 0.5 cm; tape tailor), thickness skinfolds (on the shoulder (TS), under the scapula (SS) and on the abdomen (AS) with an accuracy of 1 mm; caliper) had been done. All measurements had been made by one person.

Body composition was evaluated by use of bioimpedance analyzer (InBody 340 produced by Biospace). There was obtained the data of Total Body Water (TBW), Minerals, Protein, Body Fat Mass (BFM), Fat Free Mass (FFM), Skeletal Muscle Mass (SMM), Percent Body Fat (%BF) taking into account individual body segments.

The two groups had been compared using:

1. Student's t-test for independent samples [11]—if the compared groups were normally distributed (Lilliefors test [18]), had equal variances (Brown-Forsythe test [2]) and had similar cardinalities;
2. Welch's t-test [23]—if the assumption of equal variances was not fulfilled, but the two other above-mentioned assumptions were satisfied;
3. Mann-Whitney U test [4]—in the remaining cases.

4 Results of the Preliminary Study

The Table 1 contains the average characteristics of the statistics (mean, maximum, minimum and SD values) and the amount of change for the variables describing the structure and composition of girls' bodies in the two groups (before and after menarche). Pre–menarcheal girls were mean aged 11.8 years (min. 9.3, max. 18.0) and menstruating girls—aged 14.8 years (min. 11.2, max. 18.6). Menarcheal girls were higher (approximately 8.6 %), heavier (approximately 40 %) and had higher BMI (19.3 %) compared to pre–menarcheal girls. These girls have also more subcutaneous adipose tissue. From the three measured skinfolds, the largest increase had been recorded in the subscapular skinfold (on average by 102 %, from 7.2 to 14.5 mm) and abdominal skinfold (average of about 60 %, from 15.6 to 24.8 mm). At the same time waist and hip circumference had increased by 14 and 17 %, WHR increased by 5.5 % with an average of 0.79 to 0.83.

Comparing body composition, girls menstruating had not only signifcantly higher amount of adipose tissue in kg (Body Fat Mass, 82.6 %) and Percent Body Fat (more by 28.3 %), but also had about 28 % more Total Body Water, about 28 % more Proteins and 33 % Minerals. Comparison of the amount of muscle mass indicated an increase in Skeletal Muscle Mass by about 32 % in girls after menarche.

Analysis of changes in the body segments indicated that the biggest increase in body fat (BFM) has appeared on the trunk (about 120 %) and in the arms (78.9–80.6 %). In the case of the lower limbs, Body Fat Mass increased by 63 %. There was a smaller increase in Fat Free Mass and thus essentially muscle mass: about 30 % in the trunk, about 42.4–43.9 % increase in shoulder's muscle mass and 38 % increased musculature of the legs.

5 Discussion

The comparison of body composition of pre– and post–menarcheal girls had shown the change in all body components (Fig. 1). However, increased weight associated with the development and maturation is mainly the result of the growth of fat

Table 1 The average values of girls' bodies parameters (RA—Right Arm, LA—Left Arm, RL—Right Leg, LL—Left Leg)

Parameter	Before menarche				After menarche				Parameter
	Mean	SD	Min	Max	Mean	SD	Min	Max	Change (%)
Height*	149.83	11.18	133.00	169.60	162.67	6.38	141.50	172.00	8.57
Weight*	39.42	8.03	28.40	50.60	55.03	8.43	32.90	75.20	39.61
WC*	59.61	3.60	53.00	65.00	67.86	6.72	55.00	86.00	13.84
HC*	78.92	7.13	70.00	89.00	92.53	6.57	76.00	108.00	17.26
TS**	15.39	4.92	8.00	25.00	19.86	6.20	11.00	35.00	29.07
SS*	7.17	3.11	4.00	17.00	14.48	7.60	7.00	37.00	102.09
AS*	15.56	7.67	4.00	31.00	24.79	10.47	10.00	50.00	59.38
TBW*	22.93	4.76	17.80	32.40	29.32	3.23	18.80	35.50	27.87
Protein*	6.09	1.23	4.80	8.60	7.83	0.84	5.00	9.40	28.44
Minerals*	2.15	0.44	1.65	2.99	2.86	0.34	1.80	3.59	32.74
BFM*	8.23	2.77	4.10	15.20	15.03	5.45	7.30	28.10	82.56
FFM*	31.18	6.43	24.30	44.00	40.00	4.41	25.60	48.50	28.27
SMM*	16.44	3.82	12.40	24.00	21.66	2.63	13.10	26.80	31.77
BMI*	17.38	1.59	14.50	20.00	20.74	2.67	16.40	27.60	19.30
%BF*	20.75	4.80	13.00	30.90	26.63	6.10	16.10	40.00	28.34
FFM of RA*	1.28	0.42	0.80	2.07	1.85	0.30	0.89	2.42	43.93
FFM of LA*	1.26	0.41	0.77	2.09	1.80	0.29	0.88	2.32	42.41
FFM of Trunk*	13.37	2.96	9.90	18.90	17.40	1.98	10.60	21.10	30.15
FFM of RL*	4.62	1.37	2.85	7.24	6.37	0.90	3.42	7.94	38.05
FFM of LL*	4.61	1.37	2.87	7.32	6.38	0.88	3.46	7.90	38.30
BFM of RA*	0.55	0.16	0.30	1.00	0.99	0.42	0.50	2.10	80.56

(continued)

Table 1 (continued)

Parameter	Before menarche				After menarche				Parameter
	Mean	SD	Min	Max	Mean	SD	Min	Max	Change (%)
BFM of LA*	0.57	0.16	0.30	1.00	1.01	0.42	0.50	2.10	78.90
BFM of Trunk*	3.22	1.58	0.70	7.00	7.08	2.93	2.60	13.90	119.60
BFM of RL*	1.53	0.41	0.90	2.60	2.49	0.81	1.40	4.40	63.18
BFM of LL*	1.52	0.41	0.90	2.60	2.48	0.81	1.40	4.40	63.10

∗∗—significant at $\alpha = 0.05$
∗—significant at $\alpha = 0.01$

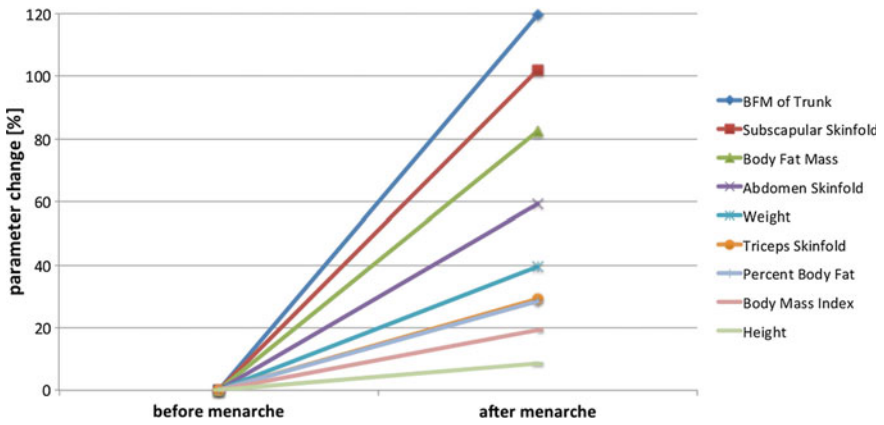


Fig. 1 Parameters change of girls' bodies composition

tissue, which was reflected in both the anthropometric parameters like waist and hip circumferences, and the thickness of skinfolds and BIA analysis.

The body composition during adolescence is an issue which human biologists are interested in. On the one hand the amount of body fat and its distribution is closely related to initiate the process of girls' maturation; on the other hand adolescence is known as a critical period for obesity development [3].

From the epidemiological point of view, it is important to indicate groups of people with higher risk of development of overweight and obesity as soon as possible. Excessive weight during pre–pubertal period, sharp gain of adiposity in the period of adolescence may be a risk factor for the development of many disease complications of obesity in adulthood.

In particular, it is known that obesity is a very important risk factor in the etiology of diseases of the cardiovascular system (hypertension, atherosclerosis, coronary artery disease, stroke) and certain metabolic diseases (diabetes type II), and therefore greatly increases the risk of premature death. From this point of view study of the connection between excess adiposity and risk of cardiovascular disease conducted in a 14-year-old youth of Lower Silesia seem to be interesting [15]. The results indicate that obese boys (BMI greater than 90 percentile) have a 2–5 times greater risk of elevated blood pressure, blood lipid and lipoprotein, compared to their peers with normal relative weight. The results of an analysis of these indicators are even more striking—obese boys have more than 16 times higher risk of raised values of three or more of the above parameters compared to their lean peers. Obese girls were burdened with more than 7 times, which generally is considered to be a risk developing cardiovascular disease in later life [15].

The nature of health problems observed in obese subjects is also dependent on the type of fat distribution in the body [14]. Visceral adiposity is (regardless of the degree of adiposity general) one of the main reasons for developing high blood pressure [1], coronary artery disease, heart failure and other cardiovascular diseases [7].

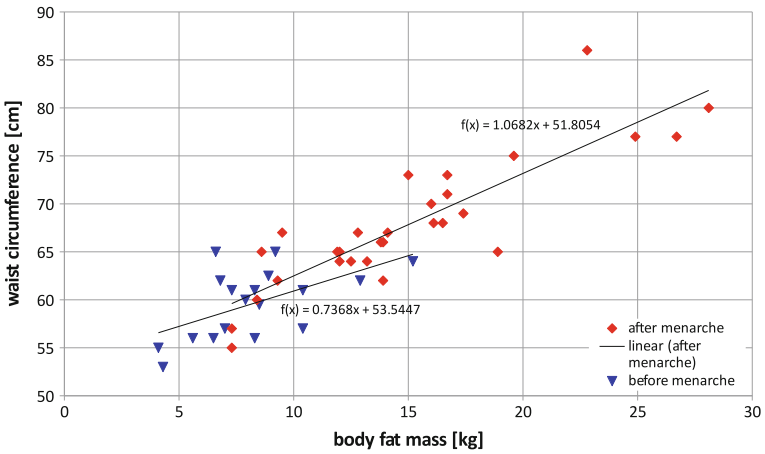


Fig. 2 The relationship between body fat mass (kg) and waist circumference (cm) in pre– and postmenarcheal girls

During puberty phenomenon of fat redistribution is observed, which involves a reduction in the level of fatness around the limbs and increase fat in the lower parts of the body. Our results of girls' body composition during puberty had confirmed a significant increase (about 120 %), in the amount of adipose tissue around the abdomen in postpubertal girls (see also Fig. 2) but at the same time increase in the amount of fat accumulated on the upper limbs (about 80 %) and the legs was observed (about 60 %).

Significant increase of fat in abdominal area of the postmenarchal girls is in line with the global trend of rising incidence of overweight and obesity, especially related to the androidal type of fat distribution [19, 20]. These changes for children and youth in XXI century in Poland are so alarming that should be continuously monitored, using all available methods, especially those of non-invasive and provide information not only about the level of adiposity general, but also about the level of adiposity of individual body segments and type of distribution of body fat.

6 Conclusions

Research methods of the assessment of changes in the human body are at the center of interest of anthropologists. Obtained results are metrological basis for the steady growth of knowledge about the dynamics of changes in the characteristics of human development in the constantly varying environmental and cultural conditions.

The material presents selected research methods for the measurement of physiological properties of a man. Due to the need to use these techniques in population studies, particularly in relation to children and adolescents, the main problem is the

friendliness of a techniques. Measurements must be non-invasive and not burdensome. Used research methods, characterized in the article, have these qualities, but taken individually have a relatively large measurement error. Their combined use provides self-correction of errors in further research, because the lack of distinctions in the one case is compensated by changes in the value of several other parameters.

Presented a package of research does not exhaust all the available or possible measurements of basic parameters of human; it is a recognized set of reference, and the detailed discussion in the science and medical equipment areas is essential in terms of illustrative and cognitive, worthy of separate study, also because of the extensivness and specificity of parameters and available measurement techniques.

References

1. Blair, S.N., Ludwig, D.A., Goodyear, N.N.: A canonical analysis of central and peripheral subcutaneous fat distribution and coronary heart disease risk factors in men and women aged 18–65 years. Hum. Biol. **60**, 111–122 (1988)
2. Brown, M.B., Forsythe, A.B.: Robust tests for equality of variances. J. Am. Stat. Assoc. **69**, 364–367 (1974)
3. Cameron, N., Demerath, E.W.: Critical periods in human growth and their relationship to diseases of aging. Yearb. Phys. Anthropol. **45**, 159–184 (2002)
4. Corder, G.W.: Foreman, D.I.: Nonparametric Statistics: A Step-by-Step Approach. Wiley (2004)
5. De Faria, E.R., De Faria, F.R., Goncalves, V.S., Franceschini, S.C., Peluzio, M.C., Sant'Ana, L.F., Priore, S.E.: Prediction of body fat in adolescents: comparison of two electric bioimpedance devices with dual-energy X-ray absorptiometry. Nutricion Hospitalaria **30**, 1270–1278 (2014)
6. De Palo, T., Messina, G., Edefonti, A., Perfumo, F., Pisanello, L., Peruzzi, L., Di Iorio, B., Mignozzi, M., Vienna, A., Conti, G., Penza, R., Piccoli, A.: Normal values of the bioelectrical impedance vector in childhood and puberty. Nutrition **16**, 417–424 (2000)
7. Ducimetiere, P., Richard, J., Cambien, F., Avous, P., Jacqueson, A.: Relationship between adiposity measurements and the incidence of coronary heart disease in middle aged men population—the Paris Prospective Study I. Am. J. Nutr. **4**, 1–8 (1985)
8. Dżygadło, B., Łepecka-Klusek, C., Pilewski, B.: Wykorzystanie analizy impedancji bioelektrycznej w profilaktyce i leczeniu nadwagi i otyłości (Use of bioelectrical impedance analysis in the prevention and treatment of overweight and obesity). Problemy Higieny i Epidemiologii **93**, 274–280 (2012). (in Polish)
9. Frisch, R.E., McArthur, J.: Menstrual cycles: fatness as a determinant of minimum weight for height necessary for their maintenance or onset. Science **185**, 949–951 (1974)
10. Garn, S.M., Lavelle, M., Rosenburg, K.R., Hawthorne, V.M.: Maturational timing as a factor in female fatness and obesity. Am. J. Clin. Nutr. **43**, 879–883 (1986)
11. Hazewinkel, M.: Student Test. Springer, Encyclopedia of Mathematics (2001)
12. Kayoung, L., Sangyeoup, L., Kim, Y., Kim, Y.: Waist circumference, dual-energy X-ray absortiometrically measured abdominal adiposity, and computed tomographically derived intra-abdominal fat area on detecting metabolic risk factors in obese women. Nutrition **24**, 625–631 (2008)
13. Kim, J., Shen, W., Gallagher, D., Jones Jr., A., Wang, Z., Wang, J., Heshka, S., Heymsfield, S.B.: Total-body skeletal muscle mass: estimation by dual-energy X-ray absorptiometry in children and adolescents. Am. J. Clin. Nutr. **84**, 1014–1020 (2006)
14. Kopelman, P.G.: Obesity as a medical problem. Nature **404**, 635–643 (2000)

15. Kozieł, S., Kołodziej, H., Lipowicz, A., Boznański, A.: Otłuszczenie, a ryzyko chorób sercowo-naczyniowych u 14-letniej młodzieży dolnośląskiej (Adiposity in the context of a risk of cardiovascular disease in a 14-year-old adolescents of Lower Silesia). Monografie Zakładu Antropologii PAN, Wrocław **19** (2000). (in Polish)
16. Lassek, W.D., Gaulin, S.J.C.: Brief communication: menarche is related to fat distribution. Am. J. Phys. Anthropol. **133**, 1147–1151 (2007)
17. Lewitt, A., Mądro, E., Krupienicz, A.: Podstawy teoretyczne i zastosowania analizy impedancji bioelektrycznej (BIA) (Theoretical basics and applications of bioelectrical impedance analysis (BIA)). Endokrynologia, Otyłość i Zaburzenia Przemiany Materii **3**, 79–84 (2007). (in Polish)
18. Lilliefors, H.: On the Kolmogorov-Smirnov test for normality with mean and variance unknown. J. Am. Stat. Assoc. **62**, 399–402 (1967)
19. McCarty, H.D., Ellis, S.M., Cole, T.J.: Central overweight and obesity in British youth aged 11–16 years: cross sectional surveys of waist circumference. Br. Med. J. **326**, 624–626 (2003)
20. Okosun, I.S., Chandra, K.M., Boev, A., Boltri, J.M., Choi, S.T., Parish, D.C., Dever, G.E.: Abdominal adiposity in U.S. adults: prevalence and trends, 1960–2000. Prev. Med. **39**, 197–206 (2004)
21. Roemmich, J.N., Clark, P.A., Weltman, A., Rogol, A.D.: Alterations in growth and body composition during puberty. I. Comparing multicompartment body composition models. J. Appl. Physiol. **83**, 927–935 (1997)
22. Rush, E.C., Puniani, K., Valencia, M.E., Davies, P.S., Plank, L.D.: Estimation of body fatness from body mass index and bioelectrical impedance: comparison of New Zealand European, Maori and Pacific Island children. Eur. J. Clin. Nutr. **57**, 1394–1401 (2003)
23. Welch, B.L.: The generalization of "student's" problem when several different population variances are involved. Biometrika **34**, 28–35 (1947)
24. World Health Organization, Branca, F., Nikogosian, H., Lobstein, T. (eds).: The challenge of obesity in the WHO European Region and the strategies for response. Summary. WHO Regional Office for Europe, Copenhagen, Denmark (2007)

Impact of Music on the Effectiveness of Performing Mathematical-Logical Tasks

Andrzej W. Mitas, Anna M. Lipowicz, Agata M. Wijata and Maria J. Bieńkowska

Abstract The article is the result of studies on the effects of music on human characteristics, with particular emphasis on the effectiveness of the performance of mathematical and logical operations. The material describes briefly some problems in the field of music therapy and presents results of a pilot study concerning possible improving of the intellectual work efficiency. A group of students who was faced with performance-related stress (standardized psychological tests and mathematical tasks) under music (Mozart and Beethoven) and no music condition was examined. Analyses showed that there was no significant difference in time of solving the tasks in silent and Mozart's music condition. Also, time of performing exercises was longer during Beethoven stimulation than during silent condition. It may be concluded that commonly applied Mozart's work does not always give the expected results and the other work can have stronger (but not necessarily better) effect on human being.

Keywords Mozart effect · Music impact · Work effectiveness · Mathematical and logical operations

1 Introduction

Music is commonly used in medicine for many centuries ranging from ancient Greece. However, only in recent years, discoveries in medicine and psychology, supported by equipment biocybernetic led to a new study on the impact of music

A.W. Mitas (✉) · A.M. Wijata · M.J. Bieńkowska
Faculty of Biomedical Engineering, Department of Informatics and Medical Equipment, Silesian University of Technology, Zabrze, Poland
e-mail: Andrzej.Mitas@polsl.pl

A.M. Lipowicz
Department of Anthropology, Wroclaw University of Environmental and Life Sciences, Wrocław, Poland

E. Piętka et al. (eds.), *Information Technologies in Medicine*, Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_34

on human psychophysical condition and the interpretation of the observed phenomena [1, 17, 18, 24, 25].

Evaluation of the influence of music therapy on human mental system can be done on several ways. Traditionally questionnaires are applied to recognize direct, although the subjective feelings of the patient (a number of standardized tools like e.g. Personal Feeling Survey which asses tension, mood and *mental clarity*) [16]. On the other hand objective condition of the patient can be tested using the biomedical equipment.

Objective evaluation of the impact of music on the physical condition of human may be carried out with measurements of parameters characterizing biomedical physiological state of the subject. One of these is the electrical activity of the brain (EEG), which enables the observation of brain stimulation using music [5]. The heart rate also slows under the impact of music, which may indicate a relaxed state of the respondents [29] and also decreases the power spectral density calculated from EMG [12]. Properly chosen music also reduces skin conductance (GSR) [26] and increases heart rate variability (HRV) [2]. The influence of music on human condition also may be assessed by measuring the concentration of certain biomarkers in the blood or saliva. For example, the concentration of immunoglobulin A in saliva decreases under the impact of relaxation music, after being increase as a result of the stressful tasks [19]. Furthermore, it was observed that music inhibits cortisol in saliva during a stressful situation [14].

Over the years, there were repeated attempts to study the music impact on improving efficiency at working and learning. Jausovec had analysed [9] influence of Mozart's music on brain activity during the learning process and demonstrated that Mozart's music, by activating task-relevant brain areas, enhances the learning of spatio-temporal rotation tasks. Simultaneously, the observed EEG patterns indicated beneficial effect of Mozart music exactly, non music in general context [9]. Also, research conducted in Estonia showed positive changes in the evaluation of public school students who listened to classical music in the educational process [13]. Similarly, students achieved significantly higher scores in tests after the lecture where at the same time listened to classical music (Mozart, Vivaldi, Bach, Tchaikovsky, Beethoven, Strauss) [3]. In spite of studies that have showed the importance of music to improve efficiency of short-term memory, especially during test examining cognitive tasks, learning and spatio-temporal abilities, so called the *Mozart effect* not yet been demonstrated or explained the mechanisms of the action of music. Therefore, the investigation of the impact of music on learning results should be continued using available equipment and knowledge as well as examining the impact of acoustic stimuli to very specific abilities [3]. In the light of this expectation the aim of this pilot study was to determine the effect of music on the effectiveness of performing mathematical-logical tasks.

2 Materials and Methods

The study group had consisted of thirty individuals (16 men and 14 women) aged from 20 to 25 years, students of Silesian University of Technology, Zabrze, Poland. The eligibility criteria had included: no diagnosed defect of hearing or vision acuity, or having a vision defect effectively corrected by glasses or contact lenses. Also they were free from overt diseases.

2.1 Selection of Sound Stimuli

Selected sound stimuli were music of two composers from the same cultural background: Mozart and Beethoven. Analysis of the impact of music on the human in any case is not simple, because music is a sound wave which amplitude varies in time and space. In most studies, more important for the selection of music for the experiment is a level of musical culture, characteristic for the cultural group (e.g. the world of culture of the East or the West), than actual impact on the psychosomatic system of man. Usually selected songs belong to so called music standards (popular and widely recognized). Each music of this type is characterized by a large number of parameters—degrees of freedom (such as rhythm, melody, harmony, etc.). It is difficult to determine a priori which of these elements of music is determinant of changes in human body [27].

The problem of selection of music stimuli is the subject of much reflections and experiments. The superficial knowledge of music history usually indicates us (when testing the effect of music on human) to choose music, which are limited catalog of popular pieces. The lack of an obvious reference to the elements of music is the result of the variability of behavioral responses only to the same mode (major or minor). Author's experience suggests tendency to select a specific mode depending on the personality. For serene person *Sonata for Two Pianos in D Major*, K. 448 Allegro con spirito by Wolfgang Amadeus Mozart is exciting, because the definition contains (spiritual—con spirito) allegro sonata, which is part of the fast and dynamic, due to the horizontal density of sounds. This piece, which expression and virtuosity stimulates the listener, can not certainly be classified as a sedative. Frequent thematic expositions, elevated mood and dynamics of the many fragments of a classical term definitively determine the dynamic pulse of the song, which was written by quite young (25 years old) composer barely holding their romantic emotions in the classical boundaries. Therefore, the first chosen stimulus was *Sonata for Two Pianos in D Major*, K. 448 Allegro con spirito by Wolfgang Amadeus Mozart. It is commonly used in experiments, which are intended to demonstrate the influence of music on human being (the *Mozart effect*) [9, 10].

However, for a human of introverted disposition (or melancholy mood) piece by Mozart may be downright annoying, for example due to mode, especially in the key of D, which is elementary in the context of tonal lying in diameter tonal.

Metro-rhythmic element of this sonata is, nota bene, comparable to the widely used form described in the research work of Beethoven—Egmont Overture.

The choice of work *Egmont*, Op. 84 by Beethoven is primarily due to its strong emotionality. Overture was composed as an illustrative music to the works of Goethe, therefore the fundamental interpretation of this work should be searched in this outstanding poetry. This overture has several music plans: the growing up, or rather accumulation of emotion, through diatonic sequences to the top of the piano scale, which are periodically followed by brilliant passages to the lowest sounds, which release tenseness and cause affection and emotion. A F-minor tonality in this music is obviously very dramatic (as in a later Chopin's concert), however, a F-major phrase, though rare, create in the overture Egmont the background of joyous expectation for happy ending, which will crown an existential despair. The same sound stimulus was used by Jackson [8]. Instead of mentioned piece of Beethoven one could suggest a well-known (unfortunately fragmentary) *Piano Sonata*, Op. 23, from the initial (strictly classical) period of the author, but the limitation to *exploited* the first part would be rather completely incompatible with a piece of Mozart and research objective, stimulate higher (nervous?) reaction. It was not until the end (unusually) allegro sonata inflames emotions listener. Complementarity stimulation is therefore a prerequisite for the study and requires careful analysis of the formal.

The idea of the experiment presented in this work was a reference to the so called *Mozart effect*, which is associated with the music of this composer (in particular, by people who know the directory selected a few pieces, too often *used* in the media, because even the ease of performing and more popularity). Using a little younger, but historically consistent Beethoven is merely a verification of the hypothesis of the uniqueness of European classical music (completely detached from music world). Research has however caused, how to choose appropriate music elements that strongly influence the listener, regardless of the level of knowledge and personal musical culture.

2.2 Procedure for Experiment

Participants had solved four tasks under three conditions: during two different music stimuli and during silence (three phases).

Tasks were divided into two groups:

- Raven's Progressive Matrices (RPM): Standard Progressive Matrices (SPM) [21] and Colored Progressive Matrices (CPM) [20];
- two arithmetic operations.

Two types of the Raven's Progressive Matrices (RPM) were used in the following order: Standard Progressive Matrices and Colored Progressive Matrices. RPM is test of nonverbal reasoning with some visual-spatial elements typically used in educational settings as standardized tool [6]. RPM is also used in studying the effect

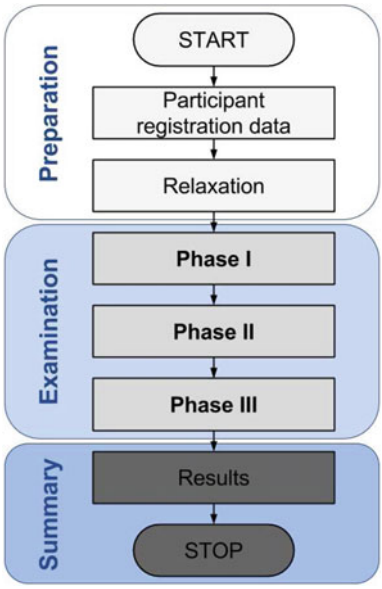


Fig. 1 The algorithm of experiment

of music on the human being [6]. The arithmetic tasks had consisted of two mathematical operations which difficulty should not be problematic for students after the basic course in mathematics at the university.

According to the above, experiment had consisted of 3 phases (Fig. 1):

- Phase I: silence;
- Phase II: *Egmont*, Op. 84, by Ludwig van Beethoven [8];
- Phase III: *Sonata for Two Pianos* in D Major, K. 448 Allegro con spirito, Wolfgang Amadeus Mozart [9, 10].

The application has been implemented for the study of time solving the tasks during acoustic stimulation. In order to standardize the measurement conditions, a fixed sequence of phases and tasks assigned to them was adopted (Fig. 1). The design of the test stand is presented in Fig. 2.

2.3 Testing Procedure

The participant of the experiment had sat in a comfortable position, allowing free access to the mouse, monitor and speakers. The participant had encouraged to adjust the volume to a comfortable level, which was clearly audible but not disturbed. The subject had been only informed about the application but the measured parameters were withhold from him. His task had been to do the tests in the shortest possible time.

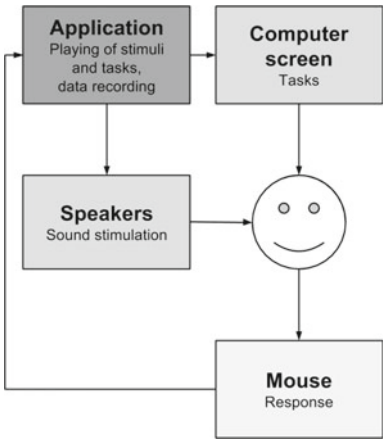


Fig. 2 Test stand schematic

Each participant's response time had been measured from the moment when the task had been displayed until the participant had given the answer. The correctness of given answers had been recorded at the same time. The user had introduced a data regarding the age, sex and occupation. Participant had pressed *Start* and taken five deep and relaxing breaths with closed eyes. The participant had started the first phase: after 30 s of silence task 1 had been displayed, which was Standard Progressive Matrices and then Colored Progressive Matrices (task 2). The next two tasks were simple arithmetic operations (tasks 3 and 4). Each task had been implemented in the form of a test with four answers. Participant had chosen the answer by pressing the appropriate button. In the phases II and III music stimulation had lasted 4 min before the presentation of the task 1 and during solving all the tasks (tasks 1–4). Performing the last of the tasks in each phase meant the completion of that phase and muting the sound file for phases II and III.

Verification of normal distribution of recorded data was performed using the Kolmogorov-Smirnov with Lillefors's amendment and the Shapiro-Wilk tests. Verification of differentiating influence of phase of the experiment on time solving tasks was carried out using dependent t-test for paired samples. Verification of influence of the music for solving tasks including the impact of gender was conducted using Kruskal-Wallis one-way analysis of variance and Dunn's post hoc method.

3 Results

In the experiment values corresponding to the times-solving task, and success or failure in addiction used the sound stimuli (phase II and III) had been recorded. In the analysis 360 dual values corresponding to time of solving the problem and the correctness of the answer had been used. Expressing formal description had obtained:

Table 1 Mean time of solving tasks in phases in seconds

	Task 1	Task 2	Task 3	Task 4	Mean	SD	Min	Max
Phase I	5.99	6.06	9.32	5.88	6.81	6.14	1.64	49.22
Phase II	10.03	8.61	9.64	4.86	8.28	9.47	1.12	54.62
Phase III	8.95	8.91	6.75	3.75	7.09	8.68	1.35	64.51

$$n_1 + n_2 + \cdots + n_i = 30[subjects] \cdot 3[phases] \cdot 4[tasks] = 360, \quad (1)$$

where:
n_i is the number of the subject, who solved 12 tasks in three phases.

Average times of solving task in subsequent phases are placed in Table 1. The average time of solving problem in phase I was 6.81 s, in phase II 8.28 s, and in phase III 7.09 s. The minimum value was 1.12 s and the maximum value of 64.51 s.

The lowest average time of solving task had been recorded for the task 2 (CPM) of each of the phases. The reason for this could be the motivation associated with the desire to completion of the experiment. Average time of solving mathematical problems had been shorter during listening to Mozart piece, which could be the result of a favorable stimulus. The average times for CPM in phase II and III had been longer than SPM as opposed to the phase I. It is possible that stimulation of sound made difficulties with the perception and processing of more colors.

Figure 3 presents the average duration of solving tasks in the various phases registered for each of the study participants, arranged in descending order separately for both sexes. In each phase, men had solved a set of deductive and mathematical tasks significantly slower as compared with women.

To check whether records were normally distributed the Kolmogorov-Smirnov with Lillefors's amendment and Shapiro-Wilk tests were performed. They had obtained p-value <0.05: recorded times of solving tasks were not normally distributed. Conducted Kolmogorov-Smirnov test with Lillefors's amendment and the Shapiro-Wilk test for individual subjects had taken the p-values >0.05, which allows the conclusion that the times recorded of solving tasks for individuals had a normal distribution.

To verified whether times of solving tasks in conditions of silence and acoustic stimulation did not differ from each other the t-test for dependent samples was used. For the adopted level of significance $\alpha = 0.05$, there were no statistically significant differences for Phase I and Phase III (Table 2). A statistically significant difference was observed when Phase I and Phase II, Phase II and Phase III were compared (Table 2). The task solving during track of Beethoven stimulation was significantly longer as compared to the other two phases of the study.

