Advances in Intelligent Systems and Computing 681

Michał Choraś
Ryszard S. Choraś *Editors*

Image Processing and Communications Challenges 9

9th International Conference, IP&C'2017 Bydgoszcz, Poland, September 2017, Proceedings

Springer

Advances in Intelligent Systems and Computing

Volume 681

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

Michał Choraś · Ryszard S. Choraś
Editors

Image Processing and Communications Challenges 9

9th International Conference, IP&C'2017
Bydgoszcz, Poland, September 2017
Proceedings

Springer

Editors
Michał Choraś
University of Science and Technology, UTP
Bydgoszcz
Poland

Ryszard S. Choraś
University of Science and Technology, UTP
Bydgoszcz
Poland

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-319-68719-3 ISBN 978-3-319-68720-9 (eBook)
DOI 10.1007/978-3-319-68720-9

Library of Congress Control Number: 2017954292

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume of AISC contains the proceedings of the International Conference on Image Processing and Communications, IP&C 2017, which will be held in Bydgoszcz, on September 13–15, 2017.

The International Conference on Image Processing and Communications started as a forum for researchers and practitioners in the broad fields of image processing, telecommunications, and pattern recognition. It is known for its scientific level, the special atmosphere, and attention to young researchers. As in the previous years, the International Conference on Image Processing and Communications IP&C 2017 was held in Bydgoszcz, Poland.

IP&C 2017 was organized by the UTP University of Science and Technology and was hosted by the Institute of Telecommunications and Computer Sciences of the UTP University.

The IP&C 2017 brought together the researchers, developers, practitioners, and educators in the field of image processing and computer networks.

The conference proceedings contain 21 papers which were selected through a strict review process. During the conference, there were also two invited talks by young outstanding researchers: Dr Adam Schmidt from Technical University Munich and Dr Rafal Kozik from UTP University of Science and Technology.

The presented papers cover all aspects of image processing (from topics concerning low-level to high-level image processing), pattern recognition, novel methods, and algorithms as well as modern communications. IP&C has been a major forum for scholars and practitioners on the latest challenges and developments in IP&C.

The organization of such an event is not possible without the effort and the enthusiasm of the people involved. The success of the conference would not be possible without the hard work of the Local Organizing Committee. Without the high-quality submissions from the authors, the success of the conference would not be possible. Therefore, we would like to thank all authors for the effort they put into their submissions.

Last but not least, we are grateful to Dr Adam Marchewka for hard work as the Publication Chair, and also to Springer for publishing the IP&C 2017 proceedings in their Advances in Intelligent Systems and Computing series.

I hope that all of the attendees found the conference informative and thought provoking.

Michał Choraś
Conference Chair

Organization

Organization Committee

Conference Chair

Michał Choraś, Poland

Honorary Chairs

Ryszard Tadeusiewicz, Poland
Ryszard S. Choraś, Poland

International Program Committee

Kevin W. Bowyer, USA
Dumitru Dan Burdescu, Romania
Christophe Charrier, France
Leszek Chmielewski, Poland
Michał Choraś, Poland
Andrzej Dąbrowski, Poland
Andrzej Dobrogowski, Poland
Marek Domański, Poland
Kalman Fazekas, Hungary
Ewa Grabska, Poland
Andrzej Kasiński, Poland
Andrzej Kasprzak, Poland
Marek Kurzyński, Poland

Witold Malina, Poland
Andrzej Materka, Poland
Wojciech Mokrzycki, Poland
Sławomir Nikiel, Poland
Zdzisław Papir, Poland
Jens M. Pedersen, Denmark
Jerzy Pejaś, Poland
Leszek Rutkowski, Poland
Khalid Saeed, Poland
Abdel-Badeeh M. Salem, Egypt

Organizing Committee

Łukasz Apiecionek, UKW Bydgoszcz, Poland
Sławomir Bujnowski, UTP Bydgoszcz, Poland
Piotr Kiedrowski, UTP Bydgoszcz, Poland
Rafał Kozik, UTP Bydgoszcz, Poland
Damian Ledziński, UTP Bydgoszcz, Poland
Zbigniew Lutowski, UTP Bydgoszcz, Poland
Adam Marchewka (Publication Chair), UTP Bydgoszcz, Poland
Beata Marciniak, UTP Bydgoszcz, Poland
Tomasz Marciniak, UTP Bydgoszcz, Poland
Ireneusz Olszewski, UTP Bydgoszcz, Poland
Karolina Skowron (Conference Secretary), UTP Bydgoszcz, Poland
Mścisław Śrutek, UTP Bydgoszcz, Poland
Łukasz Zabłudowski, UTP Bydgoszcz, Poland

Contents

Algorithms and Methods

Communications and Miscellaneous Applications

Image Processing

Binary Line Oriented Histogram

Piotr Lech(✉) and Krzysztof Okarma

Department of Signal Processing and Multimedia Engineering,
Faculty of Electrical Engineering, West Pomeranian University of Technology,
26 Kwietnia 10, 71-126 Szczecin, Poland
{piotr.lech,krzysztof.okarma}@zut.edu.pl

Abstract. The paper focuses on the presentation of the Binary Line Oriented Histogram (BLOH) algorithm, originating from the Edge Oriented Histogram (EOH). In the case of thin single-pixel lines, a problem related to the interpretation of the EOH bins may occur which can be eliminated using the proposed approach. The main goal of the conducted experimental tests, using both synthetic and natural images, is determining the proper histogram with dominating single-pixel lines of the specified orientation to the horizontal axis. The results of our tests fully confirm the usefulness of the proposed BLOH algorithm.

Keywords: Line detection · EOH · Histogram · Image analysis

1 Introduction

The Edge Oriented Histogram (EOH) algorithm is known mainly as the element of Scale Invariant Feature Transform (SIFT) descriptor or MPEG-7 standard, however it may also be used separately. One of such exemplary applications may be the determination of the Region Of Interest (ROI) in digital images. It may be especially useful in systems with visual feedback, where a simplified representation of objects as skeletons with objects' edges marked by lines may be applied in modeling of the 3D environment. Those lines can be considered as "naturally" useful for the self-localization and navigation purposes in the 3D space.

Unfortunately, in many simple microprocessor systems, e.g. used in mobile robots, processing of high resolution images can be troublesome due to high computational power demands. A critical issue in the "classical" line detection is the thresholding related to the classification of each pixel as representing the line or not. In the case of too big amount of data for processing, various data reduction methods can be applied, limiting the number of colors, image resolution or the analyzed region. In many systems those three approaches are integrated and applied in many combinations.

In this paper, a modified line detection algorithm is presented, where the ROI filter is considered as the selector of the lines with specified orientation with simultaneous reduction of image resolution and the number of colors. In many cases finding the maximum EOH values for the specified line orientation may be

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*,
Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_1

sufficient for the determination of the ROI. Some popular line detection methods are based on the edge filters such as Sobel, Prewitt, Canny or may utilize pixel intensity data e.g. based on Otsu thresholding. Another popular approach of line detection is the use of Hough transform as well as the use of morphological [3] and median filters. In such approaches the limitation to a specified orientation of lines may be easily applied.

2 Problem of Thin Lines

In most cases in natural color and grayscale images single-pixel thin lines are not common.

The results obtained after the application of the EOH algorithm for the natural image are presented in Fig. 1. The bin numbers in the histogram correspond to the following edge orientations: 1 - horizontal, 2 - vertical, 3 - 45°, 4 - 135°, 5 - non oriented.

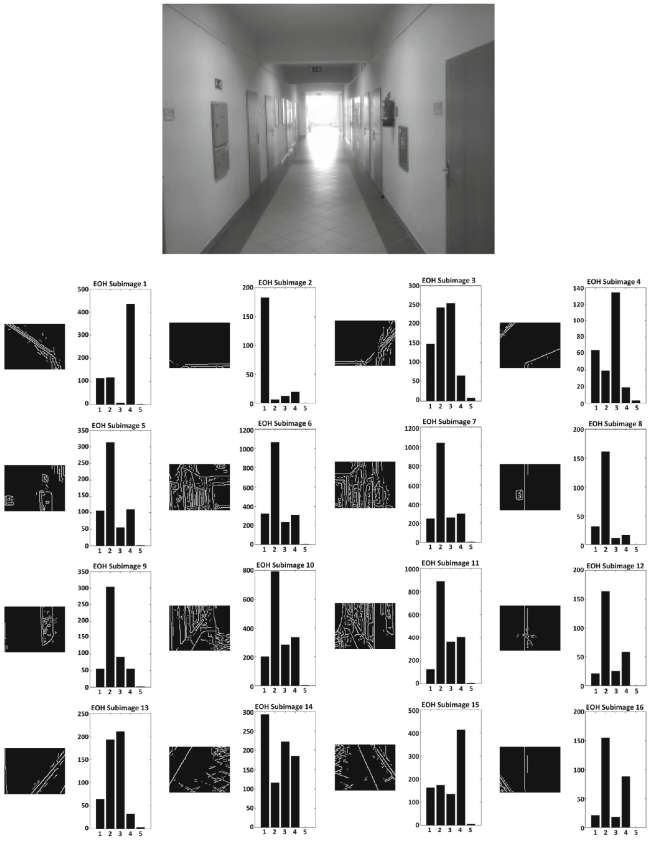


Fig. 1. Result of the EOH algorithm for gray scale image

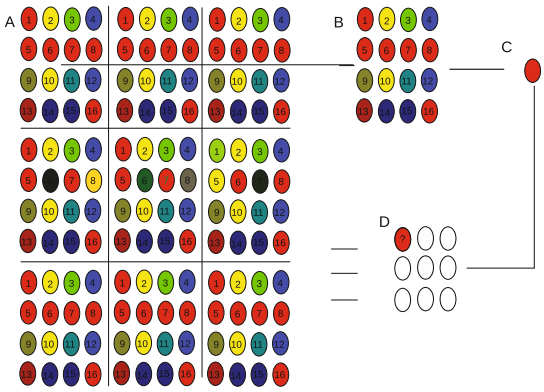


Fig. 2. Illustration of the resolution's decrease by statistical sampling

Nevertheless, a large number of such thin lines may occur in binary images obtained after the morphological thinning or the reduction of resolution, especially using the statistical sampling based methods. The presence of thin lines becomes especially important if the EOH algorithm is applied for binary images. In such cases the standard EOH algorithm produces more dominant bins and therefore an improper line detection takes place as illustrated in Fig. 3. However the histogram itself is calculated properly as it is related to the detected edges (also those which are very short) and their orientation, not to the straight lines.

Analyzing Fig. 3 the domination of the first bin for most of the histograms can be observed which suggests the presence of horizontal lines which does not occur on the original image. Such an issue is clearly visible in the histograms in Fig. 7 presenting the result of improper detection using the EOH method (middle row of the histograms) for a synthetic image used for the illustration of the problem. Due to the application of Canny edge detection [2] inside the EOH procedure, additional blurring is introduced, which causes the improper results of further stages of the algorithm and inappropriate histogram bins as the consequence.

A good example illustrating the source of thin single-pixel lines in the digital image is the decrease of the resolution using the statistical sampling presented in Fig. 2. The statistical sampling method is an alternative for the algorithms discussed in our earlier publication [6].

The image is divided into $r \times r$ pixels blocks where $r = 4, 9, 16$ etc. From each block a single pixel is randomly chosen being the element of a new image representing the whole block. Such operation can be conducted in a recurrent way in order to obtain the image of a desired reduced resolution.

3 Modification of the EOH Algorithm

In the EOH algorithm the edge orientation is evaluated by searching the maximum response over a edge filter, directly calculating the gradient in bins. Its performance is dependent on the number of bins – for the low number of them the

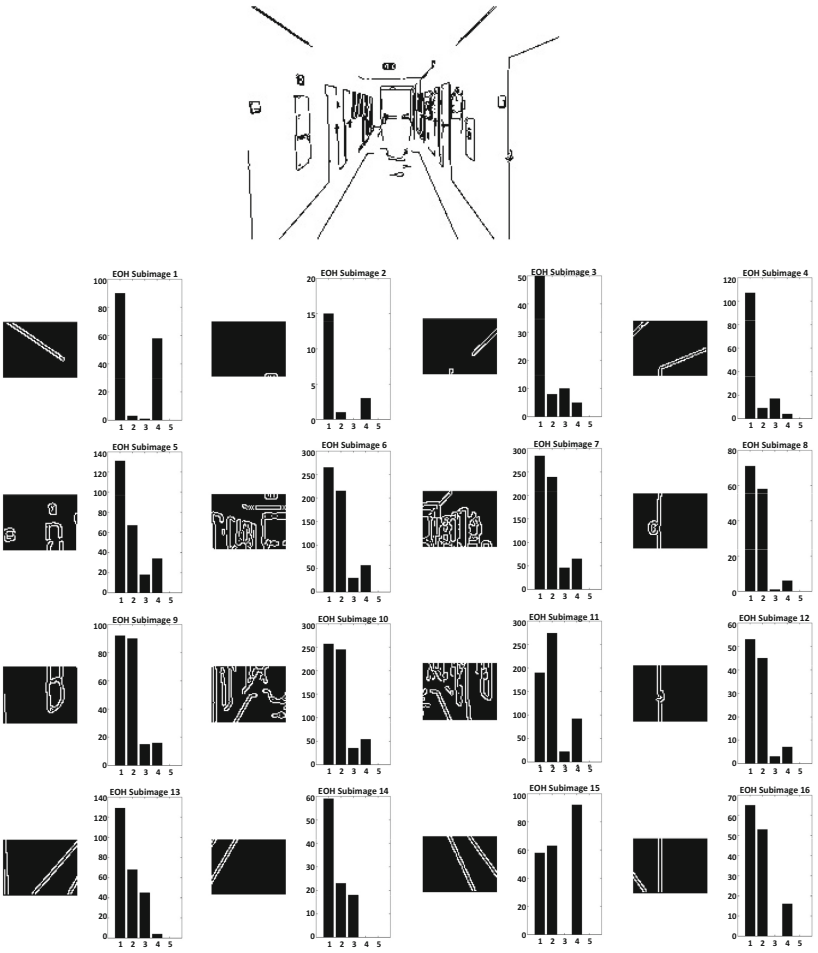


Fig. 3. Illustration of the source of the EOH algorithm's ambiguity

algorithm will be faster. The gradient magnitude and gradient orientation are calculated for each pixel and then the algorithm conducts edge detection using Canny method, allowing well visibility of lines on the image plane. The gradient orientation values are equally divided into K bins, $K = 5$ for five assumed directions (extracted with proper Sobel filter orientation) for vertical, horizontal, 45°, 135° and non-directional edges. The edge orientation histograms for each orientation bin k of cell C_i, $E_k(C_i)$, are calculated by sum of all gradient magnitudes with orientations belonging to the bin k.

The set of k EOH features for a single cell is the ratio of the bin value for single orientation to the sum of all bins values and can be expressed as:

$$EOH_{C_{i,k}} = \frac{E_k(C_i) + e}{\sum_{j=1}^{K} E_j(C_i) + e} \tag{1}$$

where e is the additional value preventing possible division by zero and unstable results for low denominator's values.

The analysis of the publicly available[1] source code [1] has led to its modification allowing a proper recognition of the single-pixel thin lines' orientation.

In the proposed BLOH method the following modifications have been introduced in comparison to the original EOH algorithm:

- analyzed image is subjected to previous binarization,
- edge detection is conducted using the Boolean approach based on the verification of accordance of the direct neighborhood of the analyzed pixel with the specified binary pattern – fragment of the code used for this purpose implemented in Octave environment is shown in Fig. 4,
- Sobel edge detection is replaced by the binary classification of the line orientation based on the accordance with the binary pattern.

The remaining part of the original EOH algorithm has not been changed. Since the new algorithm utilizes the binary data allowing the proper interpretation of thin lines, its proposed name is Binary Line Oriented Histogram (BLOH). Binary pattern representing no orientation has been simplified and defined for the image elements without neighbors in the binary image.

```
% edge detector
% BW - input binary image
rs = size(BW,1);
cs = size(BW,2);

for x = 1:rs-1
  for y = 1:cs
      BE1y(x+1,y) = (BW(x,y) == 1) & ( BW(x+1,y) == 0) ;
      BE2y(x,y)   = (BW(x,y) == 0) & ( BW(x+1,y) == 1) ;
  endfor
endfor

for x = 1:rs
  for y = 1:cs-1
      BE1x(x,y+1) = (BW(x,y) == 1) & ( BW(x,y+1) == 0) ;
      BE2x(x,y)   = (BW(x,y) == 0) & ( BW(x,y+1) == 1) ;
  endfor
endfor

BE = logical(BE1x +  BE2x + BE1y + BE2y);
%BE - output binary image with edges
```

Fig. 4. The most relevant fragment of the code implemented in Octave environment

In order to detect the edges with specified orientation each pixel of the image (together with its neighbors) has been checked for compliance with the pattern. If a specified configuration of pixels is detected according to one of the five patterns shown in Fig. 5, the pixel is classified as representing the line of respective orientation and stored. Proposed approach is much easier for implementation than the morphological edge detection [4] with sufficient quality of detection of lines.

[1] https://www.mathworks.com/matlabcentral/fileexchange/48200-ngunsu-matlab-eoh-sift.

Fig. 5. Patterns of 3 × 3 pixels, from left to right: horizontal, vertical, 45°, 135°, non oriented

4 Experimental Tests and Analysis of Results

For initial verification purposes the synthetic images containing a single line in each of the basic orientations have been created. For each of those images the

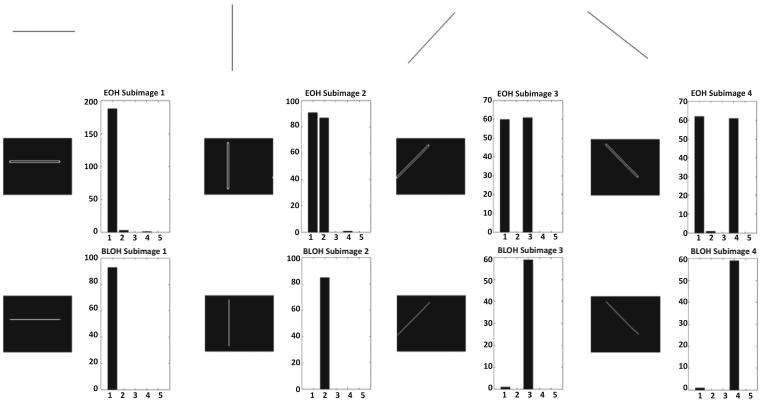


Fig. 6. Experimental results for synthetic images (top - binary image, middle - double edge problem for the EOH, bottom - result of the BLOH algorithm)

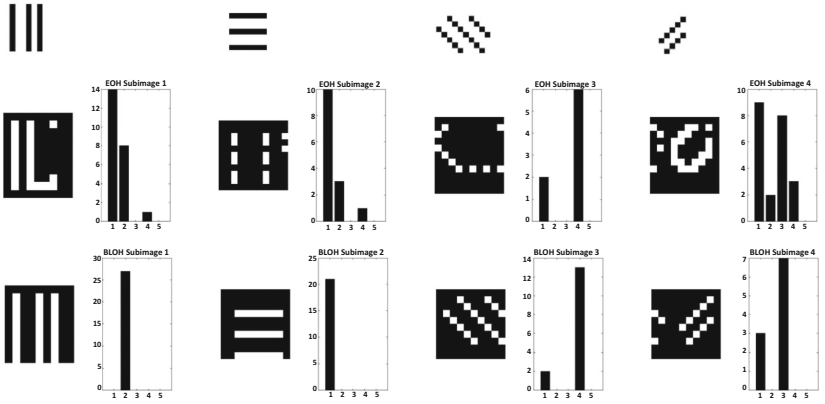


Fig. 7. Experimental results for synthetic images (top - binary image, middle - double edge problem and blurring for the EOH, bottom - result of the BLOH algorithm)

two histograms have been determined using the EOH and BLOH algorithms presented in Figs. 6 and 7.

The synthetic binary image shown in the top row of Fig. 6, containing single-pixel lines of a specified orientation, illustrates the influence of the detected double edge on the histogram results. The results of the application of the BLOH method as an alternative solution are shown in the bottom row.

The synthetic image shown in Fig. 7 contains many single-pixel thin lines oriented in predefined four directions. It has been used for the illustration of the other reason of improper histogram bins obtained by the EOH method related to the image blurring caused by the initial step of the Canny edge detection. As the BLOH algorithm is based on the other edge detection method which does not introduce such blurring, the results of its implementation are consistent with perceptual observations.

Additional tests have been conducted using the natural images captured by a camera and the results are presented in Fig. 8.

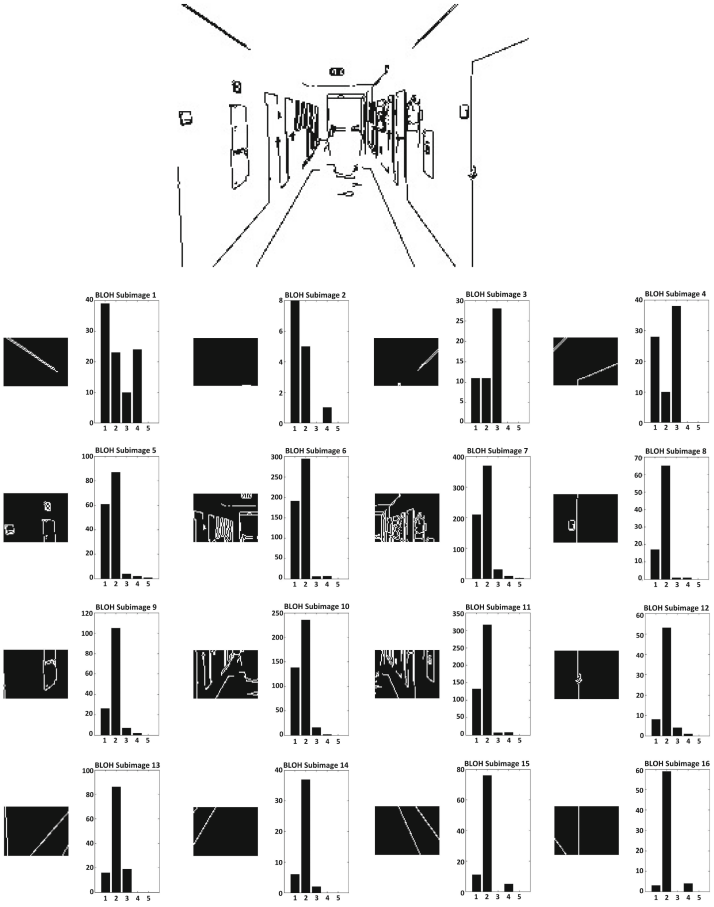


Fig. 8. BLOH experimental results for binarized natural images

5 Conclusions and Future Work

During the tests of the EOH method, some problems have occurred for the processed binary images in the case of thin single-pixel lines. Adaptation of the EOH algorithm for binary images (BLOH), preserving the general EOH idea, required the replacement of the EOH algorithm components: edge filter and line selector for a certain orientation. Comprehensive tests for binary images have shown erroneous results for the EOH method and their elimination with the BLOH algorithm fully confirming its usefulness for the images containing thin single-pixel lines. Conducted tests have proved the validity of the application of the BLOH method as the dominant orientation of the line has been properly detected in accordance to the expectations of the observers. Proposed algorithm may also be useful as a source for new classifiers devoted for binary images. As the future work the application of the BLOH algorithm for texture recognition is planned as an alternative to existing solutions [5].

References

1. Aguilera, C., Barrera, F., Lumbreras, F., Sappa, A.D., Toledo, R.: Multispectral image feature points. Sensors **12**(9), 12661–12672 (2012)
2. Canny, J.: A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. **8**(6), 679–698 (1986)
3. Iwanowski, M.: Morphological classification of binary image pixels. Mach. Graph. Vis. **18**(2), 155–173 (2009)
4. Kraft, M., Kasiński, A.: Morphological edge detection algorithm and its hardware implementation. In: Kurzynski, M., Puchala, E., Wozniak, M., Zolnierek, A. (eds.) Computer Recognition Systems, vol. 2, pp. 132–139. Springer, Heidelberg (2007)
5. Marciniak, T., Lutowski, Z., Marciniak, B., Maszewski, M.: The use of texture analysis in optical inspection of manufacturing processes. Solid State Phenom. **237**, 95–100 (2015)
6. Okarma, K., Lech, P.: Fast statistical image binarization of colour images for the recognition of the QR codes. Elektronika Ir Elektrotechnika **21**(3), 58–61 (2015)

CT–SPECT Analyzer - A Tool for CT and SPECT Data Fusion and Volumetric Visualization

Maria H. Listewnik[1], Hanna Piwowarska-Bilska[1], Krzysztof Safranow[2], Jacek Iwanowski[1], Maria Laszczyńska[3], Maria Chosia[4], Marek Ostrowski[5], Bożena Birkenfeld[1], and Przemysław Mazurek[6(✉)]

[1] Department of Nuclear Medicine, Pomeranian Medical University in Szczecin, Szczecin, Poland
[2] Department of Biochemistry and Medical Chemistry, Pomeranian Medical University in Szczecin, Szczecin, Poland
[3] Department of Histology and Developmental Biology, Pomeranian Medical University in Szczecin, Szczecin, Poland
[4] Department of Pathology, Pomeranian Medical University in Szczecin, Szczecin, Poland
[5] Department of General Surgery and Transplantology, Pomeranian Medical University in Szczecin, Szczecin, Poland
[6] Department of Signal Processing and Multimedia Engineering, West–Pomeranian University of Technology, Szczecin, 26. Kwietnia 10 St., 71126 Szczecin, Poland
przemyslaw.mazurek@zut.edu.pl

Abstract. Data fusion and specific visualization of CT and SPECT are important for diagnosis and research purposes. Selected problems are considered in the paper and are related to the developed CT–SPECT Analyzer software. Hierarchical mapping with SPECT priority for maximum value of rays is applied in this software. Three variants of color mappings are presented. Some practical aspects related to low quality of CT are considered also. The most promising is the rainbow gradient with gamma curve adjustment.

Keywords: CT · SPECT · Data fusion · Volumetric visualization

1 Introduction

Computer tomography (CT) is a standard medical procedure nowadays. CT allows the acquisition of volumetric data of patient body and it is the evolution of old X–ray technology. There are numerous challenging tasks related to CT due to phenomena related to patient body, limitations of X–ray exposition, synthesis of volume and data visualization, for example. Conventional X–ray imaging gives single view of patient body, so 2D image is obtained at high spatial 2D resolution. CT give access to arbitrary view of patient body with software based slicing technique. Multiple measurements are necessary for the reconstruction

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*, Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_2

(estimation) of single slice of patient. Multiple slices are combined into volume of patient body.

The visualization of 2D images is rather simple and numerous image enhancement algorithms are available, starting from simple like: global or local contrast enhancement, false color mapping, up to advanced like: machine vision algorithms for pattern (object) recognition and fitting to specific prior models of body structures.

Similar enhancement techniques are applied for volumetric data, but 3D visualization supports opacity, different data fusion algorithms and artificial light/shading, moreover.

Another standard technique that allows volumetric recording from patient body using gamma rays is SPECT (Single–Photon Emission Computed Tomography) [7]. The source of gamma rays is a radioisotope. CT–SPECT machine allows the acquisition of both volumes together, that are spatially aligned and calibrated. Combined CT–SPECT volume has two values for single voxel (3D pixel) corresponding to both measurements. The fusion algorithm is necessary for specific region body that enhances visibility, especially for medical research purposes.

1.1 Related Works

Data fusion of volumens is very important for medical images analysis [3,5]. There are numerous software tools for CT and SPECT data visualization. Some of them are universal and require a lot of settings, some of them are dedicated tools. The developed software is dedicated to analysis of parathyroid glands [1], so specific set of options are available, for improving interaction with operator. This set evolved during development of CT–SPECT Analyzer for improving man–machine interface toward to more intuitive and with low latency in pipeline. There are numerous variants of data visualization for volume [6] and current version of software was obtained by numerous iterations after testing visualization quality and possibilities of results interpretation.

1.2 Contribution of the Paper

This paper considers visualization pipeline for developed CT–SPECT Analyzer software. A few variants of visualization of combined CT–SPECT volume are presented, used for improving contrast of parathyroid glands.

The processing pipeline is considered in Sect. 2. Example results of visualization are provided in Sect. 3 for maximum value mode and different color mappings. Discussion of results is provided in Sect. 4. Conclusions and further work are considered in Sect. 5.

2 CT–SPECT Processing Pipeline

CT–SPECT Analyzer uses a few large frameworks, that support migration of software between Linux and Windows. The first version was designed using Linux

and later was converted to Windows due to portability of libraries. QT framework was selected as GUI and main framework, that forces structure of code and order of operations also. Most tools for volume visualization uses multiple windows with numerous options. CT–SPECT Analyzer is a single window application intentionally – sophisticated interface with multiple windows of the same object of interest requires long learning of operator. Blender, has a very efficient man–machine interface in single window mode was used as a reference. The control of camera view and object uses similar mouse interaction and some keyboard shortcuts are identical, also. QT supports e.g. SpaceNavigator 3D mouse that gives advanced control of view, moreover. Example screen of CT–SPECT Analyzer is shown in Fig. 1.

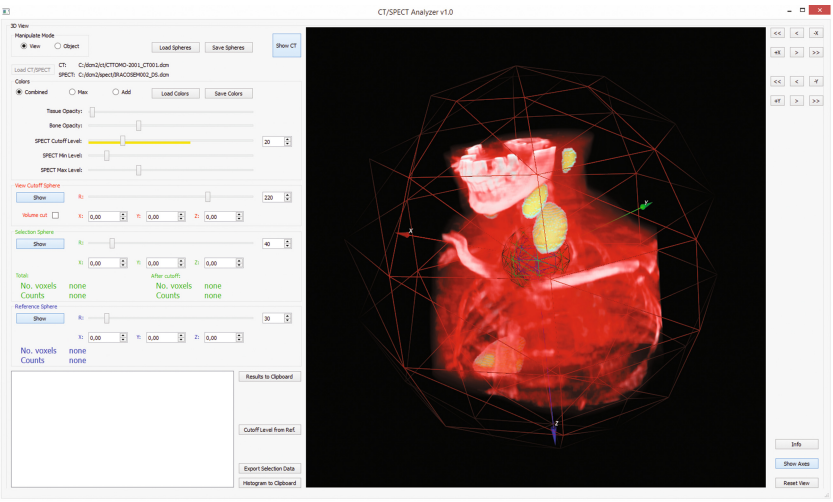


Fig. 1. Example screen of CT–SPECT Analyzer

The input data are standard DICOM files [4] that include a lot of metadata about acquisition parameters as well as patient personal data. Such data are used by CT–SPECT Analyzer, but patient data are hidden due to privacy protection requirements. DICOM files are imported using GDCM library (Grassroots DICOM), but some additional adjustments are necessary because available DICOM files do not comply standard requirements.

CT and SPECT volumes are processed using ITK [2] and VTK [6], and they are responsible for basic processing of volumes and visualization tasks respectively. Both are very complex and they offer a lot of options, but data fusion of both volumes must be developed separately. ITK and VTK are free libraries, well documented and with active development, so further improvements are expected. Schematic of data pipeline is shown in Fig. 2.

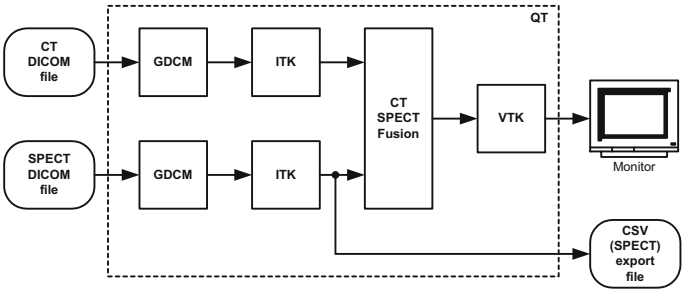


Fig. 2. Data pipeline of CT/SPECT Analyzer

2.1 Volume Rendering

Numerous volume rendering techniques are available and VTK supports a few of them. The maximal value and additive value are the most interesting. The additive value mode uses accumulation of values for particular screen pixel. The maximal value mode uses maximal value obtained for particular pixel (Fig. 3). There are some disadvantages of additive mode because color rendering does not work in this mode and some artifacts may be generated from voxels – maximal value could be as sum of different, not connected regions from biological point–of–view. During first tests maximal value mode was selected as most promising for analyzes of parathyroid glands.

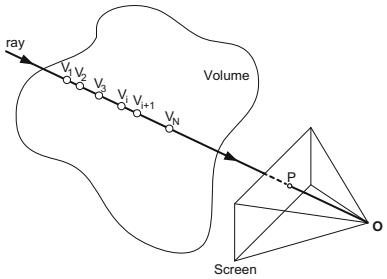


Fig. 3. Computation of pixel value

CT volume is the reference for SPECT in maximal value mode, because it contains patient skeleton and tissue. SPECT volume resolution is lower and very specific. The region of interest (ROI) for SPECT, after manual selection, is exported to a text file in CSV format for further analysis using external tools.

2.2 Selection of ROI and Data Fusion

The selection of ROI in SPECT volume is a manual process supported by CT data. The problem of selection is related to CT–SPECT volumes fusion. VTK library

allows the assignment of RGB value and opacity to particular voxel. Rendering algorithm calculates final value for particular ray and the final pixel with RGB values are achieved on computer screen. Color mapping algorithm is applied for the assignment of voxel value to color and opacity. The VTK pipeline is very fast and memory efficient, that is necessary for interactive application, but requires proper color and opacity mapping. Rendering of two volumes in the same position is not a correct solution.

There are three spheres (they are VTK objects) that define visible region and area of analysis. The selection sphere defines ROI for further analysis, because there are multiple blobs in SPECT images.

Proposed solution in CT–SPECT Analyzer uses hierarchical mapping for visualization. The most important data are SPECT, and CT data could be overlapped by SPECT. The mapping function requires input values from 0 to 255 range for the VTK interpolation algorithm using look–up table (LUT). Values from 0 to 127 are assigned to CT data and values from 128 to 255 are used by SPECT. CT volume data are mapped to particular values. It is possible for CT, because HU (Hounsfield Unit/CT number) ranges are well defined for air, water, body tissue, bones. SPECT data are contrast enhanced for improving visibility. Automatic contrast is straightforward (128 - corresponds to value 0 of SPECT; 255 - corresponds to maximal value of SPECT). Additional manual adjustments are possible using GUI sliders, that control contrast. Following formulas are applied for contrast corrections:

$$V^*_{SPECT}(x,y,z) = 127\frac{V_{SPECT}(x,y,z)}{Vmax_{SPECT}} + 128 \tag{1}$$

$$V^*_{CT}(x,y,z) = a \cdot V_{CT}(x,y,z) + b, \tag{2}$$

where V_{SPECT} and V_{CT} denote volumens, a is the scaling coefficient, and b is the bias of HU for air, and $Vmax_{SPECT}$ is the maximal measured value of SPECT. The voxel coordinates are denoted using x, y, z.

The problem of maximum value mode for rendering is related to the overlapping of CT by SPECT. Some voxels should be assigned to 0–127 range for visibility of body and bones from CT. The solution is the threshold algorithm with level T that is used for two purposes - visualization of CT and manual segmentation of ROI:

$$V^*_F(x,y,z) = \begin{cases} V^*_{SPECT}(x,y,z) : V_{SPECT}(x,y,z) > T \\ V^*_{CT}(x,y,z) : V_{SPECT}(x,y,z) \leq T \end{cases} \tag{3}$$

A higher value of threshold reduces ROI, that is possible and necessary, and improves visibility of parathyroid glands.

2.3 Color Mapping for CT

There are some expectations of users related to color mapping. Typical X–ray image is grayscale with bright area related to high density body parts like bones.

Calibrated CT volumes (in HU values) give the possibility of patient's body segmentation using predefined ranges. HU value ranges are established for typical patient body and could be applied for improving the visibility of bones. The problem is that HU ranges are related to high quality CT volumes. CT–SPECT acquisition uses CT as a reference for SPECT, and high resolution is not necessary. Faster acquisition of CT reduces resolution and introduces some reconstruction artifacts, but more important for patient, it reduces X–ray dose. The HU unit ranges for such case do not correspond to expected, so multiplication by experimentally evaluated value is used (Fig. 4).

The color gradient with three colors is used: black for air, red for tissue, and white for bones. Intermediate colors are possible depending on HU value. Opacity values for tissue and bones are controlled using sliders.

2.4 Color Mapping for SPECT

Simple gradient was assigned with two contrast colors (opposite in HS color space) during first tests. Small differences in SPECT ROI were not well visible for untrained operator and require good quality monitor and proper light conditions in room.

Alternative mapping uses rainbow gradient, but highest value of SPECT are assigned to red and magenta. Such constraint is necessary because body tissue is also red. Low values of SPECT are assigned to blue like colors so this color gives visual boundary between red color of tissue and SPECT. It is possible to disable visibility of CT on demand.

Further improving of visual contrast is possible by the application of nonlinear mapping of colors. One of the mappings is the gamma curve:

$$V_{SPECTcorr_{norm}(x,y,z)} = V_{SPECT_{norm}}(x, y, z)^{\gamma}, \tag{4}$$

where $V_{SPECT_{norm}}$ is the value from normalized range (0–1).

2.5 ROI Envelope

Additional feature that allows the better visual separation of CT and SPECT is the cyan color that is assigned to two LUT positions 127 and 128. The envelope around SPECT volume is created and is related to settings of threshold setup.

3 Example of Parathyroid Glands

The correction of HU due to low quality CT volume is shown in Fig. 4. The contrast is enhanced two times.

SPECT color mapping for maximal value mode is show in Fig. 5. The gamma coefficient is 0.5. Red–brown area is a part of CT after the enabling of volume cut using sphere. SPECT resolution is low so some artifacts are shown. Data fusion algorithm uses interpolation of SPECT and nearest neighborhood algorithms, so voxel/pixel structure is well visible. The regions are different due to not the same threshold value. This view is obtained from side view of patient.

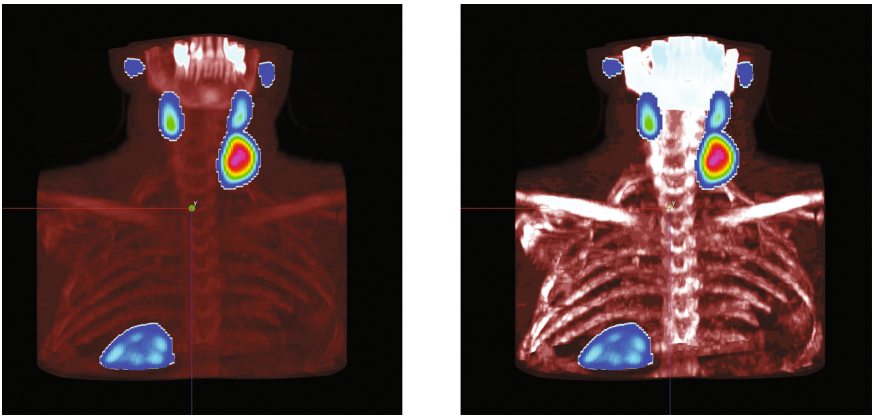


Fig. 4. Raw (left) and HU corrected (right) volume examples

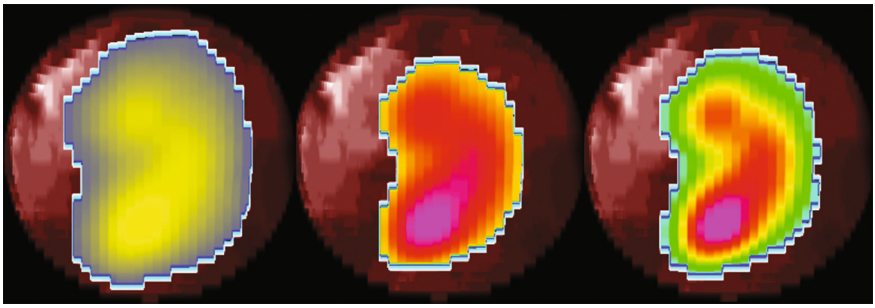


Fig. 5. Three different visualization techniques for parathyroid glands: gradient of two opposite colors (left), rainbow mapping (middle), rainbow mapping with gamma (right)

4 Discussion

Direct application of HU ranges does not provide expected visual results and most bones are low contrast for low quality CT data. Improvement of contrast gives better visual results (Fig. 4), but saturation related to high density bones is visible also (skull area). Automatic selection of HU regions is interesting alternative to experimental selection of contrast, but more different quality CT volumes are necessary.

The application of gamma curve improves visibility of details and instead a single object (mirrored C shape in Fig. 5) two or three blob regions are well visible. The application of local contrast could be applied for the further improving of details visibility, but pattern recognition approach is better for the analysis. The proposed visual contrast enhancement is sufficient for the determination of basic properties of parathyroid glands.

5 Conclusions and Further Work

Current version of CT–SPECT Analyzer support export of SPECT volume selection with mask for further processing of data using external tools, like Matlab or R. The second part of the project is Generalized Gaussian Distribution Mixture optimization tool. This part allows the estimation of parameters for selected parathyroid gland area using CUDA code. Further version of CT–SPECT Analyzer will merge optimization tool.

Acknowledgment. This study was supported by grant from budget resources for science in the years 2010–2015 as a research project No. N N402 463339 (Poland).

This work is supported by the UE EFRR ZPORR project Z/2.32/I/1.3.1/267/05 "Szczecin University of Technology – Research and Education Center of Modern Multimedia Technologies" (Poland).

We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan X GPU used for this research.

References

1. Fanti, S., Farsad, M., Mansi, L.: Atlas of SPECT-CT. Springer, Heidelberg (2011)
2. Johnson, H., McCormick, M., Ibanez, L.: The ITK Software Guide Book 1: Introduction and Development Guidelines. ITK.org (2017)
3. Ogiela, M.R., Hachaj, T.: Natural User Interfaces in Medical Image Analysis. Springer, Switzerland (2015)
4. Pianykh, O.: Digital Imaging and Communications in Medicine (DICOM). A Practical Introduction and Survival Guide. Springer, Heidelberg (2012)
5. Piórkowski, A., Kempny, A.: The transesophageal echocardiography simulator based on computed tomography images. IEEE Trans. Biomed. Eng. **60**(2), 292–299 (2013)
6. Schroeder, W., Martin, K., Lorensen, B.: The Visualization Toolkit: An Object-Oriented Approach to 3D Graphics. Prentice-Hall, Englewood Cliffs (2006)
7. Wernick, W., Asrsvold, J. (eds.): Emission Tomography: The Fundamentals of PET and SPECT. Elsevier, San Diego (2004)

Image Search Enhanced by Using External Data Sources and Reasoning

Arkadiusz Cacko(✉) and Marcin Iwanowski

Institute of Control and Industrial Electronics, Warsaw University of Technology, ul. Koszykowa 75, 00-662 Warszawa, Poland
{arkadiusz.cacko,marcin.iwanowski}@ee.pw.edu.pl

Abstract. In the paper, the concept of a system is presented, that makes use of the reasoning to perform digital image search using the information present in an annotated images combined with external text data sources in order to answer complex queries. The workflow in based the Prolog reasoning that allows for both modeling ontologies and combining the knowledge from both data sources. Prior to reasoning, both data sources are transformed into the form of Prolog predicates. An example application of the proposed concept is also presented in the paper. It deals with photographs database with annotated faces of politicians combined with the external knowledge stored in the Parliament-members database. It provides with a mechanism that supports formulating complex questions like: "Find a photograph where a politician belonging to party A stands between a woman and another political that was a member of Parliament of B-th term of office".

1 Introduction

The image search is usually performed based on either hypertext description of images or on image descriptors. In the first case, search is, in fact, a classic text-based search because both query and data that is processed are text data, the image content is not directly investigated. In the other case the image-database is processed based by computing similarities between photographs that has been supplied and is image-only-based operation. However, the data stored in text and image databases are often supplementary and search based exclusively on a single domain could not profit from all available sources.

Indexing and annotation of images in search engines is an active area of current research. The main aim of search engines is to provide the most relevant documents to the users in shortest possible time. Research in this area has been focused on finding the connection between low level image features and high level semantics. In many situations humans want to find at the photo things that computer doesn't understand because search criteria aren't directly connected with image domain content extracted by computer. That's why there is a highly desirable that systems allow for search in multiple domains. To achieve this goal it is necessary to convert the knowledge from various domains (image and text) to the form that contain a common semantic base.

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*, Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_3

The paper presents a concept of a system that allows for combining various data sources. The information that is present in annotated images from an image database is combined with the data stored in a separated relational databases containing text data. In order to overcome the difficulties related to the lack of homogeneity of information from both sources, the predicate-based data representation is used to describe all available information. From one side a set of predicates describes the content of the images. From the other, the text database is converted to another set of predicates. Also, the relative positions of image objects is modelled in stored in a separated set of predicates. Having stored the information from various sources in a homogeneous form, they are combined using reasoning techniques and new knowledge can be inferred consequently. The predicates and reasoning in the current paper are based on the Prolog logic programming language. It allows for defining ontologies referring to various data sources. A diagram presenting the proposed concept is shown in Fig. 1.

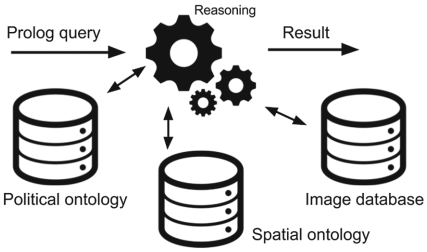


Fig. 1. Simplified system diagram

Creating photo search engine based on annotated images, enriched by ontology connected with inference process should give more valuable results than standard approaches. Inference gives a ability to discover not trivial relationship between objects presented in photos. An example application of the above concept is presented in the paper. It allows for searching the photograph database with annotated faces of politicians with use of the external knowledge from the Parliament-members database.

The paper is organized as follows. In Sect. 2, related works are described. In Sect. 3, the image annotation is discussed. Section 4 is focused on the reasoning process. In the Sect. 5, three sets of predicates are described. In Sect. 6 example Prolog queries and their results are presented. Finally, Sect. 7 concludes the paper.

2 Related Works

The complexity of image search exceeds considerably the text-search. In order to consider the content of an image, the appropriate image processing techniques are

applied. The one which is used in the current research is the image annotation [9]. During the annotation, structured data structures are created, that are further used to construct standard search queries [7]. Among other image processing techniques, that are used to extract valuable information from digital images, are feature points [6].

Image search engine could be based on structural data collected during image annotation. In the simplest case, such data may be queried by means of standard SQL-based search. This however does not allow for obtaining new information that is not directly stored in the database. In order to get such functionality, the inference methods are applied. Such an approach gives better results thanks to new knowledge discovery from database [1]. To improve further the search results, ontologies has been introduced in this domain [4]. This allows for searching supported by semantic understanding. Other approaches, in addition to allow standard searching in image content, support Natural Language Processing (NLP) in user queries formed as a standard text [3].