To investigate whether times of solving mathematical and logical tasks significantly differed depending on the phase of the experiment Kruskal-Wallis two-way analysis of variance was performed, where the independent variables were the phase of the experiment and gender. The Kruskal-Wallis test had demonstrated statistically

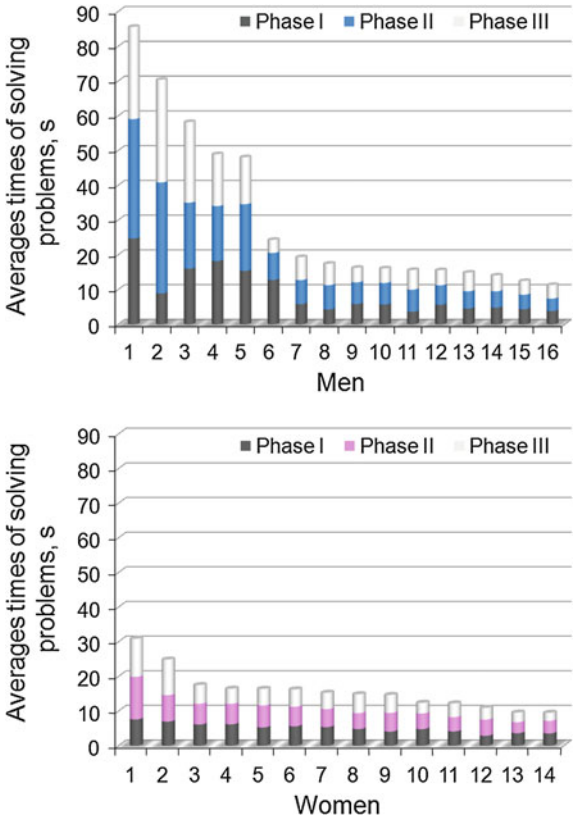


Fig. 3 Mean times of solving problems for men and women in individual phases

Table 2 Dependent t-test for paired samples—p-value

	Phase I	Phase II	Phase III
Phase I	–	0.0354	0.6846
Phase II	0.0354	–	0.0349
Phase III	0.6846	0.0349	–

significant differences between the sexes for Phase II. In the Phases I and III, times of solving for men and women were similar and statistically non-significant. Also test post-hoc Dunn had demonstrated statistically significant differences in times of solving tasks between sexes for Phase II. Statistically significant differences did not occur in the case of Phases I and III.

Number of wrong answers given during solving mathematical and logical tasks had also been calculated. Wrong answers had accounted for 6 % of all answers given during the study. The greatest number of errors had been observed in Phase II (with music by Beethoven)—accounted for approx. 44 % of all incorrect answers.

4 Discussion

In this study we have shown that the results from the tests solving and deductive mathematical tasks depend on the type of acoustic stimulus and sex of respondents (Fig. 4). Women significantly faster solved all the tasks, regardless of the type of sound stimulus. On the other hand the arithmetic and logic operation time during stimulation of piece by Beethoven was significantly longer than the task in conditions of silence and stimulation of Mozart's piece. At the same time the respondents during listening to music by Beethoven committed the most errors.

Slightly longer, but not statistically significant, the time to solve problems for men in the phase of the silence might result from the fact that area of the brain that acitivates in response to stress varies among men and women. Authors found that different parts of the brain activate with different spatial and temporal profiles for men and women when they are faced with performance-related stress [28]. So, in the case of women, centers in the brain, which are responsible for solving problems and stress response, are stimulated independent. Therefore it could be postulated that for women musical disturbances during mathematical tasks could not influence on results. Irritation of the process of problem solving under stress due to the stimulation the same region of the brain at the same time might occur in men. In the studies [11, 15] authors analyzed sex differences in stress reactivity primarily relied on measuring physiological responses to acute stressors in laboratory settings, including activities of the hypothalamic-pituitary- adrenal (HPA) axis (e.g. cortisol) and/or sympathetic nervous system (e.g. heart rate and blood pressure). In adult men, greater acute HPA and autonomic responses were observed as compared to adult women using standard psychosocial stressors such as public speaking and arithmetic tasks, which may explain the slightly longer time to solve tasks by men in the pilot study.

In the interpretation of the results may be helpful socio-biological theories and memetics (the science of dual inheritance theory, DIT). These references enable to understand evolutionary ground including influence of music on learning and memory. It should be remembered that all senses of living organisms, especially hearing and sight, irrespective of species, have the same function: to provide information about the surrounding environment, e.g. about a life-threatening dangers [7].

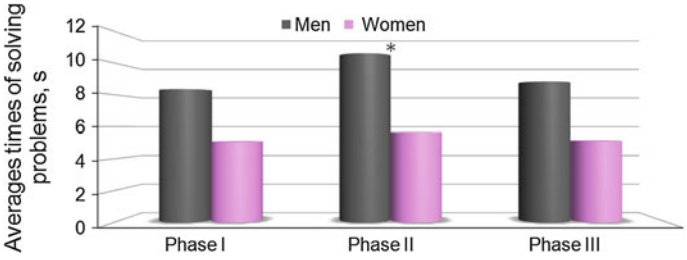


Fig. 4 Times of solving problems for men and women in individual phases; level of significance $*$ p-values <0.05

In the current world the importance of evolutionary (biological) sounds coming to ear ceases their importance, the sounds no longer carry the essential information about the threat to life (besides the risks associated with an oncoming car). They still cause deeply biological specific reactions in the brain which leads to stimulation of the hypothalamic—pituitary—adrenal axis and the autonomic nervous system. Secreted hormones (e.g. cortisol, epinephrine) cause increasing of heart rate and respiratory rate, increased blood pressure, sweating the skin.

Music, however, is more or less predicted rhythmic patterns that species *Homo* created from the beginning of its existence. It can be assumed that these were probably e.g. beating into hollow wood or singing, which became complicated and were enriched by more and more complicated in their build.

With combining the evolutionary significance of sound stimuli and the development of music it may be hypothesized that the primary role of music was not only the desire to express emotions of the creator and to experience pleasure, but also to inform about the level of safety in the surrounding environment: music rhythm as predictable in the structure and composition gave, both in prehistoric times, and gives now a feeling that the environment is safe and nothing threatens the lives of listeners [23]. In contrast, the music which is too monotonous or repetitive, carries very little essence and may bore listeners more or less. On the other hand, it was observed that music overcomplicated with a long-phrased, unpredictable, violent, with a high degree of saturation, both in amplitude and density of sound in time (e.g. the overture *Egmont* by Beethoven, used in the second phase of the experiment) may be perceived as unpleasant and disturbing [4, 22]. Such music enhances our excitation, causes negative emotions and may be subconsciously regarded as a potential indicator of a dangerous environment. It is confirmed by Schäfer et al. [23] who had investigated if music, as a form of auditory information, can trigger the experience of safety and showed that there is an optimal, subjectively preferred degree of information density of musical sounds, at which safety-related information can be processed optimally, and the preferred tempo decreased for increasingly complex rhythms. They had underlined that people still prefer music not too predictable and not too variable [23].

The results of the experiment do not entitle also easily ascertain so called *Mozart effect* in the measurement conditions (type of music, sampling and others) because the time tasks while listening to the music of this composer did not differ significantly from time performing the same tasks in silence.

To summarize the results of the study on the impact of music on humans, especially in the context of *Mozart effect*, it is worth to pay attention to some important elements of the non-deterministic character. Influence of music on the psychosomatic organism of human should be considered as a multidimensional random variable. The analysis must take into account the variability of the physical characteristics of music: the construction of formal accumulation of sounds in vertical and horizontal diastematic and characteristics of articulation that result from methods for its preparation. Accordingly to the sentence—*primum non nocere*—it must be taken into account that the inter-individual variability of reaction to stimuli, depends for example on gender, age, or social and educational conditions. It is reasonable to presume

that a level of musical culture (patient and therapist), which determines the approval of a sound set by the recipient of a certain type of personality and sensitivity has a‘particular role in music therapy.

Reflections about the influence of music should be performed at the border of so many areas, including health, education and culture. Therapeutic, which is assumed positive dimension of music, may be seen in the context of its impact on reception, perception and deductive human capabilities. However keep in mind that this effect may be negative, as demonstrated in this paper on the example of influence of *violent music* on the effectiveness of mathematical and logical tasks or even dangerous. Additional sound stimuli for people with latent form of mental deficits, may be not only unpleasant but downright harmful.

This paper is an attempt to determine the effect of music on human psychophysical condition, supplemented authorial analysis of the anthropological perspective. Evaluating the efficiency of solving simple operations, requiring the obvious involvement of the mind, in the presence of sound stimulation, provide a link with experiments designed to observe the effect of music on short-term memory. Instead about the *Mozart effect* would rather discuss in terms of the functioning of long-term memory, as well as the efficient use stored information [9].

The arbitrary selection of music that have been used in the experiment, requires a separate comment. Presented in the literature *Mozart effect* is sometimes transposed on *effect terms of the music of Mozart*. The reason for this discrepancy in definition is the impact of music on certain parameters: metro-rhythmic, harmonic, melodic and agogic (also taking into account other elements of music). Generally, it is not a case of particular composer. The discussion on the selection of specific pieces from extremely rich catalog of creativity of the most famous composers should be more focused on the selection of the elements of music, exposed in these pieces than on concrete *recognized* the work, because assessments of music critics *sometimes* overshoot the social level of applause.

References

1. Chanda, M.L., Levitin, D.J.: The neurochemistry of music. Trends Cogn. Sci. **17**, 179–194 (2013)
2. Chiu, H., Lin, L., 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. Comput. Cardiol. **30**, 469–472 (2003)
3. Dossevillea, F., Labordea, S., Scellesa, N.: Music during lectures: will students learn better? Learn. Individ. Differ. **22**, 258–262 (2012)
4. Edworthy, J., Loxley, S., Dennis, I.: Improving auditory warning design: relationship between warning sound parameters and perceived urgency. Hum. Factors **33**, 205–231 (1991)
5. Frolov, M., Milovanova, G., Mekhedova, A.: Effect of musical accompaniment on the operation performance of subjects with different anxiety levels. Hum. Physiol. **31**, 164–170 (2005)
6. Hetland, L.: Learning to make music enhances spatial reasoning. J. Aesthet. Educ. **34**, 179–238 (2000)
7. Horowitz, S.S.: The Universal Sense: How Hearing Shapes the Mind. Bloomsbury Publishing (2012)

8. Jackson, N.: The Little Book of Music for the Classroom: Using Music to Improve Memory, Motivation, Learning and Creativity (Independent Thinking Series). Crown House Publishing (2009)
9. Jausovec, N.: The influence of Mozarts music on brain activity in the process of learning. Clin. Neurophysiol. **117**, 2703–2714 (2006)
10. Jenkins, J.S.: The Mozart effect. J. R. Soc. Med. **94**, 170–172 (2001)
11. Kajantie, E., Phillips, D.I.: The effects of sex and hormonal status on the physiological response to acute psychosocial stress. Psychoneuroendocrinology **31**, 151–178 (2006)
12. 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, vol. 3, pp. 449–452 (2007)
13. Kährik, P., Leijen, A.: Developing music listening skills using active learning methods in secondary education. Procedia: Soc. Behav. Sci. **45**, 206–215 (2012)
14. Khalfa, S.: Dalla Bella, S., Roy, M., Peretz, I., Lupien, S.J.: Effects of relaxing music on salivary cortisol level after psychological stress. Ann. New York Acad. Sci. **999**, 374–376 (2003)
15. Kudielka, B.M., Kirschbaum, C.: Sex differences in HPA axis responses to stress: a review. Biol. Psychol. **69**, 113–132 (2005)
16. McCraty, R., Barrios-Choplin, B., Atkinson, M., Tomasino, D.: The effects of different types of music on mood, tension and mental clarity. Altern. Ther. Health Med. **4**, 75–84 (1998)
17. Metera, A.: Muzykoterapia. Muzyka w medycynie i edukacji (Music therapy. Music in medicine and education). Wydawnictwo Centrum Technik Nauki Metronom (2006) (in Polish)
18. Mitas, A.W.: Muzyka w aspekcie biocybernetycznym (Music in biocybernetic aspect). Muzyka w dialogu z edukacją. Wybrane konteksty aktywności i edukacji muzycznej (Music in dialogue with education. Selected contexts of music activity and education) (red. E. Kochanowska, R. Majzner). Wydawnictwo Libron (2015) (in Polish)
19. Nomura, S., Tanaka, H., Nagashima, T.: A physiological index of the mental stresses cased by the deskwork with VDT and the relaxation by music. In: Sensors and the International Conference on New Techniques in Pharma- ceutical and Biomedical Research, pp. 130–134 (2005)
20. Raven, J.C.: Coloured Progressive Matrices. Oxford Psychologists Press (1976)
21. Raven, J.C.: Standard Progressive Matrices. Oxford Psychologists Press (1976)
22. Rosenblum, L.D., Wuestefeld, A.P., Saldana, H.M.: Auditorylooming perception: influences on anticipatory judgments. Perception **22**, 1467–1482 (1993)
23. Schäfer, T., Huron, D., Shanahan, D., Sedlmeier, P.: The sounds of safety: stress and danger in music perception. Front. Psychol. **6**, 1–9 (2015)
24. Schwabe, Ch.S.: Leczenie muzyka chorych z nerwicami i zaburzeniami czynnościowymi (Music therapy in patients with neuroses and functional disorders). Wydawnictwo Lekarskie PZWL (1972) (in Polish)
25. Skotnicka, M., Mitas, A.: About the measurement methods in music therapy. Information Technologies in Biomedicine, vol. 4. Advances in Intelligent Systems and Computing, vol. 284, pp. 203–214 (2014)
26. Thayer, J.F., Levenson, R.W.: Effects of music on psychophysiological responses to a stressful film. Psychomusicology: J. Res. Music. Cognition **3**, 44–52 (1983)
27. Trost, W., Ethofer, T., Zentner, M., Vuilleumier, P.: Mapping aesthetic musical emotions in the brain. Cereb. Cortex **22**, 2769–2783 (2012)
28. Wang, J., Korczykowski, M., Rai, H., Fan, Y., Pluta, J., Gur, R.C., McEwen, B.S., Detre, J.A.: Gender difference in neural response to psychological stress. Soc. Cogn. Affect. Neurosci. **2**, 227–239 (2007)
29. Zhou, P., Sui, F., Zhang, A., Wang, F., Li, G.: Music therapy on heart rate variability. BMEI **3**, 965–968 (2010)

Model of Attenuation of Sound Stimuli in Prenatal Music Therapy

Maria J. Bieńkowska, Andrzej W. Mitas and Anna M. Lipowicz

Abstract Studies on the process of acoustic stimulation in prenatal period require to assess the beneficial condition for child development, as well as characteristic sounds which could negatively affect the development of auditory system, brain or even entire body. It can be provided with the simplified model of sound attenuation in the uterus, which can be represented by first order passive low pass filter. On the basis of total harmonic distortion analysis for a single sound, a proposal indicating the sound reduction in the intrauterine environment is formulated. It is hypothesised that the sound stimuli in the uterus is not the same sound as the music in the source.

Keywords Music therapy · Prenatal development · Sound attenuation

1 Introduction

Fetal environment, woman's and other mammal's uterus, is an unique, perfect place to live for immature, dynamically developing young organism. The inner condition has evolved in such a way, that it created an optimal safe place, which is isolated from harmful external stimuli, which could therefore damage or lead to death of the fetus. This fluid environment is characterized by almost constant temperature. The fetus has a source of supply of oxygen and nutrients provided by the mother's blood and passed through placenta, fetal waste products are also eliminated. The characteristic

M.J. Bieńkowska (✉)
Faculty of Automatic Control, Electronics and Computer Science,
Silesian University of Technology, Gliwice, Poland
e-mail: maria.bienkowska@polsl.pl

A.W. Mitas
Faculty of Biomedical Engineering,
Department of Informatics and Medical Equipment,
Silesian University of Technology, Gliwice, Poland

A.M. Lipowicz
Institute of Anthropology,
Wroclaw University of Environmental And Life Sciences, Wrocław, Poland

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_35

feature is also fetus' isolation from external environment by attenuation of stimuli quantity that reach inside the uterus. Taking into account the contemporary world, the anxiety may rise, because the noisy habitat of mothers, connected with excessive number of stimuli e.g. sound stimuli, can have harmful impact on future generation's development.

Pregnant women usually feel fetus' movements in reaction to loud noises or rhythmical sounds [11, 12]. We can assume that if fetus reacts to such a stimuli, it must receive them. Therefore, we have to be sure that the stimuli which reach fetus wouldn't be harmful for its health and development.

The studies indicate that music, or more specifically, certain sounds, have got an impact on the fetus' state. As described in [1] listening to music during pregnancy is creating neurons connections that *transport thoughts and information*. It is supposed to improve newborn's intelligence which is incorrectly linked to a Mozart's effect. The phenomena of the Mozart's effect had been defined in 1991 by Tomatis, and it had been researched in 1993 by Frances Rauscher et al. who observed the improvement in test results by people listening to this compositor's pieces while solving the task. The improvement however, was only temporary and it could also be achieved in groups listening to other composers. Nevertheless, Mozart's music is so various in form, articulation, dynamics, agogics or metro–rhythmics that it is hard to tell that the impact on human is homogenous.

Intrauterine environment, in which the fetus is developed, unlike the external one, has got little stimuli, and basic sounds that reach the fetus are coming from mother's body. They track down from cardiovascular, digestive, and respiratory system activity. Also some of the sounds are generated by mother's movement. Other stimuli are received from the outside (traffic, human voices, tv, music, etc.). However, no matter what music would be applied, we should not define such a stimulation as music (at least from the child point of view), because sound stimuli which reach the fetus does not have these features, which make music from sequence of sounds. The problem of aesthetic equivalence of music both external and fetal environment, should be resolved before the discussion about creationism triggered by aesthetic experiences will start.

A wealth of new research has indicated that sounds produced by maternal body consist mainly of low frequency components. The lowest frequency reaches the level of 60 dB (some sources indicate up to 90 dB [11]), which for higher components is decreased to 40 dB [3, 12]. Such stimuli from internal fetal environment can drown out external sounds (e.g. speech). The above–mentioned situation takes place when the level of sounds derived from maternal digestive, respiratory and cardiovascular systems excesses (in certain frequency bands) the level of external noises by at least 10 dB [11].

The development of the auditory cortex is strictly dependent on the acoustic stimuli, also fetal hearing system needs appropriate stimulation for its proper growth [13]. The research conducted on animals has shown, that diversified acoustic environments affect on the tempo of the auditory cortical development [8]. On the other hand, exposure for excessive noises has negative impact, for example, on vocal abilities or spatial orientation of birds. In humans, lack of early auditory stimulation can

lead to complete loss of functionality in brain's hearing area [9]. Simultaneously, the lack of stimuli on each stage of ontogenesis (also prenatal period) can lead to abnormalities in entire organism's development and also to fluctuation in cognitive and social processes.

Therefore, the research on the process of acoustic stimulation in prenatal period requires the assessment of the beneficial condition for child's development, as well as the characteristic sounds which could negatively affect the development of auditory system, brain or even entire body. The child, who is exposed on listening to intensive low frequency sounds in prenatal period, can get retarded in future or has problem in language abilities [13].

2 Research Objectives and Tasks

Knowledge about behavior of human, who is stimulated with sound, is still insufficient. It is due to the large number of degrees of freedom while creating the sound set process (also in the form of music), differences in sound generation (e.g. humidity which determines woodwinds instrument *tuning*), sound transfer (e.g. obese individual with relatively low muscle mass on the transmission path of the acoustic wave), current listener's disposition (e.g. sobriety, efficiency and cleanliness of the ear canal) and his music culture (the topic is quite disputable after many years of eliminating art and music at Polish schools of any degree).

In these circumstances, installation of a belt with built–in speakers in order to deliver the sound beats to the child in the womb is at least ambiguous. The piece of information about the remainings of the generated music (which selectively reach the fetus) is dependent on such factors like the thickness of maternal abdominal wall, the level of adiposity, or supposed fetus' musical sensitivity. The topic of child's aesthetic during prenatal development cannot be treated marginally if we hypothesize about superiority of classical music (especially of the one composer) to the other artistic products. Considerations about this superiority without a context (which provides knowledge in history of art and mother's music culture), may be caused by the fact that Mozart is the most famous composer in Europe. Authors are surprised that attention is concentrated on just a few pieces which are mostly played. Unpopular compositions are ignored, what should therefore result in ranking of the pieces characterized by special predispositions.

What is worth considering is the hypothesis that music which is *stripped* of details is no longer a music which we like. It might be compared with a women from a journal cover, who is not *convincing*, because of being made hairless by the influence of the computer programme (or in the case of promoting baldness she has her hair added by a computer).

First of all we should determine which part of the acoustic wave, which is hearable in maternal environment, can reach the fetus in uterus.[1]

The aim of this study was to determine characteristics of pregnant women's abdominal walls attenuation. Electrotechnical model, which was used for analysis, allows to designate frequency characteristics for sounds received in female's body. On this basis, the physical status of some musical pieces from the point of view of human susceptibility to changes (and their values) of certain musical parameters can be specified with categories of musical degree of freedom (rhythm, melody, harmony, dynamics, agogics etc.). Such formulated task is especially important in the context of danger avoidance which is connected with the impact of uncontrolled stimuli.

Apart from anything, a priori exaltation of any composer as an author of the most *influential* sounds is unjustified in the context of the possible parent's origin from the particular music culture, where European inaccurate musical scale is unknown. This exaltation would cause less chance of intellectual development for the residents of the Far East countries with their traditional musical scale. Fortunately, this is contradictory with the list of the authors' names of the most cited journals which are indexed in the Journal Citation Reports.

3 Sound Attenuation Modeling

Creating a model of sound attenuation on the path to hearing system of fetus requires several assumptions. The measurement of sound intensity in fetal environment is difficult from the technical, ethical and medical point of view. Reliable data is derived from living organism, in which work of digestive, cardiovascular and respiratory system is not disrupted. The input of a measuring device is extremely difficult in such condition, and what is more, any interference in fetal environment can lead to serious objections of doctors, ethics and ecologists. Therefore, such measurements are conducted mainly on animal models, usually on ewe. Ewes are considered as a reliable model for researchers focusing on the attenuation of the abdominal wall [2].

When hydrophone is placed in amniotic fluid or cochlear microphone is placed in the inner ear, a summary data of sound, which passed through several barriers, is received. It should be remembered that sounds coming from external environment have to pass through several barriers of different density. The first one is the air between the sound source and the mother's abdomen. The level of the sound decreases proportionally to the square distance from the source. Therefore, a silent sound which has a source far from mother's abdomen will be received with very small amplitude. We should remember that a loud sound, which ensure high amplitude level near the abdomen wall, is not a good solution. Such a sound can cause auditory damage or bad mood to the mother. Acoustic wave after passing the distance between the

[1] Team of authors refers to placing loudspeakers near to the fetus [19] with a reserve, because such kind of invasion should be preceded by unreal comparative studies of prenatal development of the same child with and without music stimulation.

source and the mother and (optionally) passing through clothes meets the barrier of an abdominal wall. It consists of several layers of various thickness and properties (skin, subcutaneous tissue, adipose tissue, muscle tissue of various type and structure, the placenta and the uterine wall). Another medium is the amniotic fluid which directly surrounds the fetus. After passing through all that barriers, the sound reaches fetal auditory system through ear canal (filled with amniotic fluid), bone conduction, or both at the same time [11, 12].

Moreover, the attenuation model should consider the gestational age, because the barriers on the sounds path change their parameters (e.g. abdominal wall's thickness is changing during pregnancy) and the auditory system develops in time. Additionally, while analysing attenuation, we should consider sound's features, such as its frequency and level at the source.

Development of fetal hearing. Sufficiently efficient hearing system that allows the fetus to hear is developed with the beginning of the third trimester of pregnancy [3]. During the remaining period of pregnancy sounds can influence the behavior of the fetus and further development of his/her nervous system. However, the fetus hears a narrower range than the adult, and his/her hearing develops gradually [12]. Gerhardt et al. had proved in their work [12] that the frequency of 500 Hz is inaudible to the fetus who is younger than 19 weeks of gestation. In 27 gestation week a fetus acquires the ability to hear lower sounds (200–500 Hz). After 30 weeks of pregnancy the fetus begins to hear high frequencies (1–3 kHz). This study had shown a gradual increase in the range and sensibility of hearing in the period from 19 to 37 weeks of fetal development.

Frequency. According to the law of mass density, high frequencies are more attenuated than the low [22]. This was confirmed in a series of studies on the suppression of sounds in utero [3, 11, 12, 21]. One of a few measurements using hydrophone placed in the woman's uterus had shown attenuation of 3.7 dB for 125 Hz to 4 kHz at 10 dB [21]. Another source [12], based on the studies conducted on sheep, had given the attenuation at low frequencies (below 200 Hz) at the level of less than 5 dB (in some cases even enhancement was noticed) increased to 20 dB for 4 kHz. Gerhardt et al. had suggested that the range of 125 Hz–2 kHz abdominal attenuation can be modeled with low–pass filter at rate 6 dB per octave. Similar range of attenuation values had been presented in [11]. The authors had pointed to the attenuation of 5 dB for sounds up to 500 Hz and 20–30 dB for higher frequencies.

Sound level at source. There are some reports indicating that higher sound level in the source does not induce higher sound level in the uterus. It was suggested that sounds with higher intensity are more attenuate than those with lower intensity. This hypothesis was verified in [18]. Stimulation of the intrauterine environment model with sound 10 dB louder than the original (100 dB) did not induce increase of attenuation. Attempts to establish the minimal sound level which could be hearable in the uterus were described in [11, 12]. According to Gerhardt et al. observations, the fetus is able to detect low–frequency components of speech only if their level in the external environment exceeds 60 dB [12]. In the another paper, Gerhardt et al. noticed that the fetus can hear sounds just as loud as a newborn child, if the sound level in the source will be about 35 dB higher for the fetus' stimulation [11].

In literature, there are presented some simplified models of sound attenuation in intrauterine environment. Among them is a diagram of attenuation level for each sound frequency [12]. This model presents a sound level measured inside the ewe's uterus, and the attenuation level recorded by a hydrophone and a cochlear microphone. The hydrophone was placed in the uterus between the mother's abdominal wall and the fetal head. The cochlear microphone (CM) was implanted in the environs of the inner ear. The attenuation on the way to this structure was determined by comparing the values recorded with the same microphone implanted in a newborn sheep. Model presented by Gerhardt et al. allows only for assessment of the frequency impact on attenuation values. Antonets et al. [6] had developed the device for non–invasive measurement of sound attenuation in a mother's womb. The apparatus consists of a soft sealed capsule filled with 150 ml of water and the hydrophone. For the tests a model of the uterus was prepared; it consisted of polyethylene capsule filled with 5 liters of water. Effects of sound frequency on attenuation level was assessed. Another physical model consists of a simple spherical ball filled with water [18]. In this model not only the influence of frequency level but also the position of the sound receiving device (position of the fetus in the uterus) and the size of the model (size of the mother's abdomen) were assessed.

None of available models take into account all variables related to hearing in maternal uterus, hence the need to create such a model is quite urgent. Physical model allows to observations of phenomena during sound transmission through layers of different density. Whereas mathematical model is not limited and allow to consider all of the variables. None of them cannot be approved as a better one, therefore physical model should be used to validate the mathematical model.

4 Sounds Analysis

According to many reports, the attenuation increases with the increase of sound frequency and it can amount 15–45 dB for frequency 2000 Hz. These differences result from various measurement methods. Lower attenuation levels were achieved for measurements in amniotic fluid, whereas maximum levels relate to data from the fetal inner ear. Changes in acoustic wave during its way from the amniotic fluid are difficult to observe and describe. Fetal hearing may occur both by bone conduction and by traditional way—an outer ear. Also according to the authors of [11], bone conduction is the dominant pathway for hearing in water environment, where the fetus lives. In the study [11] there was indicated that induction of the same response in the inner ear only by the outer ear (the remaining part of the head was covered with a soundproof helmet) needs more signal energy in the source than in the case where the whole head is exposed (the bone conduction is not disabled). Therefore, this paper considers the values of the sound attenuation in the amniotic fluid. This phenomenon is easier to measure and it is widely described in literature.

In experiments [4, 10] conducted by the use of the hydrophone which recorded the sound level in the sheep's uterus, there was indicated that low frequency sounds (250–300 Hz) were hardly attenuated and the attenuation increased with the rate of 5–6 dB per octave reaching the value of 20 dB for 4 kHz. There is no unequivocal data for attenuation of higher frequencies.

Because of the lack of detailed data about sheep's sizes, and the small number of such experiments at all, it is not possible to create an attenuation model taking into account abdominal wall's thickness or changes in fetal hearing system. We can only hypothesize about generalized model of attenuation, which is focused on a holistic body treatment.

According to above–mentioned data the simplified model of sound attenuation in the uterus can be represented by first order passive low pass filter (LPF). Filters are characterized by its cutoff frequency which corresponds with attenuation of the input by 3 dB, and by rate of frequency rolloff [23]. Filter's properties for this use should include cutoff frequency 375 Hz and attenuation of higher frequencies with the slope 6 dB per octave (amplitude is reduced by a half every time the frequency doubles).

Fourier analysis of the recorded piano sounds was carried out using MATLAB. The resulting spectrum was multiplied by an attenuation factor which was determined by the characteristics of the LPF. It was established that sounds with a frequency below 375 Hz are in 3 dB bandwidth. The cutoff frequency was connected with the factor 0.7, for frequency 500 Hz it was 0.5 etc. The analysis was conducted for frequencies lower than 4 kHz, because of the lack of consistent data for higher sound frequencies. In the next step, the total harmonic distortion (THD, (1)) for original and attenuated harmonics were designated:

$$THD = {U_1}^{-1} \sqrt{\sum_{k=2}^{n} U_k^2}, \tag{1}$$

where:
U_1—voltage of first harmonic,
U_k—voltage of k harmonic.

The algorithm of a sound analysis is presented in the Fig. 1.

In the Fig. 2 the amplitude spectra of the exemplary attenuated and original sounds are presented. We can notice that in this case all harmonics are attenuated at some level.

The sounds which occur in the nature are complex, reach with harmonic frequencies and sound to the ear naturally. Single tones tend to be artificial and are unpleasant to the ear. In order to determine the harmonic content of the original and heard in the womb sounds, total harmonic distortion was calculated for harmonics before and after attenuation. THD values for selected frequencies are presented in Table 1.

Total harmonic distortion in the case of sounds shows the richness of its timbre—the higher the THD the higher harmonic content in the sound. The data in the Table 1

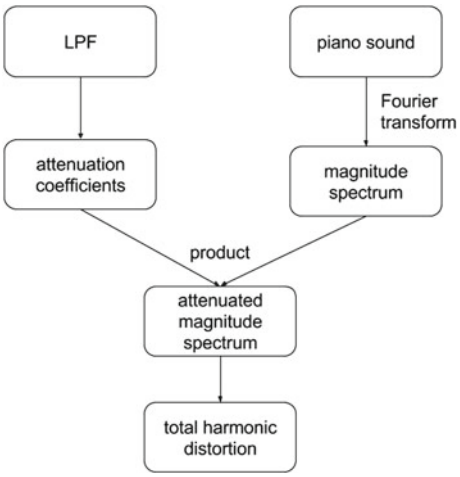


Fig. 1 The algorithm of sounds analysis

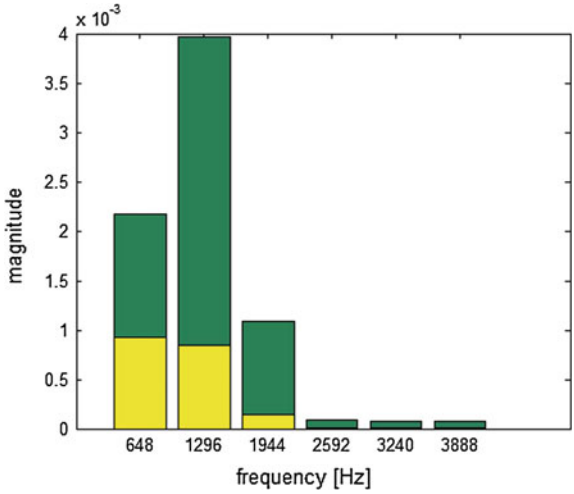


Fig. 2 Magnitude spectrum of attenuated (*yellow*) and original (*green*) sounds

illustrate a serious loss in higher harmonics in the sound that was treated with low pass filter.

On the basis of total harmonic distortion analysis for a single sound, a proposal indicating the sound reduction can be formulated. In the case of digital sound we carefully choose a bandwidth which is conditioned by Nyquist–Shannon sampling theorem. Fixing the sample rate at 44100 Hz is considered obvious, because it determines no distortion in sound transmission of human hearing range. In the analysed example, the quality of sound measured by harmonic content is much worse than in the case of music transmission with a mobile phone (maximum sample rate 6800 Hz).

Table 1 THD for selected frequencies

Sound (Hz)	Original THD	Attenuated THD	Ratio
138	1.8168	1.0207	0.5679
165	1.9214	1.1803	0.6143
222	1.2150	0.5653	0.4670
276	1.1018	0.3859	0.3502
347	0.8631	0.4187	0.4851
464	0.3699	0.1810	0.4892
551	2.1414	1.0543	0.4923
648	1.3485	0.5160	0.4888
817	0.8518	0.4250	0.4887
980	0.1313	0.0593	0.4516

5 Conclusions

There is a need to perform a set of studies which would allow to assess the impact of different parameters (such as abdominal wall's thickness, amount of amniotic fluid or changes in fetal hearing system) on sound attenuation. Such data would enable the researchers to create more accurate model of attenuation and to assess differences in sound, which are audible for fetuses, considering his/her development and mother's body composition. However, contemporary studies and proposed simplified model enable us to notice that the sound heard in the uterus is not the same music as we know.

There is no convincing evidence of the distinctive reception and especially perception of music elements. For historical reasons we can distinguish a minor third and a major third. There is no structural reasons that this interval may be superior (in physical context) to another one. Such superiority can be observed in the case of fifth clean, especially in Pythagorean tuning, where the fundamental frequency of the fifth is equal 1.5 of fundamental frequency of the basic sound.

Music (the set of sounds) reaches to fetal environment in the impaired form, therefore its structure, composition, and character are changed. For a child (fetus, neonate) this stimulus is not music in our understanding, because the child does not connect with these sounds in any historical conditioned experience. It is possible that particular impact on the child is characterized by a sound at all, but considering the number of degrees of freedom of the acoustic wave, it should be determined which of the feature that describe music has a dominant impact. Otherwise, musical stimulation in the prenatal period is completely accidental treatment that is independent of reaction type of unborn child. If we tentatively assume a possibility of autism, such musical intervenience can only harm the child, and therefore, mouth movements considered to be a symptom of hearing [19] can be only the fetal's cry for help.

It was mentioned in the Chap. 1 that one of the tasks for the research on the acoustic stimulation in prenatal period is to determine the characteristics of sounds that can negatively affect the development of auditory system, brain or even entire body. Acoustic conditions in Neonatal Intensive Care Units (NICU), which take care of prematures, are good references for these considerations. Preterm babies stay in NICU for few weeks or months. The acoustic environment of this place is full of alarms, mechanical noise and voices that could be harmful for the auditory system and the brain [7]. The fetus in the utero is protected by maternal abdominal walls which attenuate the hazardous sounds of everyday life. Such an auditory *nursery* ends when the child arrives to NICU that is rich in different sounds, including high frequencies [17]. In this new environment child can defend against the light with closed eyes, but cannot defend against the noise. Acoustic overstimulation can lead to abnormalities in child's growth and development and it can increase the risk of auditory, language and cognitive abnormalities. Preterm infants are exposed especially, because it is a critical period for still not finished neuronal development, when they are not any more protected by mother's tissues [20].

Exposure to sound has both positive and negative effects on developing preterm baby, therefore there are some recommendations for safe sound level in the NICU. Some of them are that tapping or writing on the tops of incubators and hoods are forbidden, NICU personnel and visitors should wear soft shoes. There are also indications for the sound level—it should not exceed 45 dB [5]. What is important, this recommendations are concentrated only on permissible sound level, ignoring the problem of harmful frequencies, which are inaudible for the fetus in the womb but they are present in NICU [16]. Unfortunately, current reports indicate that general recommendations are not respected and the acceptable sound levels are often exceeded [15].

Analysis of the mathematical model and performed studies form the basis for the conclusion that acoustic stimuli, which are commonly used with regulated (often excessive) intensity for the child in the prenatal period, are significantly different than the music which is the source of these stimuli. Therefore the hypothesis that such music has an impact on child's intellectual development cannot be confirmed. It is suggested to continue the studies on low frequency components impact. It is worth noticing (not demanding a special proof) dispersible impact of metro–rhythmical stimuli on human psychomotor activities.[2] Accepting the sound volume and its distance from the noise (in maternal womb) as an efficiency criterion of acoustic stimuli impact, the sound level (especially in high frequencies) should be selectively incremented. This is a modification which is unacceptable for the mother and her environment, therefore it is not a subject of interest. Separation of the mother's senses from the sound stimuli is nothing more than turning off the only mechanism for volume control which the fetus uses in not indirect but natural way.

[2]Possibility of organized marching, tapping in the even rhythm or pronouncing consecutive numbers is seriously disrupted if we receive obtrusive and loud rhythmic odd sound at the same time.

The Study Group of Neonatal Intensive Care Unit Sound and the Expert Panel convened by the Center for Physical and Developmental Environment of the High–Risk Infant on the basis of over 4000 works on auditory development of human and animal published recommendations for hearing protection of fetuses and neonates [14]:

1. *Women should avoid prolonged exposure to low-frequency (<250 Hz) sound level above 65 dB during pregnancy.*
2. *Earphones or other devices for sound production should not be used directly attached to the pregnant women's abdomen.*
3. *The voice of the mother during normal daily activities, along with the sounds produced by her body and those present in her usual surroundings, is sufficient for normal fetal auditory development. The fetus does not require supplemental stimulation. Programs to supplement the fetal auditory experience cannot be recommended.*
4. *Infant intensive care units should incorporate a system of regular noise assessment.*

Acknowledgments The work has been partially financed by Polish Ministry of Science and Silesian University of Technology statutory financial support for young researchers BKM-508/RAu-3/2016.

References

1. Abate, M.A.: Fetal exposures to sound and vibroacoustic stimulation. Child. Lit. Educ. **44**(4), 326–343 (2013)
2. Abrams, R., Gerhardt, K., Huang, X., Peters, A., Langford, R.: Musical experiences of the unborn baby. J. Sound Vib. **231**(1), 253–258 (2000). http://linkinghub.elsevier.com/retrieve/pii/S0022460X99926056
3. Abrams, R.M., Gerhardt, K.J.: The fetus: the acoustic environmental and physiological responses of the fetus. J. Perinatol., 30–35 (2000)
4. Abrams, R.M., Griffiths, S.K., Huang, X., Sain, J., Langford, G., Gerhardt, K.J.: Fetal music perception: the role of sound transmission. Music Percept. **15**(3), 307–317 (1998). http://labshare.psych.ryerson.ca/smart/PDF/Abrams_Griffiths_1998.pdf
5. American Academy of Pediatrics: Committee on Environmental Health: Noise: a hazard for the fetus and newborn. Pediatrics **100**(4), 724–727 (1997)
6. Antonets, V.A., Kazakov, V.V.: On noninvasive assessment of acoustic fields acting on the fetus. Acoust. Phys. **60**(3), 342–347 (2014). http://link.springer.com/10.1134/S1063771014030026
7. Barreto, E.D., Morris, B.H., Philbin, M.K., Gray, L.C., Lasky, R.E.: Do former preterm infants remember and respond to neonatal intensive care unit noise? Early Hum. Dev. **82**, 703–707 (2006)
8. Chang, E.F., Merzenich, M.M.: Environmental noise retards auditory cortical development. Science **300**(5618), 498–502 (2003). http://science.sciencemag.org/content/300/5618/498
9. Fisch, L.: Integrated development and maturation of the hearing system; a critical review article. Br. J. Audiol. **17**(3), 137–154 (1983). http://dx.doi.org/10.3109/03005368309107879. pMID: 6357323
10. Gerhardt, K.J., Abrams, R.M., Oliver, C.C.: Sound environment of the fetal sheep. Am. J. Obstetrics Gynecol. **162**(1), 282–287 (1990). http://www.sciencedirect.com/science/article/pii/0002937890908666

11. Gerhardt, K.J., Abrams, R.M.: Fetal hearing: characterization of the stimulus and response. Semin. Perinatol. **20**(1), 11–20 (1996)
12. Gerhardt, K.J., Abrams, R.M.: Fetal exposures to sound and vibroacoustic stimulation. J. Perinatol.: Off. J. California Perinatal Assoc. **20**, 21–30 (2000)
13. Graven, S.N., Browne, J.V.: Auditory development in the fetus and infant. Newborn Infant Nurs. Rev. **8**(4), 187–193 (2008). http://dx.doi.org/10.1053/j.nainr.2008.10.010
14. Graven, S.N.: Sound and the developing infant in the NICU: conclusions and recommendations for care. J. Perinatol.: Off. J. California Perinatal Assoc. **20**, 88–93 (2000)
15. Krueger, C., Horesh, E., Crossland, B.A.: Safe Sound Exposure in the Fetus and Preterm Infant. JOGNN. J. Obstet. Gynecol. Neonatal Nurs. **41**, 166–170 (2012)
16. Lahav, A.: Questionable sound exposure outside of the womb: frequency analysis of environmental noise in the neonatal intensive care unit. Acta Paediatr. **104**(1), 14–19 (2015). http://dx.doi.org/10.1111/apa.12816
17. Lahav, A., Skoe, E.: An acoustic gap between the NICU and womb: a potential risk for compromised neuroplasticity of the auditory system in preterm infants. Front. Neurosci. **8**(December), 1–8 (2014). http://journal.frontiersin.org/article/10.3389/fnins.2014.00381/abstract
18. Lecanuet, J.P., Gautheron, B., Locatelli, A., Schaal, B., Jacquet, A.Y., Busnel, M.C.: What sounds reach fetuses: biological and nonbiological modeling of the transmission of pure tones. Dev. Psychobiol. **33**(3), 203–219 (1998)
19. López-Teijón, M., García-Faura, A., Prats-Galino, A.: Fetal facial expression in response to intravaginal music emission. Ultrasound, 1–8 (2015). http://ult.sagepub.com/content/early/2015/09/29/1742271X15609367.abstract
20. Mcmahon, E., Wintermark, P., Lahav, A.: Auditory brain development in premature infants: the importance of early experience. Ann. New York Acad. Sci. **1252**, 17–24 (2012)
21. Richards, D.S., Frentzen, B., Gerhardt, K.J., McCann, M.E.: A.R.: sound levels in the human uterus. Obstetr. Gynecol. **80**(2), 186–190 (1992)
22. Yang, Z., Dai, H.M., Chan, N.H., Ma, G.C., Sheng, P.: Acoustic metamaterial panels for sound attenuation in the 50–1000 Hz regime. Appl. Phys. Lett. **96**(4), 041906 (2010). http://scitation.aip.org/content/aip/journal/apl/96/4/10.1063/1.3299007
23. Zumbahlen, H.: Analog filters. In: Zumbahlen, H., et al. (eds.) Linear Circuit Design Handbook, Chap. 8, pp. 581–679. Newnes, Burlington (2008). http://www.sciencedirect.com/science/article/pii/B9780750687034000080

Arterial Blood Gases Forecast Optimization by Artificial Neural Network Method

Wiesław Wajs, Piotr Wais, Marcin Ochab and Hubert Wojtowicz

Abstract Arterial blood gas sampling represents the gold standard method for acquiring patients' acid-base status. It is proposed that blood gas values could be measured using arterialized earlobe blood samples. Pulse oximetry plus transcutaneous carbon dioxide measurement is an alternative method of obtaining similar information as well. Since dynamics of biochemical changes occurring in the blood is an individual feature which changes during the healing process authors proposed forecast models developed using artificial neural networks. The networks are trained with data vectors containing short term (72 h) history windows of four blood gasometry parameters. Several different optimization algorithms are used in the training phase to create a set of models from which the best prediction model is then selected.

Keywords Arterial blood gas · Forecast · Artificial neural networks

1 Introduction

An arterial blood gas (ABG) is a blood test that is performed using blood from an artery. Measurement of arterial blood gas tensions is routinely used to assess gas exchange in patients with acute and chronic respiratory disorders. Arterial blood gas sampling represents the gold standard method for acquiring patients' acid-base status [13]. It is proposed that blood gas values could be measured using arterialized earlobe blood samples. Pulse oximetry plus transcutaneous carbon dioxide measurement is an alternative method of obtaining similar information as well. An ABG is a test that measures the arterial oxygen tension $PaCO_2$, carbon dioxide tension $PaCO_2$,

W. Wajs · H. Wojtowicz (✉)
The University of Rzeszów, Al. Rejtana 16c, Rzeszów, Poland
e-mail: hubert.wojtowicz@gmail.com

P. Wais
State Higher Vocational School in Krosno, Rynek 1, Krosno, Poland

M. Ochab
AGH University of Science and Technology, Mickiewicza 30, Kraków, Poland

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_36

and acidity pH. The ABG is one of the most common tests performed on patients in intensive care units (ICUs). The test is used to determine the pH of the blood, the partial pressure of carbon dioxide and oxygen, and the bicarbonate level. Many blood gas analyzers will also report concentrations of lactate, hemoglobin, several electrolytes, oxyhemoglobin, carboxyhemoglobin and methemoglobin. ABG testing is mainly used in pulmonology and critical care medicine to determine gas exchange which reflect gas exchange across the alveolar-capillary membrane. ABG testing also has a variety of applications in other areas of medicine. The machine used for analysis aspirates this blood from the syringe and measures the pH and the partial pressures of oxygen and carbon dioxide. The bicarbonate concentration is also calculated. These results are usually available for interpretation within five minutes. Two methods have been used in medicine in the management of blood gases of patients in hypothermia: pH-stat method and α-stat method. pH-stat: The pH and other ABG results are measured at the patient's actual temperature. The goal is to maintain a pH of 7.40 and the arterial carbon dioxide tension $PaCO_2$ at 5.3 kPa (40 mmHg) at the actual patient temperature. It is necessary to add CO_2 to the oxygenator to accomplish this goal. α-stat: The pH and other ABG results are measured at 37 °C, despite the patient's actual temperature. The goal is to maintain the arterial carbon dioxide tension at 5.3 kPa (40 mmHg) and the pH at 7.40 when measured at 37 °C. Both the pH-stat and α-stat strategies have theoretical disadvantages. α-stat method is the method of choice for optimal myocardial function. The pH-stat method may result in loss of autoregulation in the brain, coupling of the cerebral blood flow with the metabolic rate in the brain. By increasing the cerebral blood flow beyond the metabolic requirements, the pH-stat method may lead to cerebral microembolisation and intracranial hypertension [8]. A 1 mmHg change in $PaCO_2$ above or below 40 mmHg results in 0.008 unit change in pH in the opposite direction. The $PaCO_2$ will decrease by about 1 mmHg for every 1 mEq/L reduction in HCO_3^- below 24 mEq/L. A change in $[HCO_3^-]$ of 10 mEq/L will result in a change in pH of approximately 0.15 pH units in the same direction. These are typical reference ranges, although various analyzers and laboratories may employ different ranges. The pH or H^+ indicates if a patient is acidemic (pH $<$ 7.35; H+ $>$45) or alkalemic (pH $>$ 7.45; H+ $<$ 35). A low $PaCO_2$ indicates that the patient is not oxygenating properly, and is hypoxemic. Note that a low $PaCO_2$ is not required for the patient to have hypoxia. At a PaO_2 of less than 60 mmHg, supplemental oxygen should be administered. At a PaO_2 of less than 26 mmHg, the patient is at risk of death and must be oxygenated immediately. The carbon dioxide partial pressure ($PaCO_2$) is an indicator of CO_2 production and elimination: for a constant metabolic rate, the $PaCO_2$ is determined entirely by its elimination through ventilation [8]. A high $PaCO_2$ (respiratory acidosis, alternatively hypercapnia) indicates underventilation (or, more rarely, a hypermetabolic disorder), a low $PaCO_2$ (respiratory alkalosis, alternatively hypocapnia) hyper- or overventilation. The HCO_3^- ion indicates whether a metabolic problem is present such as ketoacidosis. A low HCO_3^- indicates metabolic acidosis, a high HCO_3^- indicates metabolic alkalosis. As

this value, when given with blood gas results, is often calculated by the analyzer, correlation should be checked with total CO_2 levels as directly measured. The base excess is used for the assessment of the metabolic component of acid-base disorders, and indicates whether the patient has metabolic acidosis or metabolic alkalosis. Contrasted with the bicarbonate levels, the base excess is a calculated value intended to completely isolate the non-respiratory portion of the pH change. This is the total amount of CO_2, and is the sum of $[HCO_3^-]$ and CO_2 by the formula: $tCO_2 = [HCO_3^-] + \alpha PaCO_2$, where $\alpha = 0.226\,\text{mM/kPa}$, $[HCO_3^-]$ is expressed in millimolar concentration (mM) (mmol/l) and $PaCO_2$ is expressed in kPa. Contamination of the sample with room air will result in abnormally low carbon dioxide and possibly elevated oxygen levels, and a concurrent elevation in pH. Delaying analysis, without chilling the sample, may result in inaccurately low oxygen and high carbon dioxide levels as a result of ongoing cellular respiration. The normal range for pH is 7.35–7.45. As the pH decreases (<7.35), it implies acidosis, while if the pH increases (>7.45) it implies alkalosis. In the context of arterial blood gases, the most common occurrence will be that of respiratory acidosis. Carbon dioxide is dissolved in the blood as carbonic acid, a weak acid; however, in large concentrations, it can affect the pH drastically. Whenever there is poor pulmonary ventilation, the carbon dioxide levels in the blood are expected to rise. This leads to a rise of carbonic acid, leading to a decrease in pH. The first buffer of pH will be the plasma proteins, since these can accept some H^+ ions to try and maintain homeostasis. As carbon dioxide concentrations continue to increase ($PaCO_2 > 45\,\text{mmHg}$), a condition known as respiratory acidosis occurs. The body tries to maintain homeostasis by increasing the respiratory rate, a condition known as tachypneoa. This allows much more carbon dioxide to escape the body through the lungs, thus increasing the pH by having less carbonic acid. If a patient is in a critical setting and intubated, one must increase the number of breaths mechanically. On the other hand, respiratory alkalosis ($PaCO_2 < 35\,\text{mmHg}$) occurs when there is too little carbon dioxide in the blood. This may be due to hyperventilation or else excessive breaths given via a mechanical ventilator in a critical care setting. The action to be taken is to calm the patient and try to reduce the number of breaths being taken to normalise the pH. The respiratory pathway tries to compensate for the change in pH in a matter of 2–4 h. If this is not enough, the metabolic pathway takes place. Under normal conditions, the Henderson-Hasselbalch equation will give the blood pH.

$$pH = 6.1 + \log_{10}\left(\frac{HCO_3^-}{0.03 * PaCO_2}\right), \tag{1}$$

where: 6.1 is the acid dissociation constant of carbonic acid (H_2CO_3) at normal body temperature, $[HCO_3^-]$ is the concentration of bicarbonate in the blood in mEq/L and $PaCO_2$ is the partial pressure of carbon dioxide in the arterial blood. The kidney and the liver are two main organs responsible for the metabolic homeostasis of pH. Bicarbonate is a base that helps to accept excess hydrogen ions whenever there is

acidaemia. However, this mechanism is slower than the respiratory pathway and may take from a few hours to 3 days to take effect. In acidaemia, the bicarbonate levels rise, so that they can neutralize the excess acid, while the contrary happens when there is alkalaemia. Thus when an arterial blood gas test reveals, for example, an elevated bicarbonate, the problem has been present for a couple of days, and metabolic compensation took place over a blood acedemia problem. It is much easier to correct acute pH derangements by adjusting respiration. Metabolic compensations take place at a much later stage. However, in a critical setting, a patient with a normal pH, a high CO_2, and a high bicarbonate means that, although there is a high carbon dioxide level, there is metabolic compensation. As a result one must be careful as to not artificially adjust breaths to lower the carbon dioxide. In such case, lowering the carbon dioxide abruptly means that the bicarbonate will be in excess and will cause a metabolic alkalosis. In such a case, carbon dioxide levels should be slowly diminished.

2 Data

Thanks to the Neonatal Intensive Care Unit of The Department of Pediatrics at the Jagiellonian University Medical College we were able to analyze extremely premature infants historical data. For the research the following arterial blood gasometry parameters have been selected: pH, pO_2, pCO_2, HCO_3. These parameters describe the acid-base balance and pulmonary gas exchange, which are acidity, oxygen partial pressure, carbon dioxide partial pressure and bicarbonate ions content in the blood, respectively. Oxygen partial pressure is strictly correlated with the degree of respiratory fitness. To obtain uninterrupted sequence of measurements we have chosen only data with 12 h intervals between measurements. Presented preliminary results were performed on one patient with birth weight of 860 g observed during first 25 days of life. During the whole period the values stayed in ranges:

- pH: 7–7.45,
- pO_2: 34–140 mmHg,
- pCO_2: 19–104 mmHg,
- HCO_3: 13–26 mEq/L.

Variations of blood gasometry parameters' values over time are characterized by considerable dynamism. Figure 1 shows dynamics of blood gasometry parameters.

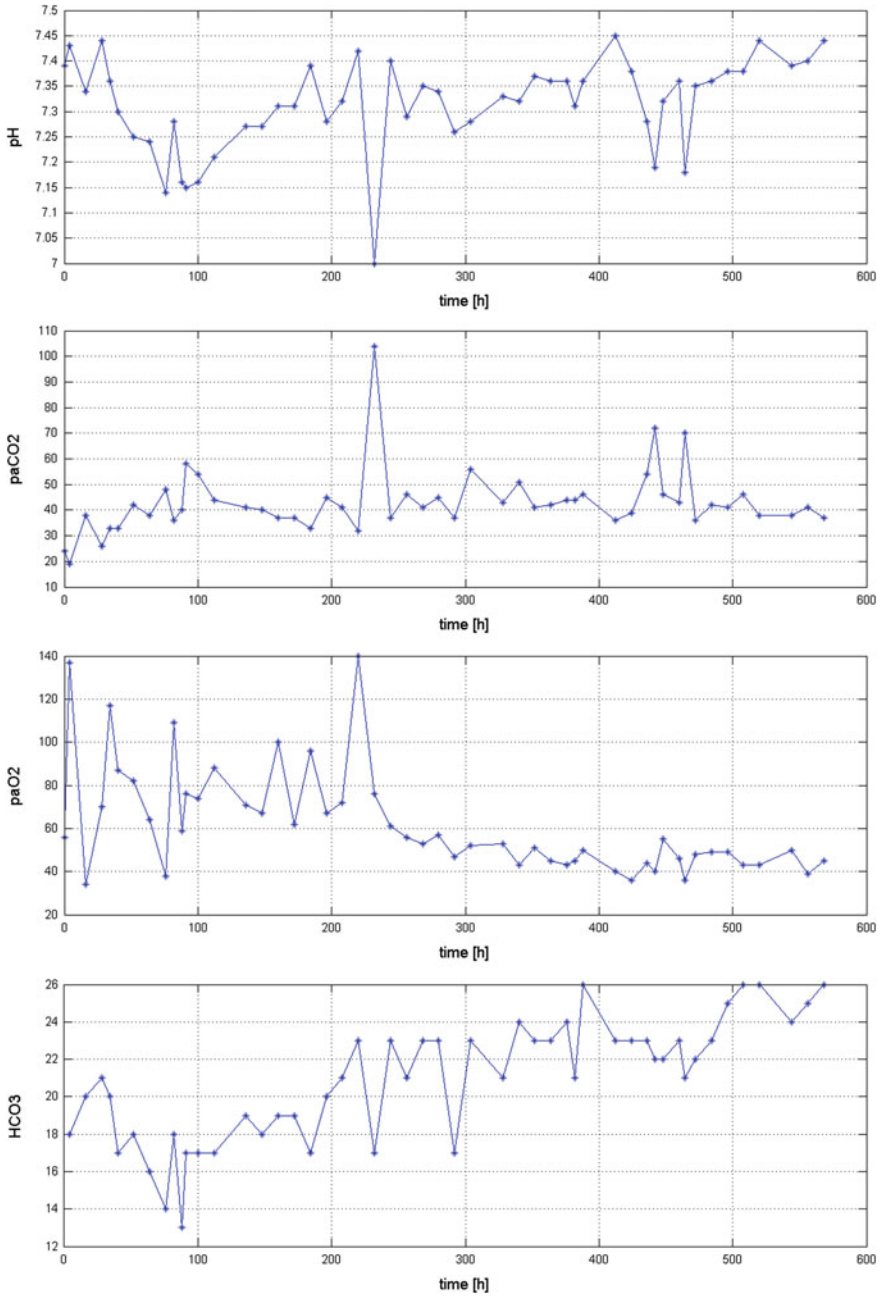


Fig. 1 Dynamics of blood gasometry parameters

3 Artificial Neural Network

The models are developed using multilayer artificial neural networks (ANNs), trained using the several algorithms. A network can have several layers. Each layer has a weight matrix W. The network shown below has $R = 4$ inputs, $S1 = 2$ neurons in the first layer and $S2 = 1$ neuron in the second layer.

$$R = \begin{bmatrix} r11 & r12 & r13 & r14 \\ r21 & r22 & r23 & r24 \\ r31 & r32 & r33 & r34 \\ r41 & r42 & r43 & r44 \end{bmatrix} = \begin{bmatrix} \mathrm{pH}_{t-3} & \mathrm{pH}_{t-2} & \mathrm{pH}_{t-1} & \mathrm{pH}_{t} \\ \mathrm{PaO2}_{t-3} & \mathrm{PaO2}_{t-2} & \mathrm{PaO2}_{t-1} & \mathrm{PaO2}_{t} \\ \mathrm{PaCO2}_{t-3} & \mathrm{PaCO2}_{t-2} & \mathrm{PaCO2}_{t-1} & \mathrm{PaCO2}_{t} \\ \mathrm{HCO3}_{t-3} & \mathrm{HCO3}_{t-2} & \mathrm{HCO3}_{t-1} & \mathrm{HCO3}_{t} \end{bmatrix} \tag{2}$$

The neuron output is given below.

$$A = \begin{bmatrix} a1 & a2 & a3 & a4 \end{bmatrix} = \begin{bmatrix} \mathrm{pH}_{t-2} & \mathrm{pH}_{t-1} & \mathrm{pH}_{t} & \mathrm{pH}_{t+1} \end{bmatrix} \tag{3}$$

The layers of a multilayer network play different roles. A layer that produces the network output is called an output layer. Multiple layer networks are quite powerful. A two layer network, where both layers are comprised of neurons with sigmoidal activation functions, can be trained to approximate any function.

We calculate an error e as the difference between the output vector s and a desired target vector pH_{t-2}, ph_{t-1}, pH_{t}, pH_{t+1} in the form:

$$e = 0.5\left((\mathrm{pH}_{t-2} - s1)^2 + (\mathrm{pH}_{t-1} - s2)^2 + (\mathrm{pH}_{t} - s3)^2 + (\mathrm{pH}_{t+1} - s4)^2\right). \tag{4}$$

The gradient of error is defined as:

$$G = \begin{bmatrix} g1 = \frac{\partial e}{\partial w1} \\ g2 = \frac{\partial e}{\partial w2} \\ \cdot \\ \cdot \\ \cdot \\ g10 = \frac{\partial e}{\partial w10} \end{bmatrix}$$

4 Levenberg-Marquardt Method

The Levenberg-Marquardt method [9, 12] addresses nonlinear least squares curve fitting problems. The Levenberg-Marquardt method is a standard technique used to solve nonlinear least squares problems. Least squares problems arise when fitting a parameterized function to a set of measured data points by minimizing the sum of squares of errors between data points and the function [11]. Nonlinear least squares problems arise when the function is not linear in parameters. Nonlinear least squares methods involve an iterative improvement to parameter values in order to reduce the

sum of squares of errors between the function and measured data points [15]. The Levenberg-Marquardt curve-fitting method is actually a combinationof two minimization methods: the gradient descent method and the Gauss-Newton method. In the gradient descent method, the sum of squared errors is reduced by updating parameters in the steepest-descent direction. In the Gauss-Newton method, the sum of squared errors is reduced by assuming the least squares function is locally quadratic, and finding the minimum of the quadratic. The Levenberg-Marquardt method acts more like a gradient-descent method when parameters are far from their optimal values, and acts more like the Gauss-Newton method when parameters are close to their optimal values. The Levenberg-Marquardt algorithm adaptively varies parameter updates between the gradient descent update and the Gauss-Newton update. where small values of the algorithmic parameter λ result in a Gauss-Newton update and large values of λ result in a gradient descent update. The parameter λ is initialized to be large so that first updates are small steps in the steepest-descent direction. If an iteration happens to result in a worse approximation, λ is increased. As the solution improves, λ is decreased, the Levenberg-Marquardt method approaches the Gauss-Newton method, and the solution typically accelerates to the local minimum. Marquardt's suggested update relationship makes the effect of particular values of λ less problem-specific, and is used in the Levenberg-Marquardt algorithm implemented in the Matlab function. The data points are denoted by:

$$(Y_i, X_{1i}, X_{2i}, \ldots, X_{mi}) \;\; i = 1, 2, \ldots, n. \tag{5}$$

The problem is to compute those estimates of parameters, which will minimize:

$$\Phi = \sum_{i=1}^{i=n} (Y_i - Y_i^*)^2, \tag{6}$$

where Y_i^* is the value of y predicted by:

$$E(y) = f(x_1, x_2, \ldots, x_m; \beta_1, \beta_2, \ldots, \beta_k), \tag{7}$$

where $x_1, x_2, \ldots, x_m$ are independent variables, $\beta_1, \beta_2, \ldots, \beta k$ are the population values of k parameters, and $E(y)$ is the expected value of the dependent variable y, at the ith data point.

Levenberg-Marquardt method defines the strategy as follows: Let $\nu > 1$. Let λ^{r-1} denote the value of λ from the previous iteration. Initially let $\lambda^0 = 10^{-2}$. Compute $\Phi(\lambda^{r-1})$ and $\Phi(\lambda^{r-1}/\nu)$, if (λ^{r-1}) is already negligible by comparison with 1.0 to the number of significant figures carried, then do test 2 or test 3 immediately without computing $\Phi(\lambda^{r-1}/\nu)$, and ignore comparisons involving $\Phi(\lambda^{r-1}/\nu)$.

1. If $\Phi(\lambda^{r-1}/\nu) \leq \Phi^r$, let $\lambda^r = \lambda^{r-1}/\nu$.
2. If $\Phi(\lambda^{r-1}/\nu) > \Phi^r$, and $\Phi(\lambda^{r-1}) \leq \Phi(\lambda^r)$, let $\lambda^r = \lambda^{r-1}$.
3. If $\Phi(\lambda^{r-1}/\nu) > \Phi^r$, and $\Phi(\lambda^{r-1}) > \Phi(\lambda^r)$, increase λ by successive multiplication by ν until some smallest w, $\Phi(\lambda^{r-1}\nu^w) \leq \Phi^r$. Let $\lambda^r = \lambda^{r-1}\nu^w$.

5 Backpropagation Method

Backpropagation, an abbreviation for *backward propagation of errors*, is a common method of training artificial neural networks used in conjunction with an optimization method such as gradient descent [1]. The method calculates the gradient of a loss function with respect to all the weights in the network. The gradient is fed to the optimization method which in turn uses it to update the weights, in an attempt to minimize the loss function. Backpropagation requires a known, desired output for each input value in order to calculate the loss function gradient. It is therefore usually considered to be a supervised learning method, although it is also used in some unsupervised networks such as autoencoders. It is a generalization of the delta rule to multilayerd feedforward networks, made possible by using the chain rule to iteratively compute gradients for each layer. Backpropagation requires that the activation function used by the artificial neurons be differentiable. The back propagation (BP) neural network algorithm is a multilayer feedforward network trained according to error back propagation algorithm and is one of the most widely applied neural network models. BP network can be used to learn and store a great deal of mapping relations of input-output model, and no need to disclose in advance the mathematical equation that describes these mapping relations. Its learning rule is to adopt the steepest descent method in which the back propagation is used to regulate the weight value and threshold value of the network to achieve the minimum sum of square error. The operations of backpropagation neural networks can be divided into two steps: feedforward and backpropagation. In the feedforward step, an input pattern is applied to the input layer and its effect propagates, layer by layer, through the network until an output is produced. The network's actual output value is then compared to the expected output, and an error signal is computed for each of the output nodes. Since all the hidden nodes have, to some degree, contributed to the errors evident in the output layer, the output error signals are transmitted backwards from the output layer to each node in the hidden layer that immediately contributed to the output layer. This process is then repeated, layer by layer, until each node in the network has received an error signal that describes its relative contribution to the overall error. Once the error signal for each node has been determined, the errors are then used by the nodes to update the values for each connection weights until the network converges to a state that allows all the training patterns to be encoded. The Backpropagation algorithm looks for the minimum value of the error function in weight space using a technique called the delta rule or gradient descent. The weights that minimize the error function are then considered to be a solution to the learning problem. The network behaviour is analogous to a human that is shown a set of data and is asked to classify them into predefined classes. Like a human, it will come up with theories about how the samples fit into the classes. These are then tested against the correct outputs to see how accurate the guesses of the network are. Radical changes in the latest theory are indicated by large changes in the weights, and small changes may be seen as minor adjustments to the theory. There are also issues regarding generalizing a neural network. Issues to consider are problems associated with under-training and over-training data. Under-training can occur when the neural

network is not complex enough to detect a pattern in a complicated data set. This is usually the result of networks with so few hidden nodes that it cannot accurately represent the solution, therefore under-fitting the data.