3 Image Annotation

Image annotation plays a crucial role in the image search process. It allows for describing the content of images so that they may be investigated in order to find desired objects. In case of photographs annotation process, objects refer to elements of scenes presented on particular photos. Depending on the type of a photo they may refer to various visually important entities. The set of object descriptors holds thus an important information about semantic content of the photo.

There are many ways to annotate images by retrieving objects from the visual scene. In a general case feature points [6] may be used but there exists many algorithms dealing with specific classes of objects. In case of photographs of humans, the *face recognition*[10] methods are applied, thanks to which it is possible to annotate a photo by indicating who is present on it, and consequently to retrieve meaningful knowledge about a visual scene of this kind.

Annotation is very useful in case of thematic collections of images. Once the object on images are annotated and thus identified, this knowledge may be combined with external data describing the image objects. In some thematic domains, the external data sources are rich and, when appropriately used, may considerably improve the process of the image search. One of such domains is a politics – various information on politicians are available from many sources. In our study we use the database of Polish politicians that is used jointly with the collection of photographs presenting them in various configurations. In this domain, there exist a huge collection of various photographs presenting politicians. Every day, in all newspapers, the news from the world of politics are exhibited. Also the Internet contains the enormous number of photos with politicians published everyday. On the current study the database of images exhibiting politicians was collected. The annotation of each photograph was performed by means of the face detection algorithm based on Haar classifiers [8]. For each

detected face, the face recognition is next performed [10] in order to assign to each face, the ID of a person that has been recognized. The final result of the photograph annotation process consists of the collection of personal ID's and positions of faces on photos, stored for further processing. An example of annotation of the single photograph of politicians is shown in Fig. 2 (left-hand side). The annotation results for all images under consideration are stored as a meta-data in the image database.

4 Reasoning

Ontology is defined as formal and declarative representation which includes specific vocabulary referring to the terms in the subject area that contains relationships among these terms. It also provides a way for representing knowledge about some topic and describing relationships between this terms. Ontologies are used across a number of domains and brings benefits in many areas [2].

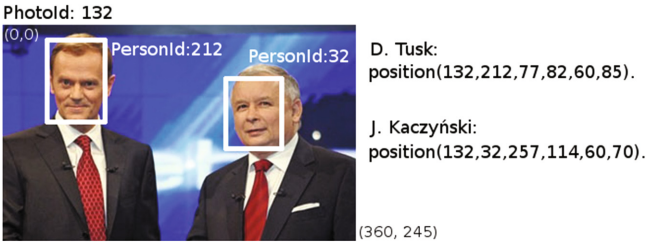


Fig. 2. Image people annotation example.

To benefit from the many different sources of knowledge, the common standard for their representation must be used. In our case, two domains are considered: facts from the external text-database and annotated images stored in the image database. The first source is the relational database while the second is represented by image meta-data generated by the image annotation process. The principal idea of the current study is to convert both data sources to the form of Prolog predicates. This form does not only introduce a common data representation but also allows for performing the *reasoning*. The latter is extremely important in this case, because it allows for inducing new knowledge based on existing information from both sources.

The external database in the current study contains the biographical information on Polish politicians. There are many sources of information on members of Polish Parliament. Web pages of Sejm and Senat (lower and upper house of Polish Parliament) contain various data for every member of parliament of the current term of office. For older terms of offices, historical publications may be helpful. Based on such a various data sources political relational database was created. A similar approach was used in many projects i.e. TrendMiner [5]

(ontology build in OWL format). This database contain information about the activity of politicians, their personal data and political history. This relational database is converted to a set of Prolog facts by a transforming every record to the predicate form.

Apart from the storing the information from external database and image database, Prolog predicates are used to describe ontology referring to spatial relations between annotated faces. Their relative positions are also used to describe and to query the photographs. Several position descriptors may be defined. Every person on the photo is located in a specific part of image. This person is also surrounded by other people. Every person may be in specific (spatial) relation to each other. For example, one man can be on the left of someone, other person may be between two people. Such relations are useful and brings additional knowledge about image content. To improve process of searching in our database, spatial predicates were introduced as the third category of predicates.

5 Predicates

The data in the proposed system come from three sources of information, each of which is converted to a set of predicates describing both facts containing the data referring to particular objects and rules that describe relations between facts:

- external database – data predicates – a set of Prolog facts and rules,
- image database – image predicates – a set of Prolog facts,
- spatial relations – spatial predicates – a set of Prolog rules.

The set of external information predicates was transformed from relational database containing facts on polish political history from 1990's to present. It contains the data on all members lower and upper house of the Polish parliament (Sejm, Senat). In our study these data are converted to a set of the following Prolog predicates (with *examples*):

- Personal data:
  person(ID,Surname,2-ndSurname,Name,Male,BirthDate,BirthPlace).
  person(1,'Jaroslaw','Aleksander','Kaczynski',M,18-06-1949,'Warszawa').
- Parliamentary groups and terms:
  group(ID,Term,House,StartDate,EndDate).
  group(27,4,Sejm,19-10-2001,18-10-2005).
- Details about the group:
  groupName(ID,GROUPID,ShortName,FullName).
  groupName(209,27,'PiS','Klub Parlamentarny Prawo i Sprawiedliwosc').
- Politician belongings to Parliamentary group:
  personInGroup(GROUPID,PersonID,StartDate,StopDate).
  personInGroup(27,1,19-10-2001,18-10-2005).

Political facts contains also details about terms of Polish parliament. Using this set of facts we are able to show political history of every person, check how long stay in politics or how often he/she was changing the party. Such knowledge about politicians can be combined with the image predicates extracted from the database of annotated photos:

- Description of a photograph:
  photo(ID,NameOfPhoto).
  photo(ID,'ziobro-tusk.jpg').
- Person on a photograph (FaceCenterX, FaceCenterY stand for the cetroid and SizeX, SizeY for the sizes of the face's bounding box; see an example on Fig. 2):
  position(PhotoId,PersonID,FaceCenterX,FaceCenterY,SizeX,SizeY).
  position(132, 32, 257 114, 60, 70).

Data predicates were created from content of political domain, *image predicates* were formed during image annotation process. The third group of predicates – *spatial predicates* – describe various possible relations between objects in visual scene. The spatial ontology is universal source of reasoning about mutual position between objects on the photos. Example spatial predicates are the following:

```
% true if Person1 is on left of Person2
onLeft(ID1, ID2, PhotoId) :- position(PhotoId,ID1,X1,_,_,_),
  position(PhotoId,ID2,X2,_,_,_), X1 < X2.
% true if Person1 is between Person2 and Person3
between(ID1, ID2, ID3, PhotoId) :-
  onLeft(ID1, ID3, PhotoId), onRight(ID1, ID2, PhotoId).
% true if Person1 is near PersonID2
near(ID1, ID2, PhotoId) :-   onLeft(ID1, ID2, PhotoId),
  not(between(_, ID1, ID2, PhotoId)).
near(ID1, ID2, PhotoId) :-onRight(ID1, ID2, PhotoId),
  not(between(_, ID1, ID2, PhotoId)).
```

6 Queries

The common data representation based on Prolog predicates allows for formulating queries on the data using also a predicate form. Depending on the complexity of a query various data sources are investigated. Some examples of queries are listed below:

- Find filenames of all photos with a person of ID=10 (only image predicates used):

```
find(Name):-position(PId,'10',_,_,_,_),photo(PId,Name).
```

- Find ID's of all photos where a person of the name Kowalski is present (image and data predicates):

```
find(PhotoId):-person(PersonId,_,_,'Kowalski',_,_,_),
  position(PhotoId,PersonId,_,_,_,_),photo(PhotoId,_).
```

– Check if a politician ID=5 was ever present on the single photograph together with a political from party B (image and data predicates):

```
check:-groupName(_,GId,'B',_),person(PId,_,_,_,_,_,_,_),
 PId\=5,personInGroup(GId,PId,_,_),group(GId,_,_,_,_,_).
```

– Check if a person ID=10 was ever staying on the left-hand side from a person of ID=20 (image and spatial predicates):

```
check:-onLeft(10, 20, PhotoId).
```

– Find the ID of a photo where politician from party A was on the right-hand side of a politician born in the city B (image, data and spatial predicates):

```
find(Id):-position(PhId,PeId,_,_,_,_), PeId \= PeId2,
 personInGroup(GId,PeId,_,_),groupName(_,GId,'A',_),
 position(PhId,PeId2,_,_,_,_,_),person(PeId2,_,_,_,_,_,_,'B').
```

An example image that is the result of a query of this type is shown in Fig. 3.

Fig. 3. Result of query "Find woman between member of PO and PSL party".

Thanks to its flexibility, Prolog allows for defining new relations as rules defined in a form of predicates. For example a rule that find a women that stands on the left-hand side from a member of a given party may be defined as:

```
womanOnLeftPersonFromParty(PhotoName, Party) :-
    person(ID1,_,_,_,_,'K',_,_,_), position(PhotoId,ID1,_,_,_,_,_),
    onLeft(ID1 ID2,PhotoId), personInGroup(GROUPID,ID2,_,_),
    groupName(_,GROUPID,Party,_), photo(PhotoId,PhotoName).
```

7 Conclusions

In the paper, a method has been proposed, that allows for combining various information sources – text and pictorial – to perform an image search that makes use of the external information that is not originally present within the images, but is

available in the external relational database. The combination of such information is possible thanks to the use of Prolog predicates. Three sets of predicates are defined: data predicates containing the information from external database, image predicated that describe the image content and spatial predicates that contain rules describing spatial relations between image objects. As an example, the political domain was chosen, where external database where the database on Polish politicians and the image database of their photographs. Search for image content enriched with non-graphical data brings valuable results. Moreover, the Prolog inference process allows for discovering new knowledge. Such a system is much more useful than standard approach where only annotated image content is used. Paper shows that combining different sources of data by means of ontology modeling using Prolog predicates improves the usability of the image database search system. As a future work, NLP (Natural Language Processing) should be considered to create automatic translation from user question to Prolog queries.

References

1. Batchelor, B.G.: Intelligent Image Processing in Prolog. Springer, New York (1991)
2. Challam, V., Gauch, S., Chandramouli, A.: Contextual search using ontology-based user profiles. In: Large Scale Semantic Access to Content (Text, Image, Video, and Sound), pp. 612–617. Le Centre de Hautes Etudes Internationale D'Informatique Documentaire (2007)
3. Dasiopoulou, S., Heinecke, J., Saathoff, C., Strintzis, M.G.: Multimedia reasoning with natural language support. In: International Conference on Semantic Computing, ICSC 2007, pp. 413–420. IEEE (2007)
4. Gupta, P., Sharma, D.A.: Context based indexing in search engines using ontology. Int. J. Comput. Appl. **1**(14), 49–52 (2010). (0975–8887)
5. Krieger, H.U., Declerck, T.: TMO—the federated ontology of the trendminer project. In: LREC, pp. 4164–4171 (2014)
6. Lowe, D.G.: Object recognition from local scale-invariant features. In: The Proceedings of the Seventh IEEE International Conference on Computer Vision, vol. 2, pp. 1150–1157. IEEE (1999)
7. Wang, C., Jing, F., Zhang, L., Zhang, H.J.: Scalable search-based image annotation. Multimedia Syst. **14**(4), 205–220 (2008)
8. Wilson, P.I., Fernandez, J.: Facial feature detection using Haar classifiers. J. Comput. Sci. Coll. **21**(4), 127–133 (2006)
9. Zhang, D., Islam, M.M., Lu, G.: A review on automatic image annotation techniques. Pattern Recogn. **45**(1), 346–362 (2012)
10. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: a literature survey. ACM Comput. Surv. (CSUR) **35**(4), 399–458 (2003)

Linguistic Description of Images Based on Fuzzy Histograms

Krzysztof Wiaderek[1(✉)] and Danuta Rutkowska[1,2]

[1] Institute of Computer and Information Sciences,
Czestochowa University of Technology, 42-201 Czestochowa, Poland
{krzysztof.wiaderek,danuta.rutkowska}@icis.pcz.pl
[2] Information Technology Institute, University of Social Sciences,
90-113 Lodz, Poland

Abstract. The paper presents a new approach for generating linguistic description of color digital images based on fuzzy histograms. The CIE chromaticity diagram is classified into fuzzy color granules and employed in order to recognize color clusters in an image or collection of images. A fuzzy inference process uses the histograms and fuzzy IF-THEN rules concerning color, location, size, and shape of the color clusters. The fuzzy histograms illustrate participation rate of pixels in the fuzzy color granules, and also within the so-called macropixels of different sizes of the same color. The histograms are derived from the matrix that expresses membership grades of pixels to the fuzzy color granules of the CIE diagram. The linguistic description is applied in image recognition and image retrieval tasks.

Keywords: Color image retrieval · CIE chromaticity diagram · Linguistic description · Fuzzy sets · Information granulation

1 Introduction

Linguistic description of color images are very useful in various applications, like image recognition and retrieval tasks, especially in the so-called intelligent systems. From artificial intelligence point of view, the CIE chromaticity diagram (see Sect. 2) is a suitable color model because it represents human perception of colors that may be employed in an intelligent image recognition system.

Fuzzy sets [14] and fuzzy granulation approach [16] are applied to the CIE chromaticity diagram, and then used in the process of fuzzy classification of colors in digital images (see Sect. 3).

Based on the classification result, a fuzzy histogram that portrays participation rates of image pixels in the fuzzy color granules of the CIE diagram is created. In addition, fuzzy histograms representing participation of pixels of the same color granule in particular fuzzy locations of the image (determined by the so-called macropixels) are generated. This is presented in Sect. 4.

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*,
Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_4

The fuzzy histograms are employed in order to produce linguistic description of an image (or images), by use of fuzzy IF-THEN rules. The inference process is explained in Sect. 5.

Examples of the linguistic description for a color digital image are presented in Sect. 6. Moreover, it is explained how such a description can be extended for a collection of images.

Section 7 includes conclusions and final remarks. Directions of further research are also outlined. This paper is a continuation of authors' previous publications [8–13].

2 CIE Chromaticity Diagram and Fuzzy Color Granules

The RGB color space is commonly applied in digital image processing. However, the RGB have some limitations – it is hardware-oriented and non-intuitive for people (they can easily learn how to use the RGB but thinking of hue, saturation and lightness seems to be more useful in practice). Transformations from the RGB color space to the XYZ and xyY, then to the CIE chromaticity diagram, are possible according to mathematical equations presented in many publications (see e.g. [2]). The CIE that stands for Commission Internationale de l'Eclairage (International Commission on Illumination) has introduced (in 1931) the chromaticity diagram based on experiments concerning human perception of colors. In addition, within the CIE diagram, color areas have been distinguished and labeled (Fig. 1).

The color regions of the CIE diagram can be viewed as fuzzy sets (see [1,14]), characterized by membership functions introduced in [8], and employed in [9–13], in order to generate linguistic descriptions of color digital images, in various aspects of image retrieval. An example of the membership functions for one of the color areas (12 - pink) is portrayed in Fig. 2. The fuzzy color regions are treated as fuzzy granules [16], so the CIE diagram has been granulated into color granules with fuzzy boundaries between them. Thus, the granulation approach for image retrieval [4], based on linguistic descriptions have been applied in the authors' publications [9–13].

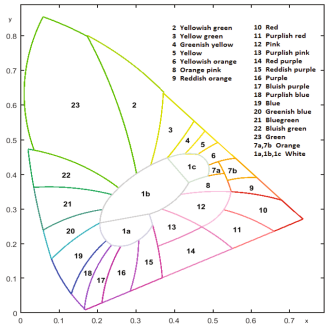


Fig. 1. CIE chromaticity diagram

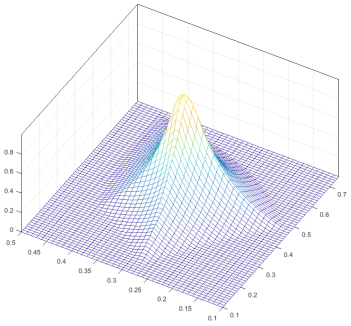


Fig. 2. Membership function of color area 12

3 Fuzzy Classification of a Digital Color Image

By use of the fuzzy granulation of the CIE diagram (see Sect. 2), a color digital image is classified into fuzzy color granules corresponding to the fuzzy regions presented in Fig. 1. The fuzzy boundaries play a role of discriminant functions that partition the space into fuzzy classes – color granules.

Figure 3 illustrates a result of the classification of the color image shown as the second last in this visualizations, as well as in Fig. 6 where it is divided into 9 squares called macropixels (introduced in [9]).

The particular visualizations portrayed in Fig. 3 correspond to the color granules of the CIE diagram. As we see in Fig. 1, the color granules are numbered and labeled by the color names – the same as the labels associated with the visualizations of the classified pixels. Each of the visualizations represent the matrix that expresses membership grades of pixels to the fuzzy color granules of the CIE diagram. Details concerning the matrix of the color granules are presented in [13].

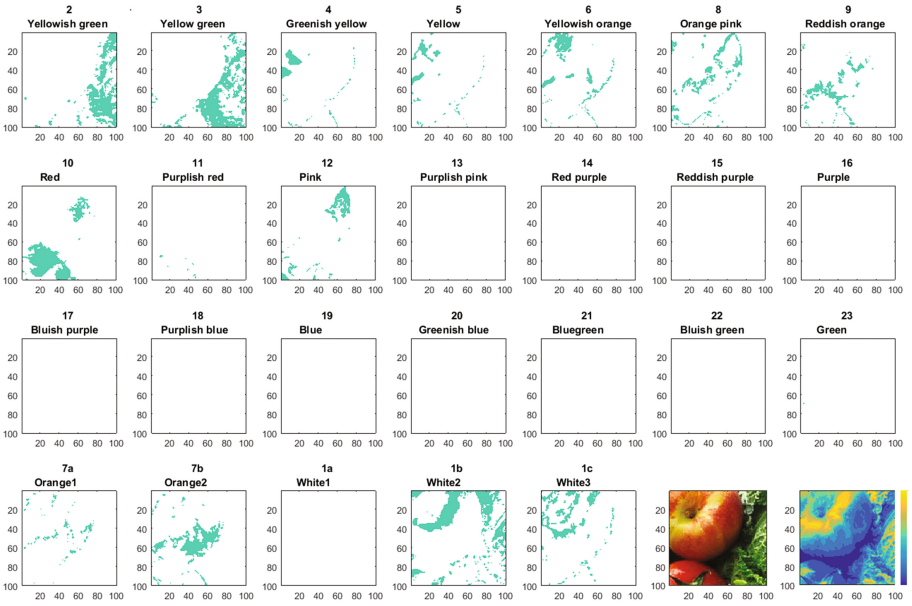


Fig. 3. Result of classification of the input image into color granules of the CIE diagram

The visualizations of the pixels of particular color in Fig. 3, according to the CIE diagram, concern pure colors (hue) without the luminance. However, it is possible to determine the luminance in the tree-dimensional CIE color space (xyY) where the x, y are two chromaticity coordinates that specify hue and saturation, and Y represents the luminance. In this way, the luminance for the image

under consideration (the second last picture in this figure) has been obtained. Although the hue with saturation is sufficient for many cases of image analysis, the luminance is important because it influences on human perception of colors. Therefore, the luminance is depicted in the last illustration.

4 Fuzzy Histograms for Linguistic Descriptions

Based on the classification result, portrayed in Fig. 3, fuzzy histograms are created. At first, participation rates of image pixels in the fuzzy color granules of the CIE diagram is considered. An example of such a histogram, for the image analyzed in Fig. 3, is presented in Fig. 4. The numbers and labels associated with the color granules of the CIE diagram, as well as the visualizations of the classified pixels in Fig. 3, correspond to the horizontal axis (see the legend). This histogram portrays participation rates of image pixels in the fuzzy color granules of the CIE diagram. The horizontal line indicates the value 0.04 (approximately) that is the average value of the participation of every color granule in the image.

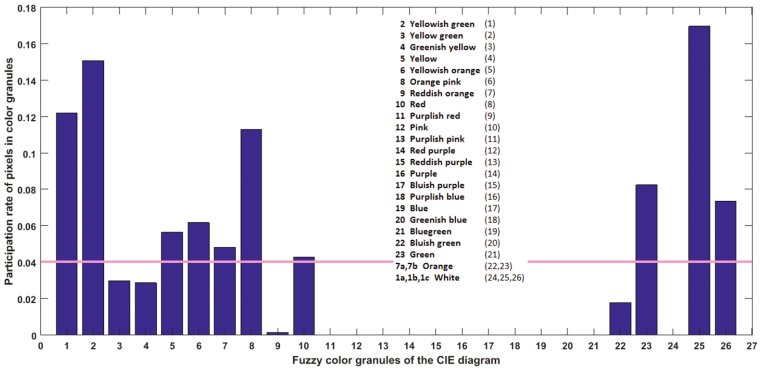


Fig. 4. Participation of the input image pixels into color granules of the CIE diagram

Let us introduce a fuzzy number (see e.g. [1]), denoted by P, representing the average value (denoted by p) as a unit of measure of participation of a color granule in an image. In Fig. 4 this unit corresponds to the participation rate indicated by the horizontal line. The fuzzy number is defined by a membership function that takes value 1 for p. The unit (*Apoc*) is applied in Sect. 5.

Apart from the histogram illustrated in Fig. 4, other fuzzy histograms – representing participation of pixels of the same color granule in particular fuzzy locations of the image – are generated. The fuzzy locations refer to the macropixels, introduced in [9] and developed in [10]. The two-dimensional fuzzy histograms are portrayed in Fig. 5, as well as in Fig. 7, with regards to the image shown in Fig. 6; see the difference in image partition into macropixels of different size.

Figure 5 concerns the color granule **Yellowish green** that is the first one in the histogram presented in Fig. 4. The macropixels of size equal to 1/9 of

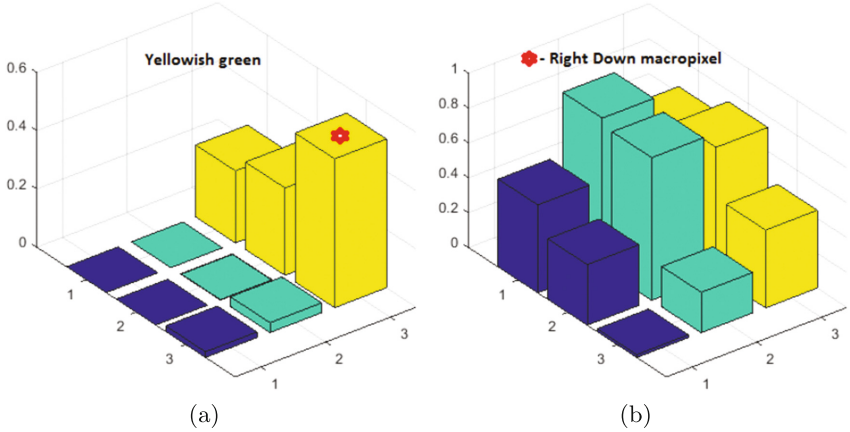


Fig. 5. (a) Color 2 in macropixels (b) Color 2 in the right down macropixel

the image (see Fig. 6) determine the locations of the image pixels of particular participation rate. Figure 5(a) directly corresponds to Fig. 6 with regard to the macropixels. As we see, the color **Yellowish green** is mostly visible at the right side of the picture (Righ Upper, Right Central, and Right Down macropixels), and also (but much less) in the Left Down and Middle Down macropixels. The maximal participation of this color is in the Right Down macropixel (the corner marked by the red star) in Fig. 5(a); see also Fig. 3.

Each of the macropixels portrayed in Figs. 5(a) and 6, are divided into smaller 9 macropixels. This is illustrated in Fig. 7 where the part representing the Right Down macropixel in Fig. 5(a) is also marked by the red stars associated with the smaller macropixels. The same participation rate within the Right Down macropixel, divided into 9 smaller ones, is shown in Fig. 5(b).

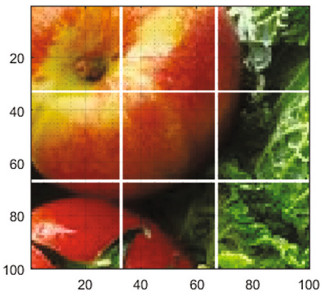


Fig. 6. An input image

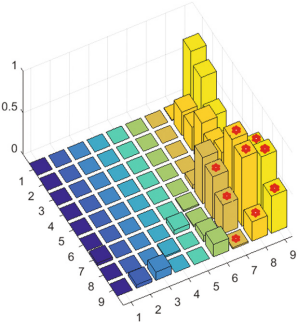


Fig. 7. Macropixels of the input image

5 Fuzzy Inference Based on the Histograms

The histograms illustrated in Sect. 4 are employed in order to generate linguistic description of the image (shown in Fig. 6) presented in Sect. 6. The linguistic description is produced by fuzzy inference and fuzzy IF-THEN rules (see e.g. [1,5,15]).

There are two reasons that the histogram portrayed in Fig. 4 is viewed as a fuzzy histogram. First, the horizontal axis expresses numbers corresponding to the fuzzy color granules. Second, we introduce the fuzzy unit (fuzzy number), called the "Average participation of color" (*Apoc*), denoted by P in Sect. 4. Thus, the vertical axis of this histogram, granulated by use of the *Apoc* unit, employs the fuzzy numbers as values $P, 2P, ..., 26P$. The maximal value of the participation rate (equal 1 in this histogram) is interpreted as 25 times P; this refers to the case when every pixel in the image belongs to the same color granule. The triangular membership functions of the fuzzy numbers have been applied; the cores (points with the membership values 1) are equal to $p, 2p, ..., 26p$, respectively.

The fuzzy IF-THEN rules, employed in the inference process, include linguistic variables corresponding to color, location, and size attributes, associated with fuzzy sets. With regard to the color, the fuzzy sets are color granules of the CIE diagram. Concerning location and size of a particular color granule in an image, fuzzy macropixels are considered with membership functions defined as shown in Fig. 8. There are trapezoidal membership functions where VS, S, M, B, VB denote *Very Small*, *Small*, *Medium*, *Big*, *Very Big* linguistic values of participation of colors in the image, respectively. In this case, the participation rate is considered focusing the attention on the areas of 1/9 of whole image. This refers to the fuzzy macropixels shown in Fig. 6.

Similar fuzzy sets as presented in Fig. 8 are employed when the image divided into smaller macropixels (see Fig. 7) is analyzed. This allows to infer more precise linguistic description by taking into account image granulation in greater detail.

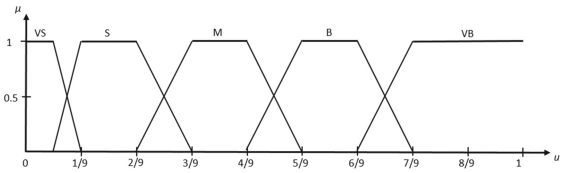


Fig. 8. Membership functions of fuzzy sets VS, S, M, B, VB

The fuzzy inference produces a decision, taking the value of color participation according to the histogram portrayed in Fig. 4, as the appropriate fuzzy number, and determining how this value (measured by the *Apoc*) matches the fuzzy IF-THEN rules.

6 Examples of Linguistic Descriptions

The fuzzy histograms presented in Figs. 4, 5, and 7 have been generated for the image shown in Fig. 6. For this example, linguistic descriptions are created by use of inference process outlined in Sect. 5.

Based on the histogram portrayed in the Fig. 4, a description concerning the participation of particular color granules in the image can be derived. The linguistic description includes, among others, the following sentences:

- No colors "Purplish pink","Red purple", "Reddish purple", "Purple", "Bluish purple", "Purplish blue", "Blue", "Greenish blue", "Bluegreen", "Bluish green", and "Green" in the image.
- The participation of "Yellowish green" and "Yellow green" is medium.

The next histograms (Figs. 5, and 7) allow to infer conclusions concerning location and size, in addition to the color, for example:

- A medium size granule of color "Yellowish green" and "Yellow green" is located in the right part.
- Small color granules "Red" and "Pink" is located in the *Middle Upper* part.

The second sequence in the first list has been inferred from the conclusion that (according to Fig. 4) the fuzzy set M (Medium) includes fuzzy number 7 *Apocs* (expressing the participation rate of both neighboring colors). The first sentence of the second list gives more information (about location). Analogously, the second sentence (in the second list) has been produced.

For a collection of images, fuzzy histograms allow to infer an example of the linguistic description, as follows:

- Most of the pictures contain "Yellowish green" and "Yellow green", "Red", and "Pink" colors.
- There are few pictures that contain medium size clusters of "Orange" color.
- There are not "Blue" color in the pictures.

This inference is realized by use of fuzzy IF-THEN rules and the operation of union of fuzzy sets (fuzzy color granules).

Apart from the color, location, and size attributes, also shapes of color granules in images can be considered, and determined by use of the fuzzy histograms of the type presented in Figs. 5 and 7. In this case, the linguistic descriptions include information concerning shapes.

7 Conclusions and Final Remarks

In this paper, a new fuzzy unit for measuring participation rate of a color in digital images is introduced, and applied in order to generate linguistic description of the image content, based on fuzzy histograms. Further research will concern shape recognition, with the luminance, as well as image understanding [6,7].

The linguistic description of a single image or a collection of images is useful in many image recognition tasks, especially in color image retrieval, see [9–13].

Although different color models are used (see e.g. [3]), it is worth emphasizing that the CIE diagram when applied in an intelligent image recognition system, is very suitable because it imitates the way of human perception of colors.

References

1. Dubois, D., Prade, H.: Fuzzy Sets and Systems: Theory and Applications. Academic Press, New York (1980)
2. Fortner, B., Meyer, T.E.: Number by Color: A Guide to Using Color to Undersdand Technical Data. Springer, New York (1997)
3. Luszczkiewicz-Piatek, M.: Image similarity in Gaussian mixture model based image retrieval. In: Choraś (ed.) Image Processing and Communication Challenges 8. Advances in Intelligent Systems and Computing, vol. 525, pp. 87–95. Springer, Cham (2016)
4. Pedrycz, W., Vukovich, G.: Granular computing in pattern recognition. In: Bunke H., Kandel A. (eds.) Neuro-Fuzzy Pattern Recofnition. World Scientific, pp. 125–143 (2000)
5. Rutkowska, D.: Neuro-Fuzzy Architectures and Hybrid Learning. Springer, Heidelberg (2002)
6. Tadeusiewicz, R., Ogiela, M.R.: Semantic content of the images. In: Image Processing & Communications Challenges, pp. 15–29. Academic Publishing House EXIT, Warsaw (2009)
7. Wei, H.: A bio-inspired integration method for object semantic representation. J. Artif. Intell. Soft Comput. Res. **6**(3), 137–154 (2016)
8. Wiaderek, K.: Fuzzy sets in colour image processing based on the CIE chromaticity triangle. In: Rutkowska, D., Cader, A., Przybyszewski, K. (eds.) Selected Topics in Computer Science Applications, pp. 3–26. Academic Publishing House EXIT, Warsaw (2011)
9. Wiaderek, K., Rutkowska, D.: Fuzzy granulation approach to color digital picture recognition. In: Artificial Intelligence and Soft Computing. LNAI, Part I, vol. 7894, pp. 412–425. Springer, Cham (2013)
10. Wiaderek, K., Rutkowska, D., Rakus-Andersson, E.: Color digital picture recognition based on fuzzy granulation approach. In: Artificial Intelligence and Soft Computing. LNAI, Part I, vol. 8467, pp. 319–332. Springer, Cham (2014)
11. Wiaderek, K., Rutkowska, D., Rakus-Andersson, E.: Information granules in application to image recognition. In: Artificial Intelligence and Soft Computing. LNAI, Part I, vol. 9119, pp. 649–659. Springer, Cham (2015)
12. Wiaderek, K., Rutkowska, D., Rakus-Andersson, E.: New algorithms for a granular image recognition system. In: Artificial Intelligence and Soft Computing. LNAI, Part II, vol. 9693, pp. 755–766. Springer, Cham (2016)
13. Wiaderek, K., Rutkowska, D., Rakus-Andersson, E.: Linguistic description of color images generated by a granular recognition system. In: Artificial Intelligence and Soft Computing. LNAI, Part I, vol. 10245, pp. 603–615. Springer, Cham (2017)
14. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)
15. Zadeh, L.A.: Fuzzy logic = computing with words. IEEE Trans. Fuzzy Syst. **4**, 103–111 (1996)
16. Zadeh, L.A.: Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets Syst. **90**, 111–127 (1997)

Using Toboggan Segmentation in Detection of Centers and Radius of Cell Nuclei

Przemysław Jacewicz(✉) and Józef Korbicz

Institute of Control and Computation Engineering, University of Zielona Góra, Zielona Góra, Poland
{P.Jacewicz,J.Korbicz}@issi.uz.zgora.pl

Abstract. The improvement of computer image analysis techniques in recent years can support the pathologist's work by the automation of nuclei segmentation, cell population count, computing statistics of morphological features or cell classification. However, due to the complexity of the image representing the cytological preparation this process does not belong to the trivial. This paper presents a proposition of detecting coordinates of centers and radius of nuclei in cytological images of biological material obtained from breast using fine needle biopsy (FNB). The solution is based on four steps - reduction of image dynamics to local minima, thresholding, calculation of potential centers and radii then selection of most probables of them.

Keywords: Nuclei segmentation · Toboggan segmentation · Cellular automata · Circle detection · Cytology

1 Introduction

Recent advances in studies and diagnoses diseases by analysis of microscopic images field have revolutionized pathology by the introduction of digital microscopy. Nowadays, more and more pathologists examine the biological material on a computer screen instead of reviewing it under a microscope. But still, they must spend many years to gain experience to become specialist in the identification of cancer cells. Fortunately along with the development of advanced vision systems and computer science, quantitative cytopathology can support the pathologist's work by the automation of nuclei segmentation, cell population count, computing statistics of morphological features or cell classification. But to accomplish these tasks we need accurate algorithm for cell or nucleus segmentation. If we look at cytological specimen under a microscope, we will see a lot of clumps of nuclei which create complex, random and heterogeneous structures without clear boundaries. Unfortunately, their segmentation is rather a challenging task.

Many publications present the possibility of using watershed methods, graph cuts, level sets, image thresholding and data clustering based methods for nuclei segmentation [3,5–7,9]. However, the problem is still open because there are

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*, Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_5

no universal segmentation methods that can be applied for images representing biological material from different tissues.

One of the most common cancers occurring in the female population worldwide is breast carcinoma. Transformation process from normal cells to the cancer cells forms through several stages. For objectively assess nuclear changes occurring during these transformational steps the most commonly approach is morphometry as a diagnostic tool. One of the most promising morphometric informations which can be used for classification is the nuclear area and diameter and its distribution in scanned plane [8].

This paper presents a proposition of detecting coordinates of centers and radius of nuclei in cytological images of biological material obtained from breast using fine needle biopsy (FNB). The solution is based on four steps - reduction of image dynamics to local minima, thresholding, calculation of potential centers and radii then selection of most probables of them.

The remainder of this paper is organized in three parts. In Sect. 2 material and motivation are described. Section 3 presents each step of proposed approach illustrated by real example. Then concluding remarks are given in Sect. 4.

2 Study Dataset and Motivation

The first step of the diagnostic process begins with the acquisition of biological material from the breast. In this case the cytological material was obtained by FNB from 50 patients of the Regional Hospital in Zielona Góra, Poland. The set contains 25 benign and 25 malignant lesions cases. Smears from the biological material were fixed in spray fixative and dyed with hematoxylin and eosin. Cytological preparations were then digitised into virtual slides using the Olympus VS120 Virtual Microscopy System. Virtual slides offer a radically improved level of detail in comparison to the images grabbed with analog camera. One can clearly observe clumps of chromatin (especially in malignant cases), nucleoli, and nuclear membranes. Each pixel in digital slide represent square of $0.17\,\mu\text{m}\times0.17\,\mu\text{m}$ on real preparation.

3 Segmentation of Cellular Nuclei by Means of Circles

The high quality material presented above causes problems in its segmentation. Due to the a lot of details in the content presented on the slides, most of commonly used approaches have problem with segmentation of some general features like edges of cells nuclei. For instance pattern representing components of the cell nuclei impeding to segment its centers and circle or ellipsis describing edge. To remedy this problem it is necessary to apply some reductions in image details – further, solution basing on a local minima will be presented.

3.1 Reduction of Image Dynamics to Local Minima

Tobogganing, introduced by Fairfield [6] and extended by Yao and Hung, over-segments an image into small regions by sliding in the derivative terrain. The basic idea is that given the gradient magnitude of an image, each pixel determines a slide direction by finding the pixel in a neighborhood with the lowest gradient magnitude. Obviously, several pixels will represent local minima in the gradient terrain (due to either image features or noise) in that they represent the minimum gradient magnitude within their own neighborhood. Pixels that "slide" to the same local minimum are grouped together, thus segmenting the image into a collection of small regions. The regions produced by tobogganing are effectively identical to the catchment basins produced by applying the popular watershed algorithm to the gradient image [2]. However, tobogganing is much more computationally efficient than the watershed algorithm.

The whole procedure starts from designation of local gradients $G(p)$ of luminance in each point p of original image I. This gradient is calculated based on luminance values in adjacent points. Those points where the value of adjacent points is greater than the central point is considered as a local minimum.

In classical toboggan approaches next we should segment the image into regions by sliding in the gradient terrain. Pixels that slide into the same local minimum are efficiently grouped into regions by assigning them a unique label. In the proposed case this step is superfluous – it turns out that the points of local minima alone are sufficient.

To determine the local minima by this manner a perfect way seems to be the utilisation of the Cellular Automata (CA) [1,10]. The recommended CA consists of a regular lattice, in which each cell is characterized by a particular state taken in a discrete finite set of admissible states. The dynamics of each cell is defined by a set of transition rules which specify the future state of this cell as a function of its previous state and the states of neighbouring cells. Thus, we can define a *cellular automaton* as the quadruple $\mathcal{A} = (\mathcal{L}, \mathcal{S}, N, f)$, where:

- $\mathcal{L}$ is a regular lattice of size corresponding to image I which consists of a mirroring boundary conditions;
- $\mathcal{S}$ is a discrete set of admissible state values; in this model, the state is characterized by the two element vector with the following components:
  - Luminance $L \in (0, 255)$: luminance of image I in pixel corresponding to CA cell;
  - Local minimum M: an boolean label which identifies designated local minima.
- The neighborhood N of size n is defined by the mapping

$$\begin{aligned} N : \mathcal{L} &\longrightarrow \mathcal{L}^n \\ c &\longrightarrow N(c) = \{c_1, c_2, \ldots, c_n\} \end{aligned} \tag{1}$$

  For this case a two most common neighbor definition – von Neumann ($n = 4$) and Moore ($n = 8$) were used.

– f is a function which defines the transition rule defined by

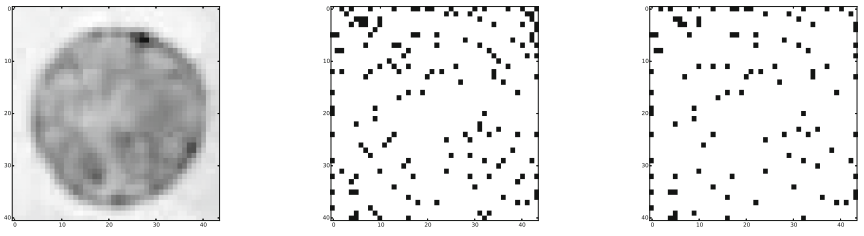
$$\begin{array}{rccc} f: & \mathcal{S}^n & \longrightarrow & \mathcal{S} \\ & s_t(N(c)) & \longrightarrow & s_{t+1}(c) \end{array} \tag{2}$$

where $s_{t+1}(c)$ is the state of the cell c at time $t+1$.
In proposed approach the following transition function were used:

$$M_{t+1}(c) = \begin{cases} true & \text{if } min(L(N(c))) > L(c) \\ false & \text{elsewhere} \end{cases} \tag{3}$$

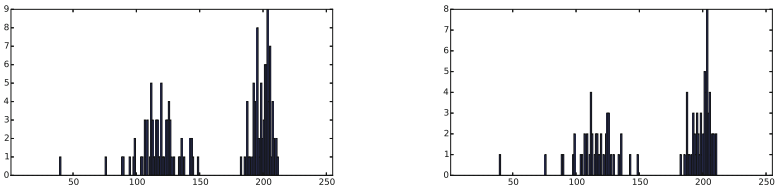
To determine local minima by the above described CA enough to evaluate only one iteration. Coordinates of local minima evaluated by this manner are shown on Fig. 1(b) for von Neumann and on Fig. 1(c) for Moore neighborhood definition – black point represent their locations.

Reduction source image to points representing local minima does not allow yet possibility of easy segmentation. To remedy this we need to apply one of thresholding methods.



(a) Input image (b) Local minima for von Neumann neighbourhood (c) Local minima for Moore neighbourhood

Fig. 1. Input image, local minima evaluated by using von Neumann and Moore neighborhood for cellular automata.



(a) von Neumann neighbourhood (b) Moore neighbourhood

Fig. 2. Histograms of luminance in local minima evaluated by CA toboggan algorithm.

3.2 Thresholding

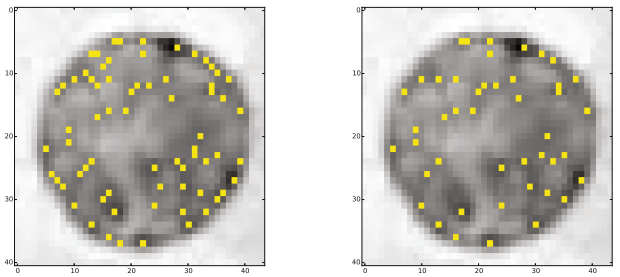
As shown on Fig. 1 locations representing local minima are insufficient for automated segmentation. Fortunately histograms of luminance in this points (see Fig. 2) shows clear two clusters representing entire cell nucleus.

For such a favorable luminance distribution, practically any of the algorithms used, should find proper thresholding value. For that case the Otsu's method were used to obtain threshold value equal 149.47 – it means that values lower than this level represents local minima inside cell nucleus but the rest will be removed.

The results of extraction of local minima representing cell nucleus area are shown in the Fig. 3.

3.3 Segmentation of Cell Nuclei by Circles

Operation of evaluation by CA toboggan algorithm only local minima representing cell nucleus reduces large amount of data do simple binary image or list of this minima coordinates. Now is the right time for segmentation of circular objects from binary image – we hope to achieve centers corresponding to location of cell nucleus and radius representing their size.



(a) von Neumann neighbourhood (b) Moore neighbourhood

Fig. 3. Local minima representing cell nucleus extracted by thresholding.

In the case described so far, showing the only one nucleus of a circular shape, it is easy to determine the center and the radius, but in the more realistic case, the local minima inside the nuclei disrupt the segmentation process by showing a lot of circles that does not exist. To solve this problem, the process of determining the circles is decided in two steps. The first is to designate all potential centers coordinates of circles passing through each combination of three points from the set of minima obtained above. Then in the next step to determine the most

probable center and radius by the kernel density estimation (KDE) algorithm [4] instead of voting in circular Hough transform algorithm.

Comparison of circles given by this approach with values returned by circles Hough's transform are set together on the table bellow:

	Neighbourhood type					
	von Neumann			Moore		
	x_c	y_c	r	x_c	y_c	r
Hough's transform	22.0	20.0	16.0	18.0	22.0	11.0
Circle eq. + KDE	21.5	22.4	17.1	19.8	22.2	17.3

However on Fig. 4 we can observe more details helpful to evaluate the proposed approach. In first row there are three images representing a circle determined by circular Hough approach on Fig. 4(a), set of potential circles centers given from circle equation for each combination of three local minima on Fig. 4(b) and a circle enumerated by using KDE algorithm on Fig. 4(c) – all for local minima calculated by using von Neumann neighborhood in CA. In a similar way the same cases for smaller collection of minima calculated by using Morre neighborhood definition in CA are shown on bottom row – Figs. 4(d), (e) and (f).

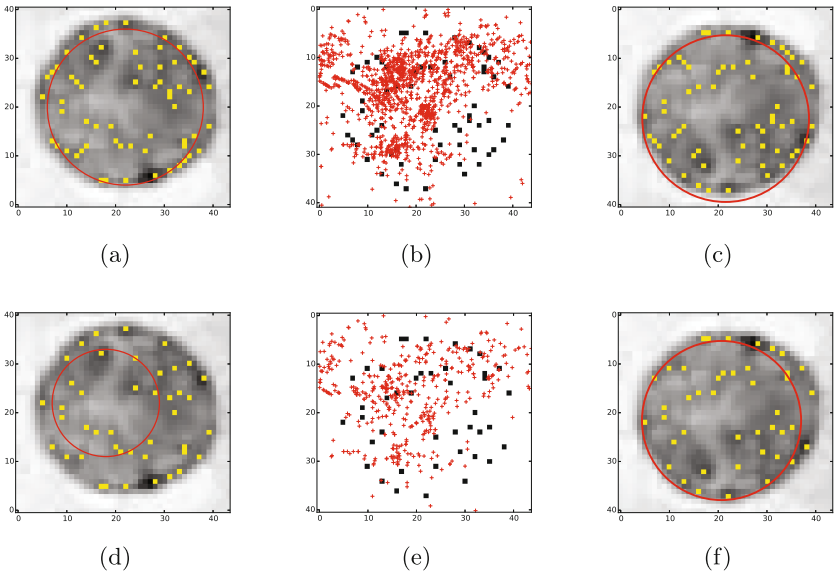


Fig. 4. Graphical representation of circle detection from local minima given form von Neumann neigh. (On top row) and from Moore neigh. (On bottom row) by using Hough's circle detection in (left column) and KDE approach (right column) executed on centers presented in central column.

4 Conclusions

Content of cytological images is highly complex and its analysis is difficult in an automated way. The crucial step of this analysis is nuclei segmentation.

Due to the a lot of details in the content presented on the modern digital slides, most of commonly used approaches have problem with segmentation of some general features like edges of cells nuclei. For instance pattern representing components of the cell nuclei impeding to segment its centers and circle or ellipsis describing edge. Thus, we can not build model able to classify cancerous nuclei. To remedy this problem we can combine different image processing algorithms [2].

The work described in this paper intends to achieve two goals: calculate centers and radius of circles possibly best describing the nuclei of cells and do this automatically without parameters. The task was done by reduce amount of information by calculating local minima of luminance using toboggan algorithm on cellular automaton. Then for select only points representing interior of cell nuclei the Otsu's method was used. From this moment all work was done on binary array representing local minima and the background. Next step is to use circular Hough transform to determine sought centers and radii of circles for each cell on preparation. Unfortunately Hough transform even in the simple case presented before, have problem with giving correctly results – see Fig. 4(d). To remedy this problem approach basing on modified circular Hough transform was given. On the beginning all circles going by each combination of threesome points on minima array were calculated. Then in contrast to Hough transform kernel density estimation (KDE) is used to determine the most probable circle on each cell.