6 Gradient Descent Method

The steepest descent method is a general minimization method which updates parameter values in the *downhill* direction, which is opposite to the gradient of the objective function. The gradient descent method converges well for problems with simple objective functions. For problems with thousands of parameters, gradient descent methods are sometimes the only viable method [7].

7 Davidon–Fletcher–Powell Method

The Davidon-Fletcher-Powell method or DFP is named after William C. Davidon, Roger Fletcher and Michael J.D. Powell. The method finds the solution to the secant equation that is closest to the current estimate and satisfies the curvature condition [3–6]. It was the first quasi-Newton method to generalize the secant method to a multidimensional problem. The basic idea behind the DFP method is a construction of a series of matrices denoted as D_k, which are approximations of inverse matrices containing second derivatives of target functions or hessians. The substitution of H^{-1} matrix is convenient due to the high computation complexity required for determining its exact form. The approximated matrices are calculated based on knowledge of changes of the objective function gradient in previous iterations of the method. Every iteration of the method relies on the fact that with a given point x_k, a direction is generated:

$$d_k = -D_k \nabla f(x_k)$$

then, in result of minimization in this direct a point x_{k+1} is obtained. In the first iteration matrix $D = I$ is used, where I is a unit matrix. In next iterations approximated matrices are calculated using a following equation:

$$D_k = D_{k-1} + \frac{s_k s_k^T}{r_k^T s_k} - \frac{D_k\, r_k\, r_k^T\, D_k}{r_k^T\, D_k\, r_k},$$

where:

$$r_k = \nabla f(x_k) - \nabla f(x_{k-1})$$
$$s_k = x_k - x_{k-1}.$$

A new value of the function is given by a following equation:

$$f(x_{k+1}) = f(x_k + step \cdot d_k).$$

The procedure finishes if a solution is found, which fulfils particular stopping criteria:

$$\sqrt{(z_k^T z_k)} < eps,$$

where:

$$z_k = -D_k \nabla f(x_k).$$

8 Results

Each patient is characterized with his own personal dynamics of biochemical processes in the arterial blood. Moreover for every patient the aforementioned dynamics is changing during the healing process. Therefore it is extremely hard to build a proper model which could reliably forecast future ABG parameters basing on a large set of historical data. To solve this problem we propose using ANN model working in a real time loop, which is retrained in every time step using only last 72 h of historical data. Each time the new model is developed using two-layer artificial neural network, with the fixed structure comprising of 2 neurons in the first layer and a 1 neuron in the second layer (Fig. 2). Each time we define the learning data as (2) for inputs, and (3) for desired outputs values; arrays values are normalized to $[-1, 1]$ range. Since we used a sigmoid as the activation function on both layers the network can be trained to approximate any function. The experiments we performed with this simple structure proved that it is capable of learning the data and accurately predict future values of blood gasometry [10, 14]. For the neural network consisting of three

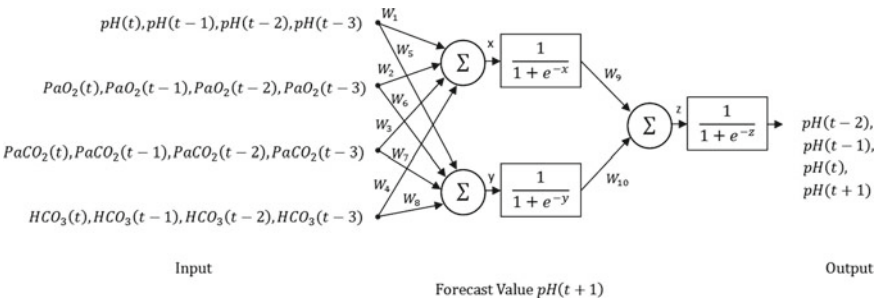


Fig. 2 Artificial neural network structure

Table 1 Statistical results of simulation

NN structure	pH	$PaCO_2$	PaO_2	HCO_3
5-3-1	75.9	90.9	68.6	86.9
5-5-1	79.3	90.0	67.6	87.7
5-10-1	79.9	91.0	67.1	86.2
5-15-1	80.8	89.6	68.5	88.2

Table 2 Optimization results

Method	Mean error for test (%)	Mean time of calculation (s)
Levenberg-Marquardt	10	0.43
Back-Propagation	15	0.02
Gradient descent	7	0.35
Davidon Fletcher Powell	6	0.17

layers the obtained results do not differ significantly from the results obtained for the neural network consisting of two layers. Table 1 contains experimental results for different three-layered network architectures. Numerical values describe percentages of satisfactory prediction results obtained in the course of statistical research for the following parameters: pH, $PaCO_2$, PaO_2, HCO_3.

For the defined artificial neural network architecture the computations were carried out in order to determine w_i weights' values. The computations were carried out using four optimization methods: Levenberga-Marquardt, Gradient descent, Backpropagation and Davidon-Fletcher-Powell. As stop condition we proposed either 100 learning epochs or the learning error value lower than 10^{-7}. Table 2 contains results for different optimization methods. Davidon-Fletcher-Powell method achieves the smallest value of mean error of the testing phase. Backpropagation method achieves the smallest value of mean time of calculations.

In the testing phase we use:

$$V = \begin{bmatrix} \mathrm{pH}_t \\ \mathrm{PaO2}_t \\ \mathrm{PaCO2}_t \\ \mathrm{HCO3}_t \end{bmatrix} \tag{8}$$

vector as the input to the already trained network to obtain the forecast of pH_{t+1} value after next 12 h on the output. The process is repeated every 12 h, when new measured data arrives. As a measure of forecast error we propose:

$$e = \frac{\mid \mathrm{pH}_{t+1} - \mathrm{pH}^*_{t+1} \mid}{\max(\mathrm{pH}) - \min(\mathrm{pH})}, \tag{9}$$

where pH_{t+1} is the expected value, H^*_{t+1} is the predicted value and pH are all the values during 25 days.

9 Summary

The prediction of blood gasometry parameters' values is a hard and complex issue. The approach proposed in this work allows continuous prediction of chosen blood gasometry parameters during the course of hospitalization and refinement of optimization methods used for the prediction of chosen parameters. In the work the architecture of two-layered neural network models was selected over the architecture of three-layered neural network models basing on the results of computations. The obtained model of two-layered neural network structure achieved better generalization properties in the testing phase. Selection of the simpler structure allowed to decrease computational complexity of optimization algorithms for the chosen group of methods.

References

1. Antoniou, A., Lu, W.: Practical Optimization: Algorithms and Engineering Applications. Springer (2007)
2. Aaron, S.D., Vandemheen, K.L., Naftel, S.A., Lewis, M.J., Rodger, M.A.: Topical tetracaine prior to arterial puncture: a randomized, placebo-controlled clinical trial. Respir Med. **97**(11), 1195–1199 (2003) (PMID 14635973)
3. Davidon, W.C.: Variable Metric Method for Minimization. A.E.C. Research and Development Report, ANL-5990 (1959)
4. Fletcher, R.: A new approach to variable metric algorithms. Comput. J. **13**, 317–322 (1970)
5. Fletcher, R.: Practical Methods of Optimization. Wiley (1987)
6. Fletcher, R., Powell, M.J.D.: A rapidly convergent descent method for minimization. Comput. J. **6**, 163–168 (1963)
7. Kelley, C.T.: Iterative Methods for Optimization. North Carolina State University, SIAM (1999)
8. Kofstad, J.: Blood gases and hypothermia: some theoretical and practical considerations. Scand. J. Clin. Lab Invest. (Suppl) **224**, 21–26 (1996) (PMID 8865418)
9. Levenberg, K.: A method for the solution of certain problems in least squares. Quart. Appl. Math. **2**, 164–168 (1944)
10. Lippman, R.P.: An introduction to computing with neural nets. IEEE ASSP Mag. 4–22 (1987)
11. Lourakis, M.I.A.: A brief description of the Levenberg-Marquardt algorithm implemented by levmar. Technical Report, Institute of Computer Science, Foundation for Research and Technology, Hellas (2005)
12. Marquardt, D.: An algorithm for least-squares estimation of nonlinear parameters. SIAM J. Appl. Math. **11**, 431–441 (1963)
13. Raoufy, M.R., Eftekhari, P., Gharibzadeh, S., Masjedi, M.R.: Predicting arterial blood gas values from venous samples in patients with acute exacerbation chronic obstructive pulmonary disease using artificial neural network. J. Med. Syst. **35**(4), 483–488 (2011)
14. Tadeusiewicz, R.: Neural Network as a tool for modeling of biological systems. Bio-Algor. Med. Syst. **11**(3), 135–144 (2015)
15. Transtrum, M.K., Machta, B.B., Sethna, J.P.: Why are nonlinear fits to data so challenging? Phys. Rev. Lett. **104**, 060201 (2010)

Coughing Record Analysis Using Matlab Signal Processing Toolbox

Lukas Behun and Milan Smetana

Abstract Coughing is well-known and theoretically described process. It can be described as a pathological phenomenon, because it is secondary phenomenon of many diseases. Mathematical analysis as well as its detection (quantification and classification) is solved for many years. This article deals with analysis of the various coughing sound recordings. Three different coughing types are analyzed under the same conditions. Matlab signal processing toolbox is used for this purpose. Theoretical background as well as obtained results are presented and discussed in the paper.

Keywords Cough · Analysis · Signal proccessing

1 Introduction

Cough is the defense mechanism of the respiratory tract. It is caused by mechanical or chemical excitation of the cough receptors in the epithelial mucosa of the larynx, trachea and bronchioles. Receptors are primary concentrated in the places, where respiratory irritation coughing reflex is triggered. It consists of a deep initial inspirium, usually by nose, followed by expiratory muscles active contraction. The final stage involves thoracic and abdominal muscles. The glottis is closed during expiratory muscles work-regime. It is possible to distinguish between compressive and expulsive phase. Compressive phase consist of the increase of intrapulmonary pressure. When a maximum intrapulmonary pressure is reached, glottis is quickly opened and a compressive phase is continuously followed by expulsive one. Due to the large pressure gradient between alveoli and atmosphere, the air flows at relatively high speed outside of the lung. This turbulent air flow is carrying the antigens to upper parts of the respiratory tract and the mouth [3].

L. Behun (✉) · M. Smetana
University of Zilina, 010 26 Zilina, Slovakia
e-mail: lukas.behun@fel.uniza.sk
URL: http://ktebi.uniza.sk/

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_37

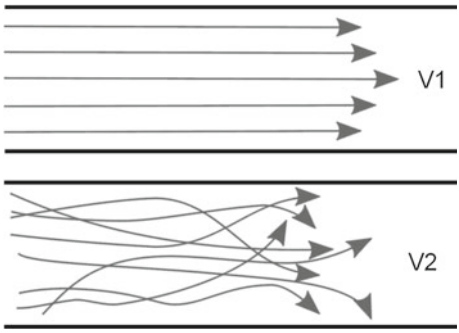


Fig. 1 Vector field of the turbulent and laminar flow [4]

A cough is represented by the sudden exclusion of air, from the respiratory tract. This process is characterized by the typical sound phenomenon: it consists of airflow through the respiratory tract. Here, due to the numerous roughness, splitting and reduction of the surface, laminar flow becomes turbulent (Fig. 1). It can be seen that the chaotic motion of the particles is presented and this is defined by the Raynold's number (R). This turbulent flow is the main reason to generation for a very specific sound, which occurs only during coughing [2, 3].

$$R = \frac{\rho\, d\, v}{\eta} \tag{1}$$

Coughing record signal is the stochastic, non-stationary signal. It is very hard to predict such type of signal, because of its characteristics (random variables, non-stationary, variable character at the same person during the day etc.). This means that the statistical parameters are function of time and may vary in very complex way [1].

2 Experimental Set-Up

All the coughing audio recordings were performed at the department of pulmonary diseases, JFMED UNIBA. Patients suffered of various respiratory diseases. All the recordings were obtained in the hospital rooms. The recording device (classical electret microphone) was located on the surface of the sternum area, for each patient. The microphone was covered by the housing, to be the surrounded noises were eliminated. Technical parameters of the recordings: sampling frequency of f = 11.025 kHz, resolution of 16-bits, length of the samples vary for every inspected patient. Every coughing sound record was separated into the several recordings, to be the useful sounds extracted. This was performed mainly to differentiation another presented sounds, such like: speech, sneezing and the noise caused by touching of the microphone by patient. The recordings were divided into two groups: one group consists

of the dry coughing and the second one of mixed coughing. This division was performed by clinician specialist with many years of medical praxis, on the field of pulmonary diseases.

The main aim of this study is to identify presence of coughing in the general sound recordings. Final algorithm is based on the comparing of the signals. Due to this fact, it was necessary to create the database of the voluntary coughing recordings, under defined conditions—different from the mentioned above. These audio recordings were created at the laboratory of the department. The recording device consists of the microphone connected to the sound card of the computer. The classical electret microphone was located in the sternum section, without any cover. Parameters of the recording were set to the same as for previous case. It was supposed that the surrounded environment has no influence to the real shape of the signal and its frequency spectrum.

3 Coughing Records Signal Processing Methods—Theoretical Background

Audio recordings were analyzed by the selected mathematical operations, in the time domain as well as in the frequency domain. The analysis results were compared and based on the specific extracted features, classified into the given groups. Followed analysis were performed: time-domain analysis, wavelet transform (WT) and power spectral density (PSD). These results are presented and discussed in this paper.

3.1 Time Domain Analysis

This method was used to detection of the 2nd coughing sound, which gives the information about the dry coughing. A time domain signals were compared each other while the voluntary sound (raw signal) and real patient sound were taken into the account. Presence of the 2nd coughing signal in the record was recognized.

3.2 Wavelet Transform

Wavelet transformation is one of the most popular of the time-frequency-transformations. Wavelet series is a representation of a square-integrale (real- or complex-valued) function by a certain orthonormal series generated by a wavelet. The fundamental idea of wavelet transforms is that the transformation should allow only changes in time extension, but not shape. This is affected by choosing suitable basis functions that allow for this. Changes in the time extension are expected to conform to the corresponding analysis frequency of the basis function. Based on the uncertainty principle of signal processing,

Useful as a wavelet function must have specific properties for signal processing. These features include [1, 5]:

(a) condition of total energy

$$\int_{-\infty}^{\infty} \psi(t)^2 dt < \infty, \tag{2}$$

(b) condition of feasibility

$$\int_{-\infty}^{\infty} \frac{\psi(\omega)^2}{\omega} d\omega < \infty, \tag{3}$$

(c) condition of zero median

$$\int_{-\infty}^{\infty} \psi(t) dt = 0. \tag{4}$$

Selection of the appropriate parent wave depends on the particular application. A very important fact is the relation of the waveform with the characteristic part of the analyzed signal. Selection and functionality of the wave depends mainly on their properties. Three types of the waves can be defined [1]: (a) complex waves (can detect oscillations and give information about amplitude and phase), (b) anti-symmetric waves (suitable for detection of gradient changes), (c) symmetric waves (has no effect on the phase shifting).

3.2.1 Continuous Wavelet Transform, CWT

CWT is a time-frequency representation of the signal with continuous or discrete time values. In comparison with the STFT (Short-Time Fourier Transform) method, the main difference is that it is not necessary to perform of Fourier Transform. Thus, no negative frequencies are presented in the resulting spectrum. The window-width is changing for individual spectral components, as they are calculated. Continuity of the CWT is not represented by the processed signal, despite of the fact that it calculates with discrete time signal. Transformation coefficients can be given for all the values of the interval (s_{min}, s_{max}) as well as for all the displacements b, where wave is moving along the entire signal. Mathematical expression of the CWT can be given using next formula [1, 6]:

$$x_m(s, b) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(t) \psi^* \left(\frac{t - b}{s} \right) dt, \tag{5}$$

where $\psi(t)$ is so called mother wavelet and the value range on given interval $(-\infty, \infty)$. Value of b is changed in the interval $< 0, \infty)$.

3.3 Power Spectral Density (PSD) Calculation

The PSD term is used when stochastic signals are processed, instead of the spectrum term. PSD describes power density distribution of the signal depending on frequency. Two different approaches may be used to calculate the PSD: (a) using Discrete Fourier Transform (direct and indirect) (b) parametric (the signal is described by the set parameters).

Periodogram: the periodogram after Schuster, known as P(f, N), can be calculated using Fourier Transform as [1, 6]:

$$P(f,N) = \frac{1}{N}\left|\sum_{n=0}^{N-1} x_n e^{\frac{-j2fn\pi}{N}}\right|^2. \tag{6}$$

Then, estimation of the PSD of the analyzed signal x_n can be expressed in the form:

$$PSD(f) = \lim_{N\to\infty} P(f,N). \tag{7}$$

Calculated PSD estimation has following properties: it may be interpreted only for signals with definite length, it is not biased, and its dispersion does not decrease with increasing length of the signal. These facts give information about inadequacy of simple periodogram as a tool for estimation and analyzing of the spectrum. Therefore, a modified or weighted periodogram can be defined [1, 5, 6]:

$$P(f,N) = \frac{\frac{1}{N}\left|\sum_{n=0}^{N-1} w_n x_n e^{\frac{-jfn2\pi}{N}}\right|^2}{\frac{1}{N}\sum_{n=0}^{N-1}|w_n|^2}. \tag{8}$$

4 Experimental Results

Coughing sound records are non-stationary signals. These signals can be analyzed using WT, PSD and many other approaches. These math transformations can characterize various signals, based on their specification, which cannot be observed from the un-processed signals at all. Obtained experimental results performed on the coughing record sounds are described in this section. Three different types of the coughing process were analyzed: voluntary, dry and mixed coughing record. Mixed cough is defined as a dry cough, but with occasionally presence of expectoration of the mucus.

4.1 Time Domain Analysis

The real shape of the three recorded signals is displayed in Fig. 2. It can be seen that the differences between individual waveforms are clearly visible. Dry coughing process

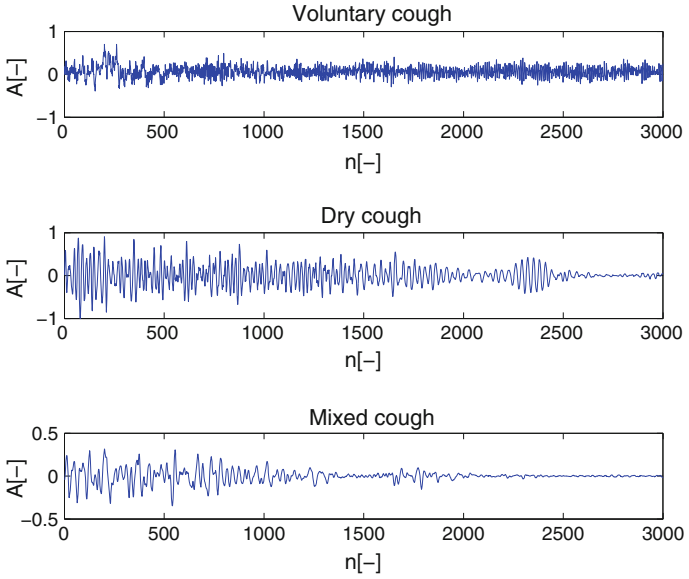


Fig. 2 Experimental results—comparison of the signals in the time domain

is characteristic due to the 2nd cough sound, in the interval within t = 181; 226 ms (n = 2000; 2500 samples). This specific feature is the main characteristic point of this type of coughing. This secondary maximum is not presented at other coughing types. Simple shape signal analysis can now give information about separation between dry and other coughing types.

4.2 Wavelet Transform Signal Analysis

The result of this mathematical operation is so called scalogram, which is percentage ratio of the power of the individual coefficients. Due to this fact, it gives unique information about the processed signals. On the other hand, evaluation of such information is quite difficult. As a reference mother wave was chosen Symlet wave, from the MATLAB software library.

Scalogram obtained for the voluntary coughing is shown at Fig. 3. Its specific characteristics are caused by this fact: first loud sound (opening of the glottis), is followed by the slow changes, what is represented by the lower frequencies. This phenomenon is represented by the higher values of the signal. Further, faster changing of the coughing signal is presented, which corresponds to the higher frequencies of the Scalogram. This means that the component of the signal is very similar to the mother wave. The details are disappearing after the first cough sound. Low-range values, where the signal is similar to the mother wavelet, are clearly defined for this type of cough.

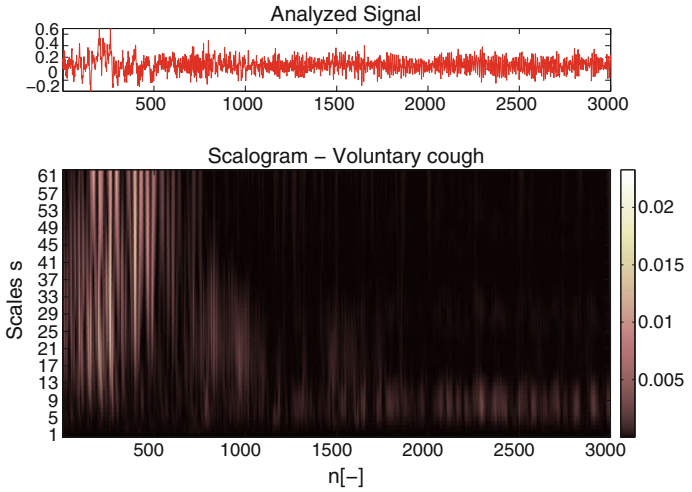


Fig. 3 Scalogram of the analyzed voluntary coughing process

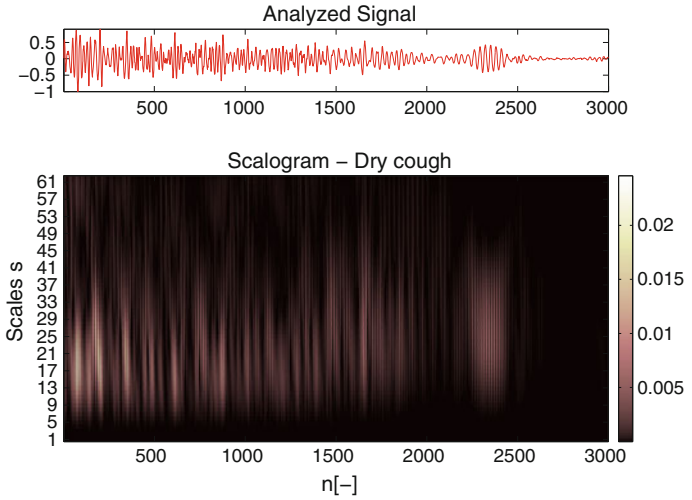


Fig. 4 Scalogram of the analyzed dry coughing process

Analyzing the dry coughing Scalogram (Fig. 4), it can be seen a low similarity to the voluntary coughing. This is caused by the fact, that the first dry cough sound is followed by the higher frequencies. In contrast, the voluntary cough is changing relatively slowly. Main difference against other cough types is the presence of the 2nd sound, in the analyzed signal. Identification of this sound type is based on the similarity of this sound phenomenon with the mother-wave.

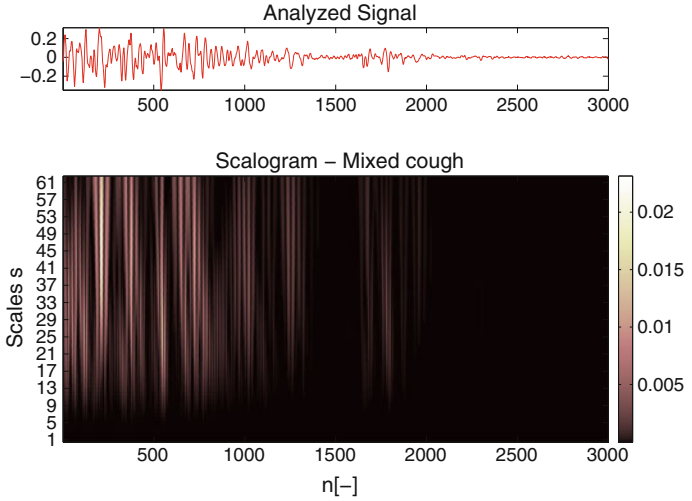


Fig. 5 Scalogram of the analyzed mixed coughing process

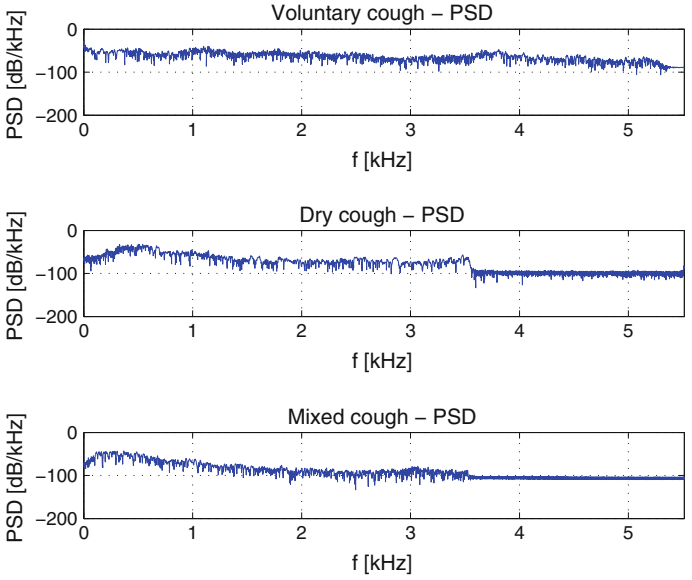


Fig. 6 Experimental results—PSD based calculations of coughing process

Scalogram of the mixed cough (Fig. 5) is specific mainly due to the distribution of the energy of the low-frequency coefficients, during the all record. It consists of mainly slowly-changing details. Part scalogram for the first cough sound is similar to other types of cough. However, it is different just the fact that it contains the said slow varying detail, that is, low frequencies at which the mother wavelet strongly

correlated to a given portion of the signal. Scalogram also identifies the 2nd cough sound, but the coefficient values are relatively low. All the CWT coefficients have lower values, which also can be taken as typical for the mixed cough.

4.3 PSD Based Coughing Signal Analysis

The result of the PSD based calculation is presented in this section, Fig. 6. Based on the results, it can be seen that the energy, as the function of time, is unequally distributed over the frequency range. Moreover, the amplitudes of the signals are also different, each other. It can be concluded that the global extremes of the time domain coughing signals correspond to their calculated PSD waveforms.

5 Conclusion

Coughing record detection and mainly its analysis is very important problem, solved for last decades. Objectification of the coughing process is also complex task, which did not be successfully solved yet. This article dealt with coughing record analysis, while three different coughing types were taken into the account. Selected mathematical methods were used to analyze the coughing record signal. The main aim was to separate individual signal types, from the medical point of view. Time-domain as well as he frequency domain analysis was presented. In total, recordings from ten patients with pulmonary disease were analyzed. It means that approximately one thousand coughing sound recording were analyzed as well. Based on the results, it can be concluded: the dry cough can be clearly separated from the mixed and voluntary coughing. Further, power spectral density calculation gives advance information about the analyzed signals, to be separated. Although, the main problem of cough process detection and analysis, lie in huge variability of the coughing signals. This one changes not only during the day between individual patients, even in one patient. Further work of the authors will focus to obtain more complex information about cough type, in the coughing record.

Acknowledgments We thank to prof. Milos Tatar, Jessenius Faculty of Medicine in Martin, for the coughing sound recordings and advices to solve the proposed problems.

References

1. Gombarska, D., Babusiak, B., Gala, M.: Signal Processing in the Medicine, 1st edn. EDIS—Publishing, Zilina (2013). 156s. ISBN: 978-80-554-0669-5
2. Jakus, J.: Biophysics of heart and blood vessels. Heart as a source of electrical potentials and as a pump. Blood flow in vessels. www.jfmed.uniba.sk/index.php?id=3646

3. Javorka a kol, K.: Medical Physiology for Medical Faculties. 1st edn. Osveta, Martin (2001). ISBN: 80-8063-023-2
4. Lamtur. https://cnx.org/resources/dc559db55e62459c13768c16214502004ae61a7a/graphics12
5. Proakis, J.G., Manolakis D.: Digital Signal Processing: Principles, Algorithms, and Applications, Prentice-Hall, Inc. (1996). ISBN 0-13-394338-9
6. Rabiner, Gold: Theory and Application of Digital Signal Processing, Prentice Hall Inc., Englewood Cliffs, 1975, ISSN 0018-9472

Part III
Medical Information Systems and Database

Integrated System for Clinical Decision Support in Emergency Stroke Care

Artur Przelaskowski, Ewa Sobieszczuk, Rafal Jóźwiak, Dominika Życka-Malesa, Ihor Mykhalevych, Katarzyna Sklinda and Antoni Sobkowicz

Abstract In this work we present a system for decision support in emergency stroke care developed as a part of integrated health-care environment. Presented system consists of four interconnected modules, each responsible for supporting medical staff in making correct decision at each step of emergency stroke care: Extra Rapid Transport module, Neurological Diagnosis module, Ischemic Stroke Confirmation module and Thrombolytic Therapy module. Initial experiments were reported with concluded clinical remarks.

Keywords Clinical decision support · Computerized emergency systems · Stroke diagnosis

1 Introduction

Acute stroke diagnosis and treatment are the problems which scientists from various branches have been struggling for years. Despite the fact that methods of all these disciplines are constantly improved, stroke still remains one of the major causes of death and long-term disability. According to the *time is brain* paradigm all stages of patient care from occurrence of stroke symptoms to appropriate treatment application should be performed as fast as possible because every second without treatment means loss of brain cells [18].

A. Przelaskowski (✉) · R. Jóźwiak · D. Życka-Malesa · I. Mykhalevych · A. Sobkowicz
Faculty of Mathematics and Information Science, Politechnika Warszawska,
Koszykowa 75, 00-662 Warszawa, Poland
e-mail: arturp@mini.pw.edu.pl

E. Sobieszczuk
Department of Neurology, Medical University of Warsaw, Banacha 1a,
02-097 Warszawa, Poland

K. Sklinda
Department of Radiology and Diagnostic Imaging,
Medical Center of Postgraduate Education, Warsaw, Poland

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_38

More than 80 % of all strokes are ischemic strokes, which occur as a result of acute reduction of blood flow through brain arteries due to their occlusion [1]. The only FDA approved treatment in case of acute ischemic stroke (AIS) is thrombolysis. The thrombolytic drugs dissolve clot in vessel and restore a blood flow in the stroke area.

Intravenous thrombolysis with rtPA (recombinant tissue plasminogen activator) for acute ischaemic stroke has been proved to be effective when given within three hours of onset of stroke symptoms but the sooner application means the better effects [4, 22]. Thrombolysis, however, in some cases may cause some serious side effects like symptomatic intracranial haemorrhage (SICH) [11]. Therefore, decision about thrombolytic drug administration requires a rapid consideration balancing benefits and risks of therapy for a particular patient. The judgment should be based on all available patient's data including results of physical examination (neurologic evaluation by one of the formal stroke scales like NIHSS is recommended), laboratory tests and brain imaging [4]. In the American Heart Association guidelines is also mentioned that brain imaging and interpretation is essential part of rapid diagnosis. In case of suspicion of stroke, it is recommended to perform brain imaging with the computed tomography (CT) or the magnetic resonance (MR). Non-contrast CT (NCCT) is more commonly used because of its advantages of being widely available, fast, inexpensive and relatively easy to interpret. It also allows exclusion of parenchymal haemorrhage which is major contraindication for the thrombolytic treatment [4, 7]. Important limitation of NCCT is poor visibility of ischemia during the first hours from the stroke symptoms onset. Signs of AIS are often indistinct and having a high intra- and inter observer variability [7]. Usage of quantitative scores like Alberta Stroke Program Early CT Score (ASPECTS) may help in image interpretation [4, 12].

The difficulty in stroke diagnosis and decision making lies in performing analysis of large amount of data of various types under time pressure and in severe conditions of emergency department. There is an obvious need for all kinds of tools that can support neurologist in decision making process. Clinical Decision Support (CDS) is rapidly growing area of research in Health Information Technology and what is important, using CDS system can generally improve process which it supports [5].

In this paper we described such system in early stage of development. It is based on clinical guidelines, neurologists experience and overall information about patient including demographic data, risk factors, laboratory tests results, medical history, stroke scales and brain imaging. System supports also NCCT-based diagnosis and integrates image interpretation with formalized neurological symptoms and a physical examination outcome.

2 Related Work

Recently a series of tools (including CDS) have been developed to assist physicians in various fields of stroke care. IStroke is [21] a tool that provides rapid screening of relevant medical history in respect of the thrombolytic therapy contraindications. Pilot study showed that extraction of important information from patients record

may be a valid element of CDS system. In [2, 9] COMPuterised decision Aid for Stroke thrombolysiS (COMPASS) was described. COMPASS calculates prediction for acute stroke outcomes (i.e. bleeding, death or extent of disability) depending on the thrombolytic drug application. Prediction is based on patients' individual characteristics. It takes into consideration results of brain imaging but does not provide aid in the stage of image interpretation. A computer aided diagnosis (CAD) system for NCCT images analysis has been presented in [19]. This is the first end-to-end automatic ASPECTS scoring system using a novel method of contralateral comparison. It supplies information about early ischemic changes on NCCT images which should facilitate decision about using the thrombolytic treatment. The system provides good accuracy with 90.2 % area under ROC curve (AUC). Important limitation is the fact that it analyzes only image data and therefore it does not consider other clinical information about the patient. In [10] the role of brain imaging in a context of recently positive clinical trials has been reviewed. In conclusions authors emphasized that although imaging remains crucial for treatment decisions, it has to be interpreted and synthesized with clinical information.

The decision-making is considered to be an evidence-based practice [6, 16]. It requires tools that link the best available evidence and clinical practice. This type of management plans are called clinical pathways. They provide recommendations, processes and timing of actions needed to achieve the objectives of care with optimal efficiency. Electronic Stroke CarePath [6] has been developed for management of ischemic stroke patients. It uses Health Information Technology for integration of standardized pathways into clinical workflows and collection of patient-reported information. It has been reported that implementation of the Electronic Stroke CarePath is associated with the improved outcomes after ischemic stroke. Should not be forgotten that using clinical pathways and guidelines entails reduction of prehospital and in-hospital delays which is essential for such a time sensitive therapy as thrombolysis is.

3 System Description

Proposed stroke care decision support system was based on a simplified clinical workflow of prehospital and in-hospital care with modeled guidelines of clue decision modules. The workflow was characterized by short description of decisive modules as follows. The modules were defined by logical and arithmetic operations applied to respective accessible input data and clinical activity characterized with accumulated pathways, guidelines and decision diagrams, processed, experimentally consulted and verified by interdisciplinary team of neurologists, radiologists, paramedics and medical engineers. Estimated output factors intended to support clinical decisions in a sequence of mutually dependent emergency activities optimized with criterion of the most effective thrombolysis regularized by strict time limits, credibility, completeness and coherence of gathered information. Serial order of the four decision-aiding modules reflects accumulated in parallel case-knowledge and increasing the weight,

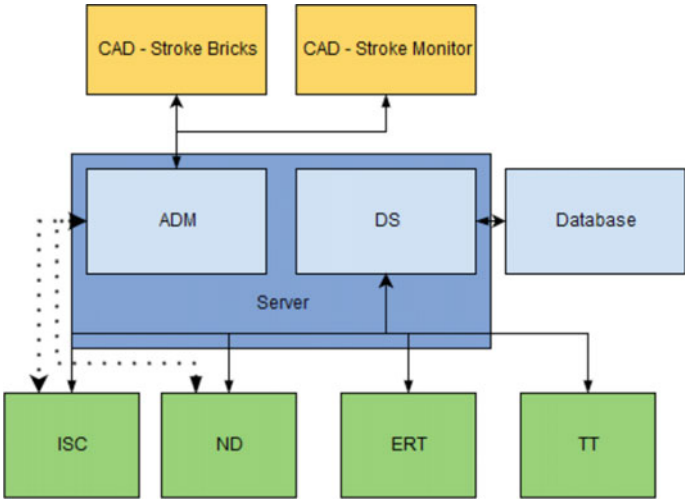


Fig. 1 Block diagram of system design

consistency and consequence of decision making. The representative schemes including decisive modules, input and output data streams and supplementary CAD tools were presented in Figs. 3, 4, 5 and 6, together with briefly characterized implemented interfaces and general concept of the realized system.

3.1 System Design

For an instance the direct to CT (DtoCT) protocol used in several British hospitals was successful in reducing pathway delays [6]. System is divided into five interconnected parts (Fig. 1). The server application that stores and retrieves patient data and four front-end interfaces which are responsible for data input and calculations for each step of stroke emergency care described in details in [15]. Front-end modules are ordered in a sequence as follows:

- ERT, responsible for Extra Rapid Transport module,
- ND, responsible for Neurological Diagnosis module,
- ISC, responsible for Ischemic Stroke Confirmation module,
- TT, responsible for Thrombolytic Therapy module.

Each module is depended on the output of the previous ones, accumulating new information and thus increasing accuracy of subsequent decisions.

3.2 Server Software

Server software itself is divided into two parts—patient data storage part and auxiliary data manipulation part. The software itself is written as a Python REST server with a SQL database and provides basic CRUD operations.

Data Storage Part (DS) is responsible for providing convenient and effective way of data manipulation needed for the CAD engine. The storage is implemented as the separate server application with the REST service as the public available API. The application is developed in the Python environment and consists of the relational database wrapped with the CRUD operations in the REST API service. Database structure is simple and contains only a single table in which patient analysis data is bound to the patient's ID.

Auxiliary Data Manipulation Part (ADM) is responsible for manipulating data before sending to front-ends. It is connected to CAD systems described in [15]—Stroke Bricks and Stroke Monitor.

3.2.1 CAD Systems

Stroke Monitor (*SM*) [14] is a computer tool supporting acute ischemic stroke diagnosis. Emergency noncontrast, routine CT scans are enhanced to increase visibility of hypodensity changes in hyperacute ischemic stroke cases. The processing algorithm of SM is based on multiscale image data processing, denoising, lesion pattern identification and description, and final extraction optimized by visualization procedure. Combining the effects of standard CT scans review with SM assistance may lead to a better diagnosis of stroke. Stroke Monitor provide a new semantic-visualization system of empowered hypodensity symptoms, realized according to elaborated and optimized four different forms of multiscale processing, localizes suggested ischemic areas in the source brain image space. Observation of suggestive ischemic density changes occurrence combined with correlation analysis of their localization related to clinical manifestation, made image content assessment and interpretation more accurate and simple. A good example of different semantic-visualization forms (semantic maps) of SM were presented in Fig. 2.

Stroke Bricks (*StBr*) [13] is a topographic division of brain tissue into the regions associated with specific clinical symptoms of ischemic stroke. It was inspired by the idea of APSECTS regions, used in quantitative evaluation of early ischemic changes within the middle cerebral artery territory on non-contrast computed tomography (NCCT). StBr, as opposed to ASPECTS, takes into consideration whole brain including both the anterior and posterior circulation. The relationship between the observed symptoms and the expected impairment of brain (expressed in StBr regions) was proposed in our solution. StBr conception allows to define the expected location and extent of ischemia, thus to reliably confirm ischemic stroke diagnosis in order to improve the accuracy of therapeutic decision for reperfusion therapy with both

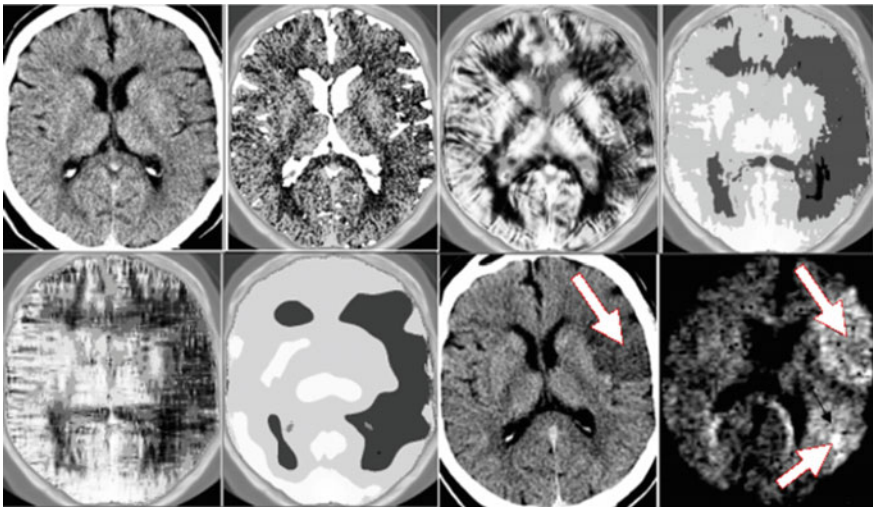


Fig. 2 The effects of SM assistance for stroke diagnosis (from *upper-left*): view of CT scan (early CT), enhanced visibility of diagnostic ROI with segmented unusual areas (*white*), four visualization forms of SM (semantic maps), follow up CT and DWI with indicated visible ischemic changes (stroke confirmation)

intravenous tPA and endovascular stroke therapies (including intra-arterial thrombolysis and mechanical thrombectomy). StBr was integrated into the diagnosis pathway of acute ischemic stroke as an additional knowledge for radiologists during NCCT interpretation, as well as a part of computer-aided diagnosis system to reduce the number of computer suggestions which are the most clinically feasible.

3.3 Front-End

Front-end modules are developed in HTML5 and JavaScript. Each module is divided into two parts—User Interface (UI) and Calculation and Communication Back-end (CaCB). UI is responsible for displaying controls to the user, while CaCB collects entered data, calculates scores and communicates with Server. Brief characteristics of these modules underlined the most important output factors forming the most important form of supporting suggestions.

ERT: Extra-rapid Transfer module is designed to be used by paramedical personnel that decides if patient qualifies for fast transport to medical facility. It allows entry of basic patient information, as well as information regarding patient status, such as sudden signs of neurological deficit, vital signs measurement and additional information regarding patient history. Module output is logical value for rapid transfer suggestion or contraindication (*ERT*) and numerical score for probability of stroke (*EPS*).

ND: Neurological Diagnosis module is designed to be used by medical personnel that decides if patient qualifies for thrombolysis. It allows entry of NIHSS and Stroke Bricks Score, reassessment of contraindications (with optional display of values entered in ERT module). Module output is a logical value for thrombolysis exclusion (*TTE*), and two numerical scores—neurological certainty of stroke (*NCS*) and normalized representation of stroke severity (*NSS*).

ISC: Ischemic Stroke Care module is designed to provide reliable confirmation of ischemic stroke presence. It allows display of NCCT imaging data (augmented by CAD procedures), and retrieves data calculated by Stroke Monitor—(via server part of the software). Module output is a logical values for exclusion of thrombolysis,

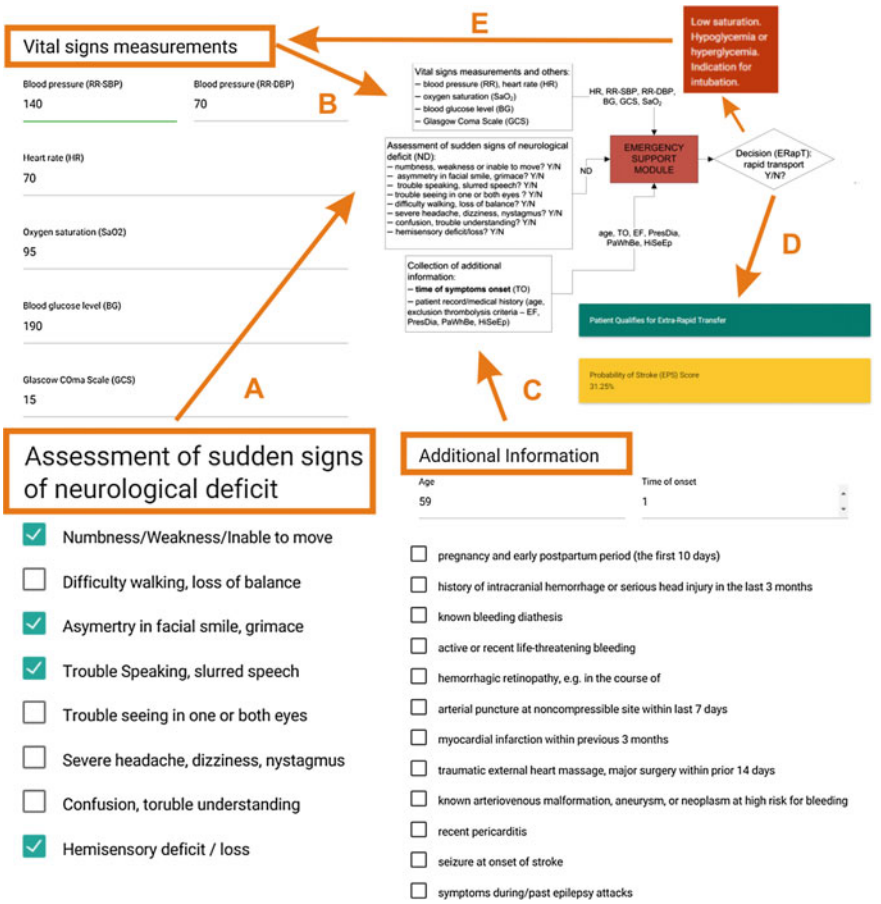


Fig. 3 GUI representing the Extra-rapid Transfer module (related to the emergency support module on functional diagram). *A* Assessment of sudden signs of neurological deficit, *B* Vital signs measurements, *C* Additional information, *D* the result of the module computed using input data (*A*–*C*). *E* Additional hints (if needed)

confirmation of ischemic stroke as logical suggestion (*ISC*) and numerical values for integrated estimate of stroke probability (*IPS*), the extent of possibly reversible ischemic changes (*ERI*) and integrated prediction of treatment outcome (*IOP*).

TT: Thrombolytic Therapy module is designed to provide information if there are any suggestions or contradictions for application of thrombolytic therapy. This module uses information from previous modules completed by actual results of laboratory examinations. The module output is numerical confirmation of the therapy utility (*UTT*) supporting final evaluation of clinicians. However, the most important factor is logical estimate of treatment admissibility (*CTA*) formulated to strengthen final therapy decision. The logical 1 (true) is reserved for expected clinical improvement with minor risk of SICH.

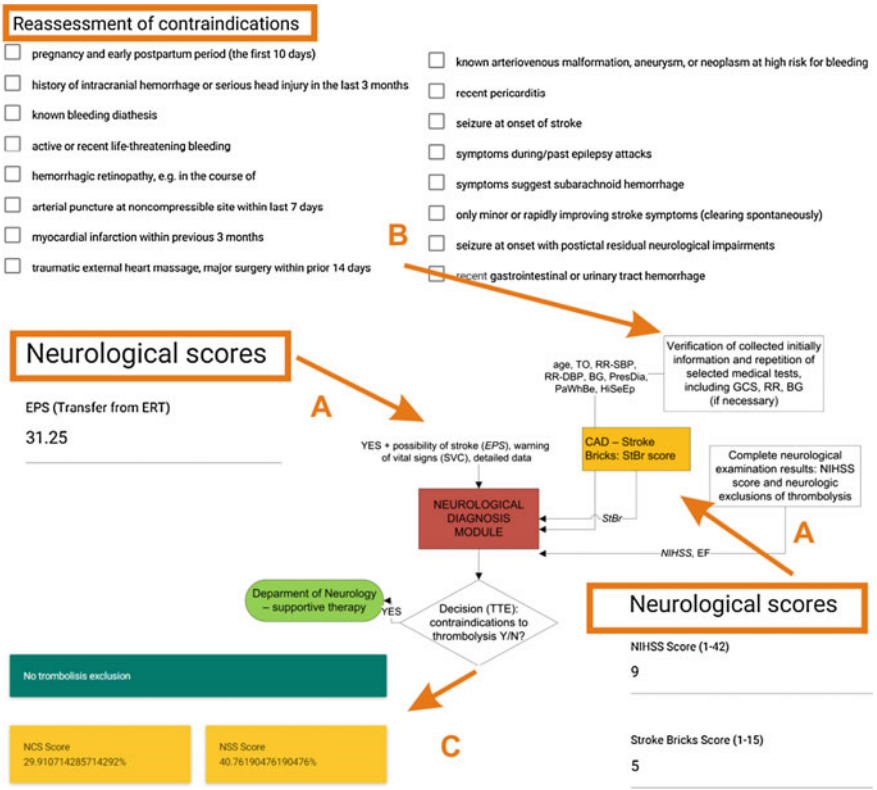


Fig. 4 GUI representing the Neurological Diagnosis module. *A* the Emergency Support module results together with Stroke Bricks and neurological examination results, *B* Revalidation of initially collected information, *C* the result of the module

3.4 Implementation

Each front-end module is designed in similar way. Patient data can be entered to the system using simple controls, such as:

- series of check-boxes, which are used for entering binary data, such as Yes/No assessment;
- text inputs, which are used for entering text and numerical data;
- number inputs, which are used strictly for numerical data input.

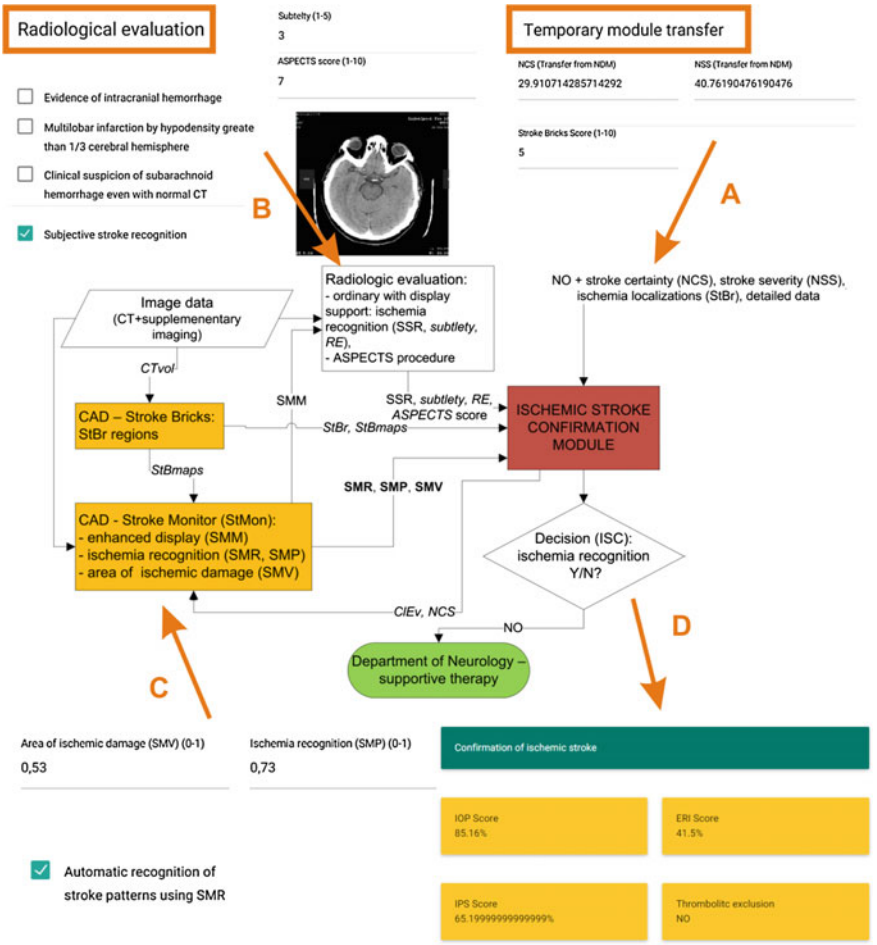


Fig. 5 GUI representing the Ischemic Stroke Confirmation module. *A* the Neurological Diagnosis module results, *B* Radiological evaluation, *C* CAD evaluation, *D* the result of the module

Scores and guidelines are calculated after clicking *Check* button, using inputted data. Calculated scores and guidelines are displayed in color coded panels:

- yellow panels are used for score display;
- green panels are used for non critical information;
- red panels contain critical information or information about contraindications.

Contraindications and critical information are presented in simple way and are easily visually distinguishable from rest of the interface so they cannot be missed.

Realized application interfaces were presented and explained in relation to the functional schemes of the system (see details in [15]) in Figs. 3, 4, 5 and 6, respectively.

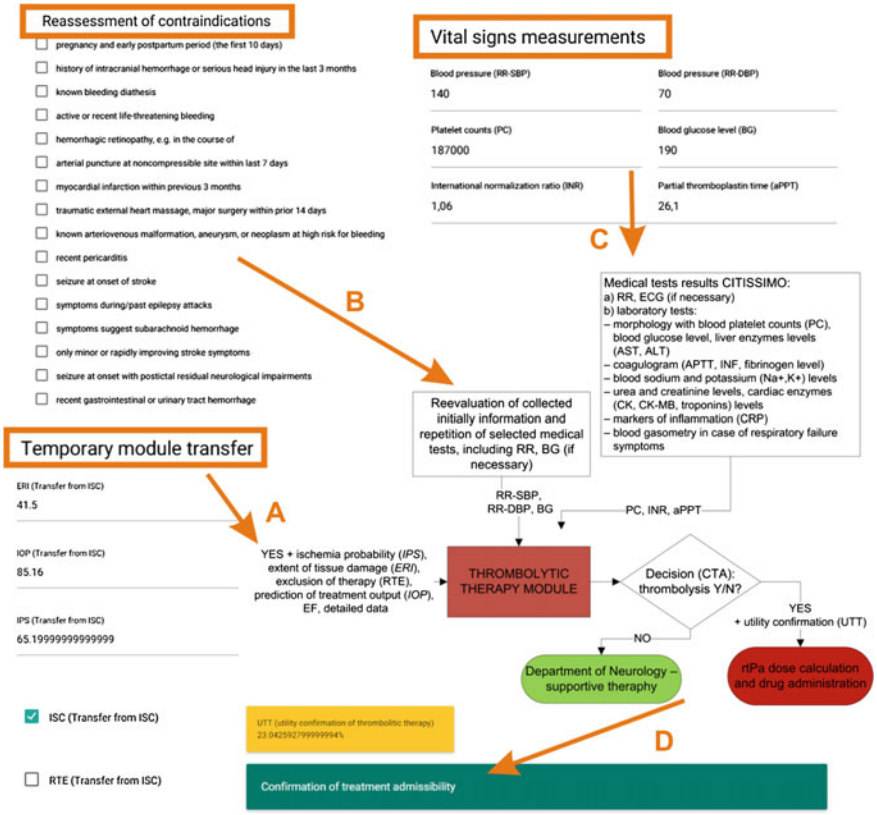


Fig. 6 GUI representing the Thrombolytic Therapy module. *A* the Ischemic Stroke Confirmation module results, *B* Revalidation of initially collected information, *C* Medical tests results (CITISSIMO), *D* the result of the module

4 Experiments

The alpha prototype of the decision support system has been preliminary tested. The aim was to study system usability and partly its feasibility. The main goals focused on verification of the system operation, removal of possible errors or defects and collection of recommendations for further improvements. On the current stage of development, the most important was to receive feedback from neurologists on validity of such a system. Initial test phase has been carried out on a limited number of cases to catch any errors before main test phase, which will be performed in future. Tests have been conducted by the group of researchers consisting of experienced neurologist and four engineers (specialists in the fields of CAD, image processing and software development).

We retrospectively analyzed 10 patients from which 6 underwent thrombolytic treatment and other 4 did not. Patient were selected from DDIS II database (Digital Database of Ischemic Stroke Cases [3]). Baseline clinical characteristics of selected test group including general patient info (A—age, G—gender, T—time from symptoms onset), vital signs measurements (SBP—systolic blood pressure, DBP—diastolic blood pressure, HR—heart rate, $Sa0_2$—oxygen saturation), laboratory test results (PC—platelet concentrates, INR—international normalized ratio, BG—blood glucose, aPPT—activated partial thromboplastin time) and patient neurological description (neurological symptoms, NIHSS—The National Institutes of Health Stroke Scale, GCS—Glasgow Comma Scale) is presented in Table 1.

The final system results were considered as a potentially useful indicator to support clinical decision-making for rtPa treatment. Especially *CTA* and *UTT* values were expected to correlate with clinical subjective assessment of therapy effects. Because of the retrospective nature of the study, available clinical outcome was treated as gold standard to verify system suggestions. Subjective assessment of treatment outcome (SATO) was determined on the basis of epicrisis (neurological condition, occurrence of SICH and NIHSS at discharge). Moreover, additional objective clinical outcome scores such as HAT [8, 20] or iScore [17] were calculated and used for system results analysis and clinical verification. The results of initial experiments are presented in Table 2.

As the results, *UTT* and other system factors such as *CTA*, *ISC* or *EOP* seem to be correlated with clinical outcome. *UTT* proved to be more sensitive parameter for predicting thrombolytic treatment success than likelihood of good outcome after rtPa (LoGO) calculated in DRAGON scale or probability of good clinical outcome (PoGO) calculated basing on iScore. Patients for whom rtPA treatment resulted as major success (total or subtotal remission) reached *UTT* near 70 % or more (e.g. case 5). Patients qualified for thrombolytic treatment who did not reach significant success or for whom thrombolytic therapy ended in deterioration (e.g. case 4), reached 50–59 % *UTT*. It is important to emphasize that likelihood of good clinical outcome (LoGO) calculated in DRAGON scale in 5 from 6 thrombolyzed patients was 63 % and did not differ between groups with poor outcome in comparison with those with good outcome.

Table 1 Baseline clinical characteristics of selected test group

General info					Vital signs				Laboratory tests				Neurological description		
Patient	No	A	G	T (h)	SBP (mmHg)	DBP (mmHg)	HR (bpm)	$Sa0_2$ (%)	PC ($10^3/\mu l$)	INR –	BG (mg/dl)	aPPT (s)	Neurological symptoms	NIHSS	GCS
rtPa	1	59	M	1	140	70	70	95	187	1.06	190	26.1	Central hypoglossal nerve palsy, hemiparesis, hemisensory loss, dysarthria	9	15
	2	42	M	0.75	129	84	77	95	242	0.97	124	28.4	Gaze palsy, expressive (motor) aphasia, sensory, aphasia, hemiparesis, hemisensory loss	16	15
	3	48	F	1.83	150	90	98	96	453	1.04	85	27.3	Gaze palsy, expressive (motor) aphasia, sensory, aphasia, central hypoglossal nerve palsy, hemiparesis	25	3
	4	71	F	0.83	170	90	58	97	255	0.99	106	29	Gaze palsy, central hypoglossal nerve palsy, hemiparesis, hemisensory loss	15	15
	5	77	F	0.83	170	80	80	98	180	0.93	130	28.3	Expressive (motor) aphasia, sensory aphasia, central hypoglossal nerve palsy, hemiparesis	11	12
	6	77	M	2	160	100	88	99	168	1.1	97	23.8	Expressive (motor) aphasia, sensory aphasia, central hypoglossal nerve palsy, hemiparesis	16	15

(continued)

Table 1 (continued)

General info					Vital signs				Laboratory tests				Neurological description		
no rtPa	7	75	F	2	130	70	62	98	304	0.91	116	25.4	Gaze palsy, expressive and sensory aphasia, hemiparesis, central hypoglossal nerve palsy, hemisensory loss	20	15
	8	81	F	3	220	120	66	97	139	1.06	89	24.1	Expressive (motor) aphasia, sensory aphasia, dysarthria, central hypoglossal nerve palsy, upper limb paresis	7	14
	9	75	F	4	190	90	70	94	430	0.97	106	26.1	Central hypoglossal nerve palsy, hemiparesis, hemisensory loss, dysarthria	8	15
	10	76	F	1.12	247	87	90	98	185	0.98	106	26.2	Central hypoglossal nerve palsy, hemiparesis, dysarthria	16	15

Patients were divided into two groups—treated with rtPa thrombolytic therapy (6 patients, No. 1–6) and treated without rtPa (supportive therapy, 4 patients, No. 7–10). The applied acronyms were described in text (mostly in Sects. 3 and 4)

Table 2 The results of initial experiments with proposed system supporting clinical decisions in emergency stroke care

Info		**System results**										
Patient	No	**ERT**		**NDM**			**ISC**				**TT**	
		ERT	**EPS**	**TE**	**NCS**	**NSS**	**ISC**	**IPS**	**ERI**	**IOP**	**CTA**	**UTT**
		(T/F)	(%)	(T/F)	(%)	(%)	(T/F)	(%)	(%)	(%)	(T/F)	(%)
rtPa	1	**T**	**66**	**F**	**61.2**	**46.4**	**T**	**77.2**	**46.5**	**50.7**	**T**	**51.3**
	2	**T**	**93**	**F**	**68.1**	**61.6**	**T**	**88**	**65**	**53**	**T**	**50.1**
	3	**T**	**91**	**F**	**75.1**	**72.0**	**T**	**69.2**	**38.5**	**58.3**	**T**	**52.2**
	4	**T**	**73**	**F**	**65.8**	**52.0**	**T**	**58.8**	**40**	**62**	**T**	**59.2**
	5	**T**	**64**	**F**	**63.9**	**49.6**	**T**	**71.2**	**45.5**	**72.9**	**T**	**68.4**
	6	**T**	**66.5**	**F**	**67.4**	**54.4**	**T**	**47.6**	**19.5**	**82.1**	**T**	**74.8**
no rtPa	7	**T**	**93**	**F**	**74.0**	**76.0**	T	**88.4**	**61.5**	**25.7**	F	**25.4**
	8	**T**	**66.5**	**F**	**60.5**	**39.2**	**T**	**80.4**	**57.5**	**48.5**	**F**	**51.0**
	9	**T**	**68.5**	**F**	**57.5**	**32.8**	**T**	**97.6**	**74**	**29.2**	**F**	**36.8**
	10	**T**	**68.5**	**F**	**60.2**	**36.8**	**T**	**100**	**65**	**49**	**F**	**51.8**
Info		Clinical outcome										
Patient	No	SATO	SICH	NIHSS	HAT		DRAGON			iSCORE		
				Out	RoS	RoF	LoGO	LoMO	PoGO	RoICH	30DM	30DD
		–	(y/n)	–	(%)	(%)	(%)	(%)	–	(%)	(%)	(%)
rtPa	1	Poor	No	10	10	7	63	10	1	7.7	8.2	74.5
	2	Avr	No	12	10	7	63	10	1	7.7	6.6	67.2
	3	Avr	No	17	15	6	50	23	0.8	11.2	26.8	88.8
	4	Avr	Yes	12	10	7	63	10	0.95	7.7	14.6	80.4
	5	Good	No	0	10	7	63	10	0.95	7.7	11.1	74.5
	6	Good	No	2	2	0	63	10	0.8	11.2	27.1	91.7
no rtPa	7	n/a	n/a	17	10	7	40	22	0.8	11.2	22.1	85
	8	n/a	n/a	6	5	3	50	23	1.1	2.9	3.5	51
	9	n/a	n/a	8	10	7	63	10	1.25	2.9	2	34.8
	10	n/a	n/a	8	10	7	50	23	1.1	7.7	7.8	67.2

System factors values (highlighted) for all system modules were presented. Most of the applied acronyms were described in text (mainly in Sects. 3 and 4). For HAT score: RoS—risk of symptomatic ICH after rtPa, RoF—risk of fatal ICH after rtPA. For DRAGON score: LoGO—likelihood of good outcome after rtPa, LoMO—likelihood of miserable outcome after rtPa, PoGO—probability of good clinical outcome. For iSCORE: RoICH—risk of SICH, 30DM—risk of 30 day mortality, 30DD—risk of 30 day disability

Patients who did not receive thrombolytic therapy had *UTT* within 25–52 %. The lowest *UTT* was calculated for patient with extensive stroke (case 7, NIHSS on admission 20 points), for whom likelihood of miserable clinical outcome calculated in DRAGON scale was very high (22 %).

Quite low *UTT* calculated for patient with moderate stroke (case 9, NIHSS on admission 8 points), who could possibly benefit from thrombolysis if only the time from onset was not so long, is highly correlated with low risk of 30-day disability without rtPa (30DD), calculated from iScore. Therefore, it was not reasonable to bear the risk of invasive treatment.