Proposed approach seems to be more resistant to significant reduction of information and works automatically with only two parameters determining minimal and maximal nuclei radius which are constant for whole digital slide.

The further studies in this area will be necessary to test the effectiveness of the method on full slides of 138 717×79 267 px to automatically calculate circles describing all cells nuclei positions and size.

Acknowledgments. The research was supported by National Science Centre, Poland (2015/17/B/ST7/03704).

References

1. Chopard, B., Droz, M.: Cellular Automata Modeling of Physical Systems. Collection Aléa-Saclay: Monographs and Texts in Statistical Physics, Cambridge (1998)
2. Choraś, R., Andrysiak, T., Choraś, M.: Integrated color, texture and shape information for content-based image retrieval, pattern analysis and applications. Pattern Anal. Appl. **10**(4), 333–343 (2007)
3. Filipczuk, P., Kowal, M., Obuchowicz, A.: Multi-label fast marching and seeded watershed segmentation methods for diagnosis of breast cancer cytology. In: Proceedings of 35th International Conference IEEE Engineering in Medicine and Biology Society, EMBC 2013. Oska, Japan (2013)

4. Friedman, J., Hastie, T., Tibshirani, R.: The Elements of Statistical Learning. Springer Series in Statistics. Springer, New York (2008)
5. Jeleń, L., Fevens, T., Krzyżak, A.: Classification of breast cancer malignancy using cytological images of fine needle aspiration biopsies. Int. J. Appl. Math. Comput. Sci. **18**(1), 75–83 (2010)
6. Kowal, M.: Computer-aided diagnosis for breast tumor classification using microscopic images of fine needle biopsy. In: Korbicz, J., Kowal, M. (eds.) Intelligent Systems in Technical and Medical Diagnostics, Advances in Intelligent Systems Computing, vol. 230, pp. 213–224. Springer, Heidelberg (2013)
7. Kowal, M., Filipczuk, P.: Nuclei segmentation for computer-aided diagnosis of breast cancer. Int. J. Appl. Math. Comput. Sci. **24**(1), 19–31 (2014)
8. Narasimha, A., Vasavi, B., Kumar, H.: Significance of nuclear morphometry in benign and malignant breast aspirates. Int. J. Appl. Basic Med. Res. **3**(1), 22–26 (2013). http://www.ijabmr.org/article.asp?issn=2229-516X;year=2013;volume=3;issue=1;spage=22;epage=26;aulast=Narasimha;t=6
9. Plissiti, M., Nikou, C., Charchanti, A.: Combining shape, texture and intensity features for cell nuclei extraction in pap smear images. Pattern Recogn. Lett. **32**(6), 838–853 (2011)
10. Wolfram, S.: Cellular Automata and Complexity: Collected Papers. Addison-Wesley, New York (1994)

Evaluation of the Pre-processing Methods in Image-Based Palmprint Biometrics

Agata Wojciechowska(✉), Michał Choraś, and Rafał Kozik

Faculty of Telecommunications, Computer Science and Electrical Engineering, University of Science and Technology, Bydgoszcz, Poland
{agata.wojciechowska,michal.choras,rafal.kozik}@utp.edu.pl

Abstract. In this paper we present the evaluation of several image pre-processing methods for palmprint biometrics. Pre-processing is a crucial step in mobile scenario dedicated to verify palmprints using images acquired by smartphones or other obile devices.

1 Introduction

Verification of people has become the interesting challenge in everyday life. The verification can be performed in mobile phones, before using ATM, at the entrance of a sport club or while getting access to the e-mail account. Such situation makes the society overloaded with logins, passwords and PIN codes. Thus, the biometric approach has recently become increasingly popular. In this case remembering passwords is not necessary, because the user is verified by physiological characteristics such as face, ear, hand, fingers, iris or by his behavior such as voice, gait or typing. Palmprints are one of physiological biometric modalities, where the verified user has to put a specific part of his body in front of the acquiring device and wait for the decision on granting or declining the access to certain resources. The following characteristics make palmprints a perfect trait for the user verification process [1]:

- palmprints are unique and they vary even for twins;
- palmprints remain unchanged while the time is passing (non-aging);
- no special device is needed for acquiring a sample;
- due to the specific structure (principal lines, ridges and wrinkles), a high resolution of images is not needed.

However, the accuracy of the palmprint based verification is still not sufficiently high even though there are plenty of papers with promising results [4,8]. It is possible to improve each step of the verification process:

- pre-processing - in this step the ROI is found and the image enhancement is performed;
- features extraction - using various algorithms the features are extracted from the image, such as color, texture or hybrid features;
- matching - in this step extracted feature vectors are matched against feature vectors (templates) from the database.

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*, Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_6

In this article we focus on evaluating several pre-processing methods in order to improve this step in the verification process. Since our goal is to construct the system for mobile biometrics based on palmprints acquired by smartphones [3], this step is very important for us. In this paper we focus on the improvements in the pre-processing step. Different methods were compared in terms of the whole system accuracy and time of computing. The paper is organized as follows: in Sect. 2 various pre-processing methods are described. Sections 2.1 and 2.2 contain the presentation of our work and the results, respectively. Conclusions are provided in afterwards.

2 Pre-processing Methods

In this section several pre-processing methods will be shortly overviewed.

2.1 Gaussian Blur

Gaussian blur is a low pass filter and it is implemented e.g. in [5]. It is calculated separately for each pixel in the image and uses the Eq. (1), where x, y are the distances to the original X and Y axis and σ is the standard deviation. Each pixel gets the value equal to the weighted average of its neighborhood.

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1}$$

The size of neighborhood (called kernel) may be modified. The bigger size is set, the more blurred the image is. The results of the Gaussian blur performed on an im-age with different kernels are presented in Fig. 1. The filter is very useful, especially when the Gaussian noise is present in the image.

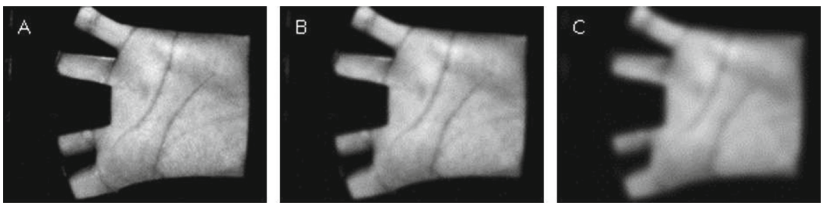


Fig. 1. Gaussian blur results: A - kernel 5×5, B - kernel 11×11 and C - kernel 29×29

2.2 Median Blur

Median blur is a filter which reduces effectively the impulsive noises (like salt and pepper noise). It is widely implemented in image processing systems e.g. in [7]. It is calculated for each pixel of the image and also depends on the chosen neighborhood size (kernel). Each pixel gets the value equal to the median value of pixels in the neighborhood. The images after the median blurring are presented in Fig. 2. Unfortunately, using this filter may affect the edges, which is visible in Fig. 2(C).

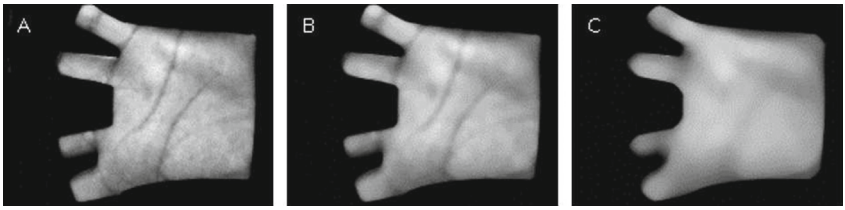


Fig. 2. Median blur results with different kernel sizes: A - size 5×5, B - size 11×11 and C - size 29×29

2.3 Bilateral Filter

Bilateral filter is the Gaussian blur modification and was implemented for instance in [6]. It uses weighted average as well, but introduces the second parameter, which modifies the Gaussian kernel shape. Each pixel from image gets the new value using Eq. (2), where W_p is the normalization factor expressed with Eq. (3), while σ_r and σ_s specify the amount of filtering for the image I, G_{σ_s} is the weight that decreases the influence of distant pixels, G_{σ_r} is the weight that decreases the influence of pixels q when their intensity values are different from I_p.

$$BF[I]_p = \frac{1}{W_p} \sum q \in SG_{\sigma_s}(||p - q||)G_{\sigma_r}(||I_p - I_q||)I_q \tag{2}$$

$$W_p = \sum q \in SG_{\sigma_s}(||p - q||)G_{\sigma_r}(||I_p - I_q||) \tag{3}$$

Although the time of the computing for bilateral filter is higher than for other filters, it preserves better edges, as it is presented in Fig. 3.

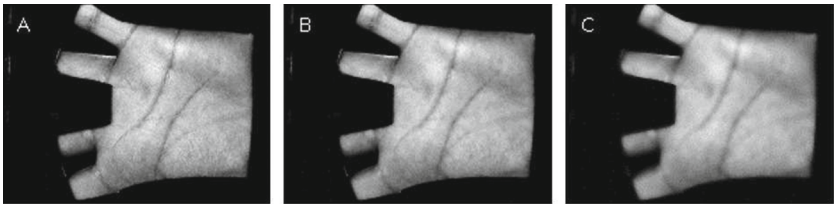


Fig. 3. Bilateral filter results with changing parameters σ_s and σ_r: A: $\sigma_s = 5$ and $\sigma_r = 5$, B: $\sigma_s = 25$ and $\sigma_r = 25$, C: $\sigma_s = 75$ and $\sigma_r = 75$

2.4 Sharpening

Sharpening may be implemented as subtracting the blurred image from the original one. For blurring it is possible to use one of above mentioned filters, but in the Fig. 4. there are presented results of sharpening based on Gaussian blur. It is visible, that the bigger is Gaussian kernel, the sharper the image is. Sharpening was evaluated by Zhang et al. in [9].

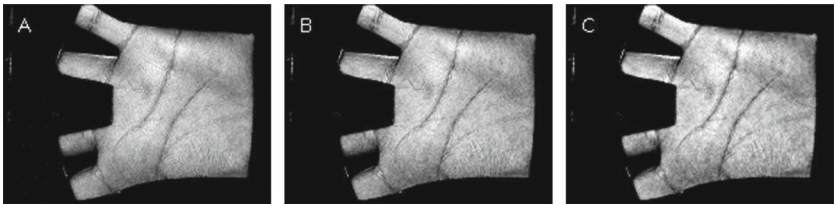


Fig. 4. Results of sharpening with different Gaussian kernel size: A - size 5×5, B - size 11×11 and C - size 29×29

3 Our Approach

The main aim of the reported research was to check, whether the change of the pre-processing method may affect the overall accuracy of the verification system and which commonly used method will be the most time-saving. To ensure stable conditions, the extraction and matching functions remained unchanged during all the tests, and those are Histogram of Oriented Gradients (HOG) for features extraction and Euclidean distance for matching.

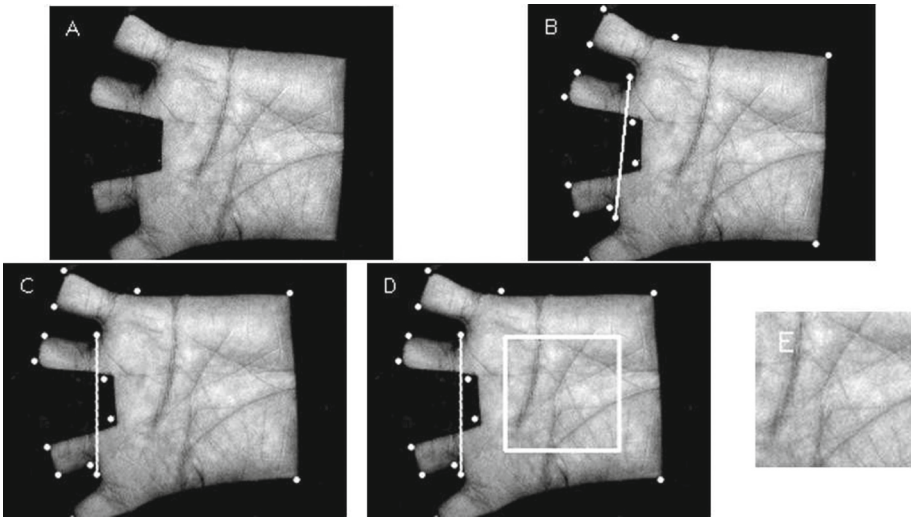


Fig. 5. ROI extraction process

The Region of Interests (ROI) setting function also remained without changes and is visualized in Fig. 5. First of all, the contours were found on the source image (A). Then the key points between fingers were extracted. Point A was set between the index finger and the middle finger, while point B between the ring finger and the little finger (B). Then, using the trigonometry function, the

angle between points was calculated and the whole image was rotated by this angle (C). The rotation was necessary due to the fact, that HOG is invariant to rotations. The last step was setting the ROI size to 128×128 (D, E). The PolyU database was used in this research [2]. It provides palmprints from 386 people, about 20 images for each person. The database is available online and is commonly used in scientific papers concerning about biometrics. For each user it was prepared a special set of samples: 3 patterns to compare against test samples, more than 10 positive test samples and 1000 negative test samples.

4 Results

All the below mentioned results were obtained on the personal computer (Intel Core i5-4210U 1.7 GHz CPU, 4-Core, 4 GB RAM, Windows 8.1-64 bit). The results of the accuracy obtained from the research are presented as confusion matrixes, which contains: True Positives (positive samples verified as positive), False Negatives (positive samples verified as negative), False Positives (negative samples verified as positive) and True Negatives (negative samples verified as negative). The perfect system has maximum TP and TN, while FP and FN should be close to zero. Apart from the accuracy, the computational time was assessed as well. First of all, the verification was performed without any special pre-processing method. Only the ROI was extracted from the sample using before mentioned algorithm. Then, popular pre-processing methods was implemented. Matrix of confusion for all methods is presented in Table 1.

Table 1. Matrix of confusion for methods

	True positives	False negatives	False positives	True negatives
Raw samples	43,3%	56,7%	24,9%	75,1%
Gaussian blur (3×3)	69,2%	30,8%	1,9%	98,1%
Gaussian blur (5×5)	78,5%	21,5%	0,5%	99,5%
Median blur (3×3)	79,2%	20,8%	3,3%	96,7%
Median blur (5×5)	79,3%	20,7%	0,7%	99,3%
Bilateral filter	63,9%	36,1%	32,5%	67,5%
Sharpening	68,3%	31,7%	12,0%	88,0%

Apart from evaluating the accuracy, the computational time was also assessed. Most methods took comparable time (Gaussian blur for both kernels, median blur for smaller size of kernel, sharpening). However, the median blur took 3.5 times longer, while the bilateral filter took 7.5 times longer.

5 Conclusions

In this paper we presented the comparison of commonly known methods of pre-processing. We implemented those methods in a verification system based on

HOG and Euclidean distance. From the experimental results we can note that the best accuracy is obtained, when the median blur (size 5×5) is implemented. However, the significantly longer computational time of median blur encourages to use the Gaussian blur (size 5×5). It is the most reasonable for this kind of system. Using this method it was possible to increase the True Positive ratio by more than 35% and the True Negative ratio by about 25%. Surprisingly, sharpening did not provide any promising results - running the verification with raw samples gave better results. In our future work we plan to use the best pre-processing methods in the mobile palmprint biometrics scenario.

References

1. Afsal, S., Jothykumar, J., Ahmed, S., Sayeed, F., et al.: A novel approach for palm print recognition using entropy information features. In: International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), pp. 1439–1442. IEEE (2016)
2. Biometrics Group: The Hong Kong Polytechnic University. http://www4.comp.polyu.edu.hk/biometrics
3. Choraś, M., Kozik, R.: Contactless palmprint and knuckle biometrics for mobile devices. Pattern Anal. Appl. **15**(1), 73–85 (2012)
4. Franzgrote, M., Borg, C., Ries, B.J.T., Bussemaker, S., Jiang, X., Fieleser, M., Zhang, L.: Palmprint verification on mobile phones using accelerated competitive code. In: 2011 International Conference on Hand-Based Biometrics (ICHB), pp. 1–6. IEEE (2011)
5. Harb, A., Abbas, M., Cherry, A., Jaber, H., Ayache, M.: Palm print recognition. In: 2015 International Conference on Advances in Biomedical Engineering (ICABME), pp. 13–16. IEEE (2015)
6. Kulkarni, S., Raut, R., Dakhole, P.: Vein pattern for personal authentication. Int. J. Electron. Commun. Soft Comput. Sci. Eng. (IJECSCSE) **4**, 1 (2015)
7. Patil, J.P., Nayak, C., Jain, M.: Palmprint recognition using DWT, DCT and PCA techniques. In: 2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), pp. 1–5. IEEE (2015)
8. Verma, S.B., Chandran, S.: Analysis of SIFT and SURF feature extraction in palmprint verification system. In: International Conference on Computing, Communication and Control Technology (IC4T), pp. 27–30. IEEE (2016)
9. Zhang, K., Huang, D., Zhang, D.: An optimized palmprint recognition approach based on image sharpness. Pattern Recogn. Lett. **85**, 65–71 (2017)

On the Way to Perfect Steganography

Artur Jakubski(✉) and Janusz Bobulski

Institute of Computer and Information Sciences, Czestochowa University of Technology, ul. J.H. Dabrowskiego 73, 42-201 Czestochowa, Poland
{artur.jakubski,januszb}@icis.pcz.pl

Abstract. The content of our paper presents the ways of hiding information with the usage of a digital pictures. A branch of science that deals with hiding messages in the wider media is called steganography. Encrypting files with sensitive data usually do not hide the existence and transmission of data. It seems therefore that a significant complement or alternative for cryptographic methods is the use of steganographic methods. Due to the rapid expansion of Internet and the associated increase in data exchange appears to be a subject of interest in this field. In the paper we present our own algorithms of hiding data in pictures, their implementations and analysis. In this paper we emphasize high quality, undetectability of the methods used. We define perfect steganography and we propose solutions to achieve this goal. The proposed solution can be used to transmit secret information and to find use in business, diplomacy, and other areas of human activity.

1 Introduction

The primary feature by which steganographic methods can be used is redundancy. In information theory, redundancy is the amount of information that exceeds the minimum required to solve a problem. Formally redundancy is referred to as the number of bits that can be omitted and is not practically noticeable. This paper focuses on the issue of redundancy in graphic files. The presented method uses the fact that some bits of the image (least significant) are imperceptible to the human eye. These bits are redundant bits [1].

The techniques used in steganography make it difficult to detect that there is a hidden message inside an innocent file. This way you not only hide the message itself, but also the fact that you are sending the message. This characteristic makes steganography the ideal science for hiding messages on the web, which is widely seen as a mass communication outlet [2].

This paper concerns steganography as a branch of science often related to cryptography. While cryptography hides the content of the information through its encryption, steganography tries to hide the fact of its occurrence. There is the following division of steganographic systems: Pure Steganography, Private Key Steganography, Public Key Steganography. Kerckhoffs' principle says that the cryptosystem should be secure even if all the details of its operation (besides the key) are known. Modification of the least significant bit is a classic representative

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*, Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_7

of methods to replace (substitution). The last bit (e.g. pixel component values) is replaced by a bit (or bits) from the message [1].

The science of discovering secret messages is called steganalysis. Steganography is considered broken even when the only thing we have is evidence that a certain document contains hidden data without actually being able to extract it. Another problem facing steganoanalysis is the extraction of hidden information. This is a much more difficult challenge than just finding out the fact of embedded information [3].

Steganography in Pictures. Steganography for digital pictures, as well as for other media, uses a carrier (in this case a picture) to forward a confidential message. Just as in other cases, in steganography it is important to prepare the carrier (media) that it does not arouse any suspicion.

In addition to normal hiding, digital images allow for the usage of specific properties of the graphic format in which the image is saved, in order to provide confidential information. One of the techniques that allow you to hide information in a picture is to modify the least significant bit (LSB - Least Significant Bit) [4,5]. As its name indicates, the technique involves the modified least significant bits of a numerical value describing the intensity of the color at the selected location of a picture.

Generally, with the LSB method in an image, you can hide any information stored in the form of consecutive bits. It is not important, whether it is a text message, the encrypted text message or other image, or any other type of binary file. Only limitation is the capacity of the media. In the example picture of size of 800×600 pixels, assuming the modification of only one bit of blue and one bit of green, you can send a secret message on a maximum of 120 000 bytes.

2 Proposed Algorithm for Steganography in Pictures

In the paper we insist on minimizing the possibility of detecting the embedding process, so one of the images is not modified though its bits are the key in the embedding process. However, if it is rejected the steganoanalyst will not be able to recover the hidden information. Another interesting issue is the selection of the pixel generator with hidden information, in such a way as to minimize the number of pixels modified. In addition, we try to minimize the size of the key that determines in which bits the image will be embedded. We use multiplicative group generator here (Z_p, p is prime). We have described it in detail in the papers [6–8].

2.1 Perfect Steganography

This subsection deals with issues related to the definition of perfect steganography. The most important requirement is that a steganographic algorithm has to be imperceptible. The fact that the image file has been modified can not be seen. The container with the embedded information must be large enough, but we do not analyze it in detail. It should also be resistant to manipulation that hinders

the extraction of embedded information; this problem is also not analyzed in our paper. We discuss only one LSB steganographic method. We are aware that this constitutes a restriction. Statistical steganalysis is the practice of detecting hidden information through applying statistical tests on image data.

Steganography that meets all of the above conditions is an example of perfect steganography.

In 1998 C. Cachin in the paper [9] proposed an information-theoretic approach to steganography. Cachin's paper, so-called perfectly secure stegosystems were defined, where messages that carry and do not carry hidden information are statistically indistinguishable. This is also analyzed in detail in [10–12]. Every steganographic scheme uses a pair of embedding and extraction mappings $Emb : C \times K \times M \rightarrow C$, $Ext : C \times K \rightarrow M$ defined on the sets of all possible covers C, stego keys K, and the set of messages M.

Mappings must be satisfied $Ext(Emb(x, k, m), k) = m$ for any message $m \in M$, cover $x \in C$, and key $k \in K$.

Assuming the cover is drawn from C with probability distribution P_c and the stego key, as well as the message, are drawn according to distributions P_k , P_m over their corresponding spaces K, M, the distribution of stego objects will be denoted as P_s. The two distributions can be compared using their Kullback-Leibler divergence (Kullback-Leibler distance or relative entropy).

$$D_{KL}(P_c \mid\mid P_s) = \sum_{x \in C} P_c(x) \log \frac{P_c(x)}{P_s(x)} \tag{1}$$

When $D_{KL}(P_c \mid\mid P_s) = 0$, then we say the stegosystem perfectly secure. If $D_{KL}(P_c \mid\mid P_s) \leq \varepsilon$ then we say the steganographic system is ε-secure [1,13]. So, the image in which we do not modify any bit is an example of perfect steganography. In the algorithm we propose, one of the images used has this property.

2.2 Algorithm

In this chapter we present the algorithm for generating an image with the hidden information in it (the text). There are two approaches to this problem. First, when a key and a text are hidden together in a picture. Secondly, when the key is outside the picture. In this case, the key can be delivered to the addressee through an encrypted connection, hidden in another image or delivered through another way. The key to our algorithm consists of a pair of numbers. The first number is a generator of multiplicative group Z_p and the second a prime number p. The group generator is found by using the following theorem.

Theorem 1. *Let $p-1$ has a decomposition into prime factors $p-1 = p_1 p_2 ... p_k$, $g \in Z_p$ is a generator of multiplicative group Z_p if and only if $g^{p-1/p_i} \neq 1$ for each $1 \leq i \leq k$.*

Theorem 2. *The number of generators of the group Z_p is $\phi(\phi(p))$. Where ϕ is the Euler function.*

In the Algorithm 2.2 an input is an array of size of k consisting of primes that are factors of $p-1$. The algorithm returns the generator of a multiplicative group Z_p.

Algorithm of generating the generator

Input $d[k]$ - array of size k prime divisors $p-1$

Output g - generator

1. Choose randomly g from the range $< 2, p-2 >$
2. For $i \leftarrow 0$ to $k-1$
   If $(g^{p-1/d[i]} = 1)$ go to step 1
3. return g

Suppose we have generated a prime number p greater than the number of bits of a bitmap and a generator of the group Z_p. Now we show an example of an algorithm to hide the text in the image. If the number of pixels in our bitmap is n, we can put in this picture $3n$ bits of text. This means that we put in the bitmap at most $3n/8$ bytes of text. In our algorithm, we will put 2 pixels in each pixel of the bitmap. The realization will take two approaches here. First, when the information will be placed in two fixed RGB colors (green, blue). Second, the information will be placed in any two of the three colors (in the sequence: red color with green, red with blue and green with blue). Let us assume then, we have k bytes of text ($k < n/4$), which we include in the bitmap.

Algorithm of hiding text in an image

Input $t[k]$ - array of text, $b[n]$ - array of values of bitmap pixels,
$c[n]$ - array of values of bitmap pixels.
Output $b[n]$ - array of values of bitmap pixels,
$c[n]$ - array of values of bitmap pixels.

1. For $i \leftarrow 0$ to $k-1$
   - a t[i] divide into four blocks of two bits (1B of text)
   - b compute the successive four values:
     $g^{4i} mod\ p, g^{4i+1} mod\ p, g^{4i+2} mod\ p, g^{4i+3} mod\ p$
   - c If the current bit of text (i-th bit) is equal to the bit of the array b defined in b at the same place in the array c, we set the value 1. Otherwise, we set the value 0.
2. return $b[n]$, $c[n]$

Note: when some of values calculated in point b is bigger than $n-1$ of the index of the last bitmap pixel), we calculate the next value.

3 Experiment

We used in our experiment images in size 800×600 pixels and 24 bits colour, see Fig. 1. One of them was used as cover image C and the other as stego keys K. The lenght of the messege was 1000 chars. According to the assumptions of the algorithm, we took the initial value of 480000 and found the nearest prime number 480003. This number must be larger than whole number of pixels in image in order to possibility of writing in all image pixels. Then, using the algorithm described earlier, we generated 138237 generators. Each of this generators was used to hide the information in the image according to the algorithm described above. The experiment was repeated for three different images (see Fig. 1),

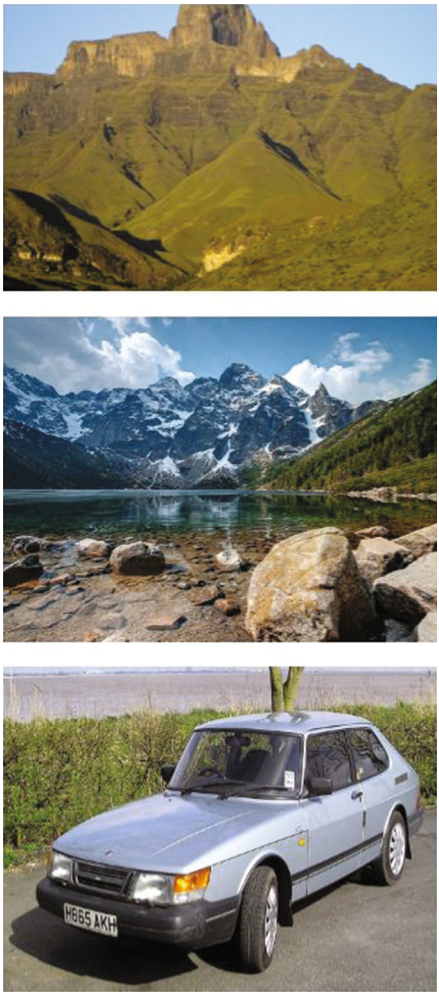


Fig. 1. Samples images used in experiment

and the results are shown in the Table 1. The results obtained were analysed for Peak Signal to Noise Ratio (PSNR) level [14]. The results are satisfactory because for different generators the PSNR value stays around 74 dB.

Table 1. Results of experiment

No. of changed bits		PSNR [dB]
Minimum	3273	73,8
Maximum	3713	74,7
Average	3493	74,3

The tests were then carried out for the percentage change in the number of bits on which the information is written. We conducted the research for different lengths of hidden message. Results are given in the Table 2. The length of the hidden information does not significantly affect the number of bits changed in the image stego keys K.

Table 2. Results of experiment for g = 156822 and different lenght of text

Text lenght [char]	Text lenght [bit]	No. of changed bits	% of changed bits	PSNR [dB]
50	350	157	45	87
100	700	306	44	84
1000	7000	3377	48	74
10000	70000	35026	50	64
100000	700000	240372	34	56
Average			44	73

4 Summary

As the study showed, the proposed algorithm is effective in hiding information and provides a high level of PSNR. In addition, changes made to image K are not noticeable (see Fig. 2). Another advantage of the algorithm is the ability to select a generator, resulting in varying levels of changes in the image K. The difference between the smallest and greatest changes is 440 bits, which represents almost 12% of the total number of changed bits. These values refer to the case in which the length of message M was 1000 characters (7000 bits).

Further work is planned to lead to the smallest changes in images. Using Markov processes, we also plan to explore other scenarios so that steganographic systems are closer to perfect steganography [11,15]. Another direction of research will be the use of the Chinese Remainder Theorem [16,17] to construct a steganographic system.

Fig. 2. Comparison images before and after hiding information

References

1. Fridrich, J.: Steganography in Digital Media: Principles, Algorithms, and Applications. Cambridge University Press, Cambridge (2016)
2. Katzenbeisser, S., Petitcolas, F.A.P.: Information Hiding Techniques for Steganography and Digital Watermarking. Artech House, Boston (2000)
3. Cox, I., Miller M., Bloom, J., Fridrich, J., Kalker, T., Digital watermarking and steganography, Morgan Kaufmann, Burlington (2008)
4. Cheddad, A., Condell, J., Curran, K., Mc Kevitt, P.: Digital image steganography: survey and analysis of current methods. Signal Process. **90**, 727–752 (2010)
5. Provos, N., Honeyman, P.: Detecting steganographic content on the internet. In: Network and Distributed System Security Symposium 2002 (2002)
6. Dyja, R., Jakubski, A.: A picture that is something more than just a picture. Scientific Issues Jan Dlugosz University in Czestochowa. Mathematics, vol. XVI, pp. 111–116 (2011)
7. Dyja, R., Jakubski, A.: Steganography in Digital Multimedia., IT Tools in Management and Education. Selected Problems, Wydawnictwo Politechniki Czestochowskiej, pp. 146–154 (2011)
8. Jakubski, A., Latasiewicz, M., Rog, P., Ruczkowski, P.: Steganograficzna metoda rozproszenia hasla, Technologie informacyjne w funkcjonowaniu organizacji. Zarządzanie z wykorzystaniem multimediów (red.) Kieltyka Leszek, Dom Organizatora, Towarzystwo Naukowe Organizacji i Kierownictwa, pp. 135–141 (2013)
9. Cachin, C.: An information-theoretic model for steganography. In: Aucsmith, D. (ed.) IH 1998. LNCS, vol. 1525, pp. 306–318. Springer, Heidelberg (1998)
10. Ryabko, B.Y., Ryabko, D.B.: Asymptotically optimal perfect steganographic systems. Probl. Inf. Transm. **45**(2), 184–190 (2009)
11. Ryabko, B.Y., Ryabko, D.B.: Constructing perfect steganographic systems. Probl. Inf. Comput. **209**(9), 1223–1230 (2011)
12. Cachin, C.: An information-theoretic model for steganography. Inf. Comput. **192**(1), 41–56 (2004)
13. Cover, T.M., Thomas, J.: A Elements of Information Theory. Wiley, New York (2012)
14. Kot, S.M., Jakubowski, J., Sokolowski, A.: Statystyka, Wyd. Difin, Warszawa (2007)
15. Bobulski, J.: Hidden Markov models for two-dimensional data. In: Burduk, R., Jackowski, K., Kurzyński, M., Woźniak, M., Żołnierek, A. (eds.) CORES 2013. AISC, pp. 141–149. Springer, Heidelberg (2013)

16. Ding, C., Pei, D., Salomaa, A.: Chinese Remainder Theorem: Applications in Computing, Coding, Cryptography. World Scientific Publishing, Singapore (1996)
17. Jakubski, A.: Selected application of the Chinese remainder theorem in multiparty computation. J. Appl. Math. Comput. Mech. **15**(1), 39–47 (2016)

PET Waste Classification Method and Plastic Waste DataBase - WaDaBa

Janusz Bobulski(✉) and Jacek Piatkowski

Czestochowa University of Technology, 73 Dabrowskiego Str.,
42-201 Czestochowa, Poland
{januszb,jacek.piatkowski}@icis.pcz.pl

Abstract. The main purpose of this work was creation of a plastic waste database of images of objects constituting the typical contents of municipal waste. This group of waste, by using methods of Computer Vision can be automatically selected on the sorting lines businesses for waste disposal. Digital images of items that will be received for processing should reflect the specific conditions of places where real objects have to be found. Thus, each thing is placed in this database should be presented in the course of several collections of images, taking into account different lighting conditions and different arrangement relative to the image recorder, and the different degree of deformation of these objects as a result of previous processes. Images created in the collection will be divided into groups based on the type of material from which individual objects were made. An second main aim of the article is to present the method of plastic waste selection based on histogram analysis. The method has to be fast so that it can be used in a waste sorting plant.

1 Introduction

Recycling is one of the most important methods of environmental protection, the purpose of which is to reduce the amount of waste stored in landfills and conservation of natural resources. The term recycling means the recovery of raw materials, which consists in the transformation of substances or materials contained in waste in the production process to obtain the substance or material on the fate of primary or other purposes, including organic recycling, excluding energy recovery. Method of processing steel to use in the 70-ies of the last century, when significantly increased the amount of waste produced. At the same time appeared the possibility of their reuse and noted that the waste put to landfill are not environmentally neutral. Some of them are extremely durable. For example, the decomposition of cotton material is 5 months, paper tissue - 3 month, aluminium cans - 100 years, plastic bag - 300 years. Taking into consideration the fact that the average man is producing the about 350 kg of waste annually, and the fact that the pace of producing rubbish much is exceeding the pace of their disintegration, creating new garbage dumps is necessary. Here a problem of high building costs and a problem of environmental, social and economic influences appear [1]. Effectively pursued politics of the waste

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*,
Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_8

management and re-using them allows us to prolong the operation of the landfill by up to 60% and reduce the need for new landfills. The recycling should be more widely applied and promoted. However in order to apply the recycling a sorting of waste is necessary on the ones being recyclable and on remaining. Mixed waste in sorting plants is segregated manual and mechanical. The need for manual sorting selectively collected fractions of municipal waste due to the lack of the corresponding transmission of information related to the rules of the system of selective collection of recyclable materials with low environmental consciousness of society. In addition, manual sorting is time consuming and expensive, so it is advisable to establish and develop methods for automatic sorting of plastic waste. The establishment of a database of images of objects constituting the typical content of the municipal waste, produced by households, i.e. packages of different kind and everyday articles are a main goal of the work. Created in the framework of the project public database of waste images, becomes the starting point for research to develop methods for automatic sorting of waste with the use of modern techniques of image processing and recognition [2].

The main motive for the creation of the work was the fact that not developed methods for automatic sorting of waste, using methods of pattern recognition. In search of reasons for this situation, it was noted that, despite the existence of several databases of images of various objects, among them no a database with waste images. Hence emerged the idea of creating a database of waste images. This database will be the basis for testing methods for automatic sorting of waste, and also will allow to compare the effectiveness of various methods. In the development of new methods of pattern recognition free access to the database of images are required. It allows you to test the method and compare the obtained results with other methods. The effectiveness of the different image recognition systems can be compared only if you use the same set of input data, which in this case are the images. In this context, the proposed project is a fundamental research, and its effect will be used as the basis for the research of other scientists.

Every year in Poland millions tons of municipal waste is created. According to Eurostat data, in 2012, households produced over 9.3 million tons of waste [3] with only slightly more than 10% of the total collected municipal waste was recycled. A huge amount of mixed waste ends up in waste management plants, which later is sorted in order to extract recyclable materials, high-energy waste, and organic disposal, which can be biodegraded through composting. This selective method for processing the input materials is a widely recognized as a way of reducing the amount of waste going to landfill. In the currently operating waste management plants the sorting process is realized in a manual-automatic way with very different amounts and kinds of devices. What also varies is the order of deployment of these devices in production lines. Depending on the degree of automation of the sorting and segregation facilities, magnetic or electromagnetic, inductive, optical and electrostatic separators are being used. However, the sorting of the waste is a complex process, and difficult to implement, due to the mechanical deformations and contaminants of delivered waste. Most modern waste collection trucks use a compaction mechanism and waste carried by

such vehicles have been affected by mechanical stress and blending, which leads to considerable difficulties in distinguishing individual objects during sorting process [4]. It is estimated that, despite a few years of the new Polish rules of selective collection and waste management, far more than half of waste still going to landfills. Meanwhile, the European Union Directive 2008/98/WE on waste management [5] and applicable Polish law [6] show that besides prevention of waste generation, the process of waste material recycling should have priority over the energy production as a result of high-energy waste combustion. It is assumed that successively in 2020, 2025 and 2030 recycling of packaging must achieve levels respectively 60%, 70% and 80%. Specialists, dealing with the issue of waste processing and disposal, agree, that increasing the level of automation and reducing the amount of work carried out manually will increase the efficiency both currently running waste management plants and plants built in the future, and also improve the purity of recyclables [7]. In the case of searching for new solutions, which can improve the processes of sorting the municipal waste and recovering the secondary raw materials, the Computer Vision technology seems to be perspective. One example of successful implementations of those techniques is biometrics, which uses human biological characteristics for identification. Those characteristics are extracted from images of human faces, fingerprints, irises, etc. [8]. The dynamic development of biometric methods, however, was possible thanks to the creation of common access databases, like e.g. the face database [9,10]. Using these databases, the scientists could test their method and compare the results. By using the same data, the comparison of different methods appeared to be credible and objective. The same mechanism could be applied to another specific group of objects such as packaging, containers or other articles of daily use, which should be properly identified in municipal waste streams - in order to ensure its effective segregation. At this point emerged the idea of creating, within the proposed project, a professional database of "valuable" objects (at least from the perspective of environmental protection), on which, as in the case of biometrics databases, various CV techniques could be tested.

2 The Plastic Waste Image Database

The WaDaBa database of plastic waste images have been created. Acquisition of the minimum of 100 objects is planned, and for each of them forty images will be done in different conditions. In the future database will be expanded because it will still appear new types of waste.

We created the research position for acquiring images of waste. To comprise it will be from: the platform, on which will be put objects - wastes, the stand with a camera, the light sources of two type: the fluorescent lamp and LED. Objects of waste used in the project were acquired from households. Everyday type of plastic waste have been used - see Fig. 1. Waste have been acquired and photographed by four months in order to recruit as the most typical kinds of the municipal waste.

Fig. 1. Plastics type marks

The object will be put on the research position and next photographed with first and second type of light. There were series carried out of 10 photographs with differ in the angle of the turnover for every object (in the vertical axis). Next the object was damaged to varying degrees: small, medium and large. For each type of destruction have been made 10 photographs. So considering all variants for every object 40 photographs were taken, multiplying it by the number of objects, 4 000 of photographs were created in the database. The parameters of photographs:

- size 1920×1277 of pixels
- resolution 300 dpi
- colour palette RBG 24 bits
- file format JPG

Concerning parameters of photographs, single images have size about 1 MB, and the entire database of 4 GB. The acquired image have been saved onto the disk under the standardized name determining parameters of acquiring the image and type of waste. The database of images is available at address http://wadaba.pcz.pl with free access. In addition, the portal will be published the results of research into automated sorting techniques that use image recognition. Sample images are shown in Fig. 2

Fig. 2. Sample images fro WaDaBa database.

It was assumed that the name of the each acquired picture should describe of photographed object in a simple form. In this way the easy choice of some interesting objects will be possible in later stages of the research. For instance, objects with the medium level of damage or all bottles with a screw caps.

It was decided that the name of the photographed object will be coded as follows:

0004_**a**01**b**05**c2d**0**e1f**0**g1h**1.jpg

where first four digits denotes the order number of the object and the next sections (a to h) described its properties such as:

a - a number denoting the plastic type;
b - colour;
c - type of light;
d - the deformation level;
e - the level of a dirt;
f - presence of a screw cap or a lid;
g - presence of a ring - characteristic for the bottles with a screw caps;
h - order number of a random position of a photographed object.

We also decided to use characters (as a name of a parameter) and digits (as a parameter value) to make the easier to understand it for the human - quickly, without a computer. In the Table 1 the established values of defined parameters are presented.

Table 1. Values of parameters used encoding of the names of photographed objects

Parameter	Parameter values
a (Plastic type)	**00** - missing or unreadable identifier, **01** - PET - polyethylene terephthalate, **02** - PE-HD - high-density polyethylene, **03** - PVC - polyvinyl chloride, **04** - PE-LD - low-density polyethylene, **05** - PP - polypropylene, **06** - PS - polystyrene, **07** - Other.
b (Colour)	**00** - transparent, **01** - white, **02** - red, **03** - green, **04** - blue, **05** - yellow, **06** - orange, **07** - purple, **08** - brown, **09** - grey, **10** - black, **99** - other.
c (Light)	**1** - scattered light of fluorescent lamp, **2** - warm LED light reflector.
d (Deformation)	**0** - undeformed, **1** - small deformation, **2** - medium deformation, **3** - high deformation
e (Dirtiness)	**0** - clean, **1** - small dirt, **2** - medium dirt, **3** - high dirt.
f (Screw cap or lid)	**0** - lack, **1** - occurs.
g (Ring)	**0** - lack, **1** - occurs.
h (Random position)	**0,..., 4.**

3 PET Classification Method

Amongst home waste they are dominating PET. Many-months observation of household rubbish confirmed it. On that account we decided that one should first select this type of plastic. With additional argument behind it, there is a fact that it is possible again to process PET and to obtain a textile e.g. fleece material [11]. At developing the algorithm developing the simplest and fastest method for it to be possible to use it in the sorting plant of rubbish on the transport belt in the real time possibly was an important establishment. In the first step, we load the image and converts it to grayscale. The edge detection is then performed to allow the object to be located. A standard mode of operation using the Canny filter was used for the detection of the edge. After locating we are allocating the image area containing the object. Next we are calculating the histogram for this part of the image. Analysis of the histogram consisting in adding up is a further step first one hundred and of one hundred last elements of the histogram. Considering the fact that PET is transparent material but the background of images it is black (similarly to the transmission belt in the sorting plant) we are comparing received earlier values of sums. For PET the value of the first sum will be greater, while for other non-transparent materials the second sum will be larger (Fig. 3). After all on this base a decision to categorize the object for the PET collection or nonPET collection are being made.

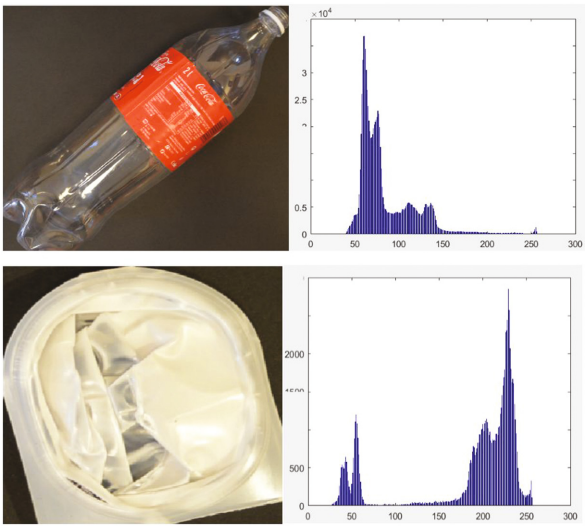


Fig. 3. Comparison of histograms

Algorithm:

- load photo
- convert to gray scale
- edge detection
- object localization
- select object
- calculate histogram
- analyse histogram
- decision: PET/notPET.

4 Experiment

The experiment was carried on our database WaDaBa. We used ten sets of data with 200 pictures each that means we used two thousand images. Table 2 presents results of experiment. These results have preliminary character for developing advanced waste selection techniques based on computer vision techniques. Analysing the current condition of the knowledge in this field we didn't come solutions of this type. From a review of the methods of waste sorting presented in [12] shows that so far computer vision methods aren't adopted in the automatic selection of wastes. In addition, we cannot compare our results to other existing methods.

Table 2. Results of experiment

Name of set	No. of images	Recogn. rate [%]	FAR [%]	FRR [%]
Set A	200	92.0	0	8.0
Set B	200	92,5	7.0	0.5
Set C	200	61,0	25.0	14
Set D	200	74,0	25.0	1.0
Set E	200	54,0	46.0	0
Set F	200	88,5	11,5	0
Set G	200	51,5	48,5	0
Set H	200	57,0	42.0	1.0
Set I	200	88,0	11.0	1.0
Set J	200	80,0	14.0	6.0
Set K	200	94,0	6.0	0
Average		75,68	21,45	2,86

5 Conclusion

The research results presented seem promising. Therefore, further research is planned on the development of more advanced methods using computer vision techniques. We analyse the types of errors we have found to improve the selection results by more accurate object extraction techniques. Improvement of the recognition level of objects can be sought by improving the feature extraction technique and operation on complex sets of features. The resulting database will be useful for testing new methods and will allow you to compare results on the same data. Therefore, it is appropriate to further expansion of the base in the future.

References

1. de Bree, M.: Waste and innovation. Delft University of Technology, Delft, the Netherlands, Thesis (2005)
2. Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and The Committee of the Regions: Taking sustainable use of resources forward - A Thematic Strategy on the prevention and recycling of waste - SEC 1681, SEC 1682 (2005)
3. Eurostat. http://ec.europa.eu/eurostat/statisticsexplained/index.php
4. Edjabou, E., Jensen, M.B., Götze, R., Pivnenko, K., Petersen, C., Scheutz, C., Astrup, T.F.: Municipal solid waste composition: sampling methodology, statistical analyses, and case study evaluation. Waste Manag. **36**, 12–23 (2015)
5. Directive of the European Parliament 2008/98/WE 19.11.2008. http://eurlex.europa.eu
6. Act of Waste 14.12.2012 (2012). http://isap.sejm.gov.pl. (in Polish)
7. Ecotechnology. http://ekotechnologie.org
8. Bobulski, J.: Multimodal face recognition method with two-dimensional hidden Markov model. Bull. Pol. Acad. Sci. Tech. Sci. **65**(1), 121–128 (2017)
9. The Facial Recognition Technology (FERET) Database. http://www.itl.nist.gov/iad/humanid/feret/feret_master.html
10. Jasinski, P., Forczmański, P.: Combined imaging system for taking facial portraits in visible and thermal spectra. Image Processing and Communication Challenges 7. Advances in Intelligent Systems and Computing, vol. 389, pp. 63–71. Springer Verlag, Cham (2016)
11. Singh, N., Hui, D., Singh, R., Ahuja, I.P.S., Feo, L., Fraternali, F.: Recycling of plastic solid waste: a state of art review and future applications. Compos. B Eng. **115**, 409–422 (2017)
12. Gundupalli, S.P., Hait, S., Thakur, A.: A review on automated sorting of source-separated municipal solid waste for recycling. Waste Manag. **60**, 56–74 (2017)

Algorithms and Methods

Estimation of Geometrical Deformations of 3D Prints Using Local Cross-Correlation and Monte Carlo Sampling

Jarosław Fastowicz(✉), Dawid Bąk, Przemysław Mazurek, and Krzysztof Okarma

Department of Signal Processing and Multimedia Engineering, Faculty of Electrical Engineering, West Pomeranian University of Technology, Szczecin, 26 Kwietnia 10, 71-126 Szczecin, Poland
{jfastowicz,orinoco,okarma}@zut.edu.pl, dawid.d.bak@gmail.com

Abstract. This paper presents a novel method for estimating the degree of deformations in 3D prints. The increased popularity of three-dimensional printing techniques introduces a necessity to create methods for quality assessment of printed materials. One of the key problems of 3D printing are strains and deformations of printed objects. This problem is determined by many factors like: printing material (filament), object geometry or temperature. The conducted research is focused on a method of automatic analysis of deformations in 3D prints based on surface scans of objects. In our research some surface scans containing varying degrees of deformations have been used for verification of obtained results. In order to evaluate the degree of deformations of materials a local cross-correlation with Monte Carlo sampling have been used. Tests carried on multiple samples have shown that the local cross-correlation technique works well when assessing the degree of deformations of printed objects on the basis of surface scans. The obtained results show that our method can be further applied for improvement of the quality of the objects received from 3D printers.

Keywords: 3D prints · Visual quality assessment · Cross-correlation · Image analysis

1 Introduction

3D printing is a new revolutionary technology which is increasingly used. Application of 3D printing in many areas of science and technology makes it much easier to build models and prototyping. Use of 3D printers has significantly shortened the time required for many engineering jobs which require complex models and parts with complex shapes. Elements that were previously required to use many machines or even had to be hand-made now can successfully be printed. Unfortunately despite the many advantages of 3D printing technology as a relatively young discipline still struggling with many difficulties. The main

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*, Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_9

problems that arise when creating objects with 3D printing methods are precise production of objects, quality of finishing and the appearance of defects. These problems are currently widely discussed and research is being carried out on their solution, similarly as the growth of image based inspection methods based e.g. on texture analysis [8].

Growing interest in vision based applications for monitoring of the 3D printers and similar purposes can be observed recently [14,15]. In Chauhan's paper [2] there has been proposed a method for automatic fault detection. Some earlier attempts have also been made by Fang [4] and Cheng [3]. One of the most interesting systems has been proposed by Straub [13], however it is far from perfection being sensitive to changing lighting conditions as well as motions of cameras and imprecise calibration. However the quality assessment issue has not been considered in those papers.

This paper, following some earlier papers [9–11], addresses the issue of defects arising during the printing process. The element that will be considered is the issue of geometry defects of printings objects and the various deformations that occur in the printing process. The main goal of this paper is to answer the question whether the use of cross-correlation based image analysis and signal processing techniques can be helpful in estimation of the degree of surface deformations of the printed object.

Due to the specific technology of the 3D printing process the final printed object has some certain recurrent features associated with the method of applying the filament in layers. Therefore it should be possible to evaluate the print defects by analyzing the disturbances in periodicity of object surface. Deformations like: convexity of object, offset layers and missing layers should be easily detected during the analysis of the surface images (captured by a camera or scanned). The purpose of our research is to determine if it is possible to detect these changes in an automated manner only on the basis of scans of analyzed surfaces using the cross-correlation based approach.

2 The Idea of the Proposed Method

In order to ensure objective results, scans of surface samples taken at different resolutions were used. To determine the degree of deformations of the samples on surface scans, the local cross-correlation method has been applied. Using this method, it is possible to estimate the distortions of the individual layers for the mutual position of the layers with respect to each other. Displaced anomalies in the predictable image structure allow the estimation of the degree of geometric deformation of the sample's surface. The use of local cross-correlation and the mutual comparison of consecutive areas of the same scan enables qualitative assessment without the need of using a perfect quality reference sample.

The cross-correlation tests for the entire surface image are time-consuming and also demanding relatively high computing power, so statistical sampling has been performed and the local cross-correlation has been calculated for randomly chosen areas. Because deformations of samples may be periodic, and could

therefore be overlooked when arbitrarily selecting sampling areas, that is why the selection of the tested areas by the Monte Carlo method has been used [12]. The use of the Monte Carlo sampling helps to avoid the problem of wrong sample selection for the correlation analysis. The choice of the Monte Carlo method has motivated by its successful applications in image processing e.g. related to inpainting [6], filtering using Random Non-Local Means [1] or retargeting [5].

In order to limit the amount of the analyzed image data, especially for high resolutions scans, we have drawn randomly a specified number of R positions of the square windows of size $L \times L$ pixels, being the local Regions Of Interest (ROI). According to the idea of the Monte Carlo method, a random number generator of uniform distribution has been applied in order to determine the positions of the ROIs, using the appropriate margins dependent on the specified size of compared windows.

Assuming the even value of R, half of ROIs can be considered as "left" and the other half as "right" samples which have been compared pairwise. For the comparison of two fragments of the image the cross-correlation has been used due to the assumption of its relatively high values for high quality 3D prints. Such a mutual comparison can lead to promising results assuming that both fragments are not shifted in phase. Since the influence of such a phase shift may lead to improper results, the additional preprocessing step has been proposed allowing the adjustment of the ROIs maximizing the phase congruency between two compared fragments of an image.

This adjustment has been made by shifting down the "right" ROI by the determined number of pixels Δ. The values of Δ for each pair of the ROIs has been calculated after resizing both ROIs into two one-dimensional vertical vectors containing the average luminance for each row of the window. Such an operation can be made e.g. using MATLAB's *imresize* function, leading to

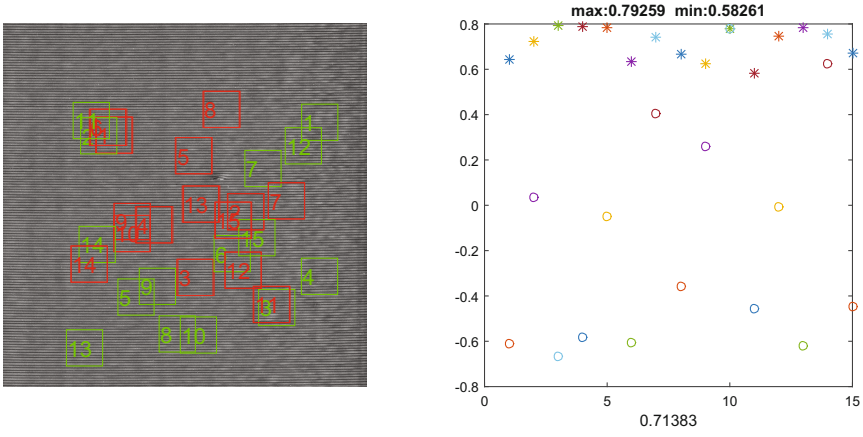


Fig. 1. Randomly chosen pairs of ROIs (left) and 2D cross-correlation values (right) for an exemplary high quality image the silver filament

the vector of L elements, followed by calculation of the cross-correlation between both vectors. Finding the first local maximum of the cross-correlation allows to determine the necessary Δ value and further adjustment of window positions. For such adjusted pairs of ROIs the 2D cross-correlation has been calculated considered as the metric corresponding to the quality of the 3D printed surface.

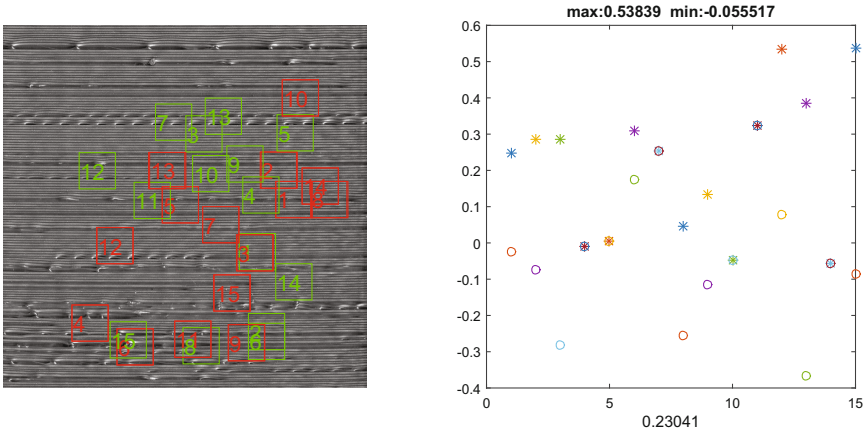


Fig. 2. Randomly chosen pairs of ROIs (left) and 2D cross-correlation values (right) for an exemplary low quality image for the silver filament

An illustration of the randomly chosen ROI pairs together with final 2D cross-correlation values obtained for 1200 dpi scans is shown in Figs. 1 and 2. The marks '**o**' on the right plots denote the raw cross-correlation values before phase adjustments and the marks '*****' correspond to the values obtained after the phase adjustment. The overall final cross-correlation values averaged for 15 ROI pairs are: 0.71 for high quality print and 0.23 for low quality print. The maximum and minimum correlation values for the adjusted ROI pairs are shown as well. As can be easily observed, for high quality surfaces there are no negative cross-correlation values obtained after the phase adjustment.

3 Experimental Results

In order to verify the proposed approach more extensive tests using different colors of filaments have been conducted. All the images of the scanned 3D printed surfaces have been converted to grayscale according to the ITU Recommendation BT.601-7 [7] using MATLAB's *rgb2gray* function. During our experiments high and low quality red, green, silver and white samples have been used as well as orange 3D prints containing high quality areas and some local distortions (considered as low quality samples in Tables 2, 3, 4, 5, 6, 7 and 8). Decrease of the surface quality of 3D prints has been obtained by forced change of the filament's

delivery speed, change of temperature as well as sudden slight motions of the table. The most representative results have been obtained for 1200 dpi scans although the use of higher resolution images has led to the same conclusions.

Experiments have been conducted for different number of ROI pairs obtained using the Monte Carlo method and the repeatable results have been achieved using 10 and more pairs. Therefore, for the presentation of results in the paper 15 ROI pairs have been assumed. Another variable influencing the results is the size of the ROI window. During the experimental tests we have assumed L equal to $\frac{1}{5}$, $\frac{1}{10}$, $\frac{1}{15}$ and $\frac{1}{20}$ of the image size (considered as the lowest value of its width and height).

The results of the mean and median 2D cross-correlation obtained for various sizes of the ROIs without and after phase adjustment are presented in Tables 1, 2, 3, 4, 5, 6, 7 and 8.

Analyzing the results, the threshold value of the correlation about 0.35 can be determined which separates high and low quality prints. Nevertheless, classification of the 3D prints can be efficiently made applying the threshold dependent

Table 1. Average and median 2D cross-correlation without phase adjustment for $L = \frac{1}{20}$ of the image size

Color		White		Red		Orange		Silver		Green	
Sample		No. 1	No. 2	No. 1	No. 2	No. 1	No. 2	No. 1	No. 2	No. 1	No. 2
High	Mean	0.271	0.255	0.279	0.307	-	-	0.477	0.471	0.379	0.348
Quality	Median	0.253	0.247	0.222	0.242	-	-	0.508	0.501	0.344	0.316
Low	Mean	0.165	0.188	0.122	0.152	0.151	0.155	0.146	0.192	0.150	0.130
Quality	Median	0.147	0.166	0.104	0.133	0.087	0.092	0.126	0.166	0.121	0.108

Table 2. Average and median 2D cross-correlation obtained after phase adjustment for $L = \frac{1}{20}$ of the image size

Color		White		Red		Orange		Silver		Green	
Sample		No. 1	No. 2	No. 1	No. 2	No. 1	No. 2	No. 1	No. 2	No. 1	No. 2
High	Mean	0.487	0.447	0.708	0.800	-	-	0.897	0.891	0.826	0.795
Quality	Median	0.513	0.480	0.753	0.819	-	-	0.913	0.900	0.837	0.813
Low	Mean	0.175	0.201	0.210	0.314	0.317	0.315	0.244	0.336	0.237	0.193
Quality	Median	0.150	0.168	0.190	0.292	0.140	0.154	0.218	0.320	0.1757	0.143

Table 3. Average and median 2D cross-correlation without phase adjustment for $L = \frac{1}{15}$ of the image size

Color		White		Red		Orange		Silver		Green	
Sample		No. 1	No. 2	No. 1	No. 2	No. 1	No. 2	No. 1	No. 2	No. 1	No. 2
High	Mean	0.255	0.243	0.255	0.303	-	-	0.477	0.465	0.379	0.343
Quality	Median	0.247	0.227	0.219	0.237	-	-	0.509	0.497	0.345	0.315
Low	Mean	0.143	0.172	0.106	0.140	0.136	0.141	0.129	0.173	0.138	0.114
Quality	Median	0.126	0.151	0.093	0.119	0.072	0.081	0.109	0.141	0.109	0.097

Table 4. Average and median 2D cross-correlation obtained after phase adjustment for $L = \frac{1}{15}$ of the image size

Color		White		Red		Orange		Silver		Green	
Sample		No. 1	No. 2	No. 1	No. 2	No. 1	No. 2	No. 1	No. 2	No. 1	No. 2
High	Mean	0.486	0.466	0.643	0.766	-	-	0.875	0.868	0.797	0.745
Quality	Median	0.498	0.492	0.686	0.790	-	-	0.893	0.884	0.812	0.763
Low	Mean	0.148	0.199	0.183	0.289	0.295	0.288	0.207	0.300	0.226	0.182
Quality	Median	0.125	0.159	0.161	0.270	0.140	0.132	0.176	0.282	0.164	0.142

Table 5. Average and median 2D cross-correlation without phase adjustment for $L = \frac{1}{10}$ of the image size

Color		White		Red		Orange		Silver		Green	
Sample		No. 1	No. 2	No. 1	No. 2	No. 1	No. 2	No. 1	No. 2	No. 1	No. 2
High	Mean	0.232	0.219	0.233	0.278	-	-	0.441	0.447	0.352	0.314
Quality	Median	0.232	0.213	0.214	0.227	-	-	0.472	0.475	0.338	0.303
Low	Mean	0.113	0.138	0.087	0.124	0.123	0.116	0.097	0.148	0.109	0.102
Quality	Median	0.096	0.130	0.074	0.108	0.068	0.068	0.079	0.123	0.087	0.083

Table 6. Average and median 2D cross-correlation obtained after phase adjustment for $L = \frac{1}{10}$ of the image size

Color		White		Red		Orange		Silver		Green	
Sample		No. 1	No. 2	No. 1	No. 2	No. 1	No. 2	No. 1	No. 2	No. 1	No. 2
High	Mean	0.463	0.435	0.518	0.695	-	-	0.819	0.826	0.740	0.662
Quality	Median	0.473	0.448	0.506	0.727	-	-	0.848	0.851	0.764	0.685
Low	Mean	0.128	0.166	0.155	0.261	0.256	0.254	0.171	0.275	0.194	0.174
Quality	Median	0.100	0.135	0.138	0.245	0.119	0.121	0.141	0.244	0.146	0.140

Table 7. Average and median 2D cross-correlation obtained without adjustment for $L = \frac{1}{5}$ of the image size

Color		White		Red		Orange		Silver		Green	
Sample		No. 1	No. 2	No. 1	No. 2	No. 1	No. 2	No. 1	No. 2	No. 1	No. 2
High	Mean	0.195	0.192	0.179	0.247	-	-	0.390	0.406	0.290	0.261
Quality	Median	0.193	0.176	0.166	0.226	-	-	0.404	0.436	0.297	0.271
Low	Mean	0.089	0.100	0.065	0.112	0.109	0.109	0.078	0.137	0.077	0.090
Quality	Median	0.072	0.090	0.056	0.098	0.073	0.089	0.062	0.121	0.062	0.074

Table 8. Average and median 2D cross-correlation obtained after phase adjustment for $L = \frac{1}{5}$ of the image size

Color		White		Red		Orange		Silver		Green	
Sample		No. 1	No. 2	No. 1	No. 2	No. 1	No. 2	No. 1	No. 2	No. 1	No. 2
High	Mean	0.357	0.378	0.393	0.551	-	-	0.673	0.721	0.571	0.521
Quality	Median	0.355	0.377	0.360	0.581	-	-	0.728	0.768	0.599	0.534
Low	Mean	0.112	0.135	0.109	0.228	0.212	0.198	0.173	0.269	0.147	0.162
Quality	Median	0.078	0.101	0.084	0.204	0.130	0.150	0.122	0.233	0.102	0.139

on the ROI size as higher cross-correlation values can be obtained for smaller regions subjected to comparison.

4 Conclusions and Future Work

Application of 2D cross-correlation for the quality assessment of 3D printed surfaces leads to very promising results. An interesting result is the possibility of reduction of computational burden due to the application of the Monte Carlo method. Nevertheless, its efficient use requires the additional step related to the phase adjustment of the positions of two compared image regions. Such operation is a crucial element of the proposed method leading to significant increase of the 2D cross-correlation values for high quality 3D prints.

The influence of image adjustment on the results obtained for poor quality prints is much smaller allowing proper automatic assessment of their quality. Increase of the density of scanning over 1200 dpi does not influence significantly on the obtained results. Nevertheless, tests conducted for the 300 dpi images have not led to satisfactory results.

One of the directions of our further research is related to the extension of the proposed method for lower resolution scans as well as the natural images captured by cameras, also with the application of non-square local ROIs and various color to grayscale conversion methods. However, a relevant limitation of the approach proposed in the paper is related to the necessity of visibility of layers and therefore its successful application for blurred images captured by a camera may be troublesome.

References

1. Chan, S.H., Zickler, T., Lu, Y.M.: Monte carlo non-local means: random sampling for large-scale image filtering. IEEE Trans. Image Process. **23**(8), 3711–3725 (2014)
2. Chauhan, V., Surgenor, B.: A comparative study of machine vision based methods for fault detection in an automated assembly machine. Procedia Manufact. **1**, 416–428 (2015)
3. Cheng, Y., Jafari, M.A.: Vision-based online process control in manufacturing applications. IEEE Trans. Autom. Sci. Eng. **5**(1), 140–153 (2008)

4. Fang, T., Jafari, M.A., Bakhadyrov, I., Safari, A., Danforth, S., Langrana, N.: Online defect detection in layered manufacturing using process signature. In: Proceedings of IEEE International Conference on Systems, Man and Cybernetics, vol. 5, pp. 4373–4378. San Diego, California, USA (1998)
5. Gallea, R., Ardizzone, E., Pirrone, R.: Monte-Carlo image retargeting. In: 2014 International Conference on Computer Vision Theory and Applications (VISAPP), vol. 1, pp. 402–408 (2014)
6. Gu, J., Peng, S., Wang, X.: Digital image inpainting using Monte Carlo method. In: 2004 International Conference on Image Processing, ICIP 2004, vol. 2, pp. 961–964 (2004)
7. International Telecommunication Union: Recommendation BT.601-7 - studio encoding parameters of digital television for standard 4: 3 and wide-screen 16: 9 aspect ratios (2011)
8. Marciniak, T., Lutowski, Z., Marciniak, B., Maszewski, M.: The use of texture analysis in optical inspection of manufacturing processes. In: Advances in Manufacturing Engineering II, Solid State Phenomena, vol. 237, pp. 95–100. TransTech Publications Ltd. (2015)
9. Okarma, K., Fastowicz, J.: No-reference quality assessment of 3D prints based on the glcm analysis. In: 2016 21st International Conference on Methods and Models in Automation and Robotics (MMAR), pp. 788–793 (2016)
10. Okarma, K., Fastowicz, J.: Quality assessment of 3D prints based on feature similarity metrics. In: Choraś, R.S. (ed.) Image Processing and Communications Challenges 8: 8th International Conference, IP&C 2016 Bydgoszcz, Poland, September 2016 Proceedings, pp. 104–111. Springer, Cham (2017)
11. Okarma, K., Fastowicz, J., Tecław, M.: Application of structural similarity based metrics for quality assessment of 3D prints. In: Chmielewski, L.J., Datta, A., Kozera, R., Wojciechowski, K. (eds.) Computer Vision and Graphics: International Conference, ICCVG 2016, Warsaw, Poland, 19–21 September 2016, Proceedings. Lecture Notes in Computer Science, vol. 9972, pp. 244–252. Springer, cham (2016)
12. Okarma, K., Lech, P.: Monte Carlo based algorithm for fast preliminary video analysis. In: Bubak, M., van Albada, G.D., Dongarra, J., Sloot, P.M.A. (eds.) Computational Science - ICCS 2008, Lecture Notes in Computer Science, vol. 5101, pp. 790–799. Springer, Heidelberg (2008)
13. Straub, J.: Initial work on the characterization of additive manufacturing (3D printing) using software image analysis. Machines **3**(2), 55–71 (2015)
14. Szkilnyk, G., Hughes, K., Surgenor, B.: Vision based fault detection of automated assembly equipment. In: Proceedings of ASME/IEEE International Conference on Mechatronic and Embedded Systems and Applications, Parts A and B. vol.3, pp. 691–697. Washington, DC, USA (2011)
15. Tourloukis, G., Stoyanov, S., Tilford, T., Bailey, C.: Data driven approach to quality assessment of 3D printed electronic products. In: Proceedings of 38th International Spring Seminar on Electronics Technology (ISSE), pp. 300–305. Eger, Hungary, May 2015

Quality Assessment of 3D Printed Surfaces in Fourier Domain

Jarosław Fastowicz(✉), Dawid Bąk, Przemysław Mazurek, and Krzysztof Okarma

Department of Signal Processing and Multimedia Engineering, Faculty of Electrical Engineering, West Pomeranian University of Technology, 26 Kwietnia 10, 71-126 Szczecin, Poland
{jfastowicz,przemyslaw.mazurek,okarma}@zut.edu.pl, dawid.d.bak@gmail.com

Abstract. In the paper the issue of quality assessment of 3D printed surfaces using image analysis is considered with particular attention paid to Fourier analysis of image fragments resized to one-dimensional vectors. Due to the application if Fourier analysis the regularity of visible patterns related to the consecutive layers of the filament can be assessed, assuming the side view of the printed surface. In order to avoid the problems of uneven lighting, our experiments have been conducted for scanned images of several 3D printed flat samples. As some of them have been contaminated by forced distortions, it is possible to classify them into two groups depending on the presence and amount of them. Due to the application of Fourier analysis some encouraging experimental results have been obtained which can be useful also for online monitoring of 3D prints quality for the images captured by cameras.

Keywords: 3D prints · Visual quality assessment · Fourier analysis · Image analysis

1 Introduction

Current progress of technology causes growing availability of relatively cheap 3D printers which can be useful in many various applications. Considering the time required for printing an object using the most popular Fused Deposition Modeling (FDM) based printers, one of the recent challenges in this area of science and technology is a reliable quality assessment of the 3D prints, preferably during the printing process. It may allow the correction of the printed surfaces in case of detected defects as well as stopping the printing in order to save the filament and time.

One of the most valuable directions of research for those purposes seems to be the application of image analysis allowing to determine the amount and location of possible distortions. Some researchers have proposed recently various solutions, useful mainly for fault detection [1,12,13]. Some older papers describing simpler solutions proposed by Fang [3] and Cheng [2] can also be considered as interesting directions of further work. The most advanced one seems

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*, Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_10

to be the system based on five cameras and Raspberry Pi modules developed by Straub [11], however its sensitivity to changing lighting conditions, camera motions and calibration errors decreases significantly its universality.

Nevertheless, according to our knowledge, there have been no attempts for visual quality assessment of 3D prints and the first papers directly related to this issue have been published during recent two or three years. The first attempts to quality assessment of 3D printed surfaces utilize time-consuming texture analysis using Gray-Level Co-occurrence Matrix (GLCM) [8], structural similarity based image quality metrics [10] and feature similarity metrics [9]. Considering the periodicity of the visible pattern caused by the presence of layers of the filament, the application of Fourier analysis may also be an interesting direction of research in this regard, similarly as used in some other applications [7]. Some similar applications based in texture analysis can also be observed in industry e.g. for milling [4].

2 The Idea of the Fourier Based Method

The main purpose of the investigated approach is the ability to estimate the quality of the 3D printed surface in no-reference way i.e. without the knowledge of the perfect quality reference print. An initial verification of the usefulness of Fourier analysis has been made assuming the choice of two different fragments of varying quality 3D printed surface. After conversion to grayscale, the analyzed fragments have been averaged into a single line perpendicular to layers of the filament. Is has been assumed that for high quality print the regular pattern should result in explicit peaks in Fourier domain whereas in the presence of structural distortions such a dominance should not take place.

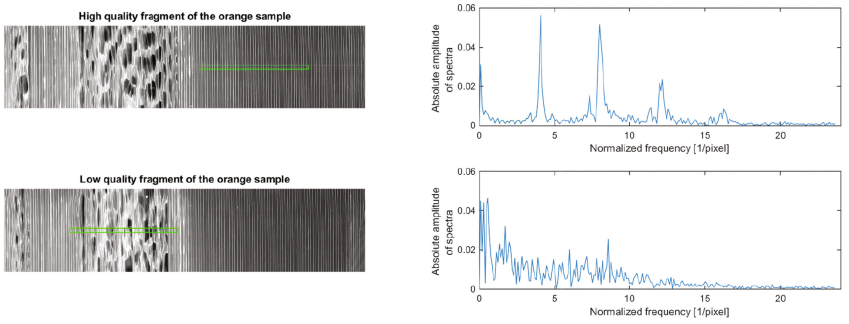


Fig. 1. Illustration of the analyzed fragments of two sides of an orange sample (left) and the obtained spectra (right)

The results of the verification of the idea are presented in Fig. 1 for an orange sample using the averaging of $N = 20$ lines whereas the results for two silver samples (also for $N = 20$ lines) are shown in Fig. 2. Both scanned images are

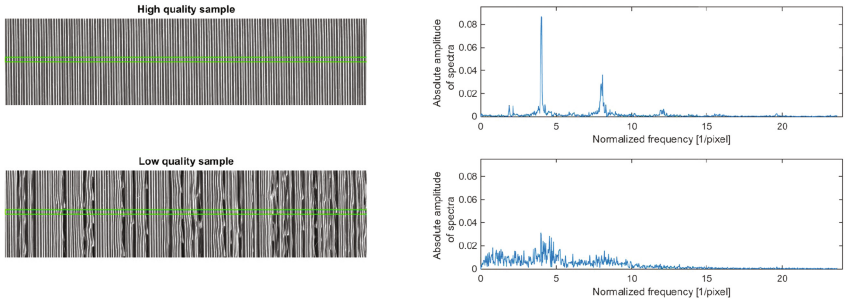


Fig. 2. Illustration of the analyzed fragments of two silver samples of different quality (left) and the obtained spectra (right)

rotated by 90° for better visibility. The upper plots and images correspond to high quality fragments of an orange sample and high quality silver one.

As can be easily noticed, the numerical value allowing to determine the quality of analyzed fragments of samples may be the integral of the spectra values calculated e.g. using the trapezoidal rule. Obtained higher values denote the low quality of the analyzed sample whereas low values are characteristic for the domination of frequencies related to the visible pattern of consecutive layers.

Quality evaluation of the whole surface of the 3D print requires more comprehensive analysis and therefore the whole analyzed image has been divided into fragments containing a specified number of N lines further averaged into single lines with noticeable oscillations. Next, Fourier transform has been applied in order to calculate the spectra for each of such obtained lines. They can be analyzed separately but the most useful information corresponds to minimum, maximum and average values of the integral of the spectra.

Since the images obtained for various colors of the filament after conversion from color to grayscale (typically ITU BT.601-7 Recommendation [5] is used) may be characterized by different overall brightness and contrast, the normalization of the luminance range has been applied in the preprocessing step. However, the luminance values obtained for the white filament have been significantly higher so the modification of the integral based proposed quality indicator has been proposed based on the use of average image luminance. The modified formula can be expressed as:

$$Q = \frac{1}{I_{avg}} \cdot log \frac{\sum_i (X_i)}{I_{avg}} \tag{1}$$

where I_{avg} is the average image luminance and X_i are the obtained values of the spectra.

3 Discussion of Results

The results obtained for tested 18 samples are illustrated in Fig. 3 where blue color denotes high quality samples and red - low quality ones. Symbols "*" stand

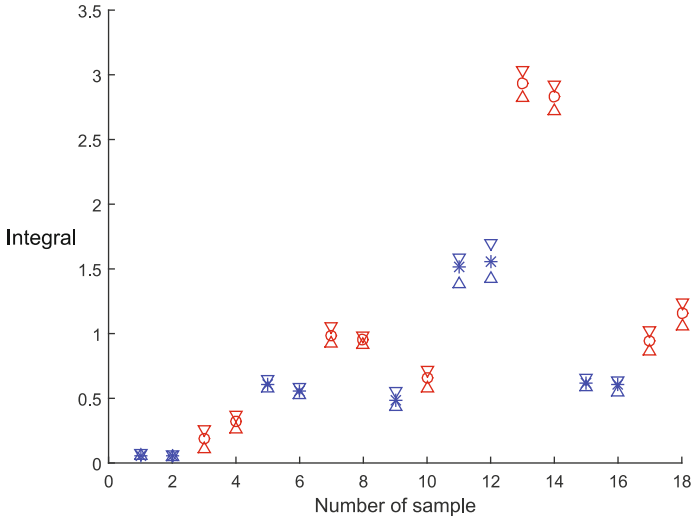


Fig. 3. Obtained values of the integral for $N = 10$ and color to grayscale conversion according to ITU BT.601-7 Recommendation - blue color denotes high quality samples whereas red - low quality ones

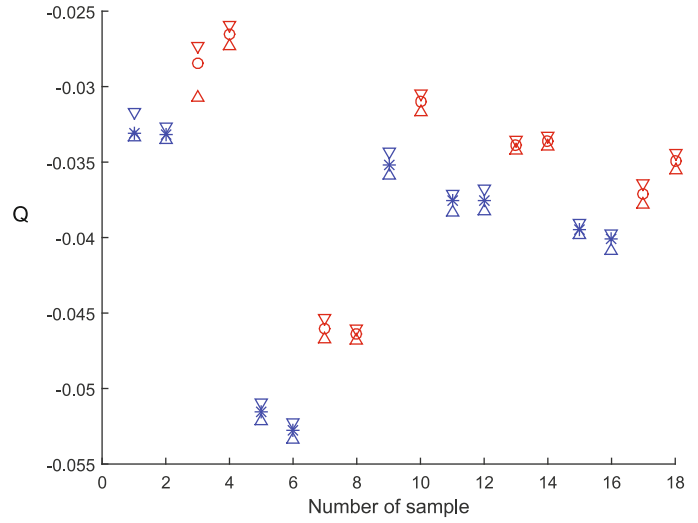


Fig. 4. Obtained values of the proposed quality indicator for $N = 10$ and color to grayscale conversion according to ITU BT.601-7 Recommendation - blue color denotes high quality samples whereas red - low quality ones

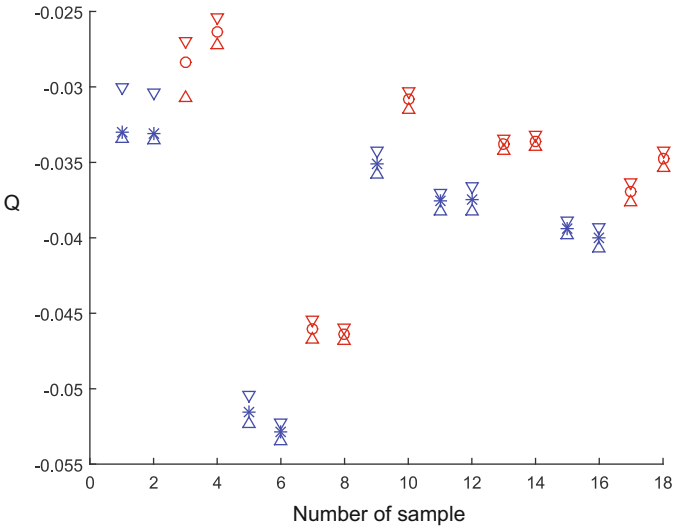


Fig. 5. Obtained values of the proposed quality indicator for $N = 5$ and color to grayscale conversion according to ITU BT.601-7 Recommendation - blue color denotes high quality samples whereas red - low quality ones

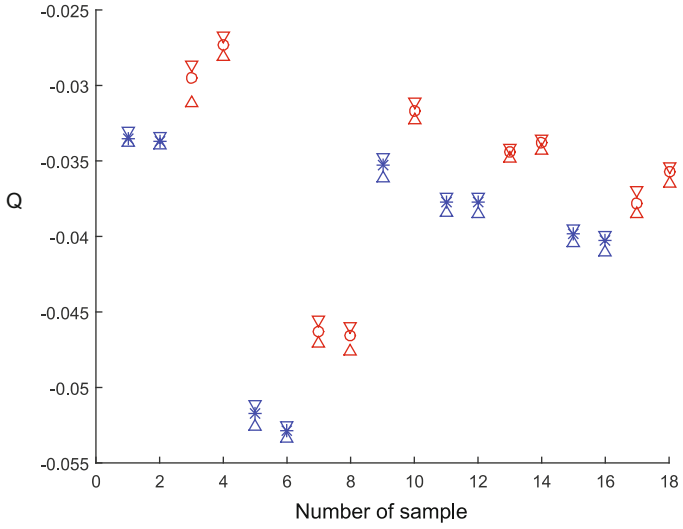


Fig. 6. Obtained values of the proposed quality indicator for $N = 50$ and color to grayscale conversion according to ITU BT.601-7 Recommendation - blue color denotes high quality samples whereas red - low quality ones

for the average values of the quality indicator obtained for high quality samples whereas "**o**" represents average values for low quality 3D prints. Triangles stand for minimum and maximum values respectively. The colors of samples are: white (1–4), red (5–8), orange (9–10), silver (11–14) and green (15–18).

As can be observed, the results of the integral are strongly dependent one the color of the sample. However, a noticeable progress can be made applying the formula (1) shown in Fig. 4 where the results are presented in the same way. Regardless of the modifications of the formula, obtained results are still dependent on the color of the sample. Nevertheless, interpretation of results is much easier and a distinct threshold can be determined between the high and low quality samples for each color of the filament. It is worth noticing that the values obtained for the red filament differ significantly from the others and therefore there is a necessity of using different threshold values.

A relevant issue is also the determination of the number of averaged lines N. As can be observed in Fig. 5, the values of N significantly smaller than the period of oscillations (in our case 12 pixels) may lead to worse classification results for white filament. However, the increase of the number of averaged lines N does not decrease the classification accuracy as shown in Fig. 6.

4 Conclusions and Future Work

Proposed approach to quality evaluation of the 3D printed surfaces has been tested using 18 samples created using various colors of filament. The most troublesome color seem to be white and the reason is related to high overall brightness of the samples. Nevertheless, applying the proposed method, promising results have been achieved, although dependent of the filament's color.

Extension of the proposed approach may lead to the detection of some periodic distortions as well as changes of thickness of the layers. Further detection of sample geometry distortions (deflection of whole or a part of sample) by comparing the frequency analysis results of several horizontal sample sections may also be possible. Due to the extraction of the dominant frequency it is also possible to determine the thickness of the layers and - indirectly - some of the parameters of the 3D printer.

One of the directions of our future research will be related to the application of the proposed method for the natural images captured by the camera. Considering the different distribution of light, the necessity of using some other color to grayscale conversion methods may occur. Another interesting idea may also be the application of the 2D FFT which has been successfully used recently for general image quality assessment purposes [6].

References

1. Chauhan, V., Surgenor, B.: A comparative study of machine vision based methods for fault detection in an automated assembly machine. Procedia Manuf. **1**, 416–428 (2015)

2. Cheng, Y., Jafari, M.A.: Vision-based online process control in manufacturing applications. IEEE Trans. Autom. Sci. Eng. **5**(1), 140–153 (2008)
3. Fang, T., Jafari, M.A., Bakhadyrov, I., Safari, A., Danforth, S., Langrana, N.: Online defect detection in layered manufacturing using process signature. In: Proceedings of IEEE International Conference on Systems, Man and Cybernetics, vol. 5, pp. 4373–4378. San Diego, California, USA (1998)
4. Herwig, J., Leßmann., Büijrger, F., Pauli, J.: Adaptive anomaly detection within near-regular milling textures. In: 2013 8th International Symposium on Image and Signal Processing and Analysis (ISPA), pp. 113–118 (2013)
5. International Telecommunication Union: Recommendation BT.601-7 - studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios (2011)
6. Khorrami, M., Azimzadeh, Z., Nabipour, S.: Reduced-reference image quality assessment based on 2-D discrete FFT and edge similarity. In: 2015 9th Iranian Conference on Machine Vision and Image Processing (MVIP), pp. 24–28 (2015)
7. Mkayes, A.A., Saad, N.M., Faye, I., Walter, N.: Image histogram and FFT based critical study on noise in fluorescence microscope images. In: 2016 6th International Conference on Intelligent and Advanced Systems (ICIAS), pp. 1–4 (2016)
8. Okarma, K., Fastowicz, J.: No-reference quality assessment of 3D prints based on the GLCM analysis. In: 2016 21st International Conference on Methods and Models in Automation and Robotics (MMAR), pp. 788–793 (2016)
9. Okarma, K., Fastowicz, J.: Quality assessment of 3D prints based on feature similarity metrics. In: Chorań, R.S. (ed.) Proceedings of Image Processing and Communications Challenges 8: 8th International Conference, IP&C 2016 Bydgoszcz, Poland, pp. 104–111. Springer, Heidelberg (2017). September 2016
10. Okarma, K., Fastowicz, J., Tecław, M.: Application of structural similarity based metrics for quality assessment of 3D prints. In: Chmielewski, L.J., Datta, A., Kozera, R., Wojciechowski, K. (eds.) Proceedings of Computer Vision and Graphics: International Conference, ICCVG 2016, Warsaw, Poland, 19–21 September 2016. Lecture Notes in Computer Science, vol. 9972, pp. 244–252. Springer, Cham (2016)
11. Straub, J.: Initial work on the characterization of additive manufacturing (3D printing) using software image analysis. Machines **3**(2), 55–71 (2015)
12. Szkilnyk, G., Hughes, K., Surgenor, B.: Vision based fault detection of automated assembly equipment. Proceedings of ASME/IEEE International Conference on Mechatronic and Embedded Systems and Applications, Parts A and B, Washington, DC, USA **3**, 691–697 (2011)
13. Tourloukis, G., Stoyanov, S., Tilford, T., Bailey, C.: Data driven approach to quality assessment of 3D printed electronic products. In: Proceedings of 38th International Spring Seminar on Electronics Technology (ISSE), Eger, Hungary, pp. 300–305, May 2015

The AdaBoost Algorithm with Linear Modification of the Weights

Robert Burduk(✉)

Department of Systems and Computer Networks, Wroclaw University of Science and Technology, Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland
robert.burduk@pwr.edu.pl

Abstract. This paper presents a new extension of the AdaBoost algorithm. This extension concerns the weights used in this algorithm. In our approach the original weights are modified, we propose a linear modification of the weights. In our study we use the boosting by the reweighting method where each weak classifier is based on the linear classifier. The described algorithm was tested on Pima data set. The obtained results are compared with the original the AdaBoost algorithm.

Keywords: AdaBoost algorithm · Ensemble of classifiers · Linear classifier

1 Introduction

Boosting is a machine learning effective method of producing a very accurate classification rule by combining a weak classifier [7]. The weak classifier is defined to be a classifier which is only slightly correlated with the true classification i.e. it can classify the object better than a random classifier. In boosting, the weak classifier is learned from various training examples sampled from the original learning set. The sampling procedure is based on the weight of each example. In each iteration, the weights of examples are changing. The final decision of the boosting algorithm is determined on the ensemble of classifiers derived from each iteration of the algorithm. One of the fundamental problems of the development of different boosting algorithms is choosing the weights and to define rules for an ensemble of classifiers. In recent years, many authors presented various concepts based on the boosting idea [6,9]. In this article we present a new extension of AdaBoost [5] algorithm in which a linear modification of the weights was applied.

This paper is organized as follows: Sect. 2 introduces the necessary terms of the AdaBoost algorithm. In the next section there is our modification of the presented algorithm. Section 4 presents the experiment results comparing AdaBoost with our modification. Finally, some conclusions are presented.

2 AdaBoost Algorithm

In the work [5] weak and strong learning algorithms were discussed. The weak algorithms can classify the object better than random, on the other hand strong

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*, Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_11

algorithms can classify the object accurately. Schapire formulated the first algorithm to "boost" a weak classifier. The main idea of boosting is to improve the prediction of weak learning algorithms by creating a set of weak classifiers which is a single strong classifier. The well-known and widely applied is the AdaBoost algorithm. Its main steps are as follows [2] (Tables 1 and 2):

Table 1. AdaBoost algorithm

1.	Initialize the weight vector $w_{1,1} = ... = w_{1,n} = 1/n$.
2.	Assign weights to the learning sample LS_n.
3.	For $t = 1, 2, ..., T$:
a.	At the base of LS_n learn the classifier Ψ_t,
b.	Calculate the classification error $e_t = \sum_{i=1}^{n} w_{t,i} * I(\Psi_t(x_i) \neq j_i)$,
c.	Calculate $c_t = \frac{\ln(1-e_t)}{e_t}$
d.	Update weights: $w_i(t+1) = \frac{w_{t,i} \exp(c_t * I(\Psi_t(x_i) \neq j_i))}{\sum_{j=1}^{n}(w_{t,j} \exp(c_t * I(\Psi_t(x_i) \neq j_i)))}, \quad i = 1, ..n,$
e.	Assign updated weights to the learning sample LS_n.
4.	Classify observation x according to the rule: $\Psi_{AdaBoost}(x) = sign\left(\sum_{t=1}^{T} c_t \Psi_b(x)\right)$.

Table 2. Notation of the AdaBoost algorithm

i	Observation number, $i = 1, ..., n$.
t	Stage number, $t = 1, ..., T$.
x_i	A p-dimensional vector containing the quantitative variables of the ith observation.
Ψ_t	The weak classifier at the tth stage.
$w_{t,i}$	The weight of the ith observation at the tth stage, $\sum_i w_{t,i} = 1$.
I	The indicator function, $I(\Psi_t(x_i) \neq j_i)$.
e_t	The classification error at the tth stage, $\sum_i w_{t,i} I_{t,i}$.
c_t	The weight of e_t.
sign(x)	$= 1$ if $x \geq 0$ and $= -1$ otherwise.

One of the main steps in the algorithm is to maintain a distribution of the training set using the weights. Initially, all weights of the training set observations are set equally. If an observation is incorrectly classified (at the current stage) the weight of this observation is increased. Similarly, the correctly classified observation receives less weight in the next step. For this reason the weak learner is forced to focus on the and examples from the training set in each next step of the algorithm. In each step of the AdaBoost algorithm the best weak classifier according to the current distribution of the observation weights is found.

The goodness of a weak classifier is measured by its error. Based on the value of this error the ratio is calculated. The final prediction of the AdaBoost algorithm is a weighted majority vote of all weak classifiers.

3 AdaBoost Algorithm with Linear Modification of the Weights

One of the main factors that have an effect on the action of the AdaBoost algorithm is the selection of weights assigned to individual elements of the learning set. Let's propose then a modification of the AdaBoost algorithm in which a linear modification of the weights is introduced. The value of factor c_t is modified in point 4d. The value of the modification depends on the iteration of the algorithm t. In experimental studies we have assumed that the value of the coefficient after modification (point 4d) is $1.25, 1.5, 1.75, 2, 2.25$ or 2.5 times higher in the first iteration compared to the original algorithm (point 4c). The proposed algorithm steps are presented in Table 3.

Table 3. Lmw-AdaBoost algorithm

1. Initialize a and b.
2. Initialize the weight vector $w_{1,1} = ... = w_{1,n} = 1/n$.
3. Assign weights to the learning sample LS_n.
4. For $t = 1, 2, ..., T$:
   a. At the base of LS_n learn the classifier Ψ_t,
   b. Calculate the classification error $e_t = \sum_{i=1}^{n} w_{t,i} * I(\Psi_t(x_i) \neq j_i)$,
   c. Calculate $c_t = \frac{\ln(1-e_t)}{e_t}$
   d. Calculate $c_t = c_t * (a * t + b)$
   e. Update weights:
      $w_i(t+1) = \frac{w_{t,i} \exp(c_t * I(\Psi_t(x_i) \neq j_i))}{\sum_{j=1}^{n}(w_{t,j} \exp(c_t * I(\Psi_t(x_i) \neq j_i)))}, \quad i = 1, ..n,$
   f. Assign updated weights to the learning sample LS_n.
5. Classify observation x according to the rule:
   $\Psi_{Lmw-AdaBoost}(x) = sign\left(\sum_{t=1}^{T} c_t \Psi_t(x)\right)$.

In the earlier work [1] we proposed the changes in weights based on interval-valued fuzzy sets and in the work [11] the linear combination of the upper and lower value of the weights to brain-computer interface was applied.

4 Experiments

To test Lmw-AdaBoost algorithm, we performed experiments on Pima data set. The feature selection process [10] was performed to indicate four most informative features for this data set. The final results are obtained via the 10-fold-cross-validation method.

The results for the twenty-five iterations of AdaBoost and proposed Lmw-AdaBoost algorithms are presented in Table 4.