The following two patients' *UTT* oscillated around 50 %. Calculated *UTT* values (51 % for case 8 and 51.8 % for case 10) do not confirm unequivocally the benefits of thrombolytic therapy. This suggested the same clinical decisions as made by experienced clinicians at neurology emergency department. Moreover, it may increase the accuracy of clinical decisions, especially at controversial points.

Additionally, three patients (not included in Table 1) with stroke mimics symptoms (patient with Transient Ischemic Attack, TIA) were also evaluated. For all TIA cases system correctly excluded thrombolysis at second stage of evaluation due to mild or rapidly improving (clearing spontaneously) stroke symptoms.

In the opinion of neurologist *UTT* may pose a good parameter supporting clinical decision-making about thrombolytic therapy. It seems to be more specific and more flexible in predicting the real effect of rtPa treatment than scales used so far. Although, to assess its specificity and sensitivity accurately, studies on larger group of patients are necessary.

5 Summary and Future Work

Conducted tests indicate that some formulas on which the system outcomes are calculated (presented in [15]) still needs to be optimized for larger training data set. For some of the input cases appropriate weights and even formulas could be modified in order to obtain more accurate clinical predictions. However, CDS system has been assessed as a useful tool which can provide potential benefits in the area of stroke diagnosis and care. What is important, it imitates clinical practice very well.

Launching of the system and conducted tests also revealed the need for a number of improvements in the form of user interface and possible expansion in the range of diverse organizational conditions and the preferences of individual clinics. In general the system was described as easy to use and intuitive. According to the neurologist's recommendations, modifications of UI (i.e. some additional buttons with extended functionality) will be implemented, and recommendation text will be made clearer to the user. For example, changing *No confirmation of treatment admissibility* to *Carefully consider using rtPa—no confirmation of treatment admissibility*. The pilot study enabled determination of some shortcomings prior to testing more of various cases. Obtained feedback revealed that the system is desirable and justifies its further development. The complexity of the system and the support provide at almost all stages of a stroke patient management was highlighted.

Acknowledgments This publication was funded by the National Science Centre (Poland) based on the decision DEC-2011/03/B/ST7/03649.

References

1. El-Koussy, M., Schroth, G., Brekenfeld, C., Arnold, M.: Imaging of acute ischemic stroke. Eur. Neurol. **72**(5–6), 309–316 (2014)
2. Flynn, D., Nesbitt, D.J., Ford, G.A., McMeekin, P., Rodgers, H., Price, C., Kray, C., Thomson, R.G.: Development of a computerised decision aid for thrombolysis in acute stroke care. BMC Med. Inf. Decis. Mak. **7**, 15:6 (2015)
3. http://www.mini.pw.edu.pl/~aidmed/zasoby
4. Jauch, E.C., Saver, J.L., Adams Jr., H.P., Bruno, A., Connors, J.J., Demaerschalk, B.M., Khatri, P., McMullan Jr., P.W., Qureshi, A.I., Rosenfield, K., Scott, P.A., Summers, D.R., Wang, D.Z., Wintermark, M., Yonas, H.: Guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke **44**(3), 870–947 (2013)
5. Jones, S.S., Rudin, R.S., Perry, T., Shekelle, P.G.: Health information technology: an updated systematic review with a focus on meaningful use. Ann. Intern. Med. **7;160**(1):48–54 (2014)
6. Kendall, J., Dutta, D., Brown, E.: Reducing delay to stroke thrombolysis—lessons learnt from the Stroke 90 Project. Emerg. Med. J. **32**(2), 100–104 (2015)
7. Kurz, K.D., Ringstad, G., Odland, A., Advani, R., Farbu, E., Kurz, M.W.: Radiological imaging in acute ischaemic stroke. Eur. J. Neurol. **23**(Suppl 1), 8–17 (2016)
8. Lou, M., Safdar, A., Selim, M., et. al.: The HAT Score: a simple grading scale for predicting hemorrhage after thrombolysis. Neurology **71**(18), 1417–1423 (2008)
9. McMeekin, P., Flynn, D., Ford, G.A., Rodgers, H., Gray, J., Thompson, R.G.: Development of a decision analytic model to support decision making and risk communication about thrombolytic treatment. BMC Med. Inf. Decis. Mak. **11;15**, 90 (2015)
10. Menon, B.K., Campbell, B.C., Levi, C., Goyal, M.: Role of imaging in current acute ischemic stroke workflow for endovascular therapy. Stroke **46**(6), 1453–1461 (2015)
11. Miller, D.J., Simpson, J.R., Silver, B.: Safety of thrombolysis in acute ischemic stroke: a review of complications, risk factors, and newer technologies. The Neurohospitalist **1**(3), 138–147 (2011)
12. Pexman, J.H., Barber, P.A., Hill, M.D., Sevick, R.J., Demchuk, A.M., Hudon, M.E., Hu, W.Y., Buchan, A.M.: Use of the Alberta stroke program early CT score (ASPECTS) for assessing CT scans in patients with acute stroke. AJNR Am. J. Neuroradiol. **22**(8), 1534–1542 (2001)
13. Przelaskowski, A., Ciszek, B., Jozwiak, R., Domitrz, I., Sobieszczuk, E.: Stroke Bricks—the segments of interests to analyze early stages of the disease evolution. Research report. Warsaw University of Technology (2015)
14. Przelaskowski, A., Ostrek, G., Sklinda, K.: Multiscale extraction of diagnostic content applied for CT brain examinations. Biochem. Biomed. Eng. **29**(4), 25–40 (2009)
15. Przelaskowski, A., Sobieszczuk, E., Spalik, P., Domitrz, I., Ostrek, G., Podsiadly-Marczykowska, T., Jóźwiak, R., Jasionowska, M.: Stroke emergency care: a system of computerized decision support for prehospital and in-hospital phases. Research report. Warsaw University of Technology (2015)
16. Rotter, T., Kinsman, L., James, E., Machotta, A., Gothe, H., Willis, J., Snow, P., Kugler, J.: Clinical pathways: effects on professional practice, patient outcomes, length of stay and hospital costs. Cochrane Database Syst. Rev. **17**(3), CD006632 (2010)
17. Saposnik, G., Fang, J., Kapral, M., Tu, J., Mamdani, M., Austin, P., Johnston, S.: On behalf of the investigators of the registry of the Canadian Stroke Network (RCSN) and the Stroke Outcomes Research Canada (SORCan) working group, The iScore predicts effectiveness of thrombolytic therapy for acute ischemic stroke. Stroke **3**(5), 1315–1322 (2012)
18. Saver, J.L.: Time is brain-quantified. Stroke **37**(1), 263–266 (2006)
19. Shieh, Y., Chang, C.H., Shieh, M., Lee, T.H., Chang, Y.J., Wong, H.F., Chin, S.C., Goodwin, S.: Computer-aided diagnosis of hyperacute stroke with thrombolysis decision support using a contralateral comparative method of CT image analysis. J. Digit. Imaging **27**(3), 392–406 (2014)

20. Strbian, D., Meretoja, A., Ahlhelm, F.J., Pitkaniemi, J., Lyrer, P., Kaste, M., Engelter, S., Tatlisumak, T.: Predicting outcome of IV thrombolysis-treated ischemic stroke patients: the DRAGON score. Neurology **78**(6), 427–432 (2012)
21. Sun, M.C., Chan, J.A.: A clinical decision support tool to screen health records for contraindications to stroke thrombolysis-a pilot study. BMC Med. Inf. Decis. Mak. **14;15**(1), 105 (2015)
22. Wardlaw, J.M., Murray, V., Berge, E., del Zoppo, G.J.: Thrombolysis for acute ischaemic stroke. Cochrane Database Syst. Rev. **7**, CD000213 (2014)

Semantics Discovering in Relational Databases by Pattern-Based Mapping to Association-Oriented Metamodel—A Biomedical Case Study

Marcin Jodłowiec and Marek Krótkiewicz

Abstract The paper presents the issue of mapping database schema within the context of possible partial loss of their semantics. Moreover, authors present methodology of mapping relational models to association-oriented model and briefly present the association-oriented metamodel itself, which is one of the co-authors proposition for novel database metamodel. It is the proposition of productive tool, which is free of other metamodels' defects. The issues described within the paper have been presented in terms of the case study, which is the typical example of modeling problems that are concerned within interest of bioinformatics.

Keywords Association-oriented metamodel · Relational model mapping · Database design patterns

1 Introduction

In this article authors elaborate upon database schema translation. This issue is present in research as well as in everyday database modeling practice for many years now and majority of database designers search for optimal solutions in this area. The traditional methodology of relational databases design assumes the following three stages of design process [2]: conceptual, logical and physical design. Each of the classical stages results in creation of database *model*. However, only the last level of them (physical) delivers the *model*, which is implementable. The transition stages between *models* obtained in particular stages are known as *translations* or *mappings*. This article is dedicated to the reverse task, i.e. the discovery of model primary semantics [7]. The aim of the project described is to propose the database model prepared by the use of association-oriented metamodel AODB, which preserves the completeness of the original semantics, being simultaneously implemented on the

M. Jodłowiec (✉) · M. Krótkiewicz
Institute of Control and Computer Engineering, Opole University of Technology, Opole, Poland
e-mail: marcin.jodlowiec@gmail.com

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_39

physical level. The association-oriented metamodel is completely novel approach to design and implement databases. The author of this metamodel is M. Krótkiewicz, one of the co-authors of this paper.

The proposed approach to database translation is based on the concept of *database design patterns* as the elementary semantic unit [8]. The substantial problem is the identification of design patterns. The design patterns allow to denominate the problem, detach it from the context and identify foremost and recurrent structural dependencies. However, the notion of *design pattern* is not clear, i.e. interpretation of what can be called the design pattern is ambiguous.

In the following sections issues regarding metamodel mapping will be presented in the context of conscientiousness with preservance of the original semantics.

2 Description of Database Design Patterns and Its Implementation in Relational Metamodel

The following section presents design patterns specified by authors, which represent the elementary semantic units of database model conceptualization [1]. Due to the large number of possible database design patterns able to be created within the presented methodology, only a few original patterns have been described. The description of patterns is purely conceptual, it defines concepts and entities occurring in a pattern and relationships between them. For each of patterns presented authors show implementation in relational metamodel [3].

2.1 Design Pattern: List

The *list* pattern is one of the simplest design patterns in terms of semantics. It implements the *master-detail* relationship, where the first element type acts as the linking participant, while the second element type acts as the linked participant. It is important, that each linking element can be connected with many linked elements. The *list* can have the following optional properties defined: *uniqueness*, *attribution*, *filling*, *navigability*, *lifetime*. Uniqueness should be considered as possibility to be linked to the same list several times. Attribution means the necessity of being linked to any *list*. Filling determines, whether the *list* has to have at least one element linked. Finally, the lifetime is a behavioral property, which determines whether deletion of one participant should affect the other one.

$$\mathbf{LIST}(\underline{ID\_LIST});\ \mathbf{ELEMENT}(\underline{ID\_ELEMENT}, ID\_LIST)$$
$$\pi_{\{ID\_LIST\}}(ELEMENT) \subseteq \pi_{\{ID\_LIST\}}(LIST) \cup \{NULL\} \tag{1}$$

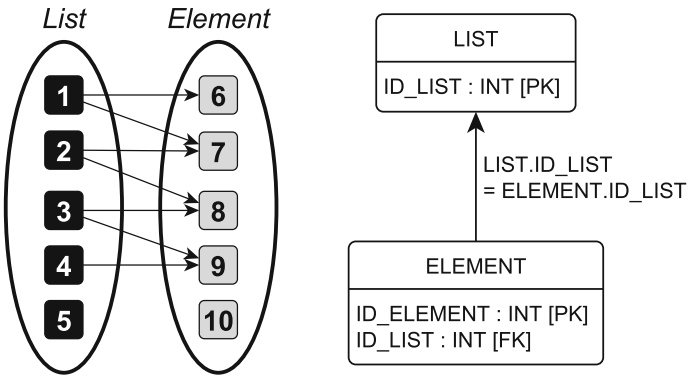


Fig. 1 Conceptual diagram representing *list* design pattern (*left*) and its implementation (*right*) in relational metamodel (1)

The most common implementation of the *list* pattern in RDB disregards the fact, that from the semantic point of view, the list itself posses the information about its elements, but not the opposite (Fig. 1).

2.2 Design Pattern: Dictionary

The *dictionary* design pattern, similar to the *list* comprises the connection of two entities. This relationship always has multiplicity defined as *one-to-many*. The idea of dictionary is to give possibility to model the reference to entity of other kind, which can be also referenced by other entities of the same kind (or other kind in other instance of design pattern). This design pattern basically allows minimization of redundancy in reality description. The fundamental issue of the dictionary pattern is the ability to provide uniqueness of dictionary identity in regard to described elements. The dictionary, in almost all cases, has to provide such uniqueness, but rarely one can identify this pattern without this restriction. However, authors do not recommend such approach due to possible difficulties in `NULL` value semantics determination (Fig. 2).

$$\mathbf{DICT}(\underline{ID\_DICT});\ \mathbf{ELEMENT}(\underline{ID\_ELEMENT}, ID\_DICT)$$
$$\pi_{\{ID\_DICT\}}(ELEMENT) \subseteq \pi_{\{ID\_DICT\}}(DICT) \tag{2}$$

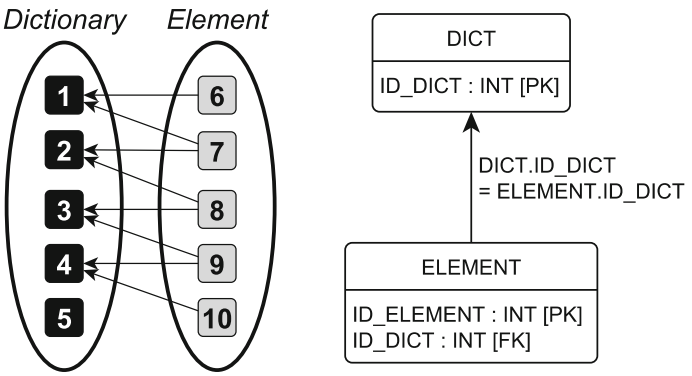


Fig. 2 Conceptual diagram representing *dictionary* design pattern (*left*) and its implementation (*right*) in relational metamodel (2)

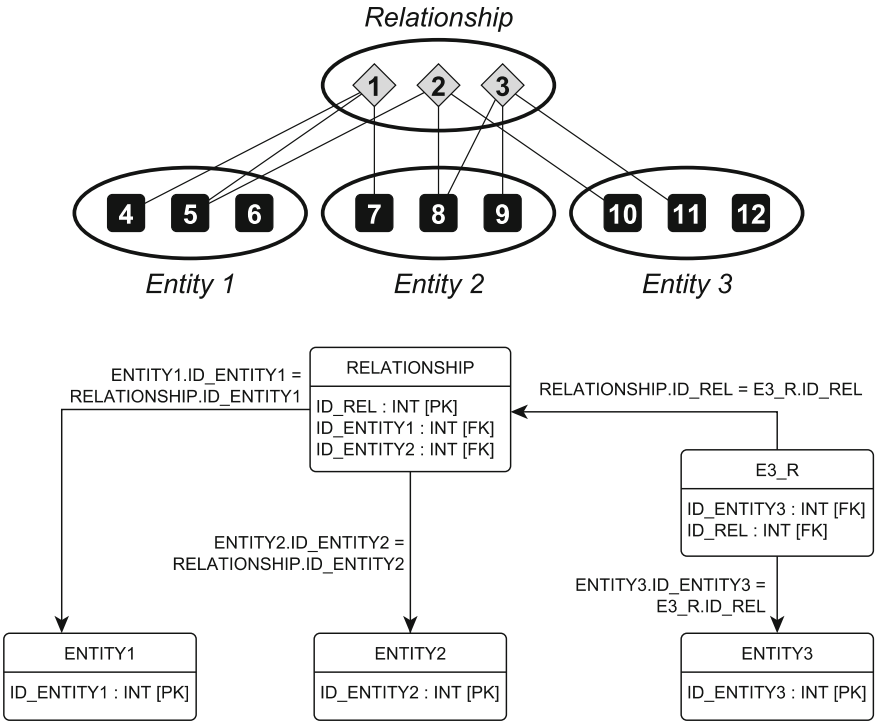


Fig. 3 Conceptual diagram representing *N-ary relationship* design pattern and its implementation in relational metamodel (3)

2.3 Design Pattern: N-ary Relationship

N-ary relationship is a specific design pattern. It represents the situation, where n entities take part in relationship. For example, medical appointment is a relationship between certain patient, doctor, it takes place in certain consulting room. For many designers this way of modeling relationships between entities is very natural, especially for those using modeling paradigms, where relationships are in the center of interest. While *list* in general models the *multiplicity* (e.g. *one-to-many*, *many-to-many*), the *n-ary pattern* models the *arity*. This pattern is generic, i.e. depending on value of n parameter, the participant cardinality in the pattern changes (Fig. 3).

$$\begin{gathered}
\textbf{ENTITY1}(\underline{ID_ENTITY1});\ \textbf{ENTITY2}(\underline{ID_ENTITY2}) \\
\textbf{ENTITY3}(\underline{ID_ENTITY3});\ \textbf{E3_R}(\underline{ID_ENTITY3}, \underline{ID_REL}) \\
\textbf{RELATIONSHIP}(\underline{ID_REL}, ID_ENTITY1, ID_ENTITY2) \\
\pi_{\{ID_ENTITY1\}}(RELATIONSHIP) \subseteq \pi_{\{ID_ENTITY1\}}(ENTITY1) \\
\pi_{\{ID_ENTITY2\}}(RELATIONSHIP) \subseteq \pi_{\{ID_ENTITY2\}}(ENTITY2) \\
\pi_{\{ID_ENTITY3\}}(E3_R) \subseteq \pi_{\{ID_ENTITY3\}}(ENTITY3) \\
\pi_{\{ID_REL\}}(E3_R) \subseteq \pi_{\{ID_REL\}}(RELATIONSHIP)
\end{gathered} \tag{3}$$

3 Association-Oriented Database Metamodel (AODB)

AODB is complete and consistent database metamodel. All of its components have been elaborated exclusively for this metamodel and are dedicated to it, i.e. it does not use any language, data storage model or other element of known database systems.

Association-oriented metamodel is based on the following primitives:

- intensional (structures): *database* (*Db*), *association* (*Assoc*), *role* (*Role*), *collection* (*Coll*), *attribute* (*Attr*),
- extensional (data): *association object* (*AssocObj*), *object* (*Obj*), *role object* (*RoleObj*).

In the extensional matter the most important categories are the following: *association* (Assoc) and *collection* (Coll). *Association* ($\Diamond$) is primitive realizing conception of relationships between data and the function of *collection* ($\Box$) is to store data.

Association is the owner of *roles*. *Roles* are lists of references to linked elements, that can be either *objects* (instances of *collection*) or *association objects*(instances of *association*). *Roles* in given *association* can have any cardinality, which means unrestricted arity of relationships. Each role is defined with number of properties, such as: identifier (name) of role unique within association, multiplicity on the side of *association*, multiplicity on the side of linked element, lifetime dependency between linking and linked elements (both directions), furthermore navigability and restriction on number of reduplication of bound elements.

Apart from standard roles, in AODB one can specify *description* role. In certain sense, one can approximately treat them like specific and redefined kind of role, which features with having no identifier, can bind only association and collection, is unidirectional, i.e. objects bound with description do not store information about it, multiplicity constraint on the side of collection is $0 \ldots 1$ and there is no mutual constraint of lifetime of bound elements.

Collection corresponds to the concept of data storage. In the intensional sense, *collection* is well defined by set of attributes, which define types of values stored in objects.

Association does not have ability to store data, and *collection* does not have possibility to create relationships, because internal structure of those primitives of association-oriented metamodel forces completely distinct way of their usage. Both categories independently are subject to the mechanisms such as inheritance. *Associations* can inherit from other *associations* and similarly trees of generalizations can be created for *collections*. Separation of data and relationships is complete, since in AODB each primitive performs only one, separate function. This means that if there are *collections* on database schema, one can conclude from the definition of association metamodel that they perform only one function strictly defined in terms of grammar. The situation known from relational or object metamodel does not occur, where *relation* or respectively *class* can perform function involving data storage (tuples of objects) and at the same time build *n*-ary relationship. AODB is completely unambiguous in terms of semantics of particular grammatical elements of metamodel. It is very significant not just whilst modeling, but also while attempting to analyze existing model or altering complex database schema.

Apart from definition of *association metamodel* ($\mathfrak{M}$), the descriptive part ($\mathfrak{D}$) and behavioral part ($\mathfrak{B}$) of AODB have been developed. The descriptive part comprise *Formal Notation* (AFN), which is strict, concise, formal and symbolical language of description of intensional and extensional part of metamodel. *Modeling Language* (AML) is graphical language of structures and data in AODB. It is fully consistent with AFN and metamodel definition. Both of the description languages namely (AFN and AML) are designed only for AODB and they are not any modification or subset of existing languages. The behavioral part of AODB contains *Query Language* (AQL) and *Data Language* (ADL). Both the languages fully correspond to metamodel and are completely original solutions for data selection or alteration problem, because they work directly on hypergraph structures, which represent the basis of AODB data model.

3.1 Modeling Language—AML

This section addresses the most important aspects of grammar and semantics of AML—Association-Oriented Modeling Language. It is a graphical language used to design database schemata in the *Association-Oriented Database Metamodel*. The graphical representation has been provided in Fig. 4. In AML, colors do not matter in syntactic terms, although using them may increase the diagram transparency.

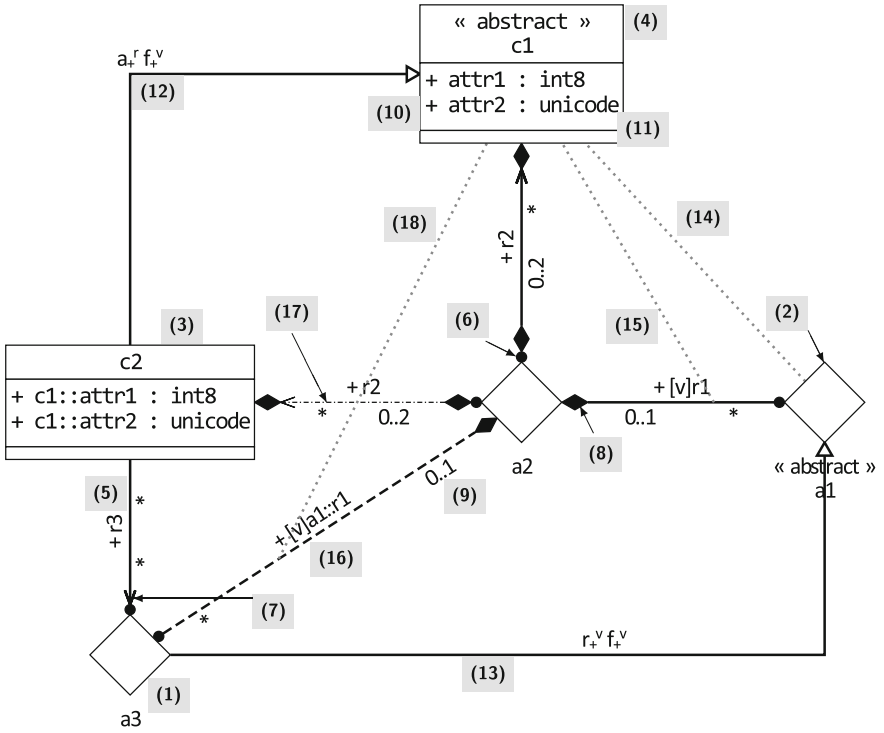


Fig. 4 Sample database schema diagram in AODB

(1) *Association* corresponds to a semantic category (*Assoc*).
(2) *Abstract Association* means that association cannot create its instances.
(3) *Collection* corresponds to a semantic category (*Coll*).
(4) *Abstract Collection* means that one collection cannot create its instances.
(5) *Role* corresponds to a semantic category (*Coll*). The graphical form in AML depends on navigability, directionality and composability.
(6) *Role Ownership*, (7) *Navigability*, (8) *Composition*, (9) *Multiplicity* is a part of the graphical form of a role (*Role*).
(10) *Attribute* is a part of the graphical form o a collection (*Coll*). Attribute (*Attr*) has a name, scope of visibility, quantity, type and default value.
(11) *Attribute Type* is a part of the graphical form of an attribute (*Attr*).
(12) *Collection Generalization* is a relationship which may link two collections.
(13) *Association Generalization* is a relationship which may link two associations. This relationship is described by an inheritance mode for roles.
(14) *Association Description* is a relationship which may link an association (*Assoc*) and a collection (*Coll*).
(15) *Role Description* is a relationship which may link a role and a collection.

(16) *Derived Role* is a relationship which may be represented in the diagram in a form similar to a role (*Role*).
(17) *Derived right to fulfill the Role* is a relationship which may be represented in the diagram analogically to the *Derived Role* case.
(18) *Derived Role Description* is a relationship which may be represented in the diagram, having an identical form as in the *Role Description*.

4 Mapping Database Design Patterns to Association-Oriented Metamodel

This section describes procedures of mapping design patterns, which have been identified in relational metamodel [7] into association-oriented metamodel. For the identified patterns authors have proposed steps, which the database designer needs to accomplish in order to create database using original structure along with its semantics. Subsequently, authors present AML diagrams of design patterns.

4.1 Mapping Procedure for List

(1) In association-oriented metamodel the linking element is *association*, therefore table LIST will be mapped to ◊LIST *association*.
(2) ELEMENT entity represents □ELEMENT *collection*, as long it has any attributes. If ELEMENT does not contain data, it can be replaced with ◊ELEMENT.
(3) If the element representing list has any additional information, the describing collection should be created: □LIST $\cdots$ ◊LIST, which will aggregate attributes representing additional information about list.
(4) Lifetime dependency on the side of linking element expressed by referential actions is mapped to *composition*, according to dependency direction.
(5) Multiplicity of list has been specified while identification as *one-to-many* or *many-to-many*. The identification defines directly the multiplicity of role implementing relationship between LIST and ELEMENT. If any other additional information in linked table have been discovered, one needs to create corresponding collection □INFO and describe role *Info* with it.
(6) If the attribute being foreign key has been specified as `NOT NULL`, the multiplicities of role are 1 instead 0..1 and 1..$*$ instead $*$ (Fig. 5).

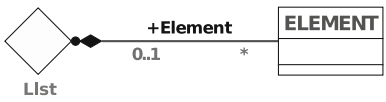


Fig. 5 Structure of *list* design pattern in association metamodel

4.2 Mapping Procedure for Dictionary

1. The DICTIONARY element always has structure of attribute set aggregate, which instance is common for many elements. Thus, the table DICTIONARY implementing the concept of dictionary is mapped to *collection* □DICTIONARY.
2. The described ELEMENT has always undefined character regarding its function. When it performs function of relationship, it is mapped to *association* ◊ELEMENT. This cannot be decided automatically and belongs to translation designer, who knows the semantics of relational model.
3. Relationships within association metamodel are implemented with the use of associations and roles. Typical mapping of relationship in *dictionary* pattern will require the following elements: *association* representing relationship between basenode ELEMENT and □DICTIONARY, which will be called ◊DICTIONARY ELEMENT. This association aggregates two roles, which targets are in both bound nodes respectively. The classical *dictionary* has multiplicity *one-to-many*, where *many* is on the □DICTIONARY side. If the foreign key attribute in relational implementation was nullable, the dictionary relationship is optional. While mapping it to AODB, multiplicity of role targeting in basenode ELEMENT will be set to 0..1 instead of 1 on the relationship side.
4. Additionally, one can determine navigability of *role* targeting dictionary element as uninavigable, in order to disable storage of information about relationships, in which dictionary element takes part. Assuming, that in practice there will not be frequent necessity to query database regarding all elements described by the same dictionary element. This solution is optimal and beneficial from the volatile or nonvolatile memory usage point of view (Fig. 6).

4.3 Mapping Procedure for N-ary Relationship

1. The elementary step is to create *association* ◊RELATIONSHIP, which implements *n*-ary relationship. This association corresponds to identified table having function of relationship. If table contains additional attributes, they should be aggregated in the describing *collection* □RELATIONSHIPDESCRIPTION $\cdots$ ◊RELATIONSHIP.
2. The mapping of tables representing the entities participating in relationship is performed analogously to the patterns described prior. For each table ENTITY_k, the table is mapped into *collection* $\square\text{ENTITY}_k$. However, it is important that association-oriented metamodel allows to link associations with other associations. If the translation engineer decides that any participant of relationship ◊RELATIONSHIP is the relationship itself, then they can perform mapping to association $\lozenge\text{RELATIONSHIPPARTICIPANT}_k$.

Fig. 6 Structure of *dictionary* design pattern in association-oriented metamodel

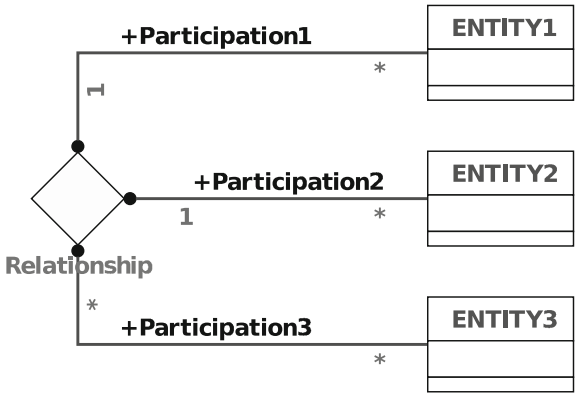


Fig. 7 Structure of *n-ary relationship* design pattern in association-oriented metamodel

3. The participation mapping of each of the elements within *n*-ary relationship is analogous to mapping of *list* pattern. The specification of multiplicities of participation in *n*-ary is based on foreign key attribute or additional table in case of *many-to-many* relationship. The multiplicities and constraints are mapped to multiplicities of *role* in implementation of this pattern in AODB (Fig. 7).

5 Case Study on Biomedical Database Sample

Following is an example of proposed methodology of mapping usage. It refers directly to database structure presented in *Bioinformatics in Tropical Disease Research. A Practical and Case-Study Approach* (2007) in chapter A02: *Understanding Database Design* [5]. There has been presented database model in Entity-Relationship metamodel [4], which involves data structures used within DNA analysis (Fig. 8). The database schema was supplemented with additional participant PROJECT of relationship HAS. The authors of the original model have shown example of relationships between entities, that clearly show that participant PROJECT exists in relationship HAS, but on the diagram describing the database schema this element has been omitted. The authors of quoted publication mentioned, that: *Be careful! A ternary relationship type cannot be substituted by three binary relationship types*. It is true, but it mostly involves E-R $\mapsto$ RDB mapping.

In this study the authors followed the thesis, that loss of original semantics in the E-R $\mapsto$ RDB transition is unavoidable not only in case of *n*-ary relationships, but also

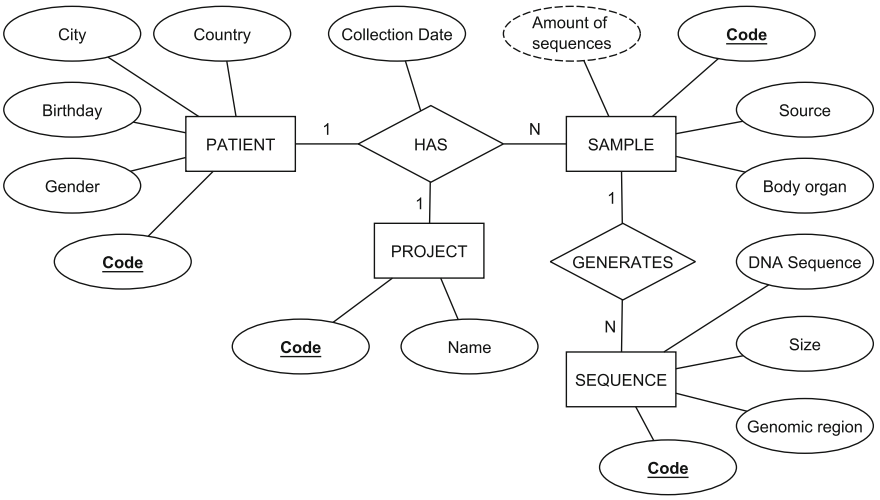


Fig. 8 The structure of database model in Entity-Relationship metamodel

in many other cases. Thus, one should not simplify conceptual models, but complement physical models with additional semantic information in form of description in natural language. Moreover, the target model proposed in this paper was presented in AODB metamodel, which introduces no issues in **physical representation of *n*-ary relationships**.