Table 4. The results of experiments

Iter.	AdaBoost	Lmw-AdaBoost					
		$1.25 * c_t$	$1.5 * c_t$	$1.75 * c_t$	$2 * c_t$	$2.25 * c_t$	$2.5 * c_t$
1	0.280	0.280	0.280	0.280	0.280	0.280	0.280
2	0.280	0.280	0.280	0.280	0.593	0.720	0.720
3	0.277	**0.277**	0.279	0.358	0.283	0.279	0.279
4	0.263	**0.262**	0.266	0.271	0.311	0.720	0.721
5	0.264	0.267	0.299	0.318	**0.263**	0.279	0.277
6	0.262	**0.249**	0.263	0.271	0.302	0.721	0.721
7	0.259	0.270	**0.251**	0.28	0.289	0.279	0.277
8	0.266	0.262	**0.254**	0.263	0.271	0.720	0.723
9	0.262	0.257	0.248	**0.246**	0.332	0.279	0.277
10	0.262	0.267	0.249	**0.246**	0.275	0.720	0.723
11	0.267	0.251	**0.250**	0.271	0.274	0.280	0.277
12	0.264	**0.254**	0.255	0.259	0.262	0.279	0.721
13	0.263	0.258	**0.251**	0.263	0.270	0.361	0.277
14	0.259	0.259	**0.250**	0.263	0.266	0.312	0.721
15	0.260	0.255	**0.254**	0.266	0.279	0.324	0.279
16	0.260	0.257	**0.249**	0.263	0.263	0.307	0.721
17	0.262	0.257	**0.251**	0.258	0.267	0.264	0.270
18	0.260	0.257	**0.251**	0.253	0.258	0.257	0.284
19	0.262	0.257	**0.251**	0.258	0.255	0.257	0.270
20	0.262	0.257	0.251	0.257	**0.244**	0.258	0.259
21	0.262	0.257	0.251	0.257	**0.249**	0.254	0.329
22	0.262	0.257	0.251	0.255	**0.250**	0.259	0.286
23	0.262	0.257	0.251	0.257	**0.249**	0.259	0.297
24	0.262	0.257	0.251	0.255	**0.246**	0.257	0.279
25	0.262	0.257	0.251	0.257	**0.245**	0.262	0.264

The best results (since the third iteration) are in bold. In general the results for the AdaBoost algorithm are worse than the proposed modifications Lmw-AdaBoost. For the first twelve iterations no clear results were obtained. In the iterations 13–19 the best is the algorithm in which the primary coefficient c_t is increased 1.5 times in the first iteration. In the last iteration this coefficient is unchanged therefore the parameters a and b are equal -0.020833333 1.541666667 respectively. In recent iterations the best is the algorithm in which parameters a

and b are equal -0.041666667 2.041666667 respectively. With such parameters in the first iteration coefficient c_t is increased 2 times. The obtained results show an improvement in the quality of the proposed modification the AdaBoost algorithm with respect to the ordinal one.

5 Conclusions

In this paper we presented the new Lmw-AdaBoost algorithm. It is a modification of the AdaBoost algorithm in which it was changed coefficient c_t. Consequently, the change affects the weights assigned to the individual learning objects. Changes compared to the original algorithm are linear. The value of the change is greater in the initial iterations.

The experiments have been carried out on Pima data sets. The aim of the experiments was to compare the proposed algorithm and the original AdaBoost algorithm. The results obtained show an improvement in the classification quality of the proposed method with respect to the original one.

Future work might include the proposed modification in other boosting algorithms such as RealAdaBoost or GentleAdaBoost as well as application of the proposed methods for various practical task [3,4,8] or testing other data sets.

Acknowledgments. This work was supported by the statutory funds of the Department of Systems and Computer Networks, Wroclaw University of Science and Technology.

References

1. Burduk, R.: The AdaBoost algorithm with the imprecision determine the weights of the observations. In: Asian Conference on Intelligent Information and Database Systems, pp. 110–116. Springer, Cham (2014)
2. Dmitrienko, A., Chuang-Stein, C., D'Agostino, R.B.: Pharmaceutical statistics using SAS: a practical guide. SAS Institute (2007)
3. Forczmański, P., Łabędź, P.: Recognition of occluded faces based on multi-subspace classification. In: Computer Information Systems and Industrial Management, pp. 148–157. Springer, Heidelberg (2013)
4. Frejlichowski, D.: An algorithm for the automatic analysis of characters located on car license plates. In: International Conference Image Analysis and Recognition, pp. 774–781. Springer, Heidelberg (2013)
5. Freund, Y., Schapire, R.E.: A desicion-theoretic generalization of on-line learning and an application to boosting. In: European Conference on Computational Learning Theory, pp. 23–37. Springer, Heidelberg (1995)
6. Freund, Y., Schapire, R.E., et al.: Experiments with a new boosting algorithm. ICML **1996**, 148–156 (1996)
7. Kearns, M., Valiant, L.: Cryptographic limitations on learning boolean formulae and finite automata. J. ACM (JACM) **41**(1), 67–95 (1994)
8. Kozik, R., Choraś, M.: The HTTP content segmentation method combined with adaboost classifier for web-layer anomaly detection system. In: International Conference on EUropean Transnational Education, pp. 555–563. Springer, Heidelberg (2016)

9. Oza, N.C.: Boosting with averaged weight vectors. Multiple Classifier Systems **2709**, 15–24 (2003)
10. Rejer, I.: Genetic algorithms for feature selection for brain-computer interface. Int. J. Pattern Recogn. Artif. Intell. **29**(5), 1559008 (2015)
11. Rejer, I., Burduk, R.: Classifier selection for motor imagery brain computer interface. In: IFIP International Conference on Computer Information Systems and Industrial Management, pp. 122–130. Springer, Heidelberg (2017)

Evaluation of Hashing Methods Performance on Binary Feature Descriptors

Jacek Komorowski(✉) and Tomasz Trzciński

Warsaw University of Technology, Warsaw, Poland
jacek.komorowski@gmail.com, t.trzcinski@ii.pw.edu.pl

Abstract. In this paper we evaluate performance of data-dependent hashing methods on binary data. The goal is to find a hashing method that can effectively produce lower dimensional binary representation of 512-bit FREAK descriptors. A representative sample of recent unsupervised, semi-supervised and supervised hashing methods was experimentally evaluated on large datasets of labelled binary FREAK feature descriptors.

Keywords: Data-dependent hashing methods · Binary feature descriptors

1 Introduction

This paper presents results of an experimental evaluation of recent data-dependent hashing methods applied to binary feature descriptors. The work was motivated by challenges in development of a real-time structure-from-motion solutions for mobile platforms with limited hardware resources. One of the key elements in a typical structure-from-motion processing pipeline is feature matching step, where correspondences between features detected on a new image and features found on previously processed images are being sought. Such correspondences are used to compute camera orientation and build a 3D model of an observed scene.

Binary feature descriptors, such as FREAK [1], are good choice for mobile solutions. They can be efficiently computed and are very compact (512 bits for FREAK versus 512 bytes for real-valued SIFT descriptor). Hamming distance between two binary feature descriptors can be quickly computed using few machine code instructions. But even if comparison of two binary descriptors is very fast, finding correspondence between thousands of features detected in a new image and millions of features on previously processed images requires significant processing power. For real-valued descriptors efficient approximate nearest neighbours search methods can be applied, such as FLANN [12]. Unfortunately methods based on clustering do not work well with binary data [13].

In this paper we investigate if hashing methods can be used to reduce dimensionality of 512-bit binary FREAK descriptors to improve feature matching performance and lower storage requirements. Additionally we want to verify if additional

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*,
Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_12

training information, if the form of landmark id linked with each descriptor, can be used to improve accuracy of searching for matching descriptors.

Authors have not encountered any results of evaluation of hashing methods on binary data. In the context of image-based information retrieval, hashing algorithms were evaluated on real-valued descriptors such as GIST or SIFT. Lack of such results motivated the research described in this paper.

The paper is structured as follows. Section 2 briefly describes hashing methods. Section 3 presents results of the experimental evaluation of representative hashing methods on large datasets of binary FREAK feature descriptors. The last section concludes the article and presents ideas for future research.

2 Overview of Hashing Methods

Hashing for similarity search is a very active area of development. Many new hashing methods were proposed in the last few years. A number of surveys documenting current state-of-the-art was published recently [15,18].

Two major categories of hashing methods can be distinguished: *data-independent* methods and *data-depended* methods. *Data-independent* methods, such as Locality Sensitive Hashing (LSH) [5], do not take into account characteristics of the input data. As such, they have inferior performance in real-live applications, where input data usually has some intrinsic characteristic which can be exploited. We focus our attention on data-dependent on *data-depended* methods, also known as *learning to hash*. These methods exploit properties of the input data to produce more discriminative and compact binary codes. Data-dependent approach can be further categorized by the level of an external supervision. Unsupervised methods use techniques as spectral analysis or kernelized random projections to compute affinity-preserving binary codes. They exploit the structure among a sample of unlabelled data to learn appropriate embeddings. Semi-supervised or supervised methods exploit additional information from annotated training data. Additional information is usually given as the similarity matrix or list of pairs of similar and dissimilar items. Semi-supervised methods assume that explicit similarity information is provided for only a fraction of an input dataset. Affinity between other elements is inferred from the distance in the input space.

Learning to hash is defined [14] as learning a **compound hash function**, $\mathbf{y} = H(\mathbf{x})$, mapping an input item $\mathbf{x}$ to a compact binary code $\mathbf{y}$, such that nearest neighbour search in the coding space is efficient and the result is a good approximation of the true nearest search result in the input space. K-bit binary code $\mathbf{y} \in \mathbb{B}^K$ for a sample point $\mathbf{x} \in \mathbb{R}^D$ is computed as $\mathbf{y} = [y_1,\ y_2, \cdots y_K,] = [h_1(\mathbf{x}), h_2(\mathbf{x}), \cdots, h_K(\mathbf{x})]$. Each h_k is a **binary hash function**, mapping elements from $\mathbb{R}^D$ to $\mathbb{B} = \{0, 1\}$ A compound hash function $H = [h_1, h_2, \cdots, h_K]$ is an ordered set of binary hash functions computing K-bit binary code.

Two most popular choices of a **hash function** are linear projection and kernel-based. Linear projection hash functions are in the form: $y = h(\mathbf{x}) = \mathrm{sgn}(\mathbf{w}^{\mathrm{T}}\mathbf{x}+\mathrm{b})$, where $\mathbf{w} \in \mathbb{R}^D$ is the projection vector and b is the bias. Kernel-based hash functions are in the form: $y = h(\mathbf{x}) = \mathrm{sgn}\left(\sum_{\mathrm{t}=1}^{\mathrm{T}} \mathrm{w_t K}\left(\mathbf{s_t}, \mathbf{x}\right) + \mathrm{b}\right)$, where K is a kernel function, $\mathbf{s}$ is a set of representative samples that are randomly chosen from the dataset or cluster centres of the dataset and w_t are weights. Other choices of hash function include spherical functions, Laplacian eigenfunctions, neural networks, decision trees-based and non-parametric functions.

3 Evaluation Results

Experiment setup. The experimental evaluation was conducted on datasets consisting of hundred thousands or more labelled 512-bit FREAK feature descriptors. Datasets were created by structure-from-motion application developed in Google Tango project.[1] Each descriptor is labelled with a corresponding landmark id. Descriptors with the same landmark id are projections of the same scene point (landmark) on different images.

Table 1 lists hashing methods evaluated in this paper. Each method was first trained on the training dataset. Learned hash functions were applied to the test dataset to generate hash codes of a different length: 32, 64, 128 and 256-bits. Search precision (*Precision@1*) using the resulting hash codes was evaluated and reported. *Precision@1* for a dataset is calculated as a mean *precision@1* when searching for nearest neighbours using a linear scan for 20 thousand elements randomly sampled from the dataset. *Precision@1* for a sampled element is 1, if its nearest neighbour (based on Hamming distance) in the entire dataset is labelled with the same landmark id. Otherwise, precision is 0.

Unsupervised hashing methods. This section contains results of an experimental evaluation of unsupervised hashing methods. The evaluation was conducted using publicly available MATLAB implementation [2]. For brief description of methods evaluated in this section refer to [2].

Hashing methods were trained using a dataset consisting of 1 267 346 FREAK descriptors. The evaluation was done on a separated test dataset consisting of 342 602 descriptors associated with 40 704 landmarks.

Figure 1 presents search precision in datasets created from the test dataset by applying seven unsupervised, data-dependent hashing methods. To baseline the results, precision of a linear search on the test dataset truncated to first k bits was evaluated. It must be noted, that discarding last 256 bits of original FREAK descriptor has little effect on the nearest neighbour search precision. *Precision@1* decreases by 2.1% from 0.962 to 0.942. The reason is likely the construction of FREAK descriptor itself, where more discriminative binary tests are used to generate first bits of the descriptor. Interestingly, for 256-bit hash codes, all evaluated hashing methods perform worse compared to naive bit truncation.

[1] See: https://get.google.com/tango/.

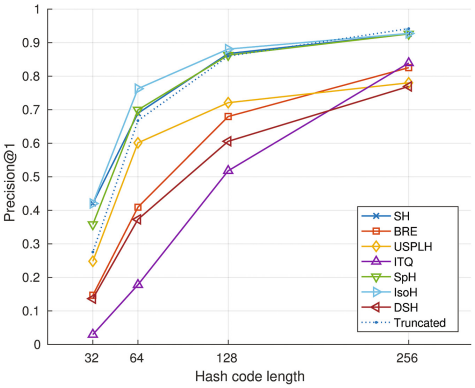
Table 1. List of evaluated hashing methods. Class: U = unsupervised, SS = semi-supervised, S = supervised.

Method	Class	Hash function
Spectral Hashing (SH) [19]	U	Laplacian eigenfunction
Binary Reconstructive Embeddings (BRE) [7]	U	Linear projection
Unsupervised Sequential Projection Learning Hashing (USPLH) [17]	U	Linear projection
Iterative Quantization (ITQ) [3]	U	Linear projection
Isotropic Hashing (IsoH) [6]	U	Linear projection
Density Sensitive Hashing (DSH) [10]	U	Linear projection
Spherical Hashing (SpH) [4]	U	Spherical function
Compressed Hashing [11]	U	Kernel-based
Harmonious Hashing (HamH) [21]	U	Kernel-based
Kernelized Locality Sensitive Hashing (KLSH) [8]	U	Kernel-based
Sequential Projection Learning (SPLH) [16]	SS	Linear projection
Bootstrap Sequential Projection Learning – linear version (BTSPLH) [20]	SS	Linear projection
Fast supervised hashing (FastHash) [9]	S	Boosted decision trees

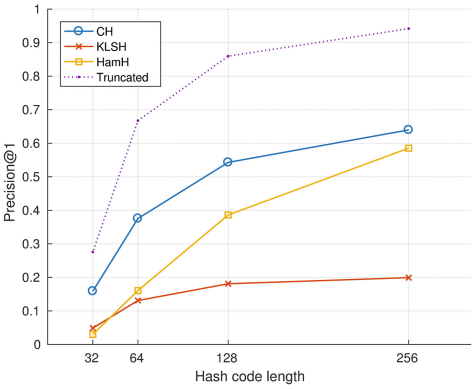
For shorter codes (128-bits and below) only three hashing methods: Isotropic Hashing [6], Spherical Hashing [4] and Spectral Hashing [19] yield better accuracy. For 128-bit codes, the advantage of best methods over a naive bit truncation is minimal (1–2%). For the shorter codes best hashing methods perform noticeably better. Best performing method is Isotropic Hashing (IsoH) [6].

Performance of kernel-based hashing methods on the binary data is very poor. Results are much worse than naive approach of truncating an original dataset to the first k bits. This is expected as kernel-based hash functions are based on the set of anchor points, that are representative samples or cluster centres for the training dataset. Clustering methods perform poorly on data from binary spaces. One of the reasons are decision boundaries in binary spaces [13], as large proportion of points in the Hamming space is equidistant from two randomly chosen anchor points.

Semi-supervised and supervised hashing methods Semi-supervised and supervised data-dependent hashing methods evaluated in this section were trained using the dataset consisting of 340 063 descriptors associated with 37 148 landmarks. The evaluation was done on the separated test dataset consisting of 342 602 descriptors associated with 40 704 landmarks.



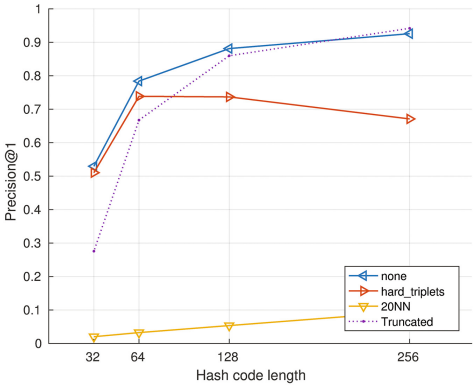
(a) Non-kernel based



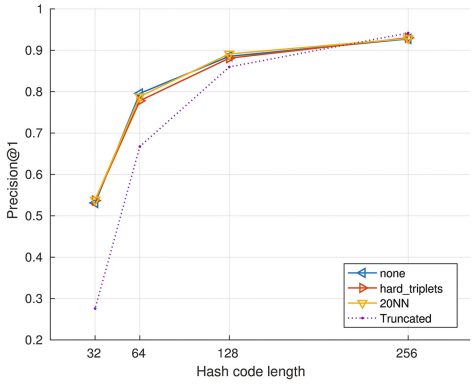
(b) Kernel based methods

Fig. 1. Search precision using hash codes generated from the test dataset by applying unsupervised hashing methods. Search precision in the test dataset truncated to first k bits is plot with a dotted line for comparison.

Figure 2 presents search precision in datasets created from the training dataset by applying semi-supervised Sequential Projection Learning (SPLH) [16] method. SPLH objective function of consists of two components: supervised empirical fitness and unsupervised information theoretic regularization. A supervised term tries to minimize empirical error on the labelled data. That is, for each bit minimize a number of instances where elements with the same label are mapped to different values and elements with different labels are mapped to the same value. An unsupervised term provides regularization by maximizing desirable properties like variance and independence of individual bits. Different lines on the plot correspond to different similarity encoding schemes. In *hard triplets* encoding, for each element from the training dataset, the closest element linked with the same landmark id is encoded as similar and the closest element



(a) $\eta = 1$



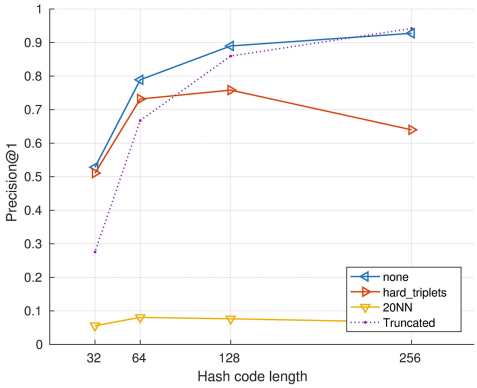
(b) $\eta = 100$

Fig. 2. Search precision using hash codes generated from the test dataset by semi-supervised SPLH method. Different subplots correspond to different weight η of an unsupervised term in the objective function. On each plot results obtained using three different similarity encodings schemes are shown. Search precision in the test dataset truncated to first k bits is plot with a dotted line for comparison.

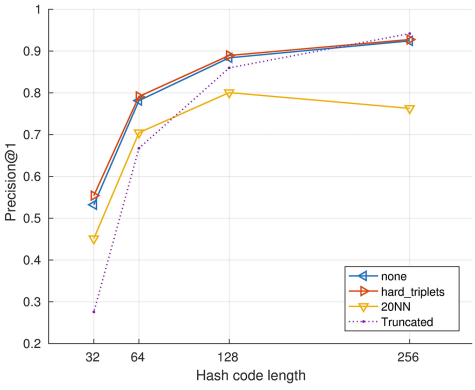
with a different landmark id is encoded as dissimilar. In *20NN* encoding, for each element from the training dataset, the similarity with its 20 nearest neighbours is encoded. For comparison, SPLH method was evaluated without any similarity information(*none*), using only unsupervised term in the optimization function. Surprisingly, using supervised information does not provide any improvement in the search accuracy. In contrary, when supervised term in the objective function has higher weight (lower η, left subplot), the results are noticeably worse, especially for longer codes. The best results are achieved when supervised information is not used at all (*none*). When unsupervised term in the objective function has higher weight (higher η, right subplot), the encoded similarity information has

little effect and the performance is the same as without using any supervised information.

Semi-supervised nonlinear hashing (BTSPLH) [20] is an enhancement of SPLH method. Instead of the boosting-like process in SPLH, authors propose a bootstrap-style sequential learning scheme to derive the hash function by correcting the errors incurred by all previously learned bits. The results of evaluation of BTSPLH method are depicted on Fig. 3. The results are very similar to previous SPLH method. When supervised term in the objective function has higher weight (lower λ, left subplot), the results using supervised information



(a) $\lambda = 1$



(b) $\lambda = 10$

Fig. 3. Search precision using hash codes generated from the test dataset by semi-supervised BTSPLH method. Different subplots correspond to different weight λ of an unsupervised term in the objective function. On each plot results obtained using three different similarity encodings schemes are shown. Search precision in the test dataset truncated to first k bits is plot with a dotted line for comparison.

(*hard_triplets* and *20NN* encoding) are poor. The best results are achieved when supervised information is not used at all.

The last method evaluated in this paper is supervised FastHash [9] algorithm. It uses a two-step learning strategy: binary code inference is followed by binary classification step using an ensemble of decision trees.

The key decision when using supervised hashing methods is a choice of similarity encoding scheme. For large datasets it's not feasible to encode similarity between all $N \times N$ pairs of elements. Figure 4(a) presents results when for each element from the training dataset, similarity for its 20 nearest neighbours,

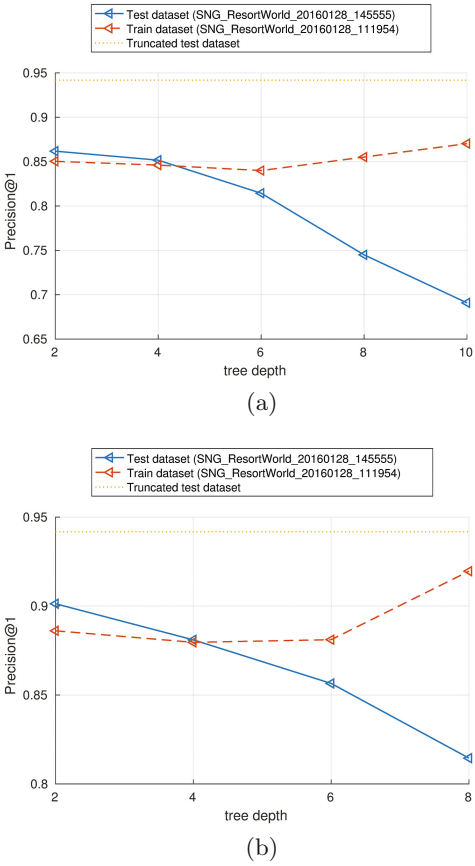


Fig. 4. Search precision using 256-bit hash codes generated from the test dataset by supervised FastHash method. For each element from the training set, similarity information was encoded for 20 nearest neighbours, all similar and different number of dissimilar elements: 100 in (a) and 300 in (b). For comparison, search precision on the training dataset is shown with a dashed line and search precision in the test dataset truncated to first k bits is plot with a dotted line.

all similar and 100 randomly chosen dissimilar elements is encoded. Figure 4(b) shows results when number of selected dissimilar elements is increased to 300. In both cases search precision is worse compared to naive bit truncation. It can be observed that encoding similarity information between more pairs improves the performance. Unfortunately, for practical reasons, it was not possible to further increase amount of supervised training data. For a training dataset consisting of over 300 thousand elements, encoding similarities between each element and over 300 other elements produces over 100 million pairs. Learning procedure requires over 10 GB of memory to efficiently process such amount of data.

In both cases severe overfitting can be observed. As decision tree depth grows and model complexity increases, the discrepancy between performance on the test and training set grows.

4 Conclusions and Future Work

Evaluation of hashing methods on large datasets of binary FREAK descriptors yield some surprising results. For 256-bit hash codes (half of the length of the original FREAK descriptor) none of the examined methods performed better compared to naive bit truncation approach. In theory, linear projection-based hash methods should be able to produce at least as good result. Yet all of the evaluated methods performed worse. For short codes, hashing methods gain advantage. For 128-bit codes Isotropic Hashing [6], the best unsupervised method, allows achieving 2% better search accuracy compared to naive bit truncation. For 64-bit codes, Isotropic Hashing produces 15% better results.

Examined semi-supervised and supervised methods were not able to benefit from the additional supervised information in the form of landmark ids linked with feature descriptors in the training set. Surprisingly, adding more supervised information in the examined semi-supervised methods produced worse results. The likely reason is, that due to limited hardware resources, similarity information can be explicitly encoded only for fraction of pairs of points from the training dataset, which leads to suboptimal performance.

Hashing methods can be beneficial when storage is a primary concern and short codes are required. They can be used to generate compact, 64-bit binary codes, form original 512-bit FREAK feature descriptors. This would decrease storage requirements four times, and increase descriptor matching speed at the expense of moderate decrease in search precision (about 15%). Further reduction to 32 bits has a detrimental effect on search precision, reducing it by over 40%.

As a future work it will be beneficial to investigate other approaches to supervised dimensionality reduction, such as Mahalanobis metric learning.

Acknowledgment. This research was supported by Google's Sponsor Research Agreement under the project "Efficient visual localization on mobile devices".

References

1. Alahi, A., Ortiz, R., Vandergheynst, P.: Freak: fast retina keypoint. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition, June 2012
2. Cai, D.: A Revisit of Hashing Algorithms for Approximate Nearest Neighbor Search. ArXiv e-prints, December 2016
3. Gong, Y., Lazebnik, S., Gordo, A., Perronnin, F.: Iterative quantization: a procrustean approach to learning binary codes for large-scale image retrieval. IEEE Trans. Pattern Anal. Mach. Intel. **35**(12), 2916–2929 (2013)
4. Heo, J.P., Lee, Y., He, J., Chang, S.F., Yoon, S.E.: Spherical hashing. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition, June 2012
5. Indyk, P., Motwani, R.: Approximate nearest neighbors: towards removing the curse of dimensionality. In: Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing. STOC 1998, NY, USA (1998). http://doi.acm.org/10.1145/276698.276876
6. Kong, W., jun Li, W.: Isotropic hashing. In: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems 25. Curran Associates, Inc. (2012). http://papers.nips.cc/paper/4846-isotropic-hashing.pdf
7. Kulis, B., Darrell, T.: Learning to hash with binary reconstructive embeddings. In: Bengio, Y., Schuurmans, D., Lafferty, J.D., Williams, C.K.I., Culotta, A. (eds.) Advances in Neural Information Processing Systems 22. Curran Associates, Inc. (2009). http://papers.nips.cc/paper/3667-learning-to-hash-with-binary-reconstructive-embeddings.pdf
8. Kulis, B., Grauman, K.: Kernelized locality-sensitive hashing for scalable image search. In: 2009 IEEE 12th International Conference on Computer Vision, IEEE (2009)
9. Lin, G., Shen, C., van den Hengel, A.: Supervised hashing using graph cuts and boosted decision trees. IEEE Trans. Pattern Anal. Mach. Intell. **37**(11) (2015). http://arxiv.org/abs/1408.5574
10. Lin, Y., Cai, D., Li, C.: Density sensitive hashing. CoRR abs/1205.2930 (2012). http://arxiv.org/abs/1205.2930
11. Lin, Y., Jin, R., Cai, D., Yan, S., Li, X.: Compressed hashing. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2013)
12. Muja, M., Lowe, D.G.: Fast approximate nearest neighbors with automatic algorithm configuration. In: VISAPP. vol. 2, no.1, pp. 331–340 (2009)
13. Trzcinski, T., Lepetit, V., Fua, P.: Thick boundaries in binary space and their influence on nearest-neighbor search. Pattern Recognit. Lett. **33**(16) (2012). http://www.sciencedirect.com/science/article/pii/S0167865512002619
14. Wang, J., Shen, H.T., Song, J., Ji, J.: Hashing for similarity search: a survey. CoRR abs/1408.2927 (2014). http://arxiv.org/abs/1408.2927
15. Wang, J., Zhang, T., Song, J., Sebe, N., Shen, H.T.: A survey on learning to hash. CoRR abs/1606.00185 (2016). http://arxiv.org/abs/1606.00185
16. Wang, J., Chang, S.F., Kumar, S.: Semi-supervised hashing for large-scale search. IEEE Trans. Pattern Anal. Mach. Intell. **34**(12), 2393–2406 (2012)
17. Wang, J., Kumar, S., Chang, S.: Sequential projection learning for hashing with compact codes. In: Fürkranz, J., Joachims, T. (eds.) Proceedings of the 27th International Conference on Machine Learning (ICML-10), Omnipress (2010). http://www.icml2010.org/papers/178.pdf

18. Wang, J., Liu, W., Kumar, S., Chang, S.: Learning to hash for indexing big data - a survey. CoRR abs/1509.05472 (2015). http://arxiv.org/abs/1509.05472
19. Weiss, Y., Torralba, A., Fergus, R.: Spectral hashing. In: Koller, D., Schuurmans, D., Bengio, Y., Bottou, L. (eds.) Advances in Neural Information Processing Systems 21, Curran Associates, Inc. (2009). http://papers.nips.cc/paper/3383-spectral-hashing.pdf
20. Wu, C., Zhu, J., Cai, D., Chen, C., Bu, J.: Semi-supervised nonlinear hashing using bootstrap sequential projection learning. IEEE Trans. Knowl. Data Eng. 25(6) (2013). http://dblp.uni-trier.de/db/journals/tkde/tkde25.html#WuZCCB13
21. Xu, B., Bu, J., Lin, Y., Chen, C., He, X., Cai, D.: Harmonious hashing. In: Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence. IJCAI 2013, AAAI Press (2013). http://dl.acm.org/citation.cfm?id=2540128.2540389

Efficient Parallelization Methods of Labeling Algorithm

Małgorzata Luchter-Boba, Piotr Łukasik, and Adam Piórkowski(✉)

Department of Geoinformatics and Applied Computer Science, AGH University of Science and Technology, A. Mickiewicza 30 Av., 30-059 Cracow, Poland
Adam.Piorkowski@agh.edu.pl

Abstract. Digital image processing is a field with broad applications. The development of technology has made it possible to introduce intelligent systems in distinctive areas such as medicine, robotics and astronomy. In this paper, the authors focus on indexing algorithms (also called labeling). Numerous studies have considered the various ways of implementing parallelization and the associated benefits. The indexing process involves assigning the same label to pixels of the same object. For the purpose of this study, a few algorithms proposed by Suzuki et al., Soh et al. and the method described by Tadeusiewicz and Korohoda were implemented. In order to parallelize the algorithms, the indexing algorithm of Niknam et al. was used and a method of partial parallelization was proposed.

Keywords: Image processing · Connected component labeling · Indexation · Labeling · Parallelization · Parallel processing

1 Introduction

Digital image processing is used in many important areas of science and technology: in medicine to diagnose diseases, in geography to analyze satellite images, and also in mineralogy, metallurgy and robotics [15]. A significant limitation of the use of digital methods in these areas is time-consuming operations such as segmentation and indexing. In order to accelerate these stages of image processing, parallelized versions of the algorithms are used in combination with the increasingly popular multiprocessor computers and even graphics cards, which have even more processors [5]. Unfortunately, the use of multiple processors is limited by certain computational problems, especially where domain decomposition is possible [8,10]. In other cases custom solutions are needed [3,6,9,14]. This paper focuses on the issue of image indexing (also known as labeling), the methods of implementing selected algorithms, and parallelization capabilities.

2 Basic Information About Indexing Images

2.1 Neighborhood

In image processing, an important concept is the neighborhood that defines the pixels adjacent to a pixel. There are neighborhoods of 4- and 8-connected pixels.

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*, Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_13

A neighborhood of 4-connected pixels includes adjacent pixels immediately above, below, and to the left and the right of the relevant pixel (central). A neighborhood of 8-connected pixels further includes pixels on the diagonals [15]. All the algorithms discussed in later chapters benefit from the definition of a neighborhood of 8-connected pixels.

2.2 Types of Indexing Algorithms

Indexing assigns the same label to adjacent pixels, resulting in groups of pixels in separate objects. (Indexing involves assigning neighboring pixels the same label, in effect leading to grouping of pixels into separate objects.) There are many ways to implement indexing. Suzuki et al. proposed the following classifications of algorithms:

1. multiple algorithms running across an image alternately from top to bottom and from bottom to top, agreeing labels until completion,
2. algorithms carrying out two passes on an image. On the first pass, pixels are assigned a temporary label. Equivalent labels supported during or after the first pass are stored in an array of connectors. During the second pass, labels are replaced with the lowest label in the table,
3. algorithms created for images represented by a hierarchical tree-like structure,
4. algorithms created for specific computer architectures [13].

There are other approaches to connected component labeling [2,4] that are not described here due to length restrictions.

2.3 The Algorithm Described by Tadeusiewicz and Korohoda (RTPK)

This is one of the easiest to implement and the basic version of the algorithm in which the image is analyzed twice using a different method on each pass [15]. The first pass starts from the top left corner and examines each pixel to assign a label, which is still subject to change. If a consecutive pixel is the same as the previous pixel, the algorithm assigns the same label, otherwise a new label is assigned which is represented by the value of a dedicated variable, which is incremented by one after each use. If a new label is obtained as a result of the operation, it is written in the connection table T and the label is stored in the

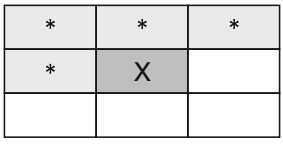


Fig. 1. Mask for the algorithm of the indexation at scanning ahead.

table as an index. The context is then the analyzed pixel, represented by mask M illustrated in Fig. 1.

For each checked pixel in the mask (symbol *), if it has a different label than the currently analyzed pixel (symbol x), the updated information on its label is already in the connection table. This operation takes place according to the formula in the connection table:

$$g(x,y) \neq F_B \; and \; g(x,y) = m, \tag{1}$$

$$T[m] = m \tag{2}$$

$$g(x,y) \neq F_B \; and \; g(x,y) = m \tag{3}$$

where:

g - the image produced during the first pass,
x, y - coordinates of the analyzed pixel,
i, j - relative coordinates mask M,
F_B - background,
m - the current label.

This means that the label is given the central element in the mask, rather than the lowest label (as in the original version). The purpose of the modification is to simplify the design code, reducing the use of temporary variables and the number of conditions. Correct indexing is also finally obtained, although the labels are not sorted.

The second pass is designed to replace the label given on the first pass according to the connection table. Therefore, it is much less computationally expensive because it relies only on the Lookup Table, which is the result of the first pass. The result of this operation is a complete indexed image.

2.4 The Algorithm Proposed by Suzuki et al.

The idea of the algorithm is based on traversing the analyzed picture forwards and backwards, but a connection table is also used in order to accelerate a solution [13]. The following is a mask for both types of searches (Fig. 2).

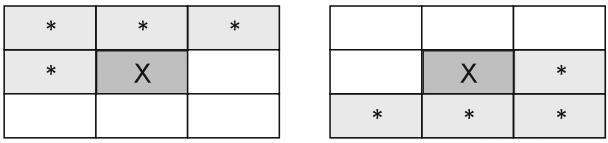


Fig. 2. Masks for scanning respectively in the front (a) and backward (b).

In the algorithm proposed by Suzuki, in the first scan of a binary image *b(x, y)*, each pixel is assigned a suitable temporary label m, according to the equation:

$$g(x,y) = \begin{cases} F_B & if\ b(x,y) = F_B \\ m, m(m = m+1)\ if\ \exists\{i,j\ \epsilon\ M\}\, g(x-i, y-j) = F_B, & (4) \\ T_{min}(x,y) & in\ other\ cases, \end{cases}$$

$$T_{min}(x,y) = min(T[g(x-i, y-j)]), \quad (5)$$

where:

$g(x, y)$ - image generated during the first pass,
F_B - background,
$T(m)$ - joins table,
i, j - relative coordinates of pixels belonging to the mask M.

During the scan, when a pixel is checked, the connection table is updated only for pixels in the mask.

$$\begin{aligned} T[m] = m \qquad & if\ \ \exists\{i,j\ \epsilon\ M\}\, g(x-i, y-j) = F_B, \\ T[g(x-i, y-j)] = T_{min}(x, y0)\ if \qquad & g(x-i, y-j) \neq F_B \end{aligned} \quad (6)$$

In subsequent scans, there is no new label; only those agreed as pixels belonging to the object are marked with the same, lowest possible label:

$$g(x,y) = \begin{cases} F_B \quad if \quad b(x,y) = F_B \\ T_{min}(x,y) \quad in\ other\ cases, \end{cases} \quad (7)$$

$$T[g(x-i, y-j)] = T_{min}(x,y)\ \ if\ g(x-i, y-j) \neq F_B. \quad (8)$$

The algorithm executes until no provisional labels change as this indicates that, the final labeled image has been obtained.

2.5 The Algorithm Proposed by Soh et al.

A different algorithm was proposed by Soh et al. [11,12]. In the first stage, each checked pixel gets a label that is the next pixel number. Subsequently, the scan continues as long as changes are still detected. For each pixel in each of the eight possible directions, the algorithm increases the range of acquired data in order to find the lowest label (Fig. 3) and stops only when it finds a background pixel.

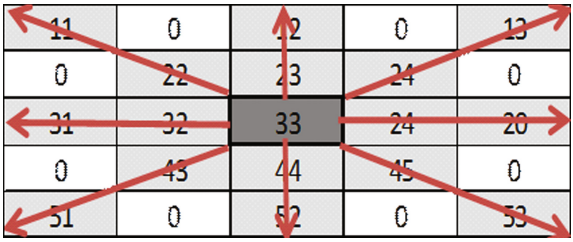


Fig. 3. Direction of propagation of the algorithm with respect to the central pixel.

The label is compared against the currently assigned label and, if it is lower, it assigns it to the current pixel. In order to facilitate parallelization, additional changes are saved in the connection table.

3 Parallelization Methods of the Indexing Algorithms

3.1 Parallelization Method of Algorithm by Niknam et al. (P1)

Parallelization of the indexing process is problematic because it has a particular sequence: the label assigned to a particular pixel depends on the existing labels and the neighborhood of the remaining pixels. Therefore, it is not possible to fully separate independent processes or divide the image into completely separate pieces (classical domain division). Niknam et al. proposed a complete parallelism of indexing and presented an example of Suzuki's algorithm [7]. An image subject to indexing is divided into vertical fragments and for each extracted fragment a separate thread is started. To properly assign a label to an object on a the edge of a fragment, threads inform each other of their positions relative to the analyzed image. This is necessary because another thread might be using the last column computed by the previous thread on the same line. For example, if the image is spread between the two threads, it initially runs for the first thread and the second waits until the analysis of the first row in the first fragment is complete Fig. 4a. Only when the first thread signals the completion of the first row and moves to the next analysis does the second thread start and begin analysis of the first row in the second fragment (Fig. 4b).

3.2 Proposed Method of Partial Parallelization of the Indexing Algorithm (P2)

In order to free them from the necessity of synchronization, the parallel and sequential stages of indexing can be divided into operations. Firstly, an image is divided into sections, the number of which varies according to how many threads are available. Subsequently, a parallelized indexing operation is performed according to the selected algorithm, as if the images were independent (Fig. 5a). Labels for different pieces at this stage cannot be repeated. For example, the range of

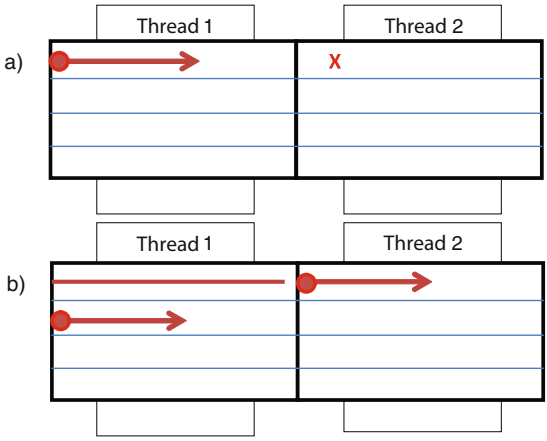


Fig. 4. Diagram illustrating parallelization indexed using continuous synchronization of concurrent threads. The other thread cannot start analysis of the first line until the first thread finishes analysis of the line (a)

labels in the labeling function may be determined, as the maximum number of objects in a given portion of an image is:

$$G_{max} = \left\lceil \frac{X}{2} \right\rceil * \left\lceil \frac{Y}{2} \right\rceil \tag{9}$$

where:

G_{max} - the maximum number of identified objects,
X - height of the image,
Y - width of the image,
$\lceil x \rceil$ - ceiling function.

Another way is to store the value of other global variable labels, which would benefit all threads iterated using atomic operations (Interlocked.Increment in the case of C#). Additionally, a connection table is available for all topics, because they operate on disjointed sets of labels, therefore there is no problem sharing global variables for each cell array.

After the first scan, the algorithm executes sequentially in order to agree on the labels that join portions of an image and save changes to the connectors table (Fig. 5b). The algorithm then checks the entire connectors table in order to detect complex changes (Fig. 5c). The final step is the substitution of labels in accordance with the array of connectors, which again is carried out in parallel (Fig. 5d).

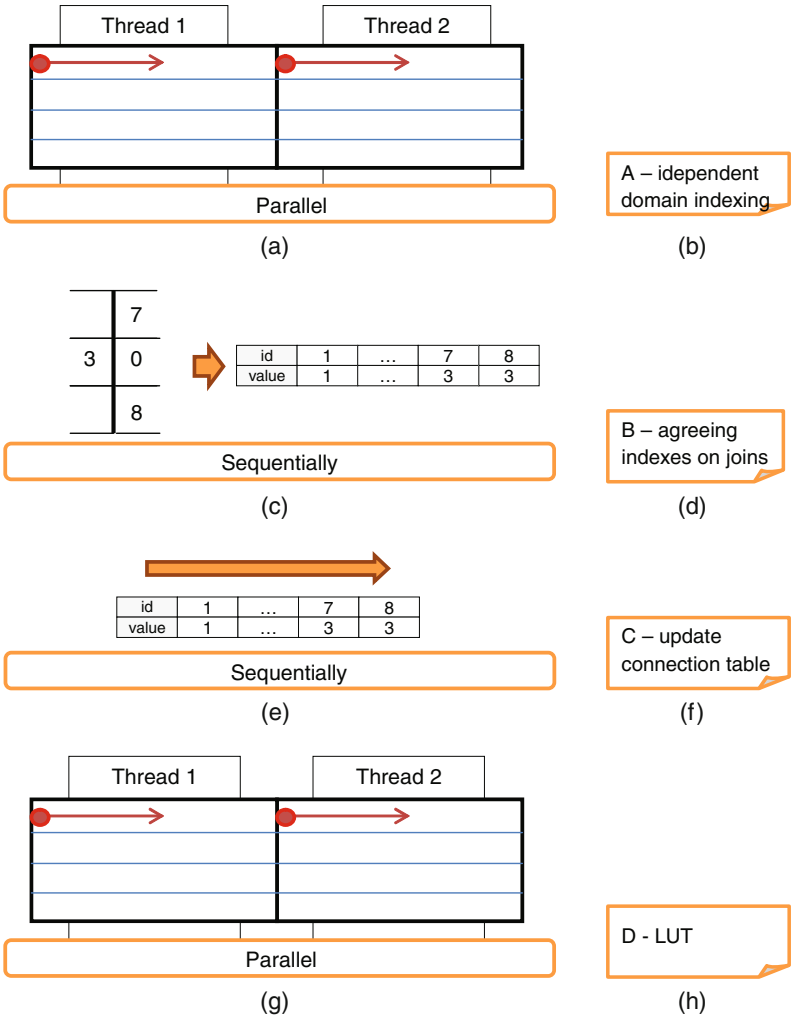


Fig. 5. The next stages of the proposed method of partial parallelization, an indexing algorithm for two processors: parallel indexing pre-separates sub-areas of the image (A), agreeing labels on connections and updating table connectors (B), re-run the table connectors (C), update labels using the updated table connectors (D).

4 Selected Issues of Parallel Programming

Parallel programming is designed to accelerate the performance of selected algorithms. Assessing the effectiveness of parallelization can be accomplished by parameters such as speedup and parallelization efficiency.

4.1 Speedup

Speedup SU_N can be used more generally to show effects on performance by comparing algorithm execution time in sequential and in parallel.

Acceleration program SU_N execution, in which part of the manual has been paralleled for N processors compared to the sequential program shows the pattern:

$$SU_N = \frac{T_s}{T_r} \tag{10}$$

where:

T_s - running time for the sequential algorithm,
T_r - runtime for the algorithm parallelized for N threads.

4.2 The Efficiency of Parallelization

The real value of acceleration is usually less than that resulting from the number of processors involved in parallel computing because of communication overheads and the sharing of storage resources or the bus. Efficiency E is an amount that reflects these issues and is given by the formula:

$$E = \frac{SU_N}{N} \tag{11}$$

In the absence of obstacles, speedup assumes a value of 1, which means the acceleration proportional to the number of processors available.

4.3 Amdahl's Law

Amdahl's law determines the maximum expected speedup if only part of the algorithm has been parallelized using N processors [1]:

$$E = \frac{1}{(1-P) + \frac{P}{N}} \tag{12}$$

where:

P - the proportion of the program that can be parallelized.

However, after appropriate transformation, the model can calculate the proportion of code which was parallelized on the basis of the measured acceleration SU_N and N processors:

$$P = \frac{\frac{1}{SU_N} - 1}{\frac{1}{N} - 1} \tag{13}$$

5 Software and Hardware

The project is written in C# using Microsoft Visual Studio 2012 and .NET 4.0. The parallel version of the algorithm uses the loop: Parallel.For available *System.Threading.Tasks* [16].

The study tested algorithms:

1. version sequence:
   - RTKP,
   - Suzuki,
   - Soh
2. in paralleled by P1:
   - Suzuki
3. in paralleled by P2:
   - RTKP,
   - Suzuki,
   - Soh

The algorithms and parallelized versions were tested on images containing horizontal stripes (Fig. 6(a)), vertical stripes (Fig. 6(b)), dots (Fig. 6(c)), diagonal stripes (Fig. 6(d)) and spirals (Fig. 6(e)), with dimensions of 600×600 px and 4000×4000 px.

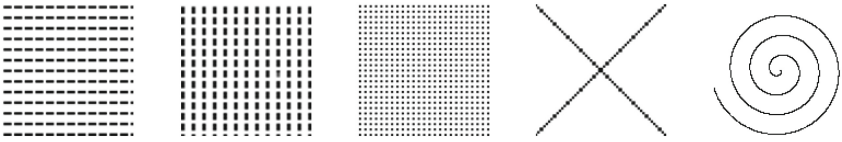


Fig. 6. Fragments of the example images used in the calculation.

Hardware parameters are described in the next section.

5.1 Hardware E1

The first computer was equipped with an Intel i7 4770, clocked at 3.40 GHz and 16 GB of RAM. This processor has 4 cores and 8 threads. For testing, we disabled Hyper-threading as it adversely affects the efficiency of parallel computing, in particular *Parallel.For* loops and other methods [16] which assume a static allocation of tasks for each core logic [10].

5.2 Hardware E2

The second computer was an Etergo Hyperion RS130 G4 server platform with two Intel Xeon E5-2690v2 CPUs clocked at 3.0 GHz and 384 GB of RAM. This processor has 10 cores and 20 threads. Hyper-threading was again disabled.

6 Analysis Research

The studies presented in the previous section were performed in environments E1 and E2, with six replicates of each image. The lowest calculation time was that which was least distorted by operating system interference. Cyclic tests were performed by utilizing the whole range of the available number of processors (E1: 1–4, E2: 1–20). The sequential version of the Suzuki algorithm was fastest, with the exception of the dots image, for which the algorithm presented by Soh et al. was slightly faster, and the RTPK algorithm was slowest. The speed of the algorithm depends not only on the size of the image, but also on its content. The Suzuki algorithm performed fastest for diagonal stripes and slowest for spirals. The algorithm by Soh et al. performed fastest with the dots and slowest with the spiral. The RTPK algorithm was fastest with horizontal strips and slowest with vertical stripes. None of the algorithms were affected by the environments, although generally the duration of the calculation performed by sequential algorithms was significantly shorter for environment E1; in some cases it took as much as twice as long.

In the context of this article, the most important outcome was positive and negative variations in algorithm efficiency when run in parallel. The included charts (Figs. 7(a) and 8) show the acceleration of image data for both environments using a given number of processors. The resulting acceleration for a specific number of processors varies depending on the algorithm used, the environment, and the content of the image. For most images, acceleration increased in line with the number of processors used, up to 8–10 processors, above which the acceleration increase began to fall. This behavior could be due to communication overheads associated with multiple threads, which prevent further increases in acceleration, or a decrease in the effectiveness of combining multiple fragments of larger images. Interestingly, in environment E1, for the same number of image fragments, there was a significant slowdown in some test cases: acceleration in the E2 environment for the dots image using the Suzuki algorithm under the parallelized process P2 using two processors (created in two parallel tasks) was 1.55, and in E1 did not exceed unity (Fig. 7(a)). As can be seen from the following diagrams, the proposed manner of parallelization (P2) of the Suzuki algorithm is more efficient than the method described by Niknam et al. (P1). According to this theory, only the dots image differs, where the algorithm parallelized by way of P1 obtained up to four times acceleration, and version P2 obtained only three times acceleration (Fig. 7(a)).

An interesting result was obtained by the parallelized version of the P2 algorithm of Soh et al. in the case of the spiral image, for which an acceleration of over 586 times was achieved when 20 processor cores were used. This efficiency significantly exceeded the expected maximum value of forty-four times acceleration (Fig. 6). The image is split into smaller fragments and then indexes are arranged on join points, which significantly simplifies calculations. This constituted the main element of the acceleration in the case of the spiral. Also, the calculated mean when using parallel algorithms varied depending on the image, the algorithm environment and the number of processors used. Because

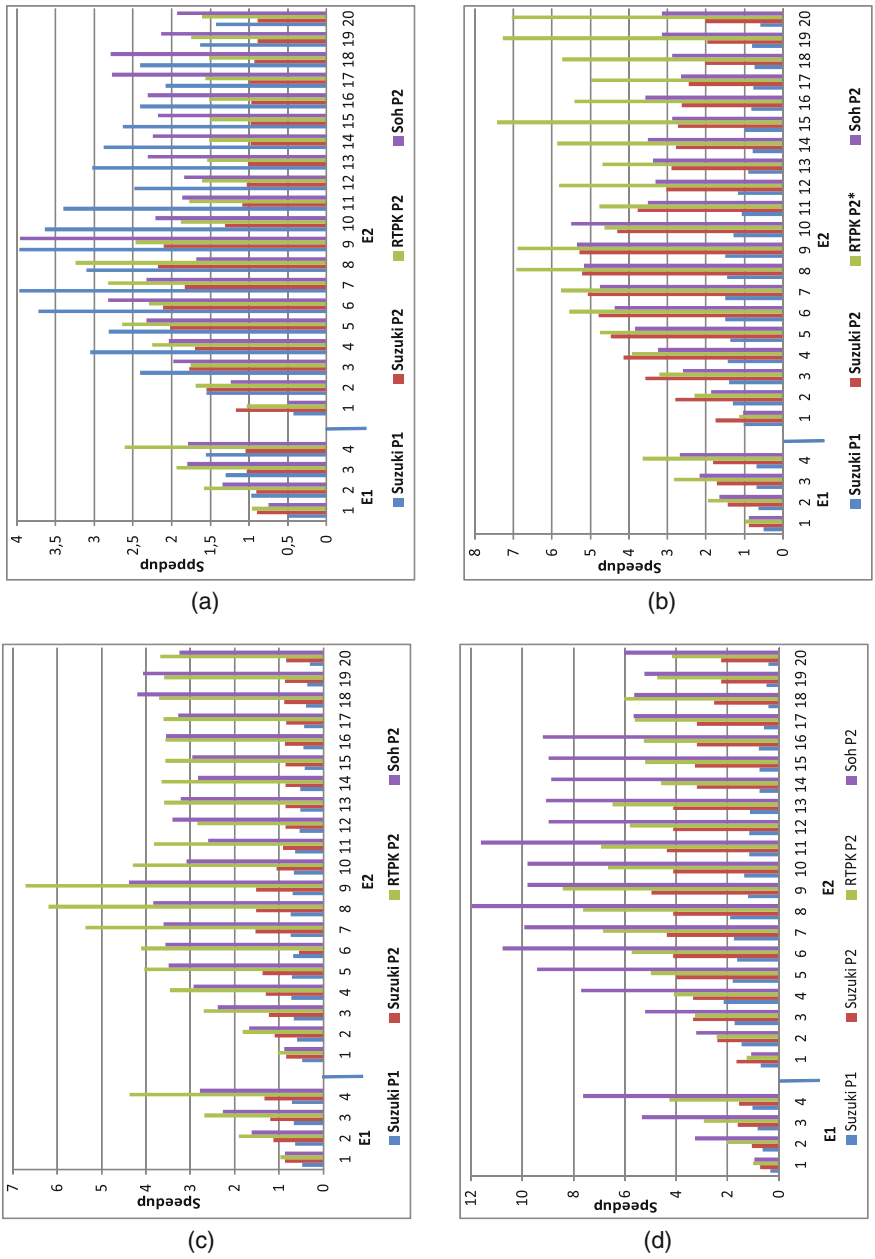


Fig. 7. Comparison of acceleration depending on increasing the number of processors in environment E1 and E2 for the dots (a), vertical stripes (b), horizontal stripes (c), and diagonal (d) images, size of 4000×4000 pixels.

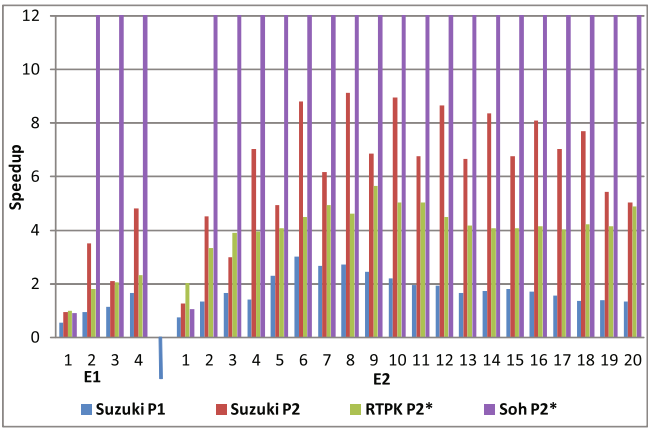


Fig. 8. Comparison of acceleration depending on increasing the number of processors in environment E1 and E2 for the spiral image measuring 4000×4000 pixels for Suzuki and for others: 600×600 pixels.

the estimate value ranged from negative to greater than unity, it is believed that parallelism is highly sensitive to the type of input data and in some specific cases the code for one processor may reach a different number of sub-function calls than with two processors, which means there is no sense in applying the mean to describe the obtained results (detailed figures in the attached tables).

7 Summary

In order to accomplish this work, three indexing algorithms were implemented: the standard described by Tadeusiewicz and Korohoda, and those of Suzuki and Soh. Then, to test the effectiveness of parallelization of labeling operations, the Nikarn concept for the Suzuki algorithm was implemented using a different partial parallelization approach for all three algorithms. The presented method of partial parallelization significantly speeds up the work of the presented examples algorithms for most of the tested images. The increase in calculation speed is observed only for a certain number of processors because too many threads running at the same time reduces the acceleration effect. In most cases, the proposed method is more efficient than the parallelization of the method presented by Niknam [7]. The results testify to the fact that the final acceleration depends not only on the algorithm used and the size of the image, but also on its content or hardware platform.

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

References

1. Amdahl, G.M.: Validity of the single processor approach to achieving large scale computing capabilities. In: Proceedings of the Spring Joint Computer Conference, 18–20 April 1967, pp. 483–485. ACM (1967)
2. Chen, C.W., Wu, Y.T., Tseng, S.Y., Wang, W.S.: Parallelization of connected-component labeling on tile64 many-core platform. J. Sig. Process. Syst. **75**(2), 169–183 (2014)
3. Czerwinski, D.: Digital filter implementation in Hadoop data mining system. In: International Conference on Computer Networks, CCIS, vol. 522, pp. 410–420. Springer, Cham (2015)
4. Gupta, S., Palsetia, D., Patwary, M.M.A., Agrawal, A., Choudhary, A.: A new parallel algorithm for two-pass connected component labeling. In: 2014 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), pp. 1355–1362. IEEE (2014)
5. Hachaj, T., Ogiela, M.: Real time area-based stereo matching algorithm for multimedia video devices. Opto-Electron. Rev. **21**(4), 367–375 (2013)
6. Mazurek, P.: Parallel distributed downsampled spatio-temporal track-before-detect algorithm. In: 2014 19th International Conference on Methods and Models in Automation and Robotics (MMAR), pp. 119–124. IEEE (2014)
7. Niknam, M., Thulasiraman, P., Camorlinga, S.: A parallel algorithm for connected component labelling of gray-scale images on homogeneous multicore architectures. J. Phys: Conf. Ser. **256**, 012010 (2010). IOP Publishing
8. Oleszko, K.: The efficiency of the code parallelization in multi core environment on the basis of image processing in 3D space. In: Elleithy, K., Sobh, T. (eds.) New Trends in Networking, Computing, E-learning, Systems Sciences, and Engineering, pp. 361–364. Springer, Cham (2015)
9. Pięta, A., Dwornik, M.: Parallel implementation of ray tracing procedure in anisotropic medium. TASK Q. **16**(1), 135–143 (2012)
10. Piorkowski, A.: Parallel processing in web-based interactive echocardiography simulators. Comput. Inf. **33**(3), 707–720 (2014)
11. Soh, Y., Ashraf, H., Hae, Y., Kim, I.: A hybrid approach to parallel connected component labeling using CUDA. Int. J. Sig. Process. Syst. **1**(2), 130–135 (2013)
12. Soh, Y., Ashraf, H., Hae, Y., Kim, I.: Fast parallel connected component labeling algorithms using CUDA based on 8-directional label selection. Int. J. Latest Res. Sci. Technol. **3**(2), 187–190 (2014)
13. Suzuki, K., Horiba, I., Sugie, N.: Linear-time connected-component labeling based on sequential local operations. Comput. Vis. Image Underst. **89**(1), 1–23 (2003)
14. Szostek, K., Leśniak, A.: Parallelization of the seismic ray trace algorithm. In: Parallel Processing and Applied Mathematics. LNCS, vol. 7204, pp. 411–418. Springer, Heidelberg (2012)
15. Tadeusiewicz, R., Korohoda, P.: Computer Analysis and Image Processing. Progress of Telecommunication Foundation Publishing House, Krakow (1997)
16. Toub, S.: Patterns of parallel programming - understanding and applying parallel patterns with the .NET Framework 4 and Visual C# (2010), version, 16 February 2010

Communications and Miscellaneous Applications

Modeling Telecommunication Networks with the Use of Reference Graphs

Sławomir Bujnowski(✉), Tomasz Marciniak, Zbigniew Lutowski, Beata Marciniak, and Daniel Bujnowski

UTP University of Sciences and Technology, Bydgoszcz, Poland
slawb@utp.edu.pl

Abstract. In this paper, an algorithm for finding graphs with parameters of Reference Graphs was presented. During the Reference Graphs search, it was found that in most cases with a specific number of nodes and equal degrees of nodes, there is more than one graph with identical baseline parameters, the diameter and average path length, and the number of these graphs increases with the increase of the number of nodes and their degree. To clarify where these differences are occurring, it was decided to define the characteristics of the comparable structures that could affect the transmission properties.

Keywords: Chordal rings · Graph theory · Reference graph · Optimization of networks

1 Introduction

What is known as a ICT (information and communication technologies) system is a group of cooperating IT devices and software providing processing, storing, as well as sending and receiving data over the telecommunications network via terminals designed for this purpose [19]. This system consists of a number of intelligent nodes, whose task is to provide users with a specific range of services, ensuring the right quality, speed, and reliability of information. In order to meet these requirements in the design and analysis of ICT systems, it is important to take into account the selection of connection topology between its network components, which significantly influences the transmission efficiency [1,15–17].

Actual telecommunication networks mainly have irregular topologies, resulting from the need to "match" these structures to the changing needs of the operator and users. The result is that such a network does not provide optimal data transfer conditions. Data transfer improvements can be achieved through the use of standardized nodes with identical network equipment, which significantly lowers the investment costs associated with the installation of the hardware, operating expenses and servicing of the aforementioned systems. The additional benefit is the symmetry of connections, routing simplicity, good scalability, simple network management and, by introducing additional bypass connections, improved reliability.

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*, Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_14

Network topologies can be described using graphs. Their vertices (nodes) are commutative modules or specialized computers, and the edges are usually two-way, independent transmission channels linking these vertices. In case of the identical technical equipment of each node, such networks may be described by regular graphs.

Fiber optic cables are most commonly used to transmit information in extended ICT networks, and these networks generally have a ring structure. The transmission characteristics of a standard ring structure are not satisfactory, therefore in order to improve it, it is modified by introducing additional inter-nodal connections called chords. The topologies obtained in such way are called Chordal Rings [1–6,8,10,11,14] (Fig. 1).

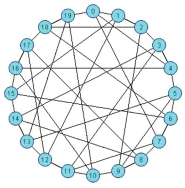


Fig. 1. An example of a fourth degree chordal ring

Definition 1. *The chordal ring is a special case of a Circulant Graph defined by the pair (p, Q), where p denotes the number of nodes, and Q the set of chords, $Q \subseteq \{1, 2, \ldots, p/2\}$. Each of the chords $q_i \in Q$ connects a pair of nodes belonging to the ring, where q_i denotes the length of the chord equal to the number of the ring edges between the nodes. The chord ring is described by the notation $G(p; q_1, \ldots, q_i)$, where $q_1 = 1 < q_2 \ldots qi$. The degree of nodes is generally equal to $d(V) = 2i$, except in case where the chord length is $p/2$; in this case p is even, and the node degree is $2i - 1$.*

The content of many publications [7,9,10] shows that the diameter and average path length has a significant influence on the transmission properties of the network modeled by graphs. The definitions of these parameters are given below.

Definition 2. *The diameter of graph $D(G)$ is the longest path among the minimum length paths connecting any two graph nodes:*

$$D(G) = \max_{v_i v_j} \{d_{\min}(v_i, v_j)\} \tag{1}$$

where v_i, v_j denote node numbers, $d_{\min}(v_i, v_j)$: minimum path length (number of edges) connecting nodes v_i, v_j.

Definition 3. *The average path length in the graph is given by the formula:*

$$d_{av} = \frac{1}{N(N-1)} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} d_{\min}(v_i, v_j) \tag{2}$$

where $d_{\min}$ *is the minimum number of edges between nodes* v_i, v_j, *where* $i \neq j$, *and* N *is the number of nodes forming the graph.*

Results of studies of this type of structures can be found in a number of publications, and on their basis the concepts of optimal and ideal graphs have been introduced [7,9,12,13,18].

Definition 4. *A regular graph with* p_i *nodes is called an ideal chordal ring of degree* $d(V)$. *It is described by the formula:*

$$n_i = 1 + \sum_{d=1}^{D(G)-1} |n_d| + \left|n_{D(G)}\right| \tag{3}$$

where n_d *denotes the number of nodes belonging to the d-th layer (the layer is a subset of nodes evenly spaced from any random source node by d edges),* $n_{D(G)}$ *denotes the number of nodes that belong to the last layer, and* $D(G)$ *is the diameter of the analyzed graph. For every* k *and* $l < D(G)$ $p_k \cap p_l = \emptyset$. *If the subset* $n_{D(G)}$ *reaches the maximum value achievable for the last layer, then such a graph is called optimal. For the ideal chord ring, the average length of the* d_{avi} *path is:*

$$d_{avi} = \frac{\sum_{d=1}^{d(G)-1} d\,|n_d| + D(G)\left|p_{d(G)}\right|}{n_i - 1} \tag{4}$$

In the optimal graph d_{avo} *is equal to:*

$$d_{avo} = \frac{\sum_{d=1}^{d(G)} dn_d}{n_o - 1} \tag{5}$$

where d - layer number, n_d *- number of nodes in the d-th layer,* n_o *- number of optimal graph nodes.*

To objectively evaluate whether the structure of inter-nodal connections has the best possible baseline parameters, the parameters of the Reference Graphs were established [7,12].

Reference Graphs are the regular structures with a predetermined number of nodes in which the diameter values and the average path lengths from any source node reach the same, theoretically calculated lower limits.

In order to determine the minimum parameter values of these graphs, a minimum spanning tree covering all n nodes of the regular graph (Fig. 2) was used. The parameters of this tree selected from any source node: the radius and average path length; correspond to values of the reference graph parameters: the diameter and average path length.

Based on the theoretical analysis, the definition of the Reference Graph was formed, taking into account the degree of its nodes.

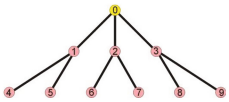


Fig. 2. An example of a tree used for the construction of the Third Degree Reference Graph

Definition 5. *Reference Graph of degree $d(V)$ is a structure with the following properties:*

- *The number of n_{dGR} nodes in the d-th of the optimal Reference Graph is given by the formula:*

$$n_{dOGR} = d(V) \cdot (d(V) - 1)^{d-1} \tag{6}$$

In the ideal graph this number applies to all layers $d < D(G)$, while in layer $d = D(G)$ this number is:

$$n_{diGR} = n_i - d(V) \cdot (d(V) - 1)^{d-2} \tag{7}$$

- *Number of nodes $n_{D(G)GR}$ in the optimal Reference Graph in a function of its diameter:*

$$n_{D(G)_{GR}} = \frac{d(V) \cdot (d(V) - 1)^{D(G)} - 2}{d(V) - 2} \tag{8}$$

- *The diameter of the optimum Reference Graph in a function of the number of nodes that make up this type of graph is expressed by the formula:*

$$D(G) = \log_{d(V)-1} \frac{n_{GR}\,(d(V) - 2) + 2}{d(V)} \tag{9}$$

whereas in the ideal graph the diameter is defined by:

$$D(G) = \left\lceil \log_{d(V)-1} \frac{n_{GR}\,(d(V) - 2) + 2}{d(V)} \right\rceil \tag{10}$$

- *The average length of the davGR path in a function of the optimum Reference Graph diameter is:*

$$d_{avo} = \frac{(d(V) - 1)^{D(G)} \cdot ((d(V) - 2) \cdot D(G) \cdot -1) + 1}{(d(V) - 2) \cdot (d(V) - 1)^{D(G)} - 1} \tag{11}$$

In the ideal graph, the average path length is described by the expression:

$$d_{avi} = \frac{d(V)\frac{(d(V)-1)^{D(G)-1} \cdot ((d(V)-2) \cdot (D(G)-1) \cdot -1)+1}{(d(V)-2)^2}}{\frac{n_{iGR}-1}{\frac{\left(n_{iGR} - n_{(D(G)-1)GR}\right) \cdot D(G)}{n_{iGR}-1}}} + \tag{12}$$

where p_{iGR} is the number of ideal graph nodes.

- *The graph is symmetric in the sense that the graph parameters remain the same regardless of from which node they are calculated.*

2 The Reference Graph Search Algorithm

The author's task was to develop a program for finding graphs with parameters of Reference Graphs. An algorithm was designed specifically for this purpose. The general working principle of the developed Optimal Reference Graphs search algorithm is illustrated by an example of a third degree graph with 10 nodes. Firstly, a prime tree (Fig. 2) is constructed from a selected node, identified as the root. The radius of the tree meets the assumed diameter-related condition resulting from the number of nodes forming the Reference Graph. In the example, the radius is 2, which means that it is possible to reach from the source node (0) to each of the other nodes by the path created by up to two edges.

In the next step, any node belonging to the first node layer is selected. A layer is defined as a set of nodes distant from the source node by k edges. A conditional test is performed to ensure that each node can be reached through a minimum path whose length is consistent with this assumption. In the example, the path lengths from node 1 to the other nodes do not meet this condition (the paths to nodes 6, 8 and 7, 9 consist of 3 edges). Additional connections are made (blue and green) - Fig. 3.

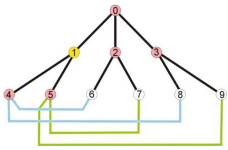


Fig. 3. First phase of the graph construction

The next node of the first layer (node 2 - Fig. 4) is selected, and a new analysis of meeting the pre-set condition for the path length is performed.

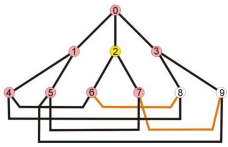


Fig. 4. Further phases of the graph construction

In this case, there were no paths consisting of two edges leading to nodes 8 and 9. Therefore, additional connections were made between nodes 6 and 8 and 7 and 9, which resulted in meeting the pre-set condition. Next, we analyze the lengths of the paths connecting the last node belonging to the first layer with the other vertices of the graph (Fig. 5).

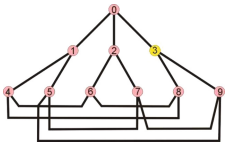


Fig. 5. Next phase of the graph search

After checking all nodes distant from the start node by one edge, the algorithm chooses nodes that belong to the next layer as the source nodes. If, in any case, the path length condition is not met, a modification of the connection is performed. Afterwards, a new analysis of meeting the conditions for the other nodes is carried out, as shown in Fig. 6.

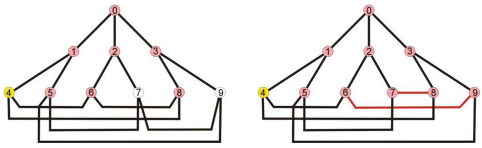


Fig. 6. Modification of the connections in order to meet the path length condition of the source node no. 4

After examining the graph from all the source nodes and finding that the distance condition is met in each case, the program terminates its operation by returning the derived topology of the Reference Graph connections.

Finding the ideal Reference Graphs is a slightly more complicated process. At first all possible simple trees are built from the node selected as the root of these trees. These trees are created up to and including penultimate layer, and the remaining nodes remain unconnected (Fig. 7).

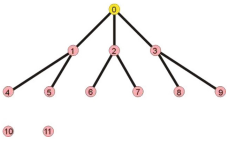


Fig. 7. The first phase of the search for the ideal Reference Graph of a third degree, consisting of 12 nodes (assuming that this graph cannot be an optimal graph)

All the possible connection combinations of nodes belonging to the last layer with nodes of the penultimate layer (Fig. 8) are checked. Repetitions, i.e. connections which have the same effect after the transformation of the structure, are skipped.

The resulting primary trees are the starting point for the further part of the application using the developed algorithm. The Reference Graph search

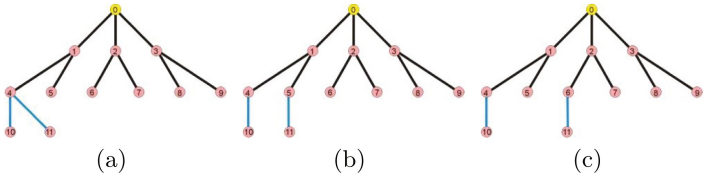


Fig. 8. Connection combinations of nodes belonging to the last layer with nodes of the penultimate layer

process starts with the first layer and, as before, the consecutive source nodes are selected. The operation of the algorithm consists of introducing additional connections between the nodes of the last and the penultimate layer, which ensures that the parameters of subgraphs obtained with this method will be consistent with the parameters of the Reference Graphs.

Figure 9 shows the graph search process for each of the proposed primary tree configurations. The Figure shows only those subgraphs, when selecting a given source node forced adding additional edges in order to meet the pre-set path length condition.

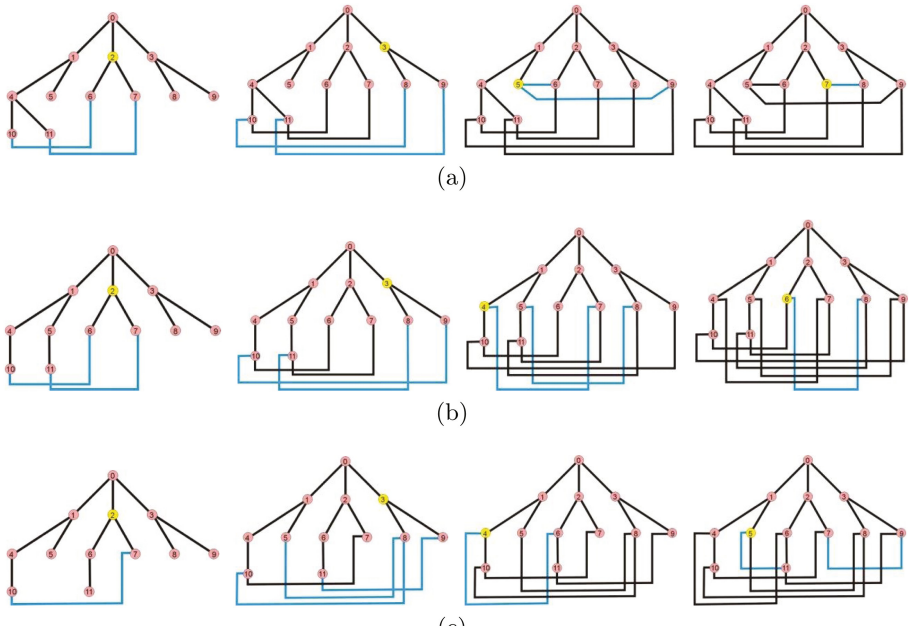
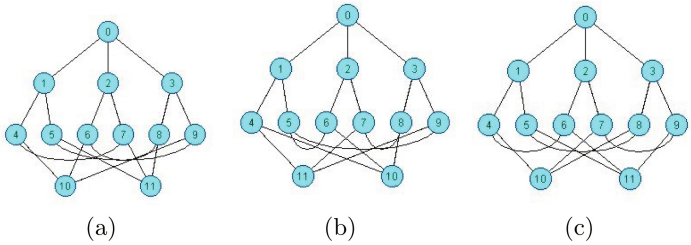


Fig. 9. Steps of graph construction

Yellow color indicates the node which, at the given stage of the graph construction, is the source node.

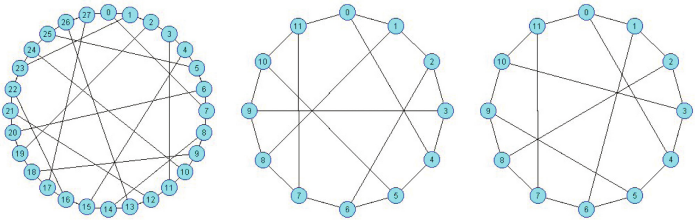
After examining the graph from the point of view of all the source nodes and finding that all the created subgraphs meet the pre-set condition, the program terminates its operation by returning the obtained node structure (Fig. 10).



(a) (b) (c)

Fig. 10. The obtained Reference Graphs topologies

An additional task fulfilled by the application, but significant from the point of view of the modeled network structure, is the transformation of the obtained graphs into the form of chord rings. After converting and renumbering, graphs shown in Fig. 11 were obtained.



(a) (b) (c)

Fig. 11. Images of Reference Graph composed of 12 nodes transformed into chord rings

During the Reference Graphs search, it was found that in most cases with a specific number of nodes and equal degrees of nodes, there is more than one graph with identical baseline parameters, the diameter and average path length, and the number of these graphs increases with the increase of the number of nodes and their degree.

For example, if the graph is composed of eight nodes, there are only two Reference Graphs of the third degree; when the number of nodes is ten, only one graph is found (Peterson's graph, which is the optimal graph) if the number of nodes is twelve, there are five of them, but with 14 nodes, 56 graphs were found. In contrast, for graphs of the fourth degree consisting of 8 nodes, the number of Reference Graphs is 12, with 10 nodes it is 684 and with 12 it is 1350.

However, a significant number of the graphs obtained in this way is isomorphic. Therefore, prior to conducting further studies on the transmission properties of the network topologies, it was necessary to eliminate the isomorphic structures from the obtained topologies, since they would assume the same transmission parameters. For this purpose, the algorithm developed by Dharwadker and Tevet was used. By using this method, it was found that among the structures shown in Fig. 10, the graphs A and C are isomorphic, whereas graph B does not have this characteristic.

A simulation program was developed in order to perform a comparative evaluation of the transmission parameters of the networks modeled with Reference Graphs. The probability of rejecting a connection was assumed as a parameter determining the transmission properties of the graphs. For this purpose Monte-Carlo method [20] was used with the following assumptions:

- Structures with the same number of nodes were compared under the condition that the bitrate of incoming and outgoing inter-nodal links is equal and constant;
- In each test, the traffic generated at each source node was assumed, and the selection of the source and target nodes to which the connections were routed, was based on uniform distribution;
- The shortest possible connection path was selected. In case of several variants to choose from, the least burdened path was selected.

The simulation of the analyzed networks was based on the following algorithm:

1. Enter an array describing the topology of the network connecting the nodes;
2. Assume a number of subscribers generating traffic in each of the nodes;
3. Enter a confidence interval;
4. Randomly, according to the normalized Poisson distribution, select the number of calls generated by the subscribers connected to the node;
5. For each connection, select the number of the source node in a pseudorandom way, based on uniform distribution;
6. For each connection, select the number of the target node in a pseudorandom way, based on uniform distribution;
7. Perform the connection;
8. Add up the global number of generated connections - $\sum GP$;
9. Calculate the number of unrealized connections - $\sum PN$;
10. Repeat steps 5 ÷ 10 until the confidence level is reached;
11. Calculate the probability of rejecting the call $p_{OW} = \frac{\sum PN}{\sum GP}$.

Exemplary results of simulation are showing on diagrams included below. Simulations results for reference graph 3rd degree with twenty eight nodes are showed in Fig. 12 for parameters $D(G) = 4$, $d_{av} = 2,78$.

Results for reference graph 4th degree with nine nodes are showed in Fig. 13 for parameters $D(G) = 4$, $d_{av} = 2,78$. $D(G) = 2$, $d_{av} = 1,5$.

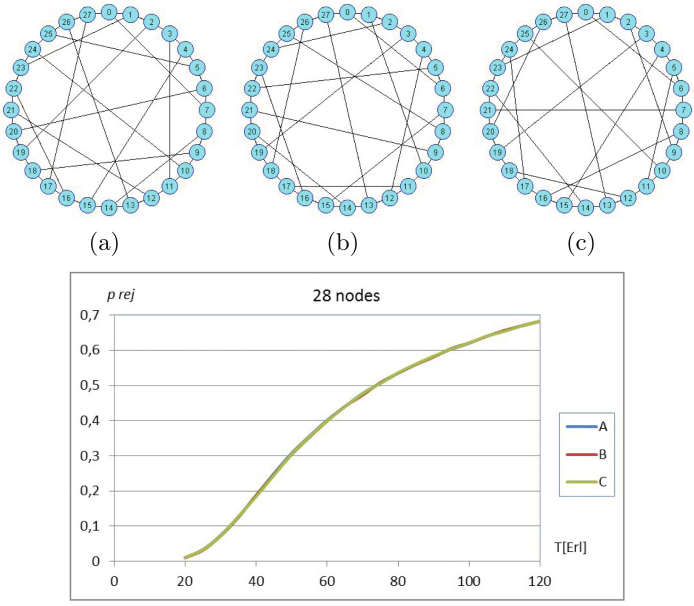


Fig. 12. Probability of rejecting as a function of traffic reference graph 3rd degree with twenty eight nodes

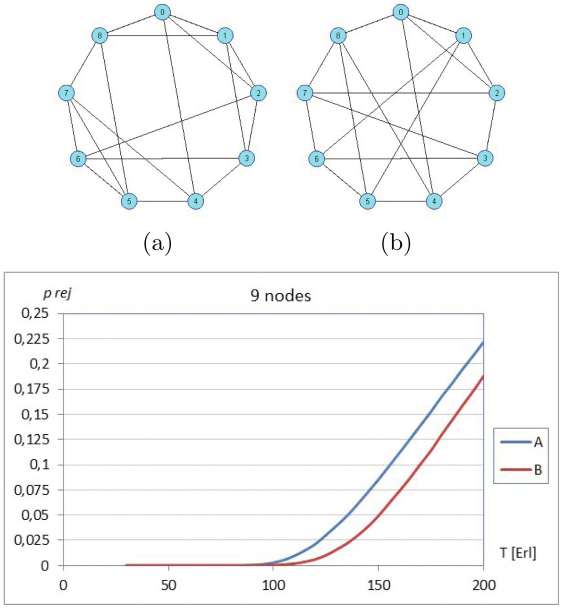


Fig. 13. Probability of rejecting as a function of traffic reference graph 4th degree with nine nodes

As obtained simulation results show, in some cases despite the same basic values, i.e. the diameters and average lengths of Reference Graphs, which consist of the same number of nodes, their transmission parameters are differ. To explain the cause of these differences, it was decided to find other parameters that could affect the transmission properties of the networks modeled with Reference Graphs.

3 Summary and Conclusions

The paper presents issues related to the study of transmission properties of networks described by Reference Graphs. The definition of these graphs and the method of their construction were given. The simulation program algorithm which allowed to investigate the transmission properties of the virtual networks described with the aforementioned graphs was presented. As a measure of these properties, the probability of rejection of the call was assumed. Based on the results obtained, it was found that reference graphs of equal number and degree of nodes differ in transmission properties. It was assumed that such a feature may be the number of uses of the individual edges of the graph links belonging to the shortest paths through which information is exchanged between the nodes or the calculated average probability of making connections through minimum length paths. The results of the analyzes will be presented in subsequent papers.

References

1. Azura, R., Othman, M., Selamat, H., Hock, P.: Modified degree six chordal rings network topology. In: Prosiding Simposium Kebangsaan Sains Mathematik Ke-16, pp. 3–5 (2008)
2. Azura, R.N.F., Othman, M., Peng, Y., Selamat, M.H.: On properties of modified degree six chordal rings network. Malays. J. Math. Sci. **4**(2), 147–157 (2010)
3. Bujnowski, S., Dubalski, B., Zabłudowski, A., Ledziński, D., Marciniak, T., Pedersen, J.M.: Comparison of modified 6 degree chordal rings. Image Process. Commun. Challenges **2**(2010), 435–446 (2010)
4. Bujnowski, S., Dubalski, B., Zabłudowski, A.: Analysis of chordal rings. Mathematical Techniques and Problems in Telecommunications. Centro International de Matematica, Tomar, pp. 257–279 (2003)
5. Bujnowski, S., Dubalski, B., Zabludowski, A.: Analysis of 4th degree chordal rings. In: Communications in Computing, pp. 318–324 (2004)
6. Bujnowski, S., Dubalski, B., Zabludowski, A.: The evaluation of transmission ability of 3rd degree chordal rings with the use of adjacent matrix. In: Proceedings of the Seventh Informs Telecommunication Conference (2004)
7. Bujnowski, S., Dubalski, B., Zabludowski, A., Ledzinski, D., Riaz, M.T., Pedersen, J.M.: Evaluation of modified degree 5 chordal rings for network topologies. In: 2010 Australasian Telecommunication Networks and Applications Conference (ATNAC), pp. 60–65. IEEE (2010)
8. Bujnowski, S., Dubalski, B., Zabłudowski, A., Pedersen, J., Riaz, T.: Analysis of degree 5 chordal rings for network topologies. Image Process. Commun. Challenges **3**, 445–457 (2011)

9. Diestel, R.: Graph Theory. Graduate Texts in Mathematics, vol. 173, 4th edn. Springer-Verlag, Heidelberg (2001)
10. Dubalski, B., Zabludowski, A., Bujnowski, S., Pedersen, J.M.: Comparison of modified chordal rings fourth degree to chordal rings sixth degree. In: 50th International Symposium on ELMAR 2008, vol. 2, pp. 597–600. IEEE (2008)
11. Farah, R., Othman, M.: In modified chordal rings degree six geometrical representation properties. In: Proceedings of Fundamental Science Congress, Kuala Lumpur, Malaysia (2010)
12. Farah, R., Othman, M., Selamat, H., Hock, P.: Analysis of modified degree six chordal rings and traditional chordal rings degree six interconnection network. In: International Conference on Electronic Design, ICED 2008, pp. 1–7. IEEE (2008)
13. Farah, R., Othman, M., Selamat, M., Rushdan, M.: In layers shortest path of modified chordal rings degree six networks. In: Proceedings of International Conference on Intelligent Network and Computing, Kuala Lumpur, Malaysia (2010)
14. Farah, R., Othman, O., Selamat, M.: Combinatorial properties of modified chordal rings degree four networks. J. Comput. Sci. **6**(3), 279–284 (2010)
15. Freire, M.M., Da Silva, H.J.: Performance comparison of wavelength routing optical networks with chordal ring and mesh-torus topologies. In: International Conference on Networking, pp. 358–367. Springer (2001)
16. Lee, H., et al.: Analysis of chordal ring network. IEEE Trans. Comput. **100**(4), 291–295 (1981)
17. Pedersen, J.M., Gutierrez, J.M., Marciniak, T., Dubalski, B., Zabludowski, A.: Describing n2r properties using ideal graphs. In: Second International Conference on Advances in Mesh Networks, MESH 2009, pp. 150–154. IEEE (2009)
18. Raja, R., Shah, M., Othman, M., Selamat, M.H.: An optimum free-table routing algorithms of modified and traditional chordal ring networks of degree four. J. Mater. Sci. Eng. **4**(10), 78–89 (2010)
19. Ustaw, D.: Obwieszczenie marszałka sejmu rzeczypospolitej polskiej z dnia 10 stycznia 2014 r. w sprawie ogłoszenia jednolitego tekstu ustawy - prawo telekomunikacyjne. Dziennik Ustaw 243(000) (2014)
20. Zieliński, R.: Metody Monte Carlo. Wydawnictwa Naukowo-Techniczne (1970)

QoS Concept for Protecting Network Resources - A Case Study

Łukasz Apiecionek(✉) and Jacek M. Czerniak

Department of Computer Science, Institute of Technology,
Kazimierz Wielki University, ul. Chodkiewicza 30, 85-064 Bydgoszcz, Poland
{lukasz.apiecionek,jacek.czerniak}@ukw.edu.pl

Abstract. Distributed Denial of Service attacks remain one of the main problems of computer networks and to date there is no method for protecting network user from the source of the attack. Such attack can block network resources provided by the servers for many hours. In order to prevent such cases, it is recommended to use firewalls and IDS/IPS mechanisms. Unfortunately, they are not solving the problem. An alternative solution is the concept of Quality of Services methods, which are not flawless. While they provide the possibility for the administrators to protect their network resources during the attack, there are no guidelines concerning the mechanism of their functioning. This paper analyzes a case study of the proposed method of using QoS for network protection.

1 Introduction

It is an undeniable fact that IT systems are nowadays omnipresent and their users need a fast access to information from every part of the network. The main weakness of the networks is their susceptibility to blockage due to Distributed Denial of Service attacks. These attacks have become a serious problem as they cause network unavailability by blocking services via seizing the system resources in computers in the network until they stop working. In most cases, a user who has already started working in the system loses the connection and cannot even log out of the system. Usually the systems log out their users after the connection timeout is reached or when a broken connection is detected. DDoS attacks are nowadays a serious obstacle for the IT systems' efficient functioning and they have to be eliminated. Unfortunately, common methods of fighting the DDoS attack problems [1,3,5–8] are usually limited to using the Intrusion Detection System and Intrusion Prevention System (IDS/IPS in short) solutions. Such systems base on a description of well-known attacks or some kind of Artificial Intelligence solution which is able to learn the actions in some specific scenarios of the attack. Other solutions suggest using a firewall mounted on the network edge which can only block the incoming traffic on specific ports or IP address ranges. In the [4] the authors present the concept of using the Quality of Services mechanisms implemented in routers. This is a very interesting solution and it can provide the required tools for network administrators. Unfortunately, the administrators have to face other problems, which will be presented in this paper.

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*,
Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_15

The structure of this paper is as follows. Section 2 shortly describes the issue of the DDoS attacks and the proposed concept of using QoS method for fighting the attacks. Section 3 provides a case study of the method implementation and the problems which it causes. The final part includes a conclusion and discussion over the method.

2 Concept of QoS Method

There are numerous papers which describe the DDoS attacks [5,7]. As it can be noticed, these attacks can be performed on various system resources: TCP/IP sockets [5,6] or DNS servers. Usually the main principle is to simulate so many correct user connections that their number exceeds the actual system performance and drives it to abnormal operation. Some papers [1,3,5–8] describe methods for dealing with the DDoS attacks by means of their global detection and the necessity of cooperation between provider's network. Unfortunately, the transmission of the attackers' packets is performed through the provider's network, therefore if it cannot be blocked, it leads to data link saturation. Of course this saturation results in a lack of connection to the server. The existing method allows to limit the incoming traffic on a firewall and lets the servers deal with the already established connection. QoS methods implemented on routers measure the incoming traffic and set the priority of transmission of the packets over the network. Alternatively, QoS methods can be applied on one router, in order to protect the network resources locally [3]. QoS method bases on counting the packets, but it is unable to determine whether the packet is a part of a DDoS attack on some server or not. The solution uses SNMP protocols for obtaining additional traffic statistics [4]. The offered QoS Service for routers comprises the following steps [4]:

- the routers are collecting the statistics of the transferred traffic (1),
- the data are divided into the counters of traffic to the specific destination (2),
- the routers are exchanging their statistics over SNMP (3),
- the server which is the aim of the DDoS attack sends a SNMP message to its router which is under the attack (4),
- the routers are passing the information between each other about the IP address of the aim of the attack (5),
- basing on the routers statistics they are looking for the sources of the attack (6),
- when the sources of the attack are recognized, they become blocked (7).

These steps are presented on Fig. 1. Figure 2 shows a general scheme of the method. The Fair Queue method is used during usual operation of the device when the packets are transmitted to the receiver. If a DDoS attack is detected, the packets are transmitted to a special data stream recognition module, which compares them with the correct connection history database. The history database is supplied by the information gathered from the server. After their categorization, the packets are either transmitted to the receiver or removed by the intelligent dropping module, the work algorithm of which is presented in the next subsection.

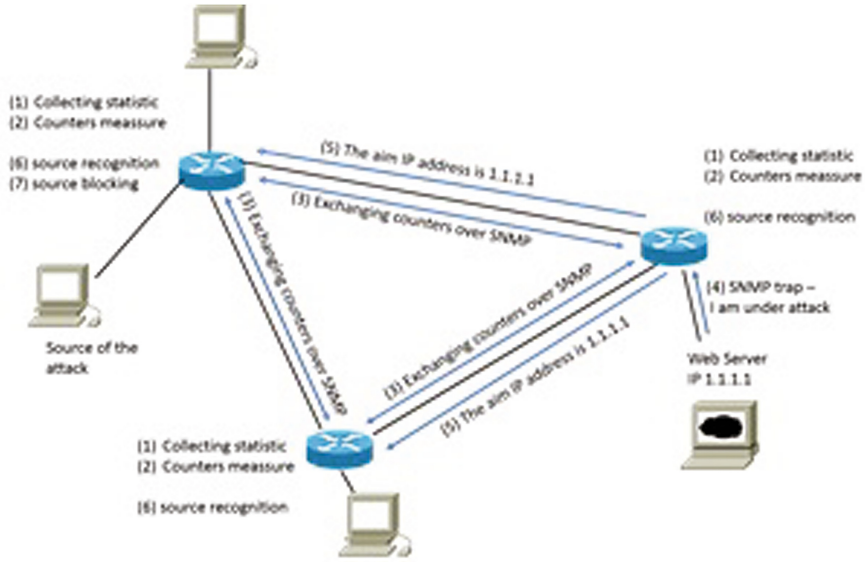


Fig. 1. QoS method concept in practice

3 Implementation Results - A Case Study

As has been already demonstrated by the authors [4], this method does not solve all problems connected with DDoS attack, but it enables the users to close their active connection when the attack starts. As it was presented, the widely used Web browsers: Internet Explorer (12%), Mozilla Firefox (27%) and Google Chrome (55%) [2], are trying to connect to the server more than once. Mozilla Firefox browser tries to establish a connection 18 times, while Google Chrome and Internet Explorer stop after the 9th connection attempt. During the implementation phase a special algorithm of dropping packets was introduced, the role of which is to filter the traffic on server's special open port and to limit it accordingly to the determined policy. At the beginning of the filtration a list of the IP addresses which communicate with the server correctly, i.e. which are not part of the DDoS attack $listIP$ is collected from the data flow history. During the filtration each packet is checked regarding its presence on the list of the valid IP addresses $listIP$. In the case of a positive verification, the packet is sent to the network, otherwise a counter of the passed packets *packet_counter* is being checked if its value is greater than the allowable packet limit *packet_limit* in a time slot t1. If the limit of the packets is exceeded, the packet is dropped - DROP. In the following time slot the number of current packets *packet_counter* is zeroed and the above mentioned filtration process is restarted. In such case the dropping module has to regulate the opening connection limit in a special intelligent way. It is made as follows:

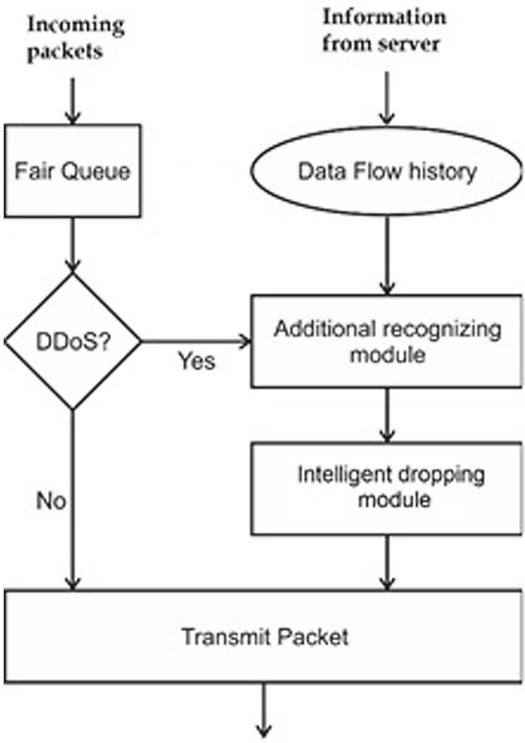


Fig. 2. Diagram of the method

- if in a given number of the subsequent time slots *subsequent_time_limit* the limit of the packets is exceeded, the module recognizes a large attack on the server, and in order to provide it some time to regain its efficiency, the limit of packets *packet_limit* allowed to passed in time slots is decreased,
- if in the following time slots the packet limit is not exceeded, the module assumes that the attack on the server started to lose its intensity, so the *packet_limit* can be increased.