The E-R diagram has been mapped to relational diagram (Fig. 9). Attributes performing dictionary function have been transformed to dictionary tables. It is particularly important in the process of instances identification in design patterns used in the model.

In the model, the following instances of design patterns described in Sect. 2, have been discovered:

- List of sequences within the sample, realized by `SAMPLE` and `SEQUENCE` tables and relationship between them realized with key attributes mechanism,
- Dictionary of body organs used within sample realized by table `BODY_ORGAN` and connected with table `SAMPLE`,
- Dictionary of genomic regions used within DNA sequence and realized by table `GENOMIC_REGION` and connected with table `SEQUENCE`,
- Dictionary of countries used within patient and sample and realized by table `COUNTRY` and connected with tables `PATIENT` and `SAMPLE`,
- *N*-ary association *has* connecting certain sample with patient in a specific project. It is realized by table `HAS` connected with table `PROJECT`, `SAMPLE`, `PATIENT`.

The mapping allowed the creation of design patterns instances implemented in association-oriented metamodel (Fig. 10). Connecting them into one single structure resulted in design of complete and functional database model. Relations, which

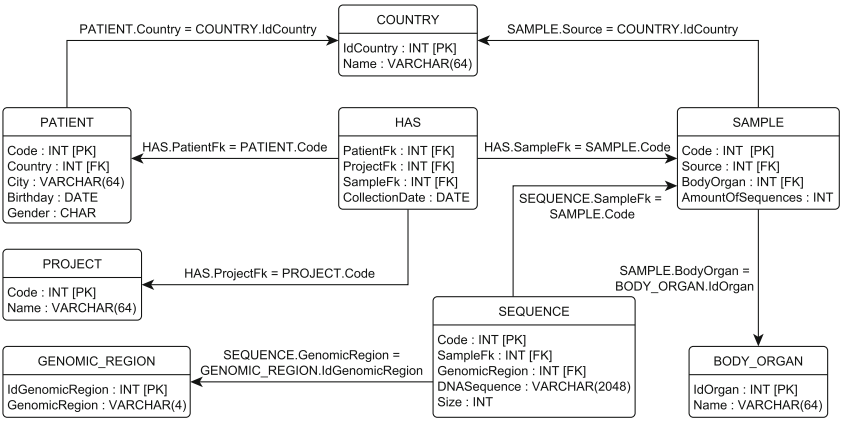


Fig. 9 Database structure in relational metamodel

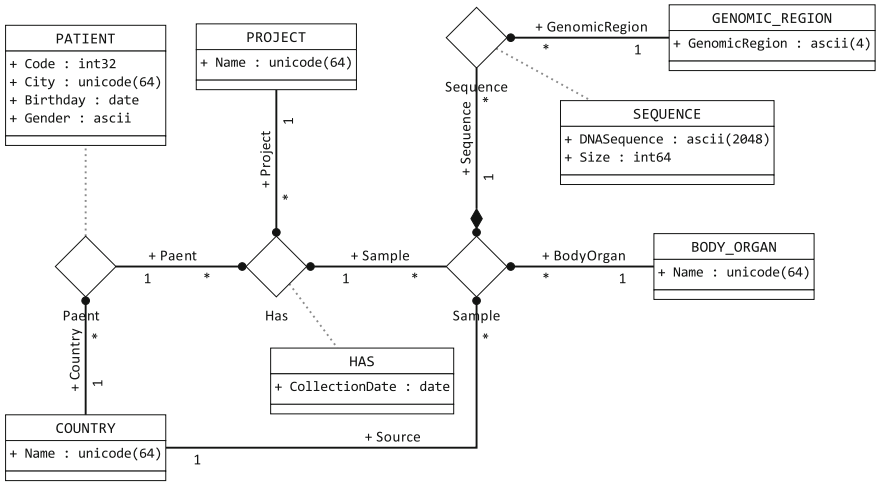


Fig. 10 The database schema in association-oriented metamodel

performed function of relationships (e.g. `Has`) have been mapped to associations. Relations performing functions of data containers have been mapped to collections. Relations performing function of relationships as well as data container have been mapped to the tandem of association and collection linked to it by description role. Moreover, during the process of schema translation, all primary keys which did not represent anything but identifier have been disregarded. The association-oriented metamodel has internally defined mechanism of identifiers and such attributes are superfluous. The derived attribute (`AmountOfSequences`) in table `Sample` also has been omitted, since AODB semantic provides this information indirectly.

6 Summary

In this article authors presented methodology for RDB $\mapsto$ AODB mapping. It comprise consistent and complete set of rules allowing unambiguously translation of relational databases into association-oriented databases. The association-oriented metamodel functions on the conceptual level as well as on the physical level without necessity of any mapping. However, due to the fact of popularity of relational metamodel in database design and large number of functioning relational databases, the problem of RDB $\mapsto$ AODB translation of database schemata seems to be very important. The case study shown covers two issues: E-R $\mapsto$ RDB mapping and RDB $\mapsto$ AODB mapping. As a conclusion, it should be noted that optimal solution is E-R $\mapsto$ AODB translation. This is due to the fact, that association-oriented metamodel realizes all the needs of modeling on the conceptual level as well as on the physical level. What is the most important is the fact that it does not require any simplifications, compromises, etc. Association-oriented metamodel is very efficient tool for database modeling and is perfectly suited to create complex database schema [6], in which semantics of model is complicated and should not be subject to any simplifications.

References

1. Casanova, M.A., Breitman, K.K., Brauner, D.F., Marins, A.L.A.: Database conceptual schema matching. Computer **40**(10), 102–104 (2007)
2. Connoly, T., Begg, C.: Database Systems. A Practical Approach to Design, Implementation, and Management, 4th edn. Pearson Education Unlimited, Essex (2005)
3. Darwen H.: The Relational Model: Beginning of an Era (2012)
4. Earp, R., Bagui, S.: Database Design Using Entity-Relationship Diagrams (2003)
5. Ferreira, J.E., Takai, O.K.: Bioinformatics in Tropical Disease Research: A Practical and Case-Study Approach. Chapter A02: Understanding Database Design (2007)
6. Krótkiewicz, M., Wojtkiewicz, K., Jodłowiec, M., Szwedziak, K.: Unified process management for service and manufacture system—Material resources. In: Proceedings—The 9th International Conference on Computer Recognition Systems. Springer Series Advances in Intelligent Systems and Computing (2016)
7. Sequeda, J., Arenas, M., Miranker, D.P.: On directly mapping relational databases to RDF and OWL. In: WWW, pp. 649–658 (2012)
8. Tan, H.B.K., Ling, T.W.: Integrated design patterns for database applications. J. Syst. Softw. **47**(2), 159–172 (1999)

Similarity Search for the Content of Medical Records

Dominik Spinczyk and Mariusz Dzieciątko

Abstract The paper presents the possibility of direct comparison of medical text content by using unstructured representation of document information in frequency matrix of terms. Dimensionality reduction is performed using Latent Semantic Indexing method. Two common metrics are used: Cosine distance and Jaccard metric. Cosine measure shows a lower sensitivity for finding similar documents. The analysis was performed using SAS Text Analytics elements on set of 400 cases of description of abdominal radiological diagnostic images.

Keywords Medical records · Text mining · Similarity search · Cosine distance · Jaccard metric

1 Introduction

Continuous development of technologies related to medical images raises challenges not only in large volume data storage but also in the field of medical diagnostics for efficient interpretation and for appropriate classification of the data. Following range of development tools is presented:

- creation of metadata,
- data indexing,
- data search,
- data compression.

D. Spinczyk (✉)
Faculty of Biomedical Engineering, Silesian University of Technology,
Roosevelta 40, Zabrze, Poland
e-mail: dspinczyk@polsl.pl

M. Dzieciątko
SAS Institute Sp z o.o., Gdańska 27/31, 01-633 Warszawa, Poland

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_40

One of the challenges is the indexing of medical images. Each medical image may be characterized by their properties, which are the values of some of its attributes:

$$feature_1 = attribute_1(object), ..., feature_k = attribute_k(object)$$

While choosing description we look for representative features characterized by reduction in the amount of memory consumed by the index, with no decrease in selectivity below a particular threshold. In the case of indexing images in general, there is a concept of semantic holes (called the semantic gap), similar images in terms of one specific objective criterion used by the indexing system are interpreted differently by human. Additionally, in medical diagnosis there exists important concept of sensitivity hole (sensory gap), resulting from the imperfections of image acquisition devices.

The MPEG multimedia databases (called Multimedia Content Description Interface) define a set of visual descriptors: color, texture and shape [7] but they are not commonly used in daily radiology practice to create indexes in existing repositories of medical images [2]. Therefore text information can be used as useful descriptive indicator. The methods, which are used for text image analysis can be divided into four groups [3, 9]:

- dictionary-based approaches,
- rule-based approaches,
- machine learning approaches,
- manual annotation.

All of the mentioned approaches have some disadvantages. Dictionary-based learning approaches tend to miss undefined terms that are not mentioned in the dictionary [6], rule-based approaches require rules that identify terms from text, and the resulting rules are often not effective in all cases [6]. Machine learning approaches generally require standard annotated training data sets which usually require big human effort to build [4].

The main motivation of this work is to propose the method which is not perfect in terms of power of discrimination but is completely automatic (what is the main advantage comparing to methods mentioned above) and can be used for big text corpuses to discover the potential group of topics. This method main idea is to directly compare the contents of text documents.

2 Materials and Methods

2.1 Representation of Information in Text Document Corpus

During this stage we can distinguish two subtasks: pre-treatment of an input corpus and determination of representation manner. The essence of the firstsubtask is to form a coherent collection of data from input corpus (removal of information formatting,

encoding and converting the standarization of the collection to a recognizable format). In terms of text structure documents can be divided into:

- Structured—atomic entries in a relational database, spreadsheet or flat file of the proper format of records,
- Semi-structured—XML (Extensible Markup Language) documents,
- Unstructured—entries of online forums, messaging, electronic mail, multimedia messaging.

Selecting the way of formal representation information from the document is the essence of the second subtask. There are two main types of representation of information from the document:

- Bag-of-words—storing information on the frequency of particular words in the processed documents,
- Structured representation—storing not only the frequency of occurrences of particular words but also the relationships between them.

In this paper we apply only bag-of-words approach because it is easier for implementation. When applied to big text corpuses it consumes less memory and is useful for the dimensionality reduction.

2.1.1 Information About the Frequency of Terms

The document is represented as a vector, where its various elements indicate the number of individual words. The transformation of the body of the input documents to this form requires the following:

- at the beginning of the process to remove punctuation and create a list of independent occurrences of words for each document,
- remove words irrelevant from the point of view of further analysis (the words they form the so-called stop-list),
- transformation of words to their basic form (called stemming). Depending on the specific language transforming words into their basic forms can be implemented using rules or dictionaries.

Words that are left after initial processing of documents are called terms. A collection of terms with all of the documents is called a dictionary. Frequency list of occurrences are then combined into a single list forming a frequency matrix:

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \dots & a_{ij} & \dots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

where:

- a_{ij}—element of the frequency matrix,
- m—the number of rows of the frequency matrix corresponds to the number of terms included in the matrix frequency,
- n—the number of columns in the array of frequency, corresponds to the number of documents of corpus.

The column number represents the index of the input document and row number represents term's index. Value of each matrix element corresponds to the number of occurences of specific term in the document. The advantage of this approach is the simplicity of implementation of the calculation and a large range of processing methods. A significant drawback of frequency matrix is limitation of information about each term to the number of its occurences. Teams matrix frequency associated with the model space vector (called vector space model) in which the document is represented. Generally, this is a multi-dimensional Euclidean space, where each axis of the matrix corresponds to specific term from frequency matrix.

It is possible to distinguish diffrent representation of frequency matrix:

- full representation—takes into account boths the location of the terms and the number of occurrences,
- boolean representation—stores only information about the occurrence of a given term in a given document ignoring the information on its frequency,
- term frequency representation—the frequency of given term are stored,
- logarithmic representation—holds a logarithmic value of frequency, zero matrix elements remain unchanged,
- weighted logarithmic representation—holds a logarithmic value of frequency, zero matrix elements remain unchanged, and the weight is in most cases the ratio of the total number of documents to documents containing that term.

Generally, the frequency term matrix can be modified by weighting function, which is defined as: $\hat{a}_{ij} = W_i L_{ij}$ where:

- W_i—term weight,
- L_{ij}—frequency weight.

2.2 *Dimensionality Reduction of the Frequency Matrix of Terms*

Before beginning the proper analysis of text documents size reduction is made for usually very large and sparse frequency matrix of terms. The statistics knowns a number of solutions to reduce the dimensionality of data:

- Principal Component Analysis,
- Factor Analysis,

- Partial Least Squares,
- Latent Semantic Indexing.

In the field of text data mining method the Latent Semantic Indexing method is used. The reduction of the dimensionality is based on a Singular Value Decomposition method of frequency matrix of terms, which decomposes the input matrix into independent linear components. The method for calculating singular value can be found in the references on algebra [1, 5, 8].

2.3 Similarity Measures for Text Documents

As mentioned earlier the transformation of the information contained in a document to a frequency matrix of terms can be treated as a vector in a vector space features of the document. Documents transformed into this representation can be compared with different metrics. While searching documents' text there exists a problem of correct query creation related to high dimensionality of features space. In our approach there is no need to create the query. In this paper two similarity measures are used: Cosine distance, Jaccard metric which were selected from the set of exisitng multiple similarity measures. Both measures preserve dimensionality but Cosine distance measure is different allows to preserve full dimensionality of features' space when the representation bag-of-words is selected. Jaccard metric is different than Cosine distance because it simplifies bag-of-words to binary representation and works on non-zero only coordinates of binary vectors.

2.3.1 Cosine Distance

One popular measure, which is used, is cosine distance:

$$dys_{cos}(\boldsymbol{doc_1}, \boldsymbol{doc_2}) = \frac{\boldsymbol{doc_1} \cdot \boldsymbol{doc_2}}{|\boldsymbol{doc_1}| \cdot |\boldsymbol{doc_2}|}.$$

When the vectors are normalized, cosine distance is defined as the scalar product of documents vectors: $dys_{cos}(\boldsymbol{doc_1}, \boldsymbol{doc_2}) = \sum_{i=1}^{m} \boldsymbol{doc_1^i}\, \boldsymbol{doc_2^i}$.

Additionally, the reader can consider whether to use all terms all terms. If this is the case, the selection criterion is the threshold of weighting function and several cases can be distinguished:

- all terms from the document,
- terms of high value empty weight of frequency,
- terms of high weight value inverse frequency term in the document,
- terms with value of weighting function under treshold.

2.3.2 Jaccard Metric

Jaccard measure is a measure of similarity, which uses the binary representation of documents vectors from the frequency matrix and uses only the non-zero coefficients. Jaccard coefficient is calculated, which is defined as the percentage of non-zero coordinates, different for the two vectors:

$$dys_{Jacc}(\boldsymbol{doc_1}, \boldsymbol{doc_2}) = \frac{|term(\boldsymbol{doc_1}) \cap term(\boldsymbol{doc_2})|}{|term(\boldsymbol{doc_1}) \cup term(\boldsymbol{doc_2})|}.$$

where:

- $term(\boldsymbol{doc_1})$—vector representing a set of terms used in the i-th document.

Then, Jaccard metric is defined as:

$$met_{Jacc}(\boldsymbol{doc_1}, \boldsymbol{doc_2}) = 1 - dys_{Jacc}(\boldsymbol{doc_1}, \boldsymbol{doc_2}).$$

2.4 Experiments

Data set consists of 400 radiological reports of abdomen studies. The average length is of about 200 words. The style of the language includes a conventional radiological terminology.

Experiment consists of the following steps:

- Text import—merging all documents into one data set
- Text parsing—finding terms in the document content and transformation into frequency matrix
- Text filtering—frequency matrix modification by waiting function, its dimensionality reduction using principal component analysis
- Analysis—Cosine and Jaccard measures are used. They are calculated on the vectors (rows) of reduced frequency matrix.
- Visualization and interpretation—for each vectors pair normalized value ([0, 1]) of each measure is calculated. Connections graph presents the pairs which are above selected threshold which is 0.95.

The calculations were made using SAS Text Analytics in Enterprise Text Miner Application.

Above analysis were performed for every speech part and for nouns only. For nouns only analysis the motivation was the assumption that nouns have more discrimination power in compare with other speech parts. When taking the nouns only we reduce dimension of features space.

3 Results and Conclusions

For the above-defined similarity measures documents of abdominal diagnostic images description were compared. Following results were presented:

- similarity intensity—metric value by definition, greater than 95 percentile of the body between the documents normalized in the range of <0; 1> (Figs. 1 and 2),
- the number of similar documents—histograms number of groups similarities documents (Fig. 3),
- connections between the documents—the ratio of similarity expressed neighbourhood identifiers of documents (Figs. 4 and 5).

The calculations were made on the reduced frequency matrix of terms according to the method of semantic analysis of hidden variables.

Figures 1 and 2 show that most of the documents have a single the most similar document what is confirmed by the histograms of groups (Fig. 3). Charts connections (Figs. 4 and 5) indicate the dominance of two-element groups similarities documents. Cosine measure shows a lower sensitivity for finding similar documents.

In the next section the similar results were presented for nouns analysis only (Figs. 6, 7, 8, 9, and 10).

Presented automatic method is useful as an introductory analysis of big text corpuses to detect existing topics. For further deeper analysis more sophisticated clusterisation and classification methods can be used, which require more customer effort.

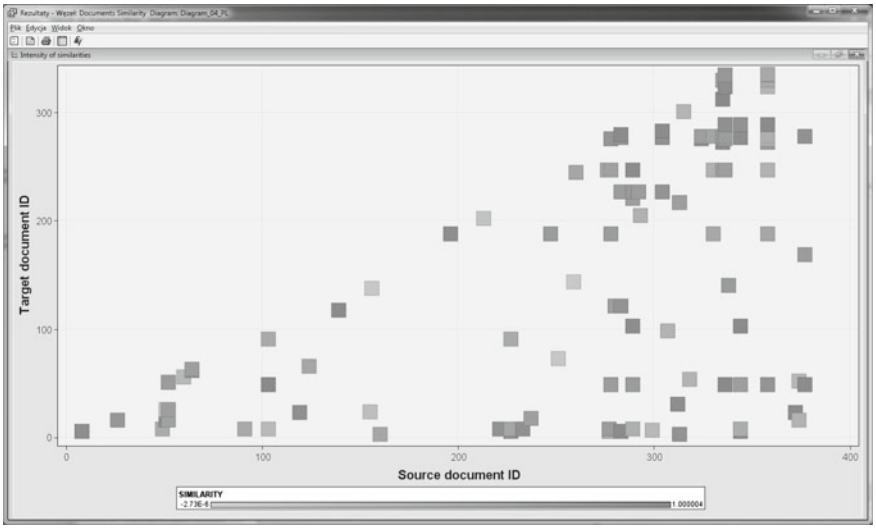


Fig. 1 Scatter plot documents' similarity groups according to the cosine distance

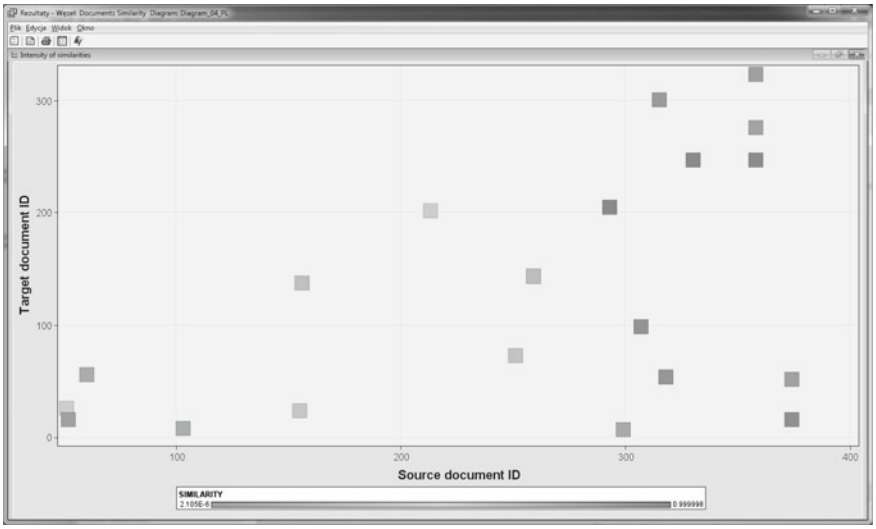


Fig. 2 Scatter plot documents' similarity groups according to the Jaccard metric

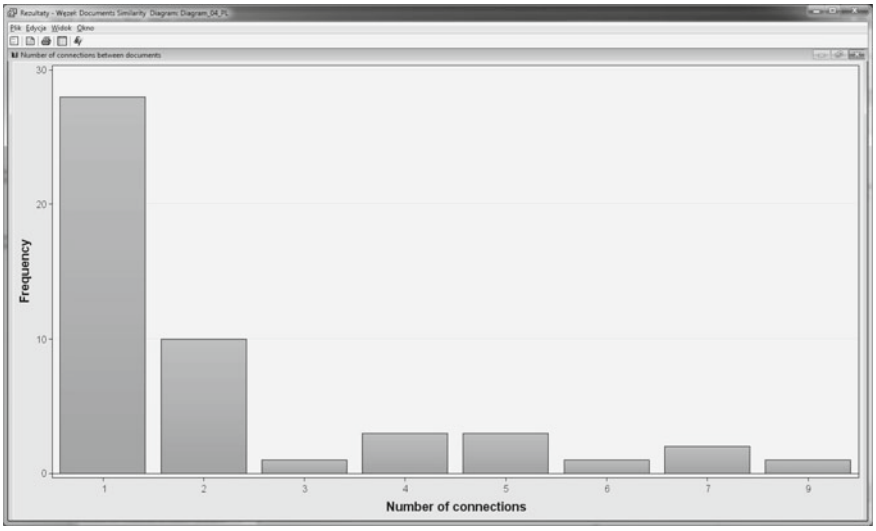


Fig. 3 Histogram of groups which include different number of connected documents according to the cosine distance

Therefore there is no reason for quantitative comparison between our and more complex methods because our method is preliminary step for the other. The proposed method which is not perfect in terms of power of discrimination but is completely automatic and can be used for big text corpuses to discover the potential group of topics. For every discovered groups the key terms are known and can help user to

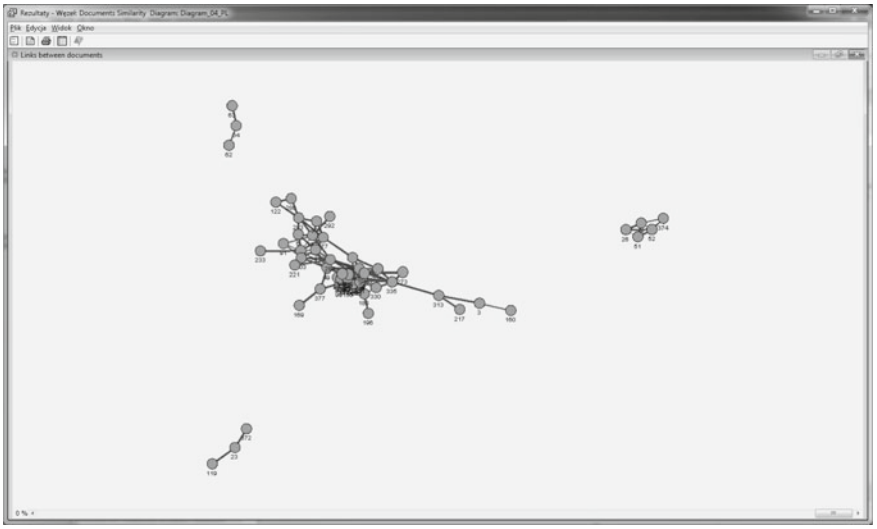


Fig. 4 Connections between radiological descriptions according to the cosine distance

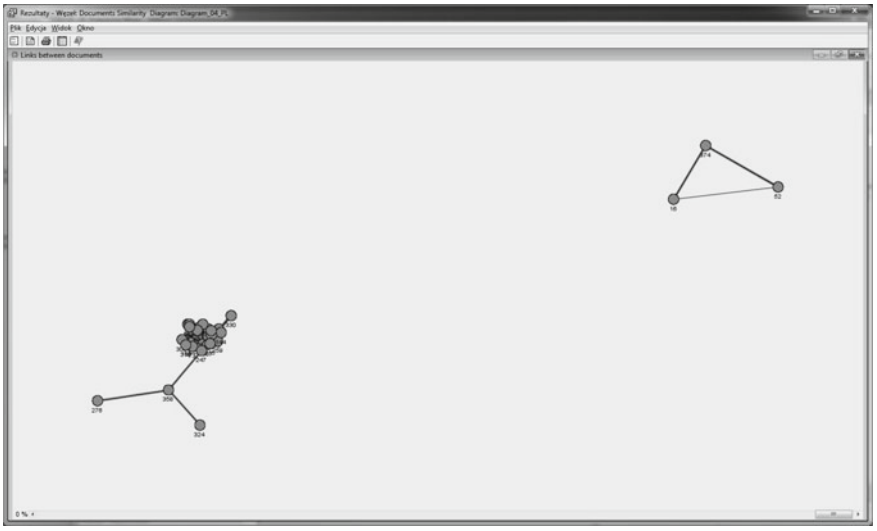


Fig. 5 Connections between radiological descriptions according to the Jaccard metric

discover the main topics in the text corpus. The number of selected groups depends on the threshold value of selected similarity measure. The optimization of selecting threshold in reference to number of groups and thoroughness of grouping will be the topic of the continuous research.

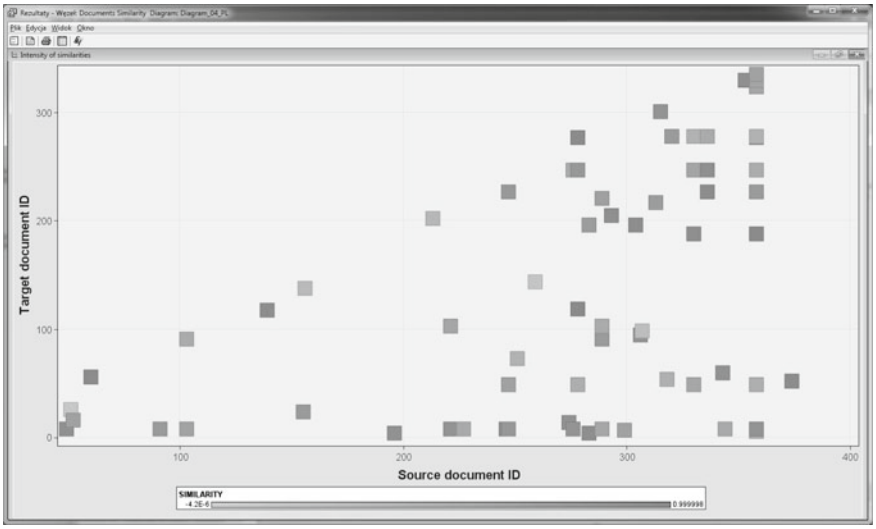


Fig. 6 Scatter plot of documents' similarity groups according to the cosine distance for nouns only

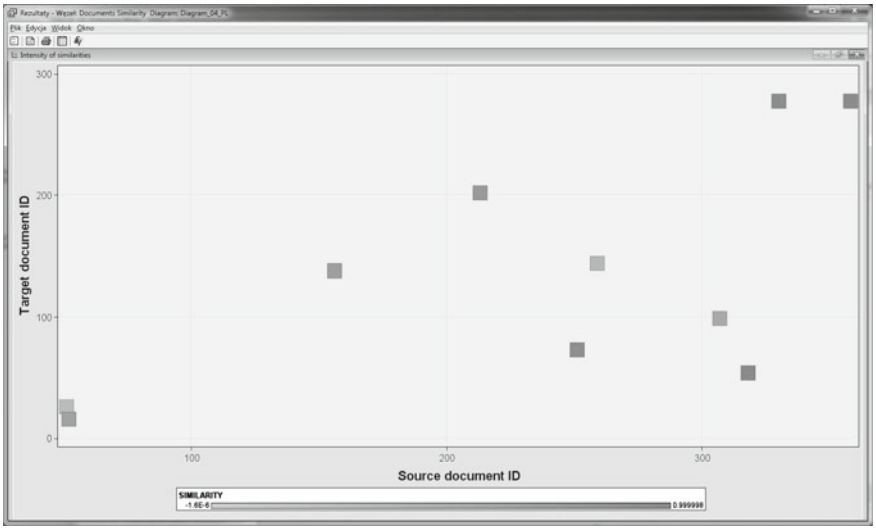


Fig. 7 Scatter plot of documents' similarity groups according to the Jaccard metric for nouns only

3.1 Conclusions

In the paper elements of SAS Text Analytics are presented to find similarity between documents of abdominal diagnostic images description.

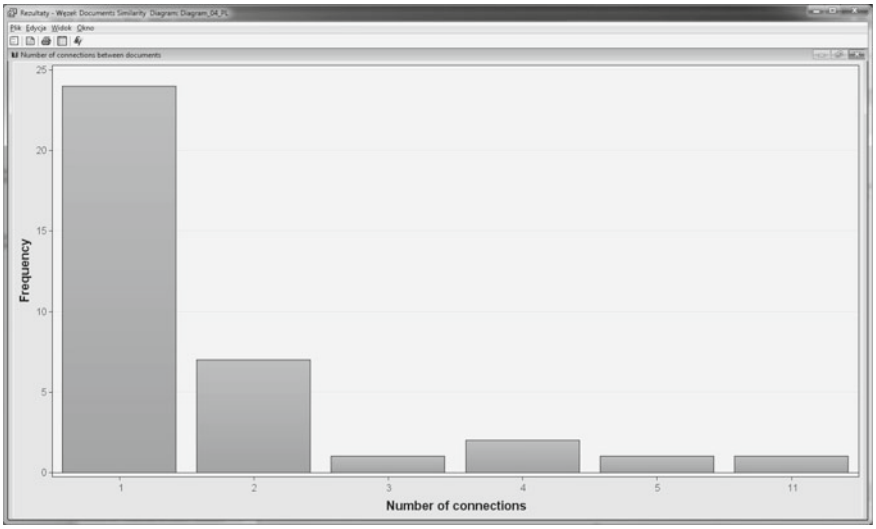


Fig. 8 Histograms groups which include different number of connected documents according to the cosine distance for nouns only

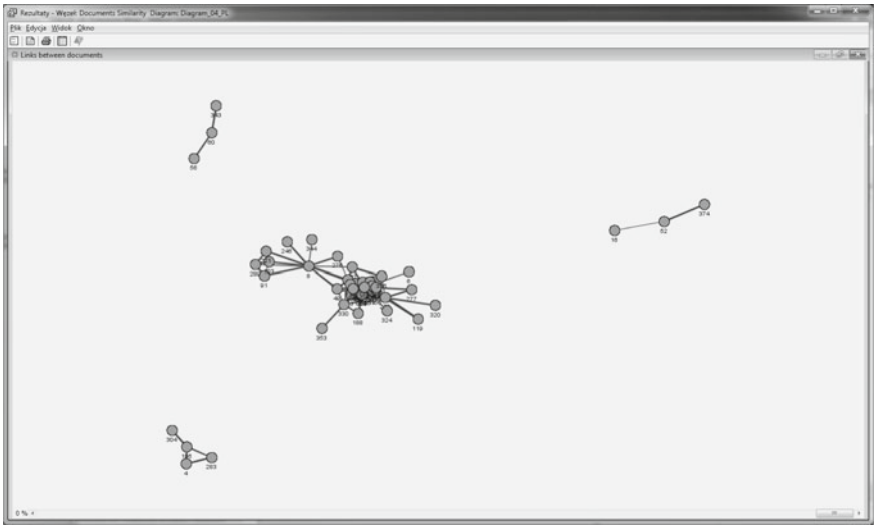


Fig. 9 Connections between radiological descriptions according to the cosine distance for nouns only

Analysis of plots connections can be used to identify similar documents, which can be useful for groups of documents with similar topic and plagiarism detection also.