Changing the packet limit is very important for the network administrator because this value allows the server to handle the incoming connections which may be potentially correct or to release the resources used incorrectly by the attacker. This process of decreasing the packet limit should depend on the server type. As it could be noticed, different server types and services may require different method and limits. Different servers handle their connections in a different way, so they may require more time or to the contrary, they may prepare the data very quickly. Thus, the network administrators should be able to customize their method of protection depending on the characteristics of the server, the way the connections are handled and the way they should be processed by the

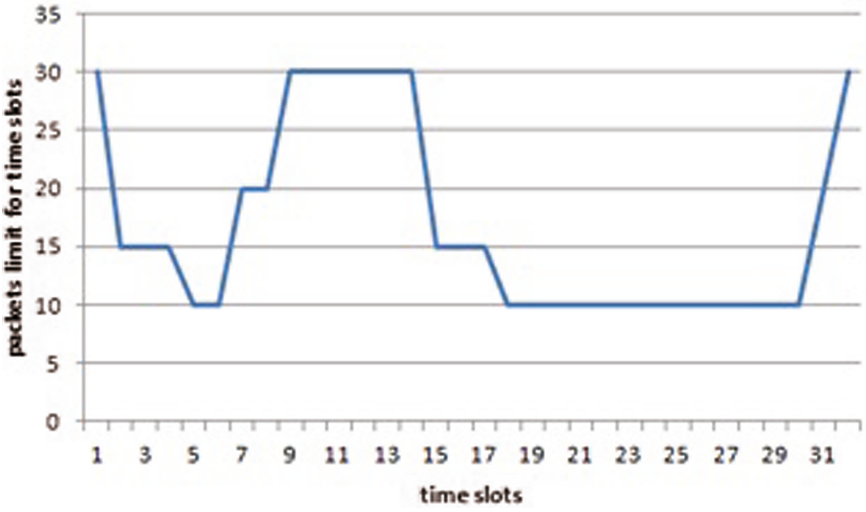


Fig. 3. Packet limit on time slots regulated by QoS method [7]

filtering module. This can be achieved by the experimental determination, however a preferable solution should be based on conscious decision of the administrator. Such decisions may be based on the server's resources, its operating system, the amount of memory and processors type. During the implementation of the method, a DDoS attack on the server was performed. The Fig. 3 presents the packet limit regulation during the timeslots. The maximum value was 30 while the minimum value was 10. Over the time those values underwent some adjustments. What is important is the fact that after the attack these values got back to their maximum.

4 Conclusions

In this article a concept of eliminating DDoS attacks [4] was discussed. While the methods suggested in the literature focus on blocking the access to the resources when the attack occurs, the QoS concept allows to work and keep the connection to the server. The main problem for the network administrators is the configuration process of the proposed method. The network administrators have to provide the characteristics of the server and the instructions on how the connection should be managed during the attack. This problem has not been considered until now. In order to prepare the firewall in the most optimal way, some additional study has to be performed. The problem discussed in this study is that the administrators have to consider their servers resources and software which is working on them and which decides how the user connection is managed. That implies an additional duty for the network administrators, but even though the configuration process can be handled automatically, basing on the experimental determination, conscious human decision remains an incomparably better solution.

References

1. Apiecionek, L., Czerniak, J.M., Zarzycki, H.: Protection tool for distributed denial of services attack. In: Beyond Databases, Architectures and Structures, BDAS 2014, vol. 424, pp. 405–414 (2014)
2. Czerniak, J.M., Zarzycki, H., Ewald, D.: AAO as a new strategy in modeling and simulation of constructional problems optimization. Simul. Model. Pract. Theory **76**, 22–33 (2017). http://www.sciencedirect.com/science/article/pii/S1569190X17300709
3. Dyczkowski, K.: A less cumulative algorithm of mining linguistic browsing patterns in the world wide web (2007)
4. Dyczkowski, K., Wygralak, M.: On triangular norm-based generalized cardinals and singular fuzzy sets. Fuzzy Sets Syst. **133**(2), 211–226 (2003)
5. Moor, D., Shannon, C., Brown, D., Voelker, G.M., Savage, S.: Inferring internet denial-of-service activity. ACM Trans. Comput. Syst. (TOCS) **24**(2), 115–139 (2006)
6. Rocky, K., Chang, C.: Defending against flooding-based distributed denial-of-service attacks: a tutorial. IEEE Commun. Mag. **40**, 42–51 (2002)
7. Schuba, C.L., Krsul, I., Huhn, M.G., Spafford, E.H., Sundaram, A.: Analysis of a denial of service attack on TCP. Computer Science Technical Reports, Paper 1327 (1996). http://Docs.Lib.Purdue.Edu/Cstech/1327
8. Stachowiak, A., Dyczkowski, K., Wojtowicz, A., Zywica, P., Wygralak, M.: A bipolar view on medical diagnosis in OvaExpert system (2016)

Acceleration of Genome Sequencing with Intelligent Cloud Brokers

Rocío Pérez de Prado[1(✉)], Sebastian García-Galán[1], Jose Enrique Muñoz-Expósito[1], and Adam Marchewka[2]

[1] Telecommunication Engineering Department, University of Jaén, Linares, Spain
rperez@ujaen.es

[2] Institute of Telecommunications and Computer Science, UTP University of Science and Technology in Bydgoszcz, Bydgoszcz, Poland

Abstract. Workflows from DNA sequencing applications have an extensive number of jobs which are reliant and that require parallel execution if high levels of performance are desired. In this work, a novel workflow broker based on expert systems is presented to accelerate workflows for DNA sequencing in cloud computing datacenters. The broker is based on the adaptation of Fuzzy Rule-Based Systems (FRBSs), which are inspired by Fuzzy Logic (FL) and rule-based systems, and as shown by simulation results, it is able to accelerate the processing of genome sequencing more efficiently than a wide range of scheduling strategies.

Keywords: Cloud computing networks · Workflows · Fuzzy Rule-Based Systems · Fuzzy Logic

1 Introduction

Cloud computing networks are distributed computing platforms which are considered as the computational paradigm pioneering the field in the last years [4]. Aside from numerous modern, business applications being conveyed, this worldview is likewise pulling in numerous established researchers to use the administrations of cloud for running substantial scale information and calculation demanding applications [11]. Further, the processing of scientific workflows is of particular relevance in some areas. A workflow can be characterized as an arrangement of assignment and conditions between the jobs that are utilized as fragments of different scientific applications. Specifically, the use of workflows is extensively used in the field of biomedicine for DNA analysis in prestigious institutions.

Epigenome Center is at present utilizing the Illumina Genetic Analyzer (GA) framework to create high throughput DNA succession information (up to 8 billion nucleotides per week) to outline epigenetic condition of human cells on an extensive scale [5]. The Center has actualized a robotized investigation pipeline utilizing Pegasus Workflow Management System to bolster these sequencing endeavors. Pegasus empowers researchers to develop work processes in unique

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*, Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_16

terms without agonizing over the points of interest of the fundamental execution condition or the particulars of the low-level determinations required by the middleware (Condor, Globus, or Amazon EC2). The Epigenomics workflow is basically an information preparing pipeline that uses the Pegasus Workflow Management System to mechanize the execution of the different genome sequencing operations. The DNA grouping information created by the Illumina-Solexa Genetic Analyzer framework is part into a few pieces that can be worked on in parallel. The information in each lump is changed over into a record organization that can be utilized by the Maq framework. Whatever is left of the operations include the sifting through of loud and defiling successions, mapping arrangements into the right area in a reference genome, creating a worldwide guide and afterward recognizing the grouping thickness at each position in the genome This work process is being utilized by the Epigenome Center in the handling of generation DNA methylation and histone change information.

The primary issue in running these work process applications is mapping the undertakings of the workflow to a proper asset in the cloud condition. In this work, a novel workflow broker for cloud computing datacenters based on expert systems is presented to accelarate workflows for DNA sequencing and Small untraslated RNA in the previous introduced project. With this aims, synthetic workflow are used. Synthetic workflow generators utilize the data accumulated from genuine executions of logical work processes in the genuine frameworks and learning of procedures behind these work processes to produce practical, manufactured work processes taking after those utilized by certifiable logical applications. Further, the code used to produce the greater part of the engineered work processes underneath, and numerous others, is accessible from the GitHub store. Results prove that the proposed broker is able to improve other scheduling strategies and significantly reduce the execution times of critical processes in genome analysis.

This work is structured as follows. In Sect. 2, the characteristics of the Epigenomics DNA sequencing workflows are presented. Next, in Sect. 3, the proposal is presented and in Sect. 4 the experimental results are discussed. Finally, in Sect. 5 the main conclusions of the paper are drawn.

2 Background

As introduced before, Epigenome Center uses the Illumina Genetic Analyzer (GA) framework to create high throughput DNA grouping information (up to 8 billion nucleotides for each week) to delineate epigenetic condition of human cells on a vast scale [5,6]. The Center has executed a robotized examination pipeline utilizing Pegasus-WMS to bolster these sequencing endeavors. A typical Epigenomic workflow can be observed in Fig. 1 where computational jobs are shown as circles. Detailed information of the jobs can be found in [12].

The workflow appeared above comprises of seven fundamental strides: 1. exchange grouping information to the bunch stockpiling framework, 2. split grouping records into various parts to be prepared in parallel, 3. change over succession documents to the suitable record design, 4. sift through uproarious and

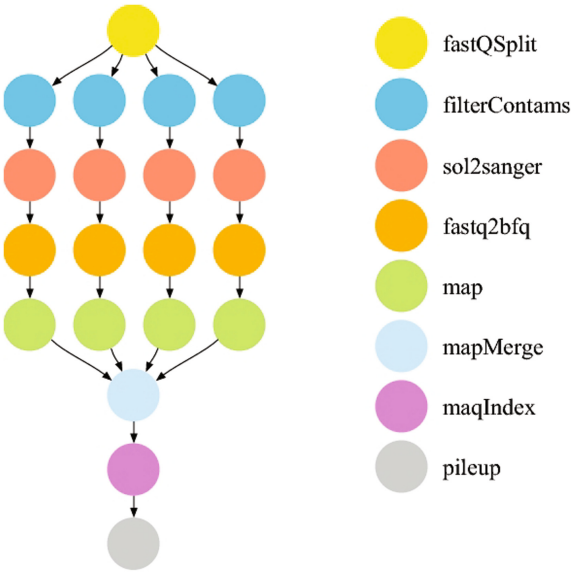


Fig. 1. Epigenomics DNA sequency workflow example [5].

debasing groupings, 5. delineate to their genomic areas, 6. blend yield from individual mapping ventures into a solitary worldwide guide, and 7. utilize arrangement maps to compute the grouping thickness at each position in the genome. The Epigenome Center is presently utilizing this work process to process its creation DNA methylation and histone change information. While the workflow as of now actualizes the base necessities to successfully break down the information, they are at present attempting to include quality control and checkpoint ventures to make the pipeline more vigorous. Tables 1 and 2 present the execution times of Epigenomics jobs and Sizes of Epigenomics data items, respectively.

Table 1. Execution times of Epigenomics jobs.

Job	Count	Mean(s)	Variance
fast2bfq	146	0.39	0.02
fastqSplit	2	42	1.8e+02
filterContams	146	1.1	0.5
map	146	9.6e+03	1.7e+07
mapMerge	3	24	33
pileup	1	3.2e+03	0
sol2sanger	146	0.24	0.01

Table 2. Sizes of Epigenomics data items.

File Type	Count	Mean(MB)	Variance
chunked_sfq	420	7.3	1.8e−01
filtered_sfq	420	5	9.6e−02
fq_format	420	3.7	5.2e−02
bfq_format	420	9.5e−01	4.5e−03
out_map	420	1	5.9e−03
merged_map	6	68	18
merged_map	1	400.44	0
indexed_map	1	20	0
pileup	1	4.4	0

In general, the execution of Epigenomics workflows is extremely demanding in terms of time and the use of distributed computing systems is necessary. Further, the way in which this type of workflows are schedule to each node is critical to accelerate the process.

3 Proposed Intelligent Broker

The processing of DNA sequencing workflows from Epigenomics requires the consideration of distributed computing systems and cloud computing is a leading platform for parallel execution of workflows. In Fig. 2, a general representation of a cloud datacenter used for biomedical applications is presented, where the computing nodes and broker can be differentiated.

The first general objective of this work is the application to Epigenomics DNA sequencing in cloud computing of FL based expert systems and in particular, the application of Mamdani FRBSs to design cloud brokers where authors have previous experience [7–10]. FRBSs are master control or broker based frameworks where FL is utilized as a method for speaking to the framework information and the association between factors [2]. These frameworks construct theirs choices in light of "IF-THEN" guidelines or rules where antecedents and consequents speak to fluffy proclamations for the factors including the framework state, and thusly, they can manage complex issues where there exist unclearness and vulnerability with a human-like thinking. Consequently, the thinking system is established on the utilization of Knowledge Bases (KBs) that incorporate the fluffy run semantics or fluffy sets, Data Bases (DBs) and etymological tenets or master learning, RBs. Besides, the general structure of these frameworks is comprised of three fundamental frameworks: fuzzification, induction and defuzzification frameworks.

In this work, the design of fuzzy rule-based broker as expert scheduling system for Epigenomics DNA sequencing that provide scheduling decisions founded

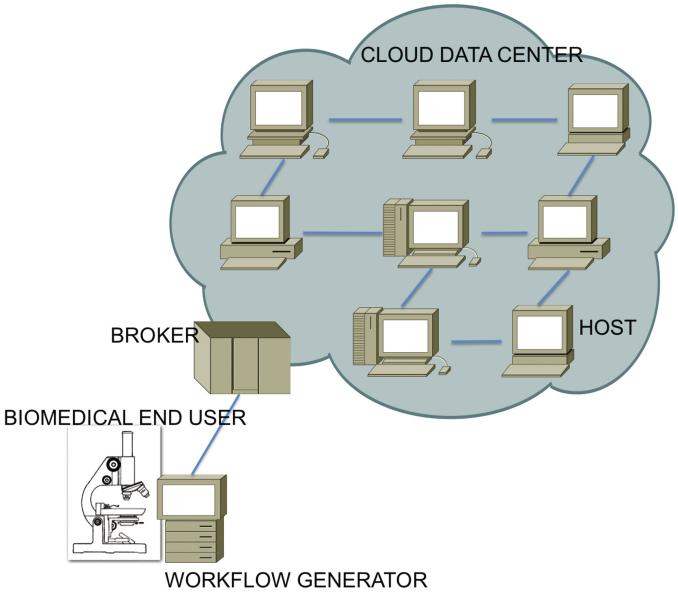


Fig. 2. Cloud datacenter structure for bioinformatics applications.

of a fuzzy characterization of the uncertain information describing the state of the cloud network used by Epigenomics and a rule-based reasoning applied to this characterization is proposed. In Fig. 3 the structure of the intelligent broker can be observed.

Basically, the point of the broker is to give an index of reasonableness to each accessible site to be utilized as a part of the present arranging. Accordingly, in each planning venture, the condition of each cloud hub must be gotten considering an arrangement of cloud info components or factors and this state in fresh configuration is to be changed into a semantic organization in the fuzzification framework for each site. Next, the information of the expert system given as fuzzy rules must be connected to this fuzzy input to get a marker of appropriateness for the considered site in a fuzzy format in the deduction or inference framework. At last, this fuzzy index must be converted into a fresh arrangement in the defuzzification framework. Once the intermediary gets all the lists for every one of the destinations, the site with a higher esteem, i.e., higher appropriateness, must be chosen for the present calendar in the DNA sequencing process.

First, in the design of the broker based on fuzzy rules, it is necessary to determine the variables describing the state of the grid as well as to specify the characteristics of Mamdani FRBSs, i.e., the knowledge representation and strategy reasoning. Hence, a characterization of the state of the cloud must be initially obtained. On the one hand, this characterization must be comprehensive enough to have a thorough knowledge of the sites state and on the other hand,

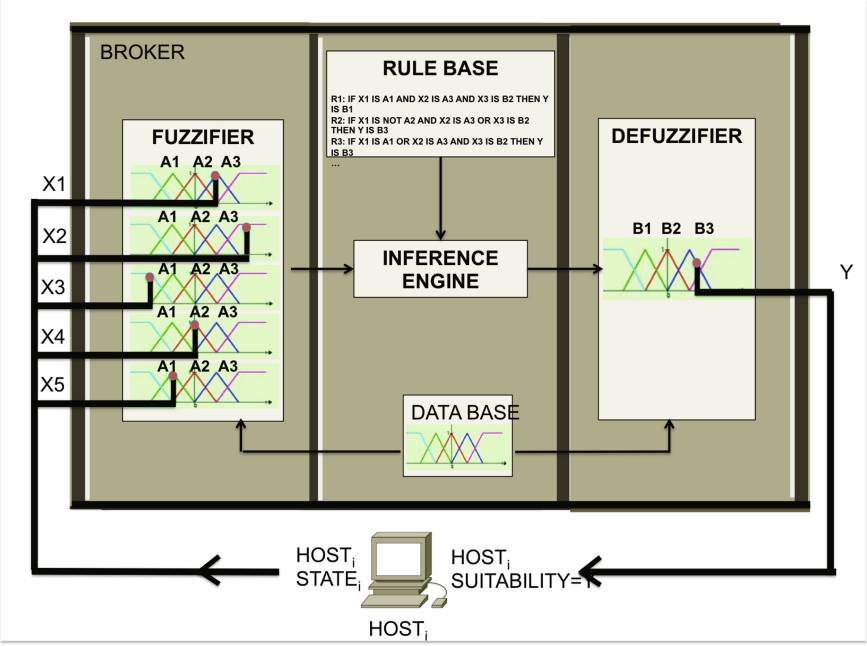


Fig. 3. General structure of the intelligent broker for DNA sequencing.

it must be as concise as possible in a way that the complexity of the KB system expert broker is not increased. The simplicity of the KB is particularly relevant considering that the computational effort of the learning processes that must be conducted to improve its quality is proportional to its complexity. In addition, the KB of the expert system must designed, i.e., the shape, the number and distribution of the fuzzy sets for each input and output variable and the structure of the rules must be determined. Also, it is necessary to specify the characteristics of the fundamental components of the Mamdani FRBSs-based broker, i.e., fuzzification, inference and defuzzification systems. Finally, the application of genetic learning strategies in FRBSs to the evolution of the expert broker is proposed in such a way that the scheduler can acquire and improve its knowledge in the form of fuzzy rules [8,9].

4 Simulation Results

Considering the large-scale infrastructure required to test the broker in real systems, the proposal has been evaluated through simulations based on WorkflowSim-DVFS [1,3]. WorkflowSim-DVFS is open-source software which is available at [3]. Specifically, the performance of the broker has been analyzed considering real conditions using an available trace of Epigenomics and the results have been compared in terms of execution time [5,6]. After the training of the intelligent broker with a genetic strategy, validation results show that

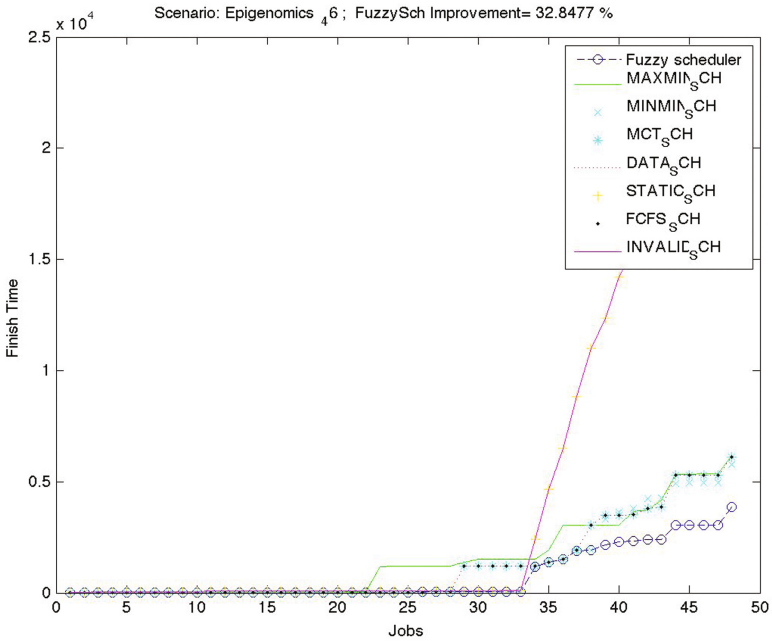


Fig. 4. Epigenomics DNA sequency time execution results with diverse scheduling techniques.

the fuzzy meta-scheduler improves the performance of a wide range of scheduling strategies: Max Min, Min Min, MCT, DATA, STATIC, FCFS and INVALID [3]. Figure 4 presents the processing evolution of jobs in terms of time for diverse well-known scheduling strategies in cloud computing and the intelligent broker for trace Epigenomics46 [5].

As shown, validation results prove that the proposed fuzzy scheduler or intelligent broker not only outperforms the rest of strategies in terms of final makespan in approximately 32%, but also the partial results for makespan are equal or lower than the rest of strategies. Hence, it has been observed that the expert broker does not only improve other classical strategies in terms of the considered optimization objective/s used for its optimization (makespan) but also, the fuzzy approach can lead to efficient results in partial makespan by offering a good harnessing of the resources in the cloud system.

5 Conclusions

This work has proposed an intelligent broker based on the adaptation of Mamdani FRBSs for DNA sequencing. The essential components that define the reasoning strategy of the broker in order to obtain the output of the broker have been explained and validated simulation results in terms of makespan have been presented for an Epigenomics real trace. As discussed, the intelligent broker does

not only outperform a wide set of scheduling strategies in cloud computing in terms of makespan but also, the evolution of the execution of partial makespan is lower, what shows the efficiency of the proposal to schedule DNA sequencing workflows.

Acknowledgments. This work was financially supported by Research Projects TEC2015-67387-C4-2 and TEC2012-38142- C04-03.

References

1. Chen, W., Deelman, E.: Workflowsim: a toolkit for simulating scientific workflows in distributed environments. In: 2012 IEEE 8th International Conference on E-Science (e-Science), pp. 1–8, October 2012
2. Cordon, O., Herrera, F., Villar, P.: Generating the knowledge base of a fuzzy rule-based system by the genetic learning of the data base. IEEE Trans. Fuzzy Syst. **9**(4), 667–674 (2001)
3. Cotes-Ruiz, I.T., Prado, R.P., Munoz-Exposito, J.E., Garcia-Galan, S., Ruiz-Reyes, N.: Dynamic voltage frequency scaling simulator for real workflows energy-aware management in green cloud computing. PLoS ONE (2017). http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0169803
4. Foster, I., Kesselman, C.: The Grid 2: Blueprint for a New Computing Infrastructure. Morgan Kaufmann Publishers Inc., San Francisco (2003)
5. Juve, G., Deelman, E., Vahi, K., Mehta, G., Berriman, B., Berman, B.P., Maechling, P.: Scientific workflow applications on Amazon EC2. In: 2009 5th IEEE International Conference on E-Science Workshops, pp. 59–66, December 2009
6. Juve, G., Chervenak, A., Deelman, E., Bharathi, S., Mehta, G., Vahi, K.: Characterizing and profiling scientific workflows. Fut. Gener. Comput. Syst. **29**(3), 682–692 (2013). Special Section: Recent Developments in High Performance Computing and Security
7. Prado, R.P., García-Galán, S., Muñoz Expósito, J.E., Yuste, A.J.: Knowledge acquisition in fuzzy rule based systems with particle swarm optimization. IEEE Trans. Fuzzy Syst. **18**(6) (2010)
8. Prado, R.P., Garcia-Galan, S., Munoz-Exposito, J.E., Yuste, A.J., Bruque, S.: Genetic fuzzy rule-based meta-scheduler for grid computing. In: Fourth International Workshop on Genetic and Evolutionary Fuzzy Systems, GEFS 2010, pp. 51–56, March 2010
9. Prado, R.P., Garcia-Galan, S., Yuste, A., Munoz-Exposito, J.: Genetic fuzzy rule-based scheduling system for grid computing in virtual organizations. Soft Comput. - A Fusion of Foundations, Methodologies and Applications **15**, 1255–1271 (2011)
10. Prado, R.P., García-Galán, S., Yuste, A.J., Muñoz Expósito, J.E.: A fuzzy rule-based meta-scheduler with evolutionary learning for grid computing. Eng. Appl. Artif. Intell. **23**(7), 1072–1082 (2010)
11. Prathibha, S.: Monitoring the performance analysis of executing workflow applications with different resource types in a cloud environment. In: 1st International Symposium on Big Data and Cloud Computing Challenges, ISBCC-2014, 27–28 March 2013. VIT University, Chennai (2013)
12. Team, P.: Scientific workflow Epigenomics. https://pegasus.isi.edu/application-showcase/epigenomics/

Learning Management Systems on Blended Learning Courses: An Experience-Based Observation

Mehmet Şükrü Kuran[1(✉)], Jens Myrup Pedersen[2], and Raphael Elsner[3]

[1] Abdullah Gul University, Kayseri, Turkey
sukru.kuran@agu.edu.tr
[2] Aalborg University, Aalborg, Denmark
jens@es.aau.dk
[3] Hamburg University of Technology (TUHH), Hamburg, Germany
raphael.elsner@tuhh.de

Abstract. This paper gives an overview of Learning Management System (LMS) features based on observations on a blended learning course under the Erasmus+ project COLIBRI. We explain the main features of LMSes under two main categories: accessibility content-related and underline the capabilities of four LMSes, Moodle, Blackboard Learn, Canvas, and Stud.IP with respect to these. We explain how these features were utilized to increase the efficiency, tractability, and quality of experience of the course. We found that an LMS with advanced features such as progress tracking, modular course support, interactive content support, and content access restriction is of paramount importance for blended learning courses.

1 Introduction

Among the various aspects of teaching a course, organization of the course materials is a component with paramount importance. A well-organized course is not only tractable and easy to manage, but also transparent for all the stakeholders involved (i.e., lecturers, students, and the administration). With the prevalent usage of ICT systems in education, new ways of presenting content to the students such as video lectures, interactive videos, discussion boards, etc. has become an integral part of many higher education courses. In order to manage this, more and more higher education institutions has started to Learning Management Systems (LMS) for the administration and organization of courses, including Moodle, Blackboard Learn, and Canvas.

As of 2017, there are more than 180 LMSes available on the market with different capabilities and functionalities [10]. In general, an LMS offers tools for course administration, assignment management, progression tracking, grading, and student collaboration. Most of the LMSes are web-based software applications with strong mobile platform support to allow higher accessibility and better suitability for e-learning programs. Although some LMSes are designed

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*,
Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_17

specifically for e-learning courses and programs or Massive Open Online Courses (MOOCs), the most prominent LMSes are designed for both classical courses and e-learning programs and contains functionalities for both.

The main contribution of this paper is an experience-based analysis of using LMSes in a blended learning course (a course with both virtual and physical learning components) to help institutions in choosing a suitable LMS and selecting relevant features for their courses. Our experience is based on the course "Future Internet Opportunities" offered by seven universities, one governmental organization, and two enterprises that work together in the Erasmus+ project COLIBRI [3]. The course is taken by students from all the universities, and consists of two virtual and two physical phases as described:

- First virtual phase: In this phase, students follow a number of modules offered over an LMS. We offer 10 different modules, where each can be taken at introductory, basic, or advanced levels. All students follow all modules at introductory level, and then choose those to be taken at basic and advanced levels. The topics include more technical topics such as wireless networks and future Internet architecture, more usability focused topics such as Information systems and Services and applications, and more business oriented topics such as Enterprise architecture and Entrepreneurship. The modules are mainly based on online learning materials (i.e., video lectures), but some also make use of peer learning activities.
- First physical phase: After the first virtual phase, the students convene in a physical location for a five-day-long midway seminar. The first days of this seminar focus on finishing the modules, along with training in group work, whereas the last days focus more on the problem based project work in groups.
- Second virtual phase: In the third phase, the students work in groups online on the projects assigned to them. They mainly utilize various virtual collaboration tools as well as the content offered in the first phase though the course LMS.
- Second physical phase: Finally, the students meet again in another location for a final five-day-long physical final seminar. In this phase, all students, supervisors and company representatives meet each other to finalize the projects and conduct the exams.

2 Learning Management Systems

For many years, LMSes are mainly hosted by the institution offering the courses in question. In the last couple of years, more prominent LMSes are being offered by the vendors as Software as a Service (SaaS) in which the LMS and all of the data is located in a cloud server managed by the vendor. Unlike classical LMS services, the vendor is responsible for managing and maintaining the LMS as well as keeping the data secure and accessible. The main advantage of the SaaS approach is greatly reducing the maintenance load of the software by the institution which is especially desirable for smaller institutions with limited IT

support capabilities. However, using a SaaS-based LMS incurs a potential security and privacy vulnerability since the institution has limited control over all of its course contents.

Among the high number of LMSes available on the market only a handful of them are widely used by institutions. According to LMS usage research in 2016, three systems in particular share the majority of the LMS market in the US, Canada region: Blackboard Learn [1], Moodle [8], and Canvas [2] with 33%, 20%, and 20% of all institutions respectively [5]. On the other hand, in the European region, Moodle dominates with a very high usage amount with 65% followed by Blackboard Learn by 12%. Although Canvas currently only has a very low percentage of usage in Europe, 33% of all newly installed LMSes in Europe is Canvas [5]. In both regions, in addition to these three LMSes, a variety of locally-developed LMSes show a high amount of usage in their home countries such as It's Learning in Norway [7], Stud.IP in Germany [11], Olat in Switzerland [9], and D2L in Canada [4].

We focus on four LMSes in this paper. First, Moodle is an open source LMS that has been initially released in 2002. It is been distributed under GNU General Public license and has been maintained by the Moodle community. Being an open source system with high customizability, it has been adapted by many institutions as the main LMS used in both classical learning, e-learning, and blended learning programs. Second, Blackboard Learn is a proprietary LMS developed and managed by Blackboard Inc. since 1997. Initially released as a locally hosted LMS, but in the last couple of years, Blackboard Learn has also been offered as a SaaS. Similar to Moodle, it offers high customizability and it is the most widely used LMS in the US, Canada region. Third, Canvas is a relatively new LMS, initially released in 2008 by Instructure. It is offered both as an open-source LMS free of charge if locally hosted or as a paid SaaS. Although a new system, it has attained a high amount of popularity over the last couple of years. Finally, we chose to add StudIP in our comparisons since it is a prominent European-originated LMS which is being used widely in Germany (Table 1).

Table 1. LMS licensing

LMS name	License (Local Hosted)	License (SaaS)
Moodle	Free (Open-Source)	N/A
Blackboard learn	Paid service	Paid service
Canvas	Free (Open-Source)	Paid service
Stud.IP	Free (Open-Source)	N/A

We investigate LMS features in two broad categories: accessibility and content. Accessibility features focus on how LMS data can be accessed and how the LMS communicates with systems outside the LMS, whereas content features deal with tools that help in presenting course content to the students. Although

a variety of LMSes are used among the partners of COLIBRI, we have utilized Moodle in the joint course because of its versatility, customizability, and being very conventional for European instructors and students.

3 Accessibility Features

LMSes are mainly accessed via a web site by all the users of the system. When locally hosted, each institution has its own web site, whereas when SaaS-based there is a single web site for the whole system. In order to increase accessibility, most contemporary LMSes also offer access through mobile applications. Each user, e.g. an instructor or a student, accesses the system via a user account associated with an e-mail address that is in many cases given and managed by the institution. Many LMSes offer an Internal Messaging (IM) system for communication between instructors and students but usually limit the messaging capability of each user based on the institution policy (e.g., an instructor can only sends messages to students taking his course, students cannot send messages to each other). An important part of the IM system is automatic forwarding to external e-mail addresses, a useful feature especially for e-learning and blended learning courses where coordination among instructors and students are conducted mainly via the LMS (Table 2).

Table 2. Accessibility features by LMS

LMS name	Internal messaging	Mobile app	Automatic IM forwarding	Content exportable	Public access
Moodle	Yes	Android, iOS	Yes	Yes	Yes
Blackboard learn	Yes	Android, iOS	Yes	Yes	Yes
Canvas	Yes	Android, iOS	Yes	Yes	Yes
Stud.IP	Yes	Android, iOS	Yes	Yes	Yes

Course materials in many LMSes can be exported or imported as a whole. A course in an LMS is generally specific to a particular semester since course materials tend to change between semesters. Preparing all the materials from scratch in every single semester introduces an additional time-consuming task for the instructor. Using such export/import tools an instructor can re-deploy a course in an upcoming semester. Also, such a feature is useful in cases where a course is to be repeated in another institute.

In some scenarios, LMSes need to communicate with external services. A good example of such an external service is the plagiarism-prevention system called Turnitin. In a text submission assignment, the LMS passes a student's submission to Turnitin which checks if the content is a copy of a text available on the Internet as a percentage value. There are many such external-system-communication-tools available for LMSes. These tools are usually handled by the Learning Tools

Interoperability (LTI) standard which works with most prominent LMSes in the market.

Lastly, some institutions want their content to be available and accessible to the general public, especially when following an open-university policy. A critical issue regarding public access to the LMS content is to be able to differentiate between the general course content and the personalized results (e.g., a student's grades and assignment). Therefore, it is desirable for LMSes to distinguish the parts to be given public access and the parts kept private.

4 Content-Related Features

The course content is composed of various items in an LMS. In the beginning, a basic module for laying out the course is described using a syllabus page which can be considered as the starting point of the course content. According to the syllabus the course content is divided into lecture-related materials (i.e., weekly lecture presentations or video lectures) and graded activities (i.e., quizzes, exams, homework assignments) (Table 3).

In both e-learning and blended learning courses, the main lecture-related material of a course is generally composed of video lectures. This can be done by both directly uploading videos to the LMS or uploading the video to commonly-used video platforms (e.g., Youtube) and putting a link to these videos in the course content. Recently via the usage of an HTML5 technology called H5P [6], some LMSes like Moodle allow interactive content such as videos with interactive components (e.g., videos with built-in quizzes) to be added to the course content. Such interactive content aims to increase the attention of the students to the course content by periodically asking the students to interact with the content instead of just passively watching.

Graded activities are another critical part of an LMS. Most LMSes offer a variety of graded assignments including true/false, multiple choice, matching, fill in the blanks, essay answer, arithmetical, file-submission questions. Usually the tools allow the instructors to add automated grading rules in some assignments, mostly for yes/no or multiple choice questions. Instructors can also add some specific comments if a particular wrong answer is given for a given question. Although, it takes time for the instructor to come up with a lot of specific comments on different incorrect answers we observe that it greatly increases the responsiveness of the system which in turn increases the course quality greatly. In addition to grading-by instructors, LMSes generally also allow peer-reviewing in assignments.

In addition to the main content-based resources, most LMSes also offer some kind of forum, discussion board support to support peer-learning. This kind of features allows students to communicate among themselves when they are stuck at a certain part and ask the help of their classmates. Forums and discussion boards are much more useful in an e-learning or blended-learning course in which the students have no-to-limited communication capabilities with their classmates.

Table 3. Content-related features by LMS

LMS name	Activity access restriction	Allows course modules	Progress tracking	Outcome management
Moodle	Date, Activity completion, Grade, Student group, Password	Yes	Yes	No
Blackboard learn	Date, Activity completion, Grade, Student group, Password	Yes	Yes	Yes
Canvas	Date, Student group, IP Address, Password	Yes	Yes	Yes
Stud.IP	Date, Student group, IP Address, Password	No	No	No

These content-related features can be considered as the main features of an LMS. Therefore, all four LMSes mentioned in this paper support these features (i.e., video lectures, quizzes, peer reviews, assignment uploads, forums). The only notable exception is the interactive content via H5P. Currently, only Moodle supports these kind of interactive components.

In some courses, the content of a given course can be asked to be divided into smaller modules. Such modular course building increases the tractability of the course as a whole and allows more granularity for course outcomes and progress tracking. However, such modular-courses might be difficult to follow by the students since they can start from any part of the course and can get overwhelmed by the content's depth. To remedy these issues, some LMSes offer access restrictions over content which allows hiding some parts of the content until a certain prerequisites are met. LMSes offer a variety of access restriction prerequisites such as dates, other activity completions, and getting at least a set amount of grades from previous graded activities.

Utilizing Progression tracking, the students can keep track of their progress in each module and the instructors can track each student's progress. Since each student will work on the modules he/she is taking in a completely asynchronous fashion than his/her classmates, such a progress tracking on an individual student basis allows a high amount of tractability on the student's progress.

Lastly, some LMSes allow an advanced module-based feature called module-based outcome management. In this feature, the instructor sets the contribution of each module to the overall outcomes of the course. At the end of the course, based on the student's grades on each module, the LMS can provide a report on each students attained level in each outcome separately. This report will provide a much more granular outcome compared to a single letter grade from a course. Among the four LMSes we are investigating in this paper, only Canvas has such a feature.

5 Experiences on LMS Features in Blended Learning

In our course offered by the COLIBRI project, we come to realize that the user interface of the LMS plays a very important part. LMSes have different design approaches in their user interface and users (both students and instructors) are used to the user interfaces of the LMS at their home university. This familiarization becomes somewhat of a problem when adapting to LMSes with different user interface design choices. If the user interface of different LMSes follow a similar design choice, this familiarization could be transformed into a benefit instead of an issue to be resolved.

We have utilized several of the features explained in the previous sections to a great benefit. One such feature was the automatic IM forwarding since the students and instructors were coming from seven different institutions having different LMSes and e-mail systems. It is our experience that it is very cumbersome for both students and instructors to be using multiple systems for different courses. An automatic IM forwarding system allowed our students and instructors to utilize the dedicated LMS of the course without the need for checking every LMS they are enrolled to individually.

One of the main content types we have utilized in the COLIBRI course has been video lectures. Based on our experience, video lectures are at highest efficiency if they are kept around 5–10 min each. In longer video lectures, it becomes harder to keep the interest of the students. In the third iteration of our course, we also experimented with interactive videos using H5P in which small questions (e.g., a multiple choice or a true/false question) can be inserted inside lectures by automatically pausing the video, showing the question, and only reverting back to the video when the student answers the question. We observe that such additional interactivity considerably increases the attention of the students to the lecture.

In addition to the video lectures, another type of content we have utilized heavily was quizzes. Generally we assigned a quiz assignment after every 1–2 video lectures to evaluate the progress of the students. Moodle allows both automated quiz grading and manual grading in assignments. For simpler questions (e.g., multiple choice questions), we chose to implement automated answer checking and commenting. Such automated grading and commenting greatly increases the responsiveness of the system to the student's activities which in turn increases the attention of the students. In other questions (e.g., written assignments) we used manual grading since written answers can vary greatly from student to student. We observe that since automated grading gives responses to the students instantly, it is very well received by the students whereas manual grading is not that well received. Consequently, we can suggest that shorter assignments aimed to increase the attention level of the students are best developed so that they allow automated grading while mini-project-like longer assignments should be kept as manual graded assignments.

One of the unique aspects of the course offered in the COLIBRI project is the fact that the course is composed of many modules but a single student is only responsible from a set of modules among all the available ones. So, in order

to facilitate such a customizable course over an LMS, we heavily utilized the module support of Moodle. All the content related to a single module is packed together and clearly separated from the rest of the content to eliminate any potential confusion of the students. In each module, the contents are laid out in a clear sequence like video lecture 1 first, then quiz 1, then video lecture 2, etc. One potential issue about such a sequenced modules is to make sure that the students are following the sequence as intended by the instructors. To enforce a mandatory flow inside each module, we utilized some date-based restrictions as well as grade-based restrictions in our course. Such restrictions eliminate the possibility of students not abiding the module flow which might be detrimental to the education value of the course.

6 Conclusion

Based on our observations of our blended learning course under the COLIBRI project, LMSes are indispensable components of blended learning courses. Advanced features such as progress tracking, course modularization, and interactive content increases the quality of blended learning courses considerably. Besides blended learning courses, the observations we have made in the paper also applies to normal course with online content such as courses using the flipped classroom paradigm. We believe that our observations are useful to both institutions that aim to develop similar blended learning courses as well as for the LMS developers in order to decide which new features to develop.

References

1. Blackboard. www.blackboard.com. Accessed 2017
2. Canvas. www.instructure.com. Accessed 2017
3. Colibri project. erasmus-colibri.eu. Accessed 2017
4. Desire2learn. www.d2l.com. Accessed 2017
5. Edutechnica, LMS data - spring 2017 data. http://edutechnica.com/2017/03/12/lms-data-spring-2017-updates/. Accessed 2017
6. H5p. https://h5p.org. Accessed 2017
7. It's learning. www.itslearning.eu. Accessed 2017
8. Moodle. www.moodle.com. Accessed 2017
9. OLAT. www.olat.org. Accessed 2017
10. Openeductaioneuropa, 183 learning management systems (plus a handy comparison guide). www.openeducationeuropa.eu/en/blogs/list-183-learning-management-systems-plus-lmss-comparison-checklist-99-features. Accessed 2013
11. Stud.ip. www.stud.org. Accessed 2017

Technology and Software for Calculating Correct Normalization of Correlation Functions

Ulkar E. Sattarova(✉)

Institute of Control Systems of ANAS,
Bakhtiyar Vahabzadeh street, 9, Baku, Azerbaijan
ulker.rzaeva@gmail.com

Abstract. With the traditional approach, it is assumed that the transition to normalized estimates of the correlation functions of noisy signals makes it possible to simplify calculations, move to a dimensionless value, and thus compare processes with different values of the variances. As this paper shows, the obtained results still contain significant noise-induced errors even when the classical conditions are fulfilled. The technology for corrected calculation of normalized correlation functions of noisy signals is proposed in the paper, which allows us to significantly reduce noise-induced errors, even if the classical conditions do not hold.

1 Introduction

In modern information systems, correlation analysis is used to solve problems of diagnostics, identification, forecasting, optimization, control, etc. In order to solve such problems, information systems have to ensure precise calculation of the estimates of statistical characteristics of technological parameters and find their own and cross-correlation matrices.

It is assumed in literature that with the traditional approach, the classical conditions are fulfilled for the real signal $g(t)$ consisting of the useful signal $X(t)$ and the random noise $\varepsilon(t)$. It means that the process under investigation is a standard ergodic process, the random noise has zero mathematical expectation $m_\varepsilon = o$, the uncorrelated values of the samples $\frac{1}{N}\sum_{i=1}^{N} \overset{\circ}{\varepsilon}(i\Delta t)\,\overset{\circ}{\varepsilon}((i+\mu)\,\Delta t) = o$ when $\mu \neq o$, the useful signal $X(t)$ and the noise $\varepsilon(t)$ obey the normal distribution law with no correlation between them [3,11]. However, as shown in [1,2], when these conditions hold, the estimates of the autocorrelation functions of noise signals have no errors at all time shifts μ, except for zero time shift $\mu = o$, when the estimate consists of the sum of the estimates of the autocorrelation function of the useful signal and the noise variance. Meanwhile, the estimates of the cross-correlation functions do not contain noise-induced errors at all time shifts μ. According to the existing literature [3,11], in order to eliminate the effects of the noise on the estimate of the autocorrelation function at zero time shift $\mu = o$, it is appropriate to proceed to the normalized estimates of correlation functions. It is known that standardization or normalization reduces the values

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*,
Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_18

of all transformed variables to the single range of values by expressing them through the relationship of these values to some value that reflects certain properties of a concrete attribute [4]. At the same time, the normalization process indeed allows us to independently compare cross-correlations of absolute values of data [6]. In addition, the normalization of the indicators is often required, and this is convenient in order to proceed to dimensionless variables [9]. Also in most cases, random functions are approximated by a normalized correlation function [10]. It is customary to immediately proceed to normalized correlation functions. Normalized correlation functions are assumed to be convenient, because their values do not exceed unity [8]. It is also known that to estimate the degree of dependence of the cross sections of a random function it is more convenient to use the normalization of the correlation function [7]. Also, it is generally accepted that the normalized correlation function is used to be able to compare processes with different values of variances [5].

However, it is natural that in the case when the classical conditions [1–3,11] are not satisfied, the magnitude of the error from the effect of the noise on the estimates of the normalized correlation functions will be even more significant, and this will lead to inadequate solutions to the above problems. In this regard, there is a need for the development of the information technology focused on the calculation of "corrected" normalized correlation functions. This article is devoted to solving this problem.