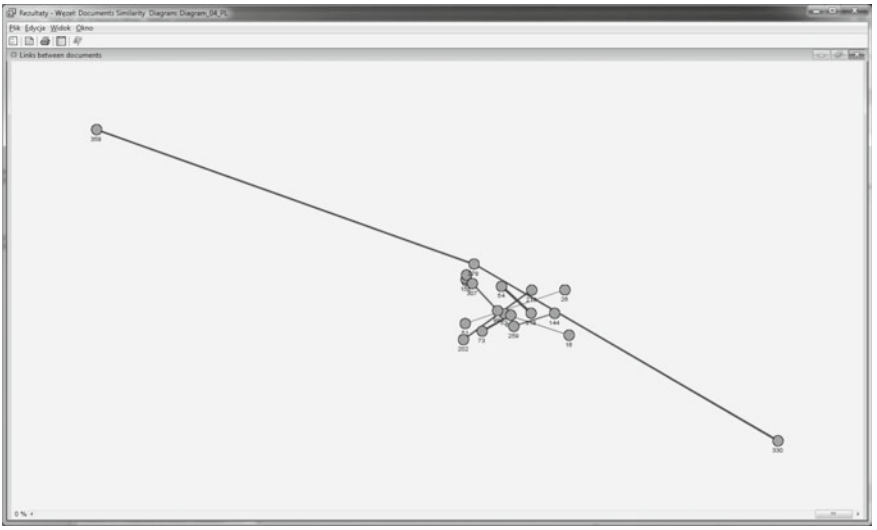


Fig. 10 Connections between radiological descriptions according to the Jaccard metric for nouns only

The user interested in using other measures of similarity can find their definitions in the description of the SAS "rocedure distance", such as "Gower's similarity", "Covariance", "Correlation", "Squared correlation", "dot (inner) product coefficient", "Overlap similarity" "Simple matching coefficient", "Hamann coefficient", the "Roger and Tanimoto," and others.

Acknowledgments The study was supported by National Science Center, Poland, Grant No UMO-2012/05/B/ST7/02136.

References

1. Albright, R.: Taming Text with the SVD. SAS Institute White Paper (2004)
2. Boniński P.: Metody indeksowania obrazów medycznych na potrzeby radiologii cyfrowej. PhD thesis. Warsow University of Technology (2007)
3. Kawa, J., Juszczyk, J., Pyciński, B., Badura, P., Piętka, E.: Radiological atlas for patient specific model generation. Adv. Intell. Syst. Comput. **84**, 69–84 (2014)
4. Krallinger, M., Vasquez, M., Leitner, F., Salgado, D., Chatr-Aryamontri, A., Winter, A. et al.: The protein-protein interaction tasks of BioCreative III: classification ranking of articles and linking bio-ontology concepts to full text. BMC Bioinformatics **12** (Suppl. 8), S3 (2011)
5. Meyer, C.: Matrix Analysis and Applied Linear Algebra. SIAM, Philadelphia (2000)
6. Rebholz-Schuhmann, D., Jepes, A., Li, C., Kafkas, S., Lewin, I., et al.: Assesment of NER solutions against the first and second CALBC Silver Standard Corpus. J. Biomed. Semantics **2**(Suppl. 5), S11 (2011)
7. Skarbek W.: Indeksowanie multimediów. Lecture materials (2005)

8. Vandenberghe, L.: Applied Numerical Computing (lecture). http://www.seas.ucla.edu/~vandenbe/103/reader.pdf (2011)
9. Zhu, F., Patumcharoenpol, P., Zhang, Ch., Yang, Y., et al.: Biomedical text mining and its applications in cancer research. J. Biomed. Inf. **46**, 200–211 (2013)

Application of Preprocessing Methods to Imbalanced Clinical Data: An Experimental Study

Szymon Wilk, Jerzy Stefanowski, Szymon Wojciechowski, Ken J. Farion and Wojtek Michalowski

Abstract In this paper we describe an experimental study where we analyzed data difficulty factors encountered in imbalanced clinical data sets and examined how selected data preprocessing methods were able to address these factors. We considered five data sets describing various pediatric acute conditions. In all these data sets the minority class was sparse and overlapped with the majority classes, thus difficult to learn. We studied five different preprocessing methods: random under- and oversampling, SMOTE, neighborhood cleaning rule and SPIDER2 that were combined with the following classifiers: *k*-nearest neighbors, decision trees and rules, naive Bayes, neural networks and support vector machines. Application of preprocessing always improved classification performance, and the largest improvement was observed for random undersampling. Moreover, naive Bayes was the best performing classifier regardless of a used preprocessing method.

Keywords Clinical data · Class imbalance · Data difficulty factors · Preprocessing methods · Classification performance

1 Introduction

Clinical data pose a significant challenge for data mining due to such factors as missing or imprecise values, noisy or inconsistent observations and uneven distribution of patients from different decision classes [4, 12]. Usually, the number of patients from a critical class, who require special management, is much smaller than the number

S. Wilk (✉) · J. Stefanowski · S. Wojciechowski
Poznan University of Technology, Poznan, Poland
e-mail: szymon.wilk@cs.put.poznan.pl

K.J. Farion
Children's Hospital of Eastern Ontario, Ottawa, Canada

W. Michalowski
University of Ottawa, Ottawa, Canada

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2_41

of patients from remaining classes (later in the text we refer to this critical class as to *minority* class, and to the remaining classes as to *majority* classes)—such situation is known as *class imbalance*.

Class imbalance limits the ability of constructed classifiers to make accurate predictions regardless of the applied learning method [6], and many methods for dealing with this problem have been already proposed (see [8] for their review). They can be divided into two general categories: *data-level* and *algorithm-level*. The former methods preprocess data, e.g., by resampling, to change the distribution of classes, while the latter adjust the learning process, so they can be applied directly to imbalanced data. Although implementing modifications on the algorithmic level can potentially lead to more extensive improvement in classification performance, preprocessing methods are still a dominant approach.

However, class imbalance itself is not the only one or main problem. There are other factors related to data distribution that combined with class imbalance can seriously deteriorate classification accuracy, especially for the minority class [17]. These factors, often referred to as *data difficulty factors*, include rare sub-concepts [10], overlapping regions of the minority and majority classes [20], or multiple minority class examples located inside the majority classes [17].

Napierala and Stefanowski proposed in [16] to approximate the above data difficulty factors using the local characteristic of the minority class examples. They distinguish between *safe*, *borderline*, *rare* and *outlier* examples. Safe examples lie inside the minority class and are surrounded mostly by neighbors from this class; borderline examples are located close to decision boundaries and thus their neighborhood is a mixture of the majority and minority class examples; rare examples form small islands (2–3 examples) inside the majority classes; finally outliers are isolated examples *thrown* into the majority classes. Borderline, rare and outlier examples are considered as *unsafe*, as they are more difficult to learn. The paper [16] also describes an automatic approach to label the minority class examples with their types that relies on examining the distribution of classes in their local neighborhoods.

In this paper we describe an experimental study aimed at identifying data difficulty factors in imbalanced clinical data sets, and at evaluating the impact of preprocessing methods on the performance of classifiers learned from these data sets. Specifically, this study answers the following research questions:

- What are the data difficulty factors (in terms of the types of minority class examples) encountered in the analyzed clinical data sets?
- How do the preprocessing methods improve the performance of classifiers, especially with respect to the minority class?
- What are the best (in terms of the improved performance) combinations of preprocesssing methods and classifiers?

The data sets used in this study describe pediatric patients managed for acute conditions in the emergency department (ED) of the Children's Hospital of Eastern Ontario (CHEO). Some of these conditions are highly prevalent (e.g., asthma exacerbation) and require significant resources for proper management. Others are less

frequent (e.g., scrotal pain), however, their misdiagnosis may have serious health-related and legal consequences. For all conditions, quick and accurate recognition of patients from the critical class is necessary to provide appropriate and timely management. At the same time, these conditions are associated with serious class imbalance—addressing this problem together with associated data difficulty factors should help construct more accurate decision models and tools for regular use in the emergency setting.

2 Related Work

The most popular resampling methods are random *oversampling*, which replicates examples from the minority class, and random *undersampling* which randomly eliminates examples from the majority classes until a required degree of balance between classes has been reached. However, undersampling may potentially remove some important examples and oversampling may lead to overfitting [13]. Thus, recent research focuses on particular examples, taking into account information about their distribution in the attribute space [8].

Kubat and Matwin claim in [13] that characteristics of mutual positions of examples is a source of difficulty when learning from imbalanced data. They introduced a method called *one-sided sampling* (OSS) which filters the majority classes in a focused way [13]. It is based on distinguishing different types of learning examples: *safe*, *borderline* and *noisy*. They proposed to use Tomek links (two nearest examples having different class labels) to identify and delete the borderline and noisy examples from the majority classes. The critical analysis of OSS inspired the development of other informed preprocessing methods.

Neighborhood cleaning rule (NCR) represents another approach to the focused removal of examples from the majority classes [14]. It deals with the local data characteristics by applying the *edited nearest neighbor rule* (ENNR) to the majority classes [26]. ENNR first looks for a specific number of k-nearest neighbors ($k = 3$ is recommended in [14]) of a *seed* example, and uses their labels to predict the class label of this seed. In case of a wrong prediction, the neighbors from the majority classes are removed from the learning set.

The best known informed sampling method is *Synthetic Minority Over-sampling TEchnique* (SMOTE) [3]. It is also based on the k-nearest neighborhood, however, this neighborhood is exploited to selectively oversample the minority class by creating new *synthetic* examples [3]. SMOTE treats each minority class example as seed and finds its k-nearest neighbors from the minority class. Then, according to the user-defined oversampling ratio o_r, SMOTE randomly selects o_r of these k neighbours and introduces new examples along the lines connecting the seed example with the selected neighbors. Although SMOTE was successfully applied to many problems, including medical ones (see experiments in [2, 17]), some of its new generalizations are better suited to deal with the data difficulty factors being considered [19].

SPIDER2 is a hybrid method that selectively filters out harmful examples from the majority class and amplifies difficult minority class examples [22]. In the first stage it applies ENNR to distinguish between safe and unsafe examples (depending how k neighbors reclassify the given seed example from the minority class). Outliers or neighbors from the majority classes that misclassify the seed example are either removed or relabeled. The remaining unsafe minority class examples are additionally replicated depending on the number of neighbors from the majority classes.

In all of these methods the k-nearest neighbourhood is calculated with the *Heterogeneous Value Difference Metric* (HVDM) [25]. HVDM aggregates normalized distances for both numeric and nominal attributes, and for the latter it employs the *Value Difference Metric* (VDM) by Stanfill and Waltz [25].

Results of experiments presented in [15, 17] showed that when using decision trees and rule-based classifiers SPIDER2 and SMOTE were more accurate than random oversampling for data sets containing many unsafe types of the minority class examples. The most recent experimental study [16] demonstrated that undersampling like NCR was particularly useful when borderline examples dominated in the minority class.

3 Methods

3.1 Data Sets

The data sets considered in this study and their brief characteristics are given in Table 1. They describe pediatric patients with various acute conditions managed in the ED at CHEO. The AP and AE2 data sets were collected prospectively using either a mobile system or paper forms, and the other data sets were transcribed retrospective from charts. Moreover, the AE1 and AE2 data sets describe the same problem, however, they could not be combined together due to the differences in definitions of some of the attributes (e.g., their values were collected at different time points in the patients' care).

Table 1 Characteristics of the considered data sets

Data set	Clinical problem	# examples (minority)	Imbalance ratio	# attributes (numeric)
AP	Abdominal pain	457 (48)	0.11	13 (3)
HP	Hip pain	412 (46)	0.11	20 (4)
SP	Scrotal pain	409 (56)	0.14	14 (3)
AE1	Asthma exacerbations (2004)	362 (59)	0.16	32 (11)
AE2	Asthma exacerbations (2007)	240 (21)	0.09	42 (9)

All data set are imbalanced—the class imbalance ratio, defined as the ratio of examples from the minority class to all examples, ranges from 0.09 to 0.16. The minority class is also the critical class, indicating patients who required special attention (e.g., specialist consultation, advanced laboratory investigations or intense and prolonged treatment). Initially, the data sets included three classes, and for the sake of analysis, the two classes corresponding to less urgent and intense management were combined into a single majority class.

Some of the data sets were initially fairly incomplete—especially SP and AE1 (both collected retrospectively). Therefore, before the analysis we removed attributes with more than 50 % of missing values (Table 1 gives the numbers of attributes after the removal). There were 15 such attributes in SP and 10 in AE1. On the other hand the prospective data sets were complete—no attribute was removed from AE2 and a single attribute was removed from AP, confirming an advantage of prospective data collection.

In all of data sets, most attributes are nominal and correspond to symptoms and signs checked by physicians. Numeric attributes represent age and vital signs, such as temperature, heart rate or oxygen saturation.

3.2 Experimental Design

The experimental design covered two phases, corresponding to the main research questions formulated for our study. In the first phase we employed the labeling technique from [16] to establish the distribution of minority class example types in the considered data sets, and thus to identify the major data difficulty factors. In the second phase we evaluated the impact of selected preprocessing methods on the performance of selected classifiers, with special focus on the minority class.

We selected the following preprocessing methods for our experiment (in brackets we give their symbols used further in the text): no preprocessing (none) as the baseline, random undersampling (RU), random oversampling (RO), SMOTE (SM), NCR, and SPIDER2 (SP2). Following our experience and suggestions from [24], the RU, RO and SM methods were set to produce a balanced distribution of classes in resulting data sets. Moreover, SM and SP2 were used with $k = 5$ nearest neighbors (this value is suggested for SM by its authors; its suitability was also confirmed by our earlier experiments with SP2), and the latter was parametrized for extended amplification of the minority class examples and for relabeling of the majority class examples instead of removing them (these options worked best in our previous studies—see [17] for details).

These methods were combined with the following classifiers frequently considered in clinical problems [1] (all implemented in WEKA[1]): k-NN with $k = 1$ and 3 neighbors (1NN and 3NN, respectively, see [16] for justification), PART decision rules (PART), C4.5 decision trees (C45), naive Bayes (NB), neural networks with

[1] http://www.cs.waikato.ac.nz/ml/weka/.

radial basis functions (RBF), and support vector machines with radial basis function kernel (SVM). We selected such a kernel in SVM for consistency with our earlier experiments [16]. For PART and C45 we employed pruning—although unpruned classifiers are generally suggested for imbalanced data, our preliminary calculations revealed that unpruned decision rules and trees performed comparable or worse than pruned ones (this is consistent with [9] where the authors claim that pruning has the same effect as preprocessing). Moreover, parameters for RBF and SVM were optimized for each data set by using a simple grid search (systematic exploration of possible combinations of parameter values [21]) over original (i.e., not preprocessed) data sets. For the remaining classifiers we used default values of their parameters.

The classification performance was evaluated using the following measures, appropriate for imbalanced data: sensitivity and specificity for the minority class and their geometric mean (G-mean) that assesses their balance. We did not use the AUC (area under the Receiver Operating Characteristic curve) measure, as most of the classifiers selected in our study gave deterministic predictions. Here we should also note that by the minority classes we always considered the least prevalent class in the original data set (that was also critical from the clinical perspective) and we did not change it after preprocessing, even though several methods heavily modified the distribution of classes.

The above measures were estimated using a stratified 10-fold cross validation repeated 10 times to reduce the variance of results. In each pass of the validation loop, preprocessing methods were applied only to the learning set—the testing set was not modified. Finally, to gain better insight into how individual preprocessing methods affected the performance of considered classifiers, we applied non-parametric Friedman test (with $\alpha = 0.05$) that globally compared the performance of multiple combinations of preprocessing methods and classifiers over multiple data sets (see [16] for more details).

4 Results

Table 2 shows the distributions of the minority class example types in the original data sets. In all these data sets there was a large portion of borderline (24–63 %) and outlier examples (25–52 %), which indicates the minority class examples were "mixed" with the majority class ones. This, combined with small and very small portions of safe examples (14 and 29 % for SP and AP, respectively; 2–7 % for HP, SP and AE1—much lower than the portion of rare examples), implies the considered data sets were difficult for learning (see the discussion in [16]).

Due to space limitations, we focus on sensitivity when presenting detailed results from the second phase of the experiment; later we comment on the other measures (specificity and G-mean). Table 3 presents sensitivity for combinations of considered preprocessing methods and classifiers. In each row of this table the best value is marked with bold font, and the second best—with italics. Below we summarize our most important observations:

Table 2 Distribution of example types (%)

Data set	Safe	Borderline	Rare	Outlier
AP	29	38	8	25
HP	7	28	15	50
SP	4	53	11	32
AE1	2	63	10	25
AE2	14	24	10	52

- Performance on the AP data set, which was predicted to be the easiest to learn due to the largest proportion of safe examples (Table 2), was indeed better than the performance on the other data sets. This is especially visible for the baseline. Preprocessing diminished this difference, however, it still could be observed. This confirms the usefulness of the labeling technique from [16] to distinguish between less and more difficult data sets.
- Application of preprocessing methods always improved performance classification performance in comparison to the baseline. The extent of improvement was dependent on the classifier—large improvement was demonstrated for the 1NN, 3NN, C45 and PART classifiers, moderate improvement for RBF, and only small improvement for NB. Poor improvement for NB was related to its stellar baseline performance that narrowed the margin for improvement in comparison with other classifiers.
- As already stated, NB was the best performing classifier. The only one serious competitor was RBF. On the one hand, its good performance could be attributed to optimization of parameters. On the other hand, parameters were optimized on the original data sets, and RBF also performed well on preprocessed data, which confirms the robustness of such optimization.
- RU was clearly the best preprocessing method. Only in a few cases was it superseded by SP2. Interestingly, these were cases where SP2 was combined with NB, which may indicate synergy between these two methods. We should also note that in cases where RU led to the best classification performance, the difference between RU and the second best preprocessing method was much larger than the difference between the best method and RU, where RU was second or third.

Table 4 reports the results of the Friedman test based on sensitivity. For each classifier, the *null* hypothesis (i.e., that all preprocessing methods performed equally) was rejected. According to the critical difference between ranks from the Nemeneyi test (3.4), the differences between best performing methods were not significant; however, the best methods were significantly better than the worst ones. Table 4 also confirms our earlier observations. RU was the best method for all classifiers, except NB. Other promising methods were SP2 (which performed best with NB) and RO—both were awarded with high average ranks for all classifiers. Here we should also note surprisingly poor performance of SM in comparison to results reported in [16]. This could be explained by the prevalence of symbolic attributes in the considered

Table 3 Sensitivity for combinations of preprocessing methods and classifiers

Data set	Classifier	None	RU	RO	SM	NCR	SP2
AP	1NN	0.4300	**0.7500**	0.4300	0.5220	*0.5635*	0.5005
	3NN	0.4385	**0.7390**	*0.6495*	0.5365	0.5330	0.6230
	C45	0.3680	**0.7610**	0.5140	0.5005	0.5455	*0.5710*
	PART	0.4375	**0.7595**	0.5170	0.5255	*0.5340*	0.5325
	NB	0.7160	*0.7990*	0.7875	0.6770	0.7490	**0.8135**
	RBF	0.5130	**0.7860**	0.7645	0.6535	0.6685	*0.7405*
	SVM	0.5020	**0.7935**	0.7880	0.6150	0.5770	*0.7640*
HP	1NN	0.2035	**0.6035**	0.2035	*0.3315*	0.3040	0.2035
	3NN	0.1205	**0.6025**	*0.4300*	0.3630	0.2095	0.4280
	C45	0.2690	**0.7170**	*0.4965*	0.3865	0.3365	0.4780
	PART	0.2875	**0.6955**	*0.5115*	0.3585	0.3370	0.4840
	NB	0.7535	0.8480	*0.8510*	0.5645	0.7660	**0.8615**
	RBF	0.5475	**0.7920**	0.7145	0.4245	0.5865	0.6840
	SVM	0.5100	**0.7210**	0.4985	0.4445	0.5340	0.4970
SP	1NN	0.2743	**0.6307**	0.2743	0.3950	*0.4743*	0.2793
	3NN	0.2440	**0.6590**	*0.5553*	0.5240	0.4617	0.5513
	C45	0.3990	**0.6203**	0.5523	0.3950	0.4550	*0.5883*
	PART	0.3893	**0.6637**	0.5487	0.3597	0.4683	*0.5760*
	NB	0.4343	**0.7797**	0.7203	0.4077	0.5187	*0.7220*
	RBF	0.3913	**0.6977**	0.4920	0.4070	0.4743	*0.5220*
	SVM	0.3293	**0.6597**	0.3813	0.3350	*0.4163*	0.3947
AE1	1NN	0.2743	**0.5903**	0.2743	*0.4570*	0.3957	0.2760
	3NN	0.1623	**0.6327**	0.5097	*0.5277*	0.3163	0.4860
	C45	0.1847	**0.6080**	*0.3910*	0.2913	0.3097	0.3617
	PART	0.2553	**0.6330**	0.3723	0.2823	0.3497	*0.3953*
	NB	0.4897	*0.7143*	0.6833	0.4680	0.5803	**0.7167**
	RBF	0.4343	*0.6940*	0.6683	0.4763	0.5203	**0.7080**
	SVM	0.3217	**0.6170**	0.3147	0.3583	*0.4080*	0.3720
AE2	1NN	0.1000	**0.5867**	0.1000	*0.3217*	0.1317	0.1000
	3NN	0.0900	**0.7133**	0.4200	*0.4417*	0.1500	0.3750
	C45	0.1733	**0.6733**	*0.3933*	0.1500	0.2683	0.3300
	PART	0.2617	**0.6767**	*0.3767*	0.2817	0.3483	0.3400
	NB	0.7117	**0.7967**	0.7267	0.2400	0.7467	*0.7533*
	RBF	0.5317	**0.7917**	0.7367	0.2500	0.6800	*0.7533*
	SVM	0.4117	**0.5950**	0.3200	0.3433	*0.3533*	0.2900

data sets—while SM is able to handle symbolic data, it does it in a very simple way (i.e., by using most frequent local values in newly created examples), thus limiting the chances for improvement.

For completeness, we also applied the Friedman test to rank classifiers applied to original data sets (no preprocessing) and to classifiers combined with RU (the

Table 4 Ranking of preprocessing methods for specific classifiers (based on sensitivity)

Method	Avg. rank
(a) 1NN	
RU	6.0
SM	4.6
NCR	4.4
SP2	2.0
RO	2.0
None	2.0
(b) 3NN	
RU	6.0
RO	4.6
SM	3.8
SP2	3.6
NCR	2.0
None	1.0
(c) C45	
RU	6.0
SP2	4.4
RO	4.4
NCR	3.0
SM	1.8
None	1.4
(d) PART	
RU	6.0
SP2	4.2
RO	4.0
NCR	3.4
SM	2.2
None	1.0
(e) NB	
SP2	5.6
RU	5.2
RO	4.0
NCR	3.2
None	2.0
SM	1.0

(continued)

Table 4 (continued)

Method	Avg. rank
(f) RBF	
RU	6.0
SP2	4.8
RO	4.4
NCR	3.0
SM	1.6
None	1.4
(g) SVM	
RU	6.0
NCR	4.2
SP2	3.0
RO	2.8
None	2.6
SM	2.4

Table 5 Ranking of classifiers methods for selected preprocessing methods (based on sensitivity)

Classifier	Avg. rank
(a) NONE	
NB	7.0
RBF	5.8
SVM	4.6
PART	3.6
C45	3.0
1NN	2.4
3NN	1.6
(b) RU	
NB	7.0
RBF	5.8
SVM	4.0
PART	4.0
C45	2.8
3NN	2.8
1NN	1.6

best preprocessing method). The results are given in Table 5 and again they confirm our initial observation—the top ranks were assigned to NB and RBF respectively. Moreover, the ordering of classifiers according to ranks is quite stable—the only difference could be observed for 1NN an 3NN that swapped their positions.

Table 6 Impact of preprocessing methods on the size and class imbalance of the considered data sets

Data set	NCR		SP2	
	# examples (minority)	Imbalance ratio	# examples (minority)	Imbalance ratio
AP	397 (48)	0.12	601 (201)	0.33
HP	367 (46)	0.13	670 (305)	0.46
SP	344 (56)	0.16	665 (316)	0.48
AE1	294 (59)	0.20	620 (323)	0.52
AE2	207 (21)	0.10	344 (126)	0.37

Observations for G-mean are similar to those for sensitivity. Obtained results emphasized benefits of preprocessing (as for sensitivity, it always led to improvement in comparison to the baseline) and identified RU and NB as the most promising preprocessing method and classifier, respectively. However, unlike for sensitivity, the synergy between SP2 and NB was less visible, and NCR became a viable alternative for SP2 when combined with NB. Observations for specificity are complementary to those for sensitivity and G-mean. In most cases, application of preprocessing methods deteriorated the performance; only in a few cases was it preserved, in comparison to the baseline. As expected, the largest deterioration was observed for those methods that led to the best improvement of sensitivity, especially for RU (however, observed improvement of G-mean for these methods implied the loss on specificity was compensated by the gain on sensitivity). From this perspective, NCR and SP2 performed better than RU, as their negative impact on specificity was smaller (this more *balanced* behavior was consistent with our results reported in [17]).

Finally, we checked how preprocessing methods changed the class imbalance ratio in the considered data sets. Specifically, we focused on NCR and SP2, as the other methods were set to produce a balanced class distribution. The results are given in Table 6. SP2 introduced more extensive changes than NCR—specifically it *strengthened* the minority class by introducing copies of existing examples and by relabeling examples from the majority class, which led to either almost balanced (for AP and AE2) or balanced (for HP, SP and AE1) distribution of classes. This behavior of SP2 corresponds to suggestions in [24] and confirms the validity of our parametrization of RU, RO and SMOTE that enforced balanced distributions in preprocessed sets.

5 Conclusions

In the study described in this paper we analyzed data difficulty factors encountered in five imbalanced clinical data sets (all describing pediatric acute conditions managed in the ED) and examined how they could be addressed with common data

preprocessing methods. Specifically, we evaluated how application of these methods affected the performance of classifiers often recommended for clinical problems [1].

All the considered data sets were characterized by large portions of borderline and outlier examples in the minority class (which was clinically also the critical class). This implies the minority class was dispersed and mixed with the majority class, thus difficult to learn, what was confirmed by the poor performance of classifiers on original (not processed) data sets.

Experimental evaluation demonstrated that preprocessing always improved classification performance (in terms of sensitivity and G-mean); the largest gain was observed for classifiers combined with random undersampling. Moreover, the best performance on both original and preprocessed data sets was observed for the naive Bayes classifier. These findings are consistent with our experience with using naive Bayes alone [7], combining random undersampling with naive Bayes [11, 12], as well as with other experimental and theoretical studies on random undersampling [5, 6, 23] and naive Bayes [18].

Here we would like to briefly comment on the viability of the naive Bayes classifier in clinical applications. On the one hand, it represents captured knowledge in the form of a priori probabilities that are definitely less comprehensible than decision rules or trees. On the other hand, Bayesian reasoning forms the foundations of clinical decision making (see, for example, [18]), and as such should be familiar to clinicians.

There were several limitations associated with our study. First, the small number of data sets affected the results of Friedman test and weakened its conclusions. Moreover, the data sets described problems from a single (and specific) clinical domain, and results may not be directly transferable to other clinical domains (e.g., chronic diseases in elderly). However, given the consistently good performance of random undersampling and naive Bayes, we think this combination offers a reasonable starting point to analyze other imbalanced clinical data sets.

Acknowledgments The first three authors would like to acknowledge support by the Polish National Science Center under Grant No. DEC-2013/11/B/ST6/ 00963.

References

1. Bellazzi, R., Zupan, B.: Predictive data mining in clinical medicine: current issues and guidelines. Int. J. Med. Inf. **77**(2), 81–97 (2008)
2. Chawla, N.: Data mining for imbalanced datasets: an overview. In: Maimon, O., Rokach, L. (eds.): The Data Mining and Knowledge Discovery Handbook, pp. 853–867. Springer (2005)
3. Chawla, N., Bowyer, K., Hall, L., Kegelmeyer, W.: SMOTE: synthetic minority over-sampling technique. J. Artif. Intell. Res. **16**, 341–378 (2002)
4. Cios, K., Moore, G.: Uniqueness of medical data mining. Artif. Intell. Med. **26**, 1–24 (2002)
5. Drummond, C., Holte, R.: C4.5, class imbalance, and cost sensitivity: why under-sampling beats over-sampling. In: Proceedings of the Workshop on Learning from Imbalanced Data Sets, ICML 2003, pp. 1–8 (2003)
6. Drummond, C., Holte, R.: Severe class imbalance: Why better algorithms aren't the answer. In: Proceedings of the 16th European Conference ECML 2005, pp. 539–546, Springer (2005)

7. Farion, K., Wilk, S., Michalowski, W., O'Sullivan, D., Sayyad-Shirabad, J.: Comparing predictions made by a prediction model, clinical score, and physicians: pediatric asthma exacerbations in the emergency department. Appl. Clinic. Inform. **4**(3), 376–391 (2013)
8. He, H., Ma, Y.: Imbalanced Learning: Foundations, Algorithms and Applications. Wiley (2013)
9. Hoens, T., Chawla, N.: Imbalanced datasets: from sampling to classifiers. In: He, H., Ma, Y. (eds.) Imbalanced Learning: Foundations, Algorithms and Applications. Wiley, pp. 43–59 (2013)
10. Japkowicz, N.: Class imbalance: are we focusing on the right issue. In: Proceedings of the 2nd Workshop on Learning from Imbalanced Data Sets, ICML 2003, pp. 17–23 (2003)
11. Klement, W., Wilk, S., Michalowski, W., Matwin, S.: Classifying severely imbalanced data. In: Proceedings of the 24th Canadian Conference on Artificial Intelligence, Canadian AI 2011, pp. 258–264. Springer (2011)
12. Klement, W., Wilk, S., Michalowski, M., Farion, K., Osmond, M., Verter, V.: Predicting the need for CT imaging in children with minor head injury using an ensemble of naive bayes classifiers. Artif. Intell. Med. **54**(3), 163–170 (2012)
13. Kubat, M., Matwin, S.: Addressing the curse of imbalanced training sets: one-sided selection. In: Proceedings of the 14th International Conference ICML 1997, pp. 179–186 (1997)
14. Laurikkala, J.: Improving identification of difficult small classes by balancing class distribution. In: Proceedings of the 8th Conference AIME 2001. Volume 2101 of LNCS, pp. 63–66. Springer (2001)
15. Napierala, K., Stefanowski, J.: The influence of minority class distribution on learning from imbalance data. In: Proceedings of the 7th Conference HAIS 2012. Volume 7209 of LNAI, pp. 139–150. Springer (2012)
16. Napierala, K., Stefanowski, J.: Types of minority class examples and their influence on learning classifiers from imbalanced data. J. Intell. Inform. Syst. (2016, to appear)
17. Napierala, K., Stefanowski, J., Wilk, S.: Learning from imbalanced data in presence of noisy and borderline examples. In: Proceedings of the 7th International Conference RSCTC 2010. Volume 6086 of LNAI, pp. 158–167. Springer (2010)
18. Sajda, P.: Machine learning for detection and diagnosis of disease. Annu. Rev. Biomed. Eng. **8**, 537–565 (2006)
19. Saez, J., Luengo, J., Stefanowski, J., Herrera, F.: Addressing the noisy and borderline examples problem in classification with imbalanced datasets via a class noise filtering method-based resampling technique. Inform. Sci. **291**, 184–203 (2015)
20. Sanchez, V.G.J., Mollineda, R.: An empirical study of the behavior of classifiers on imbalanced and overlapped data sets. In: Proceedings of the 12th Iberoamerican Conference on Progress in Pattern Recognition, Image Analysis and Applications, pp. 397–406. Springer (2007)
21. Staelin, C.: Parameter selection for support vector machines. Technical Report HPL-2002-354 (R.1). HP Laboratories, Israel (2003)
22. Stefanowski, J., Wilk, S.: Selective pre-processing of imbalanced data for improving classification performance. In: Proceedings of the 10th International Conference DaWaK 2008. Volume 5182 of LNCS, pp. 283–292. Springer (2008)
23. Wallace, B., Small, K., Brodley, C., Trikalinos, T.: Class imbalance, redux. In: Proceedings of the 11th IEEE International Conference on Data Mining, pp. 754–763 (2011)
24. Wei, Q., Dunbrack, R.: The role of balanced training and testing data sets for binary classifiers in bioinformatics. PLoS ONE **7**(8), e67863 (2013)
25. Wilson, D., Martinez, T.: Improved heterogeneous distance functions. J. Atif. Intell. Res. **6**, 1–34 (1997)
26. Wilson, D., Martinez, T.: Reduction techniques for instance-based learning algorithms. Mach. Learn. J. **38**, 257–286 (2000)

Author Index

E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 471,
DOI 10.1007/978-3-319-39796-2