2 Problem Statement

It is shown in literature that the normalized autocorrelation function is calculated from the formula [3,11]:

$$r_{\mathring{X}\,\mathring{X}}(\mu) = R_{\mathring{X}\,\mathring{X}}(\mu) \Big/ D(x) \tag{1}$$

where $X(t)$ is the useful signal, $D(x)$ is its variance, $R_{\mathring{X}\,\mathring{X}}(\mu)$ is the correlation function.

$$R_{\mathring{X}\,\mathring{X}}(\mu) = \frac{1}{N} \sum_{i=1}^{N} \mathring{X}(i\Delta t)\, \mathring{X}((i+\mu)\,\Delta t)$$

where $\mu = o$, $\mu = \Delta t$, $\mu = 2\Delta t$, $\mu = 3\Delta t$, ... The normalized autocorrelation function of the noisy signal consisting of the random useful signal $X(t)$ and the noise $\varepsilon(t)$

$$g(t) = X(t) + \varepsilon(t) \tag{2}$$

is calculated from the formula:

$$r_{\mathring{g}\,\mathring{g}}(\mu) = R_{\mathring{g}\,\mathring{g}}(\mu) \Big/ D(g) \tag{3}$$

Similarly, the formula for calculating the estimates of the normalized cross-correlated functions $r_{\mathring{x}_1\,\mathring{x}_2}(\mu)$ of the useful signals $X_1(t)$ and $X_2(t)$ looks as follows:

$$r_{\mathring{X}_1\,\mathring{X}_2}(\mu) = R_{\mathring{X}_1\,\mathring{X}_2}(\mu) \Big/ \sqrt{D(x_1) \cdot D(x_2)} \tag{4}$$

The normalized cross-correlation function of the noisy signals $g_1(t)$ and $g_2(t)$ consisting of the random useful signals $X_1(t)$ and $X_2(t)$ and the noises $\varepsilon_1(t)$ and $\varepsilon_2(t)$, respectively,

$$g_1(t) = X_1(t) + \varepsilon_1(t), \tag{5}$$

$$g_2(t) = X_2(t) + \varepsilon_2(t), \tag{6}$$

is calculated from the formula:

$$r_{\overset{\circ}{g}_1 \overset{\circ}{g}_2}(\mu) = R_{\overset{\circ}{g}_1 \overset{\circ}{g}_2}(\mu) \Big/ \sqrt{D(g_1) \cdot D(g_2)} \tag{7}$$

The following equalities hold true between the estimates of the normalized autocorrelation function $r_{\overset{\circ}{g}\overset{\circ}{g}}(\mu)$ of the noisy signal $g(t)$ and the estimates of the normalized autocorrelation function $r_{\overset{\circ}{X}\overset{\circ}{X}}(\mu)$ of the useful signal $X(t)$, as well as between the estimates of the normalized cross-correlation function $r_{\overset{\circ}{g}_1 \overset{\circ}{g}_2}(\mu)$ of the noisy signals $g_1(t)$ and $g_2(t)$ and the estimates of the normalized cross-correlation function $r_{\overset{\circ}{X}_1 \overset{\circ}{X}_2}(\mu)$ of the useful signals $X_1(t)$ and $X_2(t)$ [1, p. 172] and [2, p. 112]:

$$r_{\overset{\circ}{g}\overset{\circ}{g}}(\mu) \approx r_{\overset{\circ}{X}\overset{\circ}{X}}(\mu), \tag{8}$$

$$r_{\overset{\circ}{g}_1 \overset{\circ}{g}_2}(\mu) \approx r_{\overset{\circ}{X}_1 \overset{\circ}{X}_2}(\mu). \tag{9}$$

However, comparing experimentally expression (1) with expression (3), and expression (4) with expression (7), we can see that the estimates of the normalized auto- and cross-correlation functions of the useful signals differ significantly from the estimates of normalized auto and cross-correlation functions of the noisy signals, i.e. [1,2,4,6]

$$r_{\overset{\circ}{g}\overset{\circ}{g}}(\mu) \neq r_{\overset{\circ}{X}\overset{\circ}{X}}(\mu) \tag{10}$$

$$r_{\overset{\circ}{g}_1 \overset{\circ}{g}_2}(\mu) \neq r_{\overset{\circ}{X}_1 \overset{\circ}{X}_2}(\mu) \tag{11}$$

This paper seeks to develop a technology that allows us to obtain corrected estimates of the normalized correlation functions of noisy signals that ensure the following equalities:

$$r^R_{\overset{\circ}{g}\overset{\circ}{g}}(\mu) \approx r_{\overset{\circ}{X}\overset{\circ}{X}}(\mu) \tag{12}$$

$$r^R_{\overset{\circ}{g}_1 \overset{\circ}{g}_2}(\mu) \approx r_{\overset{\circ}{X}_1 \overset{\circ}{X}_2}(\mu) \tag{13}$$

The algorithm for calculating the noise variance was made using the following formula (14). In [1,2], it is shown that the noise variance of the noisy signal $g(i\Delta t)$ can be calculated from the expression:

$$\begin{aligned} D(g) &= R_{\overset{\circ}{g}\overset{\circ}{g}}(\mu = 0) = \tfrac{1}{N} \sum_{i=1}^{N} \overset{\circ}{g}(i\Delta t)^2 = \tfrac{1}{N} \sum_{i=1}^{N} \left(\overset{\circ}{X}(i\Delta t) + \overset{\circ}{\varepsilon}(i\Delta t) \right)^2 \\ D_\varepsilon &= \tfrac{1}{N} \sum_{i=1}^{N} \begin{bmatrix} \overset{\circ}{g}^2(i\Delta t) + \overset{\circ}{g}(i\Delta t)\, \overset{\circ}{g}((i+2)\Delta t) \\ -2\, \overset{\circ}{g}(i\Delta t)\, \overset{\circ}{g}((i+1)\Delta t) \end{bmatrix} \end{aligned} \tag{14}$$

3 Technology for Calculating the Corrected Estimates of Normalized Correlation Functions

In the following paragraphs, we will discuss a new technology for calculating the corrected estimates of normalized correlation functions, using the calculation of the estimates of the noise variances D_ε. For this purpose, the following technology is proposed for calculating the corrected estimates of the normalized autocorrelation and cross-correlation functions of noisy signals. We calculate the following:

1. estimates of the auto-correlation function of the centered noisy signal

$$\overset{\circ}{g}(t) : R_{\overset{\circ}{g}\overset{\circ}{g}}(\mu) = \frac{1}{N}\sum_{i=1}^{N} \overset{\circ}{g}(i\Delta t)\,\overset{\circ}{g}((i+\mu)\,\Delta t), \overset{\circ}{g}(t) = g(t) - m_g; \tag{15}$$

2. estimates of the normalized auto-correlation function of the noisy signal

$$\begin{aligned} \overset{\circ}{g}(t) : r_{\overset{\circ}{g}\overset{\circ}{g}}(\mu) &= R_{\overset{\circ}{g}\overset{\circ}{g}}(\mu) \Big/ D(g) \\ &= \left(\frac{1}{N}\sum_{i=1}^{N} \overset{\circ}{g}(i\Delta t)\,\overset{\circ}{g}((i+\mu)\,\Delta t)\right) \Big/ \frac{1}{N}\sum_{i=1}^{N} \overset{\circ}{g}(i\Delta t)\,\overset{\circ}{g}(i\Delta t) \end{aligned} \tag{16}$$

3. the estimates of the variance D_ε of the noise $\varepsilon(i\Delta t)$:

$$D_\varepsilon = \frac{1}{N}\sum_{i=1}^{N}\left[\overset{\circ}{g}^{2}(i\Delta t) + \overset{\circ}{g}(i\Delta t)\,\overset{\circ}{g}((i+2)\,\Delta t) - 2\,\overset{\circ}{g}(i\Delta t)\,\overset{\circ}{g}((i+1)\,\Delta t)\right] \tag{17}$$

4. corrected estimates of the normalized auto-correlation function when $\mu \neq o$:

$$r^{R}_{\overset{\circ}{g}\overset{\circ}{g}}(\mu) = R_{\overset{\circ}{g}\overset{\circ}{g}}(\mu) \Big/ (D(g) - D_\varepsilon) \tag{18}$$

where

$$D(g) = \frac{1}{N}\sum_{i=1}^{N} \overset{\circ}{g}^{2}(i\Delta t) \tag{19}$$

or

$$\begin{aligned} r^{R}_{\overset{\circ}{g}\overset{\circ}{g}}(\mu) = &\left(\frac{1}{N}\sum_{i=1}^{N} \overset{\circ}{g}(i\Delta t)\,\overset{\circ}{g}((i+\mu)\,\Delta t)\right) \Big/ \\ &\frac{1}{N}\sum_{i=1}^{N}\left[2\,\overset{\circ}{g}(i\Delta t)\,\overset{\circ}{g}((i+1)\,\Delta t) - \overset{\circ}{g}(i\Delta t)\,\overset{\circ}{g}((i+2)\,\Delta t)\right] \end{aligned} \tag{20}$$

5. corrected estimates of the normalized auto-correlation function

$$r^{R}_{\overset{\circ}{g}\overset{\circ}{g}}(\mu) = \begin{cases} 1 & \text{when } \mu = 0 \\ \dfrac{\frac{1}{N}\sum\limits_{i=1}^{N}\left[\overset{\circ}{g}(i\Delta t)\,\overset{\circ}{g}((i+\mu)\Delta t)\right]}{\frac{1}{N}\sum\limits_{i=1}^{N}\left[2\,\overset{\circ}{g}(i\Delta t)\,\overset{\circ}{g}((i+1)\Delta t) - \overset{\circ}{g}(i\Delta t)\,\overset{\circ}{g}((i+2)\Delta t)\right]} & \text{when } \mu \neq 0 \end{cases} \tag{21}$$

at any value of the time shift.

The technology for calculating the robust estimates of the normalized cross-correlation functions of the noisy signals is as follows:

1. estimates of the cross-correlation function of the noisy signals $\overset{\circ}{g}_1(t)$ and

$$\begin{gathered}\overset{\circ}{g}_2(t): R_{\overset{\circ}{g}_1\overset{\circ}{g}_2}(\mu)=\frac{1}{N}\sum_{i=1}^{N}\overset{\circ}{g}_1(i\Delta t)\overset{\circ}{g}_2((i+\mu)\Delta t),\\ \overset{\circ}{g}_1(t)=g_1(t)-m_{g_1},\overset{\circ}{g}_2(t)=g_2(t)-m_{g_2}\end{gathered}\tag{22}$$

2. estimates of the normalized cross-correlation function of the noisy signals:

$$\begin{aligned}r_{\overset{\circ}{g}_1\overset{\circ}{g}_2}(\mu)&=\frac{R_{\overset{\circ}{g}_1\overset{\circ}{g}_2}(\mu)}{\sqrt{D(g_1)\cdot D(g_2)}}\\&=\frac{\frac{1}{N}\sum\limits_{i=1}^{N}\overset{\circ}{g}_1(i\Delta t)\overset{\circ}{g}_2((i+\mu)\Delta t)}{\sqrt{\left(\frac{1}{N}\sum\limits_{i=1}^{N}\overset{\circ}{g}_1(i\Delta t)\overset{\circ}{g}_1(i\Delta t)\right)\cdot\left(\frac{1}{N}\sum\limits_{i=1}^{N}\overset{\circ}{g}_2(i\Delta t)\overset{\circ}{g}_2(i\Delta t)\right)}}\end{aligned}\tag{23}$$

3. the estimates of the variance D_{ε_1} of the noise $\varepsilon_1(i\Delta t)$:

$$\begin{aligned}D_{\varepsilon_1}=\frac{1}{N}\sum_{i=1}^{N}\Big[&\overset{\circ}{g}_1(i\Delta t)\overset{\circ}{g}_1(i\Delta t)+\overset{\circ}{g}_1(i\Delta t)\overset{\circ}{g}_1((i+2)\Delta t)\\&-2\overset{\circ}{g}_1(i\Delta t)\overset{\circ}{g}_1((i+1)\Delta t)\Big]\end{aligned}\tag{24}$$

4. the estimates of the variance D_{ε_2} of the noise $\varepsilon_2(i\Delta t)$:

$$\begin{aligned}D_{\varepsilon_2}=\frac{1}{N}\sum_{i=1}^{N}\Big[&\overset{\circ}{g}_2(i\Delta t)\overset{\circ}{g}_2(i\Delta t)+\overset{\circ}{g}_2(i\Delta t)\overset{\circ}{g}_2((i+2)\Delta t)\\&-2\overset{\circ}{g}_2(i\Delta t)\overset{\circ}{g}_2((i+1)\Delta t)\Big]\end{aligned}\tag{25}$$

5. corrected estimates of the normalized cross-correlation function will be as follows at any value of μ:

$$r^{R}_{\overset{\circ}{g}_1\overset{\circ}{g}_2}(\mu)=R_{\overset{\circ}{g}_1\overset{\circ}{g}_2}(\mu)\Big/\sqrt{(D(g_1)-D_{\varepsilon_1})}\sqrt{(D(g_2)-D_{\varepsilon_2})}\tag{26}$$

where

$$\begin{aligned}D(g_1)&=\frac{1}{N}\sum_{i=1}^{N}\overset{\circ}{g}_1(i\Delta t)\overset{\circ}{g}_1(i\Delta t),\\D(g_2)&=\frac{1}{N}\sum_{i=1}^{N}\overset{\circ}{g}_2(i\Delta t)\overset{\circ}{g}_2(i\Delta t),\end{aligned}\tag{27}$$

or

$$\begin{aligned}r^{R}_{\overset{\circ}{g}_1\overset{\circ}{g}_2}(\mu)=&\frac{\frac{1}{N}\sum\limits_{i=1}^{N}\overset{\circ}{g}_1(i\Delta t)\overset{\circ}{g}_2((i+\mu)\Delta t)}{\sqrt{\frac{1}{N}\sum\limits_{i=1}^{N}2\overset{\circ}{g}_1(i\Delta t)\overset{\circ}{g}_1((i+1)\Delta t)-\overset{\circ}{g}_1(i\Delta t)\overset{\circ}{g}_1((i+2)\Delta t)}}\\&\cdot\frac{1}{\sqrt{\frac{1}{N}\sum\limits_{i=1}^{N}2\overset{\circ}{g}_2(i\Delta t)\overset{\circ}{g}_2((i+1)\Delta t)-\overset{\circ}{g}_2(i\Delta t)\overset{\circ}{g}_2((i+2)\Delta t)}}\end{aligned}\tag{28}$$

4 Technology of Computational Experiments

To validate the new technology for correct calculation of normalized correlation functions, a large number of computational experiments have been carried out using the latest version of MATLAB Signal Processing Toolbox - R2017a.

When performing computational experiments, useful signals $X(t)$, noises $\varepsilon(t)$ and noisy signals $g(i\Delta t) = X(i\Delta t) + \varepsilon(i\Delta t)$ with given characteristics were generated for estimating the autocorrelation functions. For each generated useful signal $X(t)$, the condition of constancy of the mathematical expectation was verified, i.e., the satisfaction of the equality

$$m_x(T_1) \approx m_x(T_2) \approx \ldots \approx m_x(T_n) \approx m_x$$

or the inequality

$$m_x(T_1) \neq m_x(T_2) \neq \ldots \neq m_x(T_n) \neq m_x$$

Then, the estimates of the normalized autocorrelation functions $r_{\overset{\circ}{X}\overset{\circ}{X}}(\mu)$ the useful signals calculated from formula (1), the autocorrelation functions $r_{\overset{\circ}{g}\overset{\circ}{g}}(\mu)$ of the noisy signals were calculated from formula (16), the variance D_ε of the noise $\varepsilon(t)$ from formula (14), and finally, the corrected estimates of the normalized autocorrelation function $r^R_{\overset{\circ}{g}\overset{\circ}{g}}(\mu)$ obtained from formula (21) were computed, and the estimates of the normalized correlation functions were compared. To this end, the following were calculated: values of the relative errors of the estimates of the normalized autocorrelation functions $r_{\overset{\circ}{g}\overset{\circ}{g}}(\mu)$ and the corrected estimates of $r^R_{\overset{\circ}{g}\overset{\circ}{g}}(\mu)$ of the noisy signals from the expressions:

$$\begin{aligned} \Delta r_{\overset{\circ}{X}\overset{\circ}{X}}(\mu) &= \frac{\left|r_{\overset{\circ}{g}\overset{\circ}{g}}(\mu) - r_{\overset{\circ}{X}\overset{\circ}{X}}(\mu)\right|}{\left|r_{\overset{\circ}{X}\overset{\circ}{X}}(\mu)\right|} * 100 \\ \Delta r^R_{\overset{\circ}{X}\overset{\circ}{X}}(\mu) &= \frac{\left|r^R_{\overset{\circ}{g}\overset{\circ}{g}}(\mu) - r_{\overset{\circ}{X}\overset{\circ}{X}}(\mu)\right|}{\left|r_{\overset{\circ}{X}\overset{\circ}{X}}(\mu)\right|} * 100\%. \end{aligned} \tag{29}$$

When performing computational experiments, useful signals $X_1(t)$, $X_2(t)$, noises $\varepsilon_1(t)$, $\varepsilon_2(t)$ and noisy signals $g_1(i\Delta t) = X_1(t) + \varepsilon_1(t)$, $g_2(i\Delta t) = X_2(t) + \varepsilon_2(t)$ were generated for estimating the cross-correlation functions. For each generated useful signal $X_1(t)$, $X_2(t)$, the condition of constancy of the mathematical expectation was verified, i.e., the satisfaction of the equalities

$$m_{x_1}(T_1) \approx m_{x_1}(T_2) \approx \ldots \approx m_x(T_n) \approx m_{x_1},$$

$$m_{x_2}(T_1) \approx m_{x_2}(T_2) \approx \ldots \approx m_{x_2}(T_n) \approx m_{x_2},$$

or the inequalities

$$m_{x_1}(T_1) \neq m_{x_1}(T_2) \neq \ldots \neq m_x(T_n) \neq m_{x_1},$$

$$m_{x_2}(T_1) \neq m_{x_2}(T_2) \neq \ldots \neq m_{x_2}(T_n) \neq m_{x_2},$$

Then, the estimates of the normalized cross-correlation functions $r_{\mathring{X}_1\mathring{X}_2}(\mu)$, $r_{\mathring{g}_1\mathring{g}_2}(\mu)$ of the useful signals $X_1(t)$, $X_2(t)$ and the noisy signals $g_1(t)$, $g_2(t)$ were calculated from expressions (4) and (23); the values of the variances D_{ε_1}, D_{ε_2} of the noises $\varepsilon_1(t)$, $\varepsilon_2(t)$ from formulas (24) and (25); the robust estimates of the normalized cross-correlation function $r^R_{\mathring{g}_1\mathring{g}_2}(\mu)$ from formula (26) and a comparative analysis was performed. To this end, the following were calculated:

1. values of the relative errors of the estimates of the normalized cross-correlation functions $r_{\mathring{g}_1\mathring{g}_2}(\mu)$ of the noisy signals from the expression:

$$\Delta r_{\mathring{X}_1\mathring{X}_2}(\mu) = \left| r_{\mathring{g}_1\mathring{g}_2}(\mu) - r_{\mathring{X}_1\mathring{X}_2}(\mu) \right| \Big/ \left| r_{\mathring{X}_1\mathring{X}_2}(\mu) \right| * 100\% \tag{30}$$

2. values of the relative errors of the robust estimates of the normalized cross-correlation functions $r^R_{\mathring{g}_1\mathring{g}_2}(\mu)$ from the expression:

$$\Delta r^2_{\mathring{X}_1\mathring{X}_2}(\mu) = \left| r^2_{\mathring{g}_1\mathring{g}_2}(\mu) - r_{\mathring{X}_1\mathring{X}_2}(\mu) \right| \Big/ \left| r_{\mathring{X}_1\mathring{X}_2}(\mu) \right| * 100\%$$

5 Results of the Computational Experiments

Results of the computational experiments for the estimates of normalized auto-correlation functions.

5.1 The First Type of Experiments

Useful signal $X(t)$ is generated as a sum of harmonic oscillations and the classical conditions are fulfilled for it. Noise obeys different distribution laws with the mathematical expectation $m_\varepsilon \approx o$.

Experiment N1.
Useful signal $X(i\Delta t) = 40\sin(i\Delta t) + 100$. Noise $\varepsilon(t)$ obeys the normal distribution law with the variance $D(\varepsilon) \approx 90$. Thus, a useful signal and a noise are generated that satisfy the classical conditions.

Experiment N2.
Useful signal $X(i\Delta t) = 40\sin(i\Delta t) + 25\cos(0.5 \cdot i\Delta t) + 120$. Noise $\varepsilon(t)$ obeys the binomial distribution law with the variance $D(\varepsilon) \approx 130$.

Experiment N3.
Useful signal $X(i\Delta t) = 40\sin(i\Delta t) + 25\cos(0.5 \cdot i\Delta t) + \sin(10 \cdot i\Delta t) + 100$. Noise $\varepsilon(t)$ obeys the exponential distribution law with $D(\varepsilon) \approx 229$.

Experiment N4.
Useful signal $X(i\Delta t) = 40\sin(i\Delta t) + 25\cos(0.5 \cdot i\Delta t) + 3\sin(2 \cdot i\Delta t) + 115$. Noise $\varepsilon(t)$ obeys beta-distribution with $D(\varepsilon) \approx 136$.

Experiment N5.
Useful signal $X(i\Delta t) = 40\sin(i\Delta t) + 25\cos(0.5 \cdot i\Delta t) + 3\sin(2 \cdot i\Delta t) + 115$. Noise $\varepsilon(t)$ obeys gamma-distribution with $D(\varepsilon) \approx 52$.

5.2 The Second Type of Experiments

Useful signal $X(t)$ is generated as a sum of harmonic oscillations and the condition of constant mathematical expectation is violated for it and inequalities hold. The distribution law of the noise $\varepsilon(t)$ is other than normal, mathematical expectation $m_\varepsilon \approx o$. Thus, the classical conditions are violated for the useful signal and the noise.

Experiment N6.
Useful signal $X(i\Delta t) = 50\sin(i\Delta t) + 5\cos(0.5 \cdot i\Delta t) + 2\sin(i\Delta t) + \cdots + 5\cos(13 \cdot i\Delta t) + 15$. Noise $\varepsilon(t)$ obeys the lognormal distribution with $D(\varepsilon) \approx 665$.

Results of the computations for Experiment N1 and Experiment N6 are given in Table 1. Similar results have been obtained for Experiments N2, N3, N4, N5.

Results of the computational experiments for the estimates of normalized cross-correlation functions.

Table 1. Results of computational experiments

	$r_{\dot{X}\dot{X}}(\mu)$	$r_{gg}(\mu)$	$r^{R}_{gg}(\mu)$	$\Delta r_{\dot{X}\dot{X}}(\mu)$	$\Delta^{R} r_{\dot{X}\dot{X}}(\mu)$
1	2	3	4	5	6
Experiment N1 $D(\varepsilon)$=89,7642 D_ε=80,6617	1	1	1	0	0
	0,9995	0,9014	0,9914	9,814	0,8115
	0,998	0,8936	0,9828	10,4641	1,5265
	0,9956	0,8982	0,9879	9,7756	0,7692
	0,9921	0,8898	0,9786	10,3116	1,3587
	0,9877	0,8861	0,9746	10,2834	1,3277
	0,9823	0,8878	0,9764	9,6176	0,5954
	0,9759	0,8778	0,9654	10,0523	1,0735
	0,9686	0,8727	0,9598	9,8999	0,9059
	0,9603	0,8711	0,9581	9,2873	0,2322
Experiment N6 $D(\varepsilon)$=106,59 D_ε=101,35	1	1	1	0	0
	0,9915	0,7713	0,9938	22,2141	0,2278
	0,9701	0,7664	0,9875	20,9924	1,8019
	0,9457	0,7496	0,9658	20,7376	2,1302
	0,9294	0,7479	0,9637	19,5237	3,6944
	0,9277	0,725	0,9342	21,8451	0,7033
	0,9394	0,7325	0,9438	22,0279	0,4677
	0,956	0,7573	0,9758	20,7912	2,0612
	0,9664	0,7573	0,9758	21,6355	0,9734
	0,9617	0,7543	0,9719	21,5688	1,0592

5.3 The First Type of Experiments

Useful signals $X_1(t)$, $X_2(t)$ are generated as sums of harmonic oscillations and the classical conditions are fulfilled for them. Noises $\varepsilon_1(t)$, $\varepsilon_2(t)$ obey different distribution laws with mathematical expectations $m_{\varepsilon_1} \approx o$, $m_{\varepsilon_2} \approx o$.

Experiment N7.
Useful signals $X_1(i\Delta t) = 40\sin(i\Delta t)+100$, $X_2(i\Delta t) = 40\sin(1.15\cdot i\Delta t)+200$. Noises $\varepsilon_1(t)$, $\varepsilon_2(t)$ obey the normal distribution law with variances $D(\varepsilon_1) \approx 89.76$, $D(\varepsilon_2 \approx 430)$. Thus, useful signals and noises are generated that satisfy the classical conditions.

Experiment N8.
Useful signals $X_1(i\Delta t) = 40\sin(i\Delta t) + 100$, $X_2(i\Delta t) = 40\sin(1 \cdot i\Delta t) + \cdots + 20\cos(0.5 \cdot i\Delta t) + 120$. Noise $\varepsilon_1(t)$ obeys the binomial distribution law with $D(\varepsilon_1) \approx= 130.5$, noise $\varepsilon_2(t)$ obeys the normal distribution law with $D(\varepsilon_2) \approx= 359$.

Experiment N9.
Useful signals $X_1(i\Delta t) = 40\sin(1\cdot i\Delta t)+25\cos(0.5\cdot i\Delta t)+\sin(10\cdot i\Delta t)+100$, $X_2(i\Delta t) = 40\sin(i\Delta t) + 100$. Noise $\varepsilon_1(t)$ obeys the exponential distribution law with $D(\varepsilon_1) \approx 228.6$, noise obeys the normal distribution law with $D(\varepsilon_1) \approx 359$.

Experiment N10.
Useful signals $X_1(i\Delta t) = 40\sin(1\cdot i\Delta t)+25\cos(0.5\cdot i\Delta t)+\cdots+3\sin(2\cdot i\Delta t)+115$, $X_2(i\Delta t) = 40\sin(i\Delta t) + \sin(10 \cdot i\Delta t) + 100$. Noise $\varepsilon_1(t)$ obeys gamma-distribution with $D(\varepsilon_1) \approx 52.6$, noise $\varepsilon_2(t)$ obeys the binomial distribution with $D(\varepsilon_2) \approx 123$.

Table 2. Results of computational experiments

	$r_{X_1X_2}(\mu)$	$r_{g_1g_2}(\mu)$	$r^R_{g_1g_2}(\mu)$	$\Delta r_{X_1X_2}(\mu)$	$\Delta r^R_{X_1X_2}(\mu)$
1	2	3	4	5	6
Experiment N7 $D(\varepsilon_1)$ =89,76; D_{ε_1} =81,41 $D(\varepsilon_2)$ =429,9; D_{ε_2} =368,75	0,8283	0,6107	0,7779	26,2702	6,0869
	0,8245	0,6261	0,7975	24,0634	3,2759
	0,8206	0,6055	0,7713	26,2078	6,0073
	0,8166	0,6038	0,769	26,0655	5,826
	0,8126	0,6101	0,7771	24,9205	4,3677
	0,8084	0,5919	0,7539	26,7841	6,7414
	0,8043	0,6003	0,7647	25,3553	4,9214
	0,8	0,5925	0,7547	25,9421	5,6689
	0,7957	0,5965	0,7598	25,032	4,5097
	0,7913	0,5817	0,7409	26,4951	6,3733
Experiment N11 $D(\varepsilon_1)$ =228,64; D_{ε_1} =225,9 $D(\varepsilon_2)$ =51; D_{ε_2} =46,11	0,889	0,815	0,8958	8,3156	0,7633
	0,8889	0,8125	0,893	8,5867	0,4653
	0,8885	0,8177	0,8986	7,9755	1,137
	0,888	0,8127	0,8932	8,4801	0,5825
	0,8873	0,8164	0,8972	7,9904	1,1207
	0,8863	0,8095	0,8896	8,6735	0,3698
	0,8852	0,8114	0,8918	8,3321	0,7451
	0,8838	0,8053	0,8851	8,8814	0,1415
	0,8823	0,8033	0,8829	8,9437	0,0729
	0,8805	0,8033	0,8828	8,7648	0,2695

5.4 The Second Type of Experiments

Useful signals $X_1(t)$, $X_2(t)$ are generated as sums of harmonic oscillations and the condition of constant mathematical expectation is violated for them. Distribution laws of the noises $\varepsilon_1(t)$, $\varepsilon_2(t)$ are other than normal.

Experiment N11.
Useful signals $X_1(i\Delta t) = 50\sin(1 \cdot i\Delta t) + 5\cos(0.5 \cdot i\Delta t) + 2\sin(i\Delta t) + 5\cos(13 \cdot i\Delta t) + 15$, $X_2(i\Delta t) = 40\sin(1 \cdot i\Delta t) + 25\cos(0.5 \cdot i\Delta t) + 3\sin(2 \cdot i\Delta t) + 115$ The condition of constant mathematical expectation is violated for the useful signal $X_1(t)$. Noise $\varepsilon_1(t)$ obeys the exponential distribution with $D(\varepsilon_1) \approx 228.64$, noise $\varepsilon_2(t)$ - gamma-distribution with $D(\varepsilon_2) \approx 52,6$.

Results of the computations for Experiment N7 and Experiment N11 are given in Table 2. Similar results have been obtained for Experiments N8, N9, N10.

6 Conclusion

The following conclusions have been made on the basis of the results of the comparative analysis.

1. The estimates of the noise variances D_ε, D_{ε_1}, D_{ε_2} of the noisy signals $g(t)$, $g_1(t)$, $g_2(t)$, calculated from expressions (14) practically match the given estimates of the variances $D(\varepsilon)$, $D(\varepsilon_1)$, $D(\varepsilon_2)$ of the noises $\varepsilon(t)$, $\varepsilon_1(t)$, $\varepsilon_2(t)$ (column 1 of Tables 1 and 2):
$$D_\varepsilon \approx D(\varepsilon), D_{\varepsilon_1} \approx D(\varepsilon_1), D_{\varepsilon_2} \approx D(\varepsilon_2) \tag{31}$$
2. The estimates of the normalized correlation functions $r_{\mathring{g}\mathring{g}}(\mu)$, $r_{\mathring{g}_1\mathring{g}_2}(\mu)$ of the noisy signals differ significantly from those of the normalized correlation function $r_{\mathring{X}\mathring{X}}(\mu)$, $r_{\mathring{X}_1\mathring{X}_2}(\mu)$ of the useful signals $X(t)$, $X_1(t)$, $X_2(t)$, (columns 2, 3 of Tables 1 and 2).
3. The relative errors $\Delta r_{\mathring{X}_1\mathring{X}_2}(\mu)$, $\Delta r_{\mathring{X}\mathring{X}}(\mu)$ of the estimates of the normalized correlation functions $r_{\mathring{g}\mathring{g}}(\mu)$, $r_{\mathring{g}_1\mathring{g}_2}(\mu)$ of the noisy signals vary in different experiments from 8% to 30% and higher (column 5 of Tables 1 and 2).
4. The relative errors $\Delta r_{\mathring{X}_1\mathring{X}_2}(\mu)$, $\Delta r_{\mathring{X}\mathring{X}}(\mu)$ the corrected estimates of the normalized correlation functions $\Delta r^R_{\mathring{X}\mathring{X}}(\mu)$, $\Delta r^R_{\mathring{X}_1\mathring{X}_2}(\mu)$, are practically equal to zero or decrease from 26–30% to 5–6% even when the classical conditions are violated (column 6 of Tables 1 and 2).
   Thus, the use of the developed robust technology makes it possible to obtain corrected estimates of normalized auto- and cross-correlation functions. Moreover, the proposed technology ensures the robustness of estimates even in such a complex case, when the classical conditions are satisfied neither for useful signals nor for noise.

References

1. Aliev, T.: Robust Technology with Analysis of Interference in Signal Processing. Kluwer Academic/Plenum Publishers, New York (2003)
2. Aliev, T.: Digital Noise Monitoring of Defect Origin. Springer, Boston (2007)
3. Bendat, J.S., Piersol, A.G.: Engineering Applications of Correlation and Spectral Analysis. Wiley-Interscience, New York (1980). 315 pages
4. Bureeva, N.: Mnogomernyy statisticheskiy analiz s ispolzovaniem ppp 'statistica'. Nizhniy Novgorod (2007). [in Russian]
5. Horsthemke, W., Lefever, R.: Noise-induced transitions in physics, chemistry, and biology. Springer (1984)
6. Ifeachor, E.C., Jervis, B.W.: Digital Signal Processing: A Practical Approach. VILYAMS Publishing, Moscow, St. Petersburg, Kyiv (2004)
7. Karpov, F.: Calculation of Urban Electrical Distribution Networks. Energia Publishing House, Moscow (1968). [in Russian]
8. Kuzmin, I., Kedrus, V.: Fundamentals of Information Theory and Coding. The head publishing house of the publishing association 'VISHA SCHOOL', Kiev (1977). [in Russian]
9. Magomedovna, A.: Textbook on Multidimensional Statistical Methods, Makhachkala (2012)
10. Syritsyn, T.: Reliability of hydraulic and pneumatic lines. Mechanical Engineering, vol. 249 (1981). [in Russian]
11. Ventsel, E.: Theory of Probability. Science, Moskow (1969)

Running Sports Decision Aid Tool Based on Reinforcement Learning Approach

Rafał Kozik(✉), Joanna Morzyńska, and Michał Choraś

Institute of Telecommunications and Computer Science,
UTP University of Science and Technology in Bydgoszcz, Bydgoszcz, Poland
rkozik@utp.edu.pl

Abstract. Recently, an increasing number of people is interested in healthy lifestyle and gets involved in a variety of sports activities such as jogging, nordic walking, cycling, hiking or swimming. In order to better plan and track everyday training, people use a wide variety of smart mobile systems on their smartphones, smartbands and smartwatches. However, many of training mobile systems have their functional limitations and in many cases do not allow the user to perform in-depth data analysis in order to optimise the training. Therefore, in this paper, we propose adaptive decision aid tool that supports running sports practitioners in their daily training activities. This tool aims at suggesting the user the strategies for next training session (e.g. the maximum number of training days that can be skipped, duration of the training, the distance that should be covered) in order to meet the goals (in the experiments we have considered winning the one-year running competition). From the scientific standpoint, the tool adapts reinforcement learning in order to propose the runner suggestions that will allow for improving performance.

Keywords: User support · Reinforcement learning · Markov decision process · Running sports · Decision aid

1 Introduction

The appropriate training and exercises intensity are the key to achieve goal and to improve the performance. However, lack of basic knowledge about training methodology can lead to serious injuries, discouragement and lack of gain. Therefore, lots of people are using training mobile systems to track their progress, fulfil training goals, and monitor achievements. Nonetheless, many mobile training systems have limited functionality (especially in contrast to the amount of collected data) and do not allow for in-depth data analysis that will support the sport's practitioners in decision making. In many cases, these tools have hard-coded training scenarios, measured data is sometimes incomplete, and many assumptions or suggestions are too generic. Moreover, recently gaining in polarity so called competition-based social running in an opinion of many experts is

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*,
Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_19

not a good approach to training, because it makes people train faster and harder than they should.

Understanding the collected data is an important aspect when it comes to optimisation of training activities in order to achieve better results in shorter time. In many cases, visualising the rough data (e.g. amount of burnt calories or covered distance) is not enough for inexperienced runners to assess the progress.

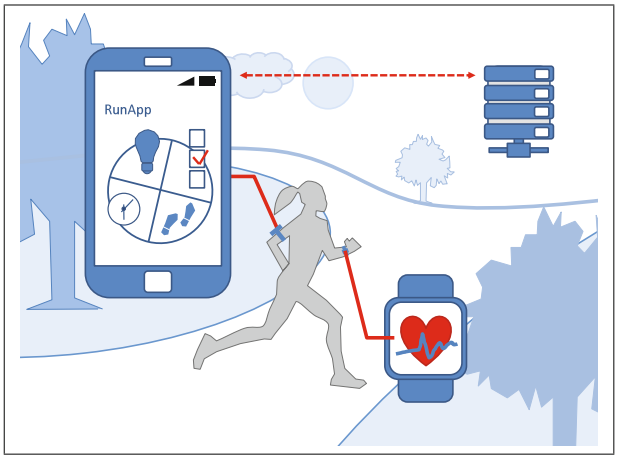


Fig. 1. High level overview of the proposed RunApp mobile training system. The mobile application communicates with server (upper left corner) in order to synchronise data (e.g. user questionnaires, feedback and recorded data), delegate core computations and retrieve suggestions regarding training

Therefore, in this paper, we propose the adaptive decision aid tool that supports running sports practitioners in their daily training activities. This tool is based on reinforcement learning that looks for optimal training policy which will allow its user to achieve satisfactory results in constrained environment (e.g. concerning weather condition, user mood, preferences, etc.).

2 Related Work

Nowadays, the wearable computing and smart devices have become popular tools commonly used by runners to monitor everyday training progress. In principle, the smartphones equipped with flexible operating systems and a wide variety of sensors have eliminated the obstacles to create mobile training systems. Therefore, there are plenty of different solutions varying in concept and complexity. For instance, there are systems such as Endomondo [1] supporting different types of sport activities (e.g. jogging, biking, hiking, etc.) enabling users to record key training characteristics (e.g. speed, burnt calories, time, track, etc.), share training details via social media, have insight into training details of other people, etc.

However, this topic is also the area of interest for many researchers. The variety of statistical and machine learning techniques have been used in order to support people practising sports and to improve their achievements and performance. In [6] authors have proposed a decision-making system for a multi-step training scenario. The system is based on dynamic programming optimisation and Markov Decision Process formulation. Another system for runners has been proposed in [4]. The goal of the system is to facilitate social running in a city park by planning the route, supporting local running communities, and promoting a healthy lifestyle. In the literature, there are also examples of adapting machine learning and data mining techniques also for other sports than jogging. For instance, in [5] authors have proposed MONEYBaRL system which uses reinforcement learning to exploit baseball pitchers decision making. In principle, the proposed algorithm allows for finding optimal pitcher-specific and general (against a collection of pitchers) batting strategies. The more holistic approach has been presented in [2], where authors proposed recommender system for running professionals and amateurs providing a wide range of personalised information concerning both workouts and diets. This approach is based on a combination of expert knowledge (maintained as an ontology) with a recommender system.

All the above-mentioned works show that different sports activities can be supported with modern wearable devices and advanced data mining techniques. However, when it comes to running sports, it seems that the problem of optimising different runners behaviours has not been well addressed yet. Therefore, in our approach, we focus on machine learning techniques that would allow the runner to (i) evaluate the training progress in different time spans as well as to (ii) propose an optimal strategy for achieving the training goal.

3 Proposed System

As it is shown in Fig. 1, the proposed system consists of several elements that enable the user to track the training progress. The runner is equipped with the mobile device with pre-installed software that facilitates bi-directional communication with the server, reads current position of the user with GPS sensor, measures an amount of steps by means of an accelerometer and provides the user with guidelines. In particular, the software allows the runner to select the appropriate type of training and have insights into current progress and recommendation. All relevant measurements, as well as user feedback (after training session we ask the user to fill in a short questionnaire about the training aspects), are uploaded to the server.

The server-side acts as a cloud storage but it also is delegated to run more complex analysis on the collected data. One element of the analysis is the evaluation of the training session. Here we apply machine learning and data mining algorithms to learn the classifier which will tell the user that he or she is progressing or not. Usually, the evaluation is not straightforward (e.g. few bad workouts does not imply that whole training program is bad) as the runner's achievements are influenced by external factors (e.g. weather, mood, diet). Therefore, we ask the user to fill in a questionnaire after the training.

The second element of the analysis concerns the training strategies. There are lots of experts opinions how well-organised jogging training plan should look like. However, the amount of information is usually overwhelming for the inexperienced runners. Therefore, here we adopt one of the reinforcement learning technique to analyse the data collected from different users in order to identify optimal approaches for achieving different training goals.

4 Training Evaluation and Decision Support Algorithm

The aspiration to achieve better running results should proceed in a reasonable way, through gradual extension of distances covered and workout duration, taking organism regeneration into consideration. In result, the runner's condition could be improved and the risk of injury reduced. Due to the fact that users do not have to make progress in every training, the collected data were considered on a weekly basis.

A runners progress was estimated based on the increase in distance and time as well as on the number of runs which were registered in the following weeks. The implementation of machine learning algorithms in the application allows detecting lack of running progress and searching reasons of that on an ongoing basis. It improves the way of preparing personalised tips which help users to develop a physical condition.

4.1 Optimisation of Jogging Strategy

In order to capture the everyday behaviour of runners, we model different training strategies as Markov Decision Process (MDP). More precisely, let $\{X^{(t)}\}$ be a Markov process with finite state space $S = \{E_1, ..., E_n\}$ of n elements (states), where each state represents set of measurements. The number of measurements depends on the scenario. In principle, these contain such information as covered distance, elapsed time, burnt calories, etc. In our case, we assume that transitional probabilities between states are necessarily stationary. In other words, $p_{E_i \to E_j} = P(X^{(t)} = E_j | X^{(t-1)} = E_i)$ is different for different t. Considering an example where runner starts a one-year competition and in time t_0 covers

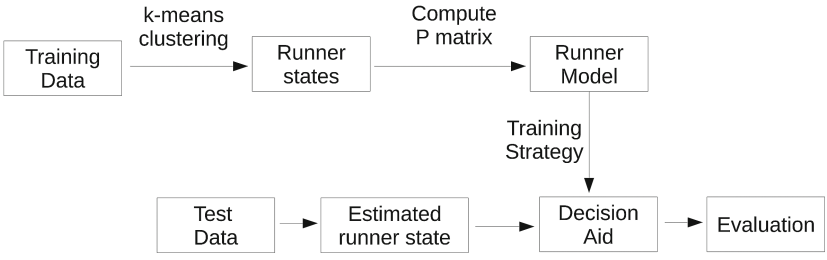


Fig. 2. Proposed algorithm overview

distance d_0. Under that circumstances, we may not expect that the runner will cover the same distance during each training session, as we must consider that at some point the performance will improve (as the runner will become more advanced) or deteriorate (e.g. due to tiredness) (Fig. 2).

An example of graph presenting possible states and transitions has been presented in Fig. 3. In fact, it is modified Markov transition diagram, since (for readability reasons) the transitions probabilities have been represented with varying thickness of graph edges (the thicker the edge is the more probable the transition).

The classical MDP is defined as 5-tuple (S, A, P, R, γ), where S indicates mentioned above finite state space, A indicates finite action space, P transition probabilities, R a pay-off function (an intermediate reward), and γ the discount factor (importance between future rewards and present rewards).

We define optimal policy π as a chain of actions $\{a_0, ..., a_n\}$ (that particular runner can take), which will maximise the expected value function V at every state of a Markov process. In other words $\pi(s)$ will tell us what is the optimal action in a state s. The value function is described by equation:

$$V^{\pi}(s) = R(s, \pi(s)) + \gamma \sum_{s' \in S} P(s'|s, \pi(s)) V^{\pi}(s') \tag{1}$$

In that sense, the optimal value function V^* of a state is the reward for that state, plus the discounted expected reward for following the optimal policy from this state:

$$V^*(s) = \max_{a \in A} \{R(s, a) + \gamma \sum_{s' \in S} P(s'|s, a) V^*(s')\} \tag{2}$$

In the same way, we can calculate the optimal policy in state $\pi^*(s)$, using the following formula

$$\pi^*(s) = \operatorname*{argmax}_{a \in A} \{R(s, a) + \gamma \sum_{s' \in S} P(s'|s, a) V^*(s')\} \tag{3}$$

In our case, in order to estimate the P we have used publicly available historical data obtained from Endomondo web portal. The details about the collected measurements have been presented in Sect. 5.

5 Experiments and Results

In this section, we have described two experiments. One is related to an evaluation of training performance by means of machine learning algorithms. For that purpose, we have used the collected measurements of different characteristics measured during the training and accompanied with user feedback in form of filled-in questionnaires. The second experiments concern evaluation of the algorithm for training strategy optimisation.

Currently, our system is under development and the number of active users is still limited to run reasonable data analysis in order to find optimal training

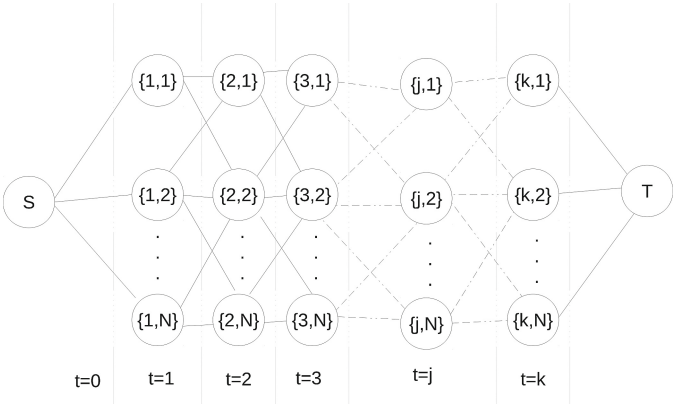


Fig. 3. Example of users states transitions over time windows (for readability reasons some edges have been removed)

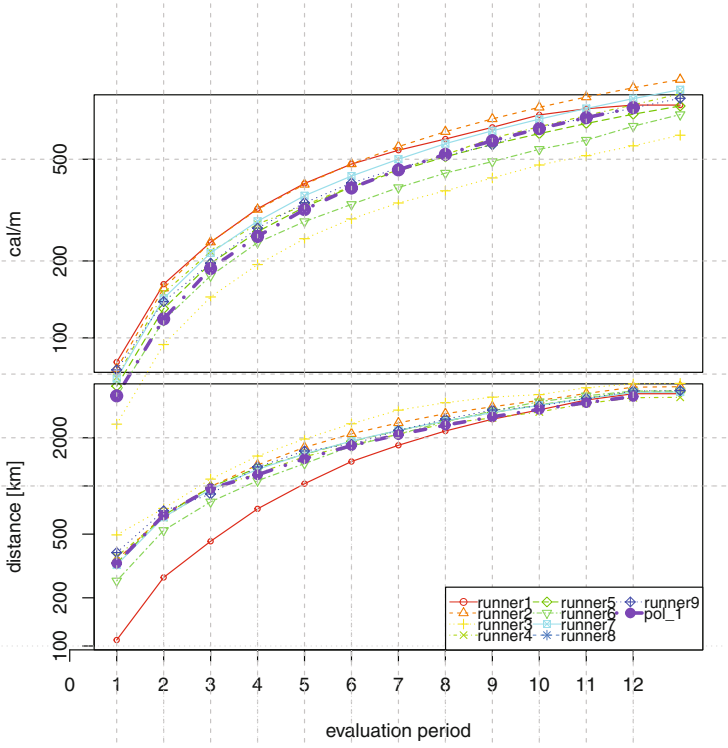


Fig. 4. Cumulative amount of calories burnt per meter (top) and cumulative distance measured for different runners (bottom) during different evaluation periods

strategy. Therefore, to prove the correctness of our assumptions, we used publicly available data of 10 competitors (participating in common jogging challenge) that have been recording their training progress during the period of one year.

Table 1. Final characteristics measured for different runners

Runner	Total distance [km]	Total Calories	Time [h]	Avgerage cal/m
pol_2	5937,9	430890,5	654,9	68,1
pol_3	4436,5	261934,9	461,4	58,7
runner3	4372,1	205302,0	422,7	47,8
runner2	4169,3	331855,0	358,8	78,9
runner8	3946,4	257455,3	403,0	66,6
runner6	3927,0	218219,3	293,9	57,5
runner5	3880,5	243200,0	363,7	62,1
runner7	3861,6	278856,6	303,2	72,0
runner1	3769,8	246340,7	402,1	62,7
pol_1	3739,9	249175,1	334,9	66,0
runner4	3580,8	243601,7	359,5	69,4
runner9	3546,3	127971,0	439,7	35,5

5.1 Training Performance Evaluation

In the experiment, the effectiveness of running progress estimation was tested through machine learning algorithms which are implemented in well-known WEKA software [3]. A ten-fold cross validation was used to assess the quality of the generated prediction models. The results obtained have shown that Bayes Net and SMO are the best classifiers for the research of running progress on the basis of a training set. For the first of them, the accuracy of 94,1% was obtained and for the second – 91,2%. The Bayes Net algorithm forecasts sporting results improvement better than SMO (accuracy is 96% for Bayes Net and 88% for SMO). However, SMO is the only one (among tested classifiers) which detects a lack of progress faultlessly. In that case, Bayes Net ensures 88,9% of accuracy. In the Fig. 5 the ROC (receiver operating characteristic) curves are shown for chosen algorithms. They inform how correctly classifiers separate positive and negative class observations. Following the scale of the evaluation of classification accuracy based on the fields under a ROC curve, we can claim that Bayes Net and SMO divide instances with high effectiveness. From among the tested algorithms, Random Forest is characterised by high accuracy (91,2%) of progress prediction. However, it does not manage forecasting lack of progress as well (77,8%). In the application algorithms, KStar and Filtered Classifier should not be implemented. It is justified by low, indicating high forecasting

randomness, the accuracy of classification of observations which do not prove sports development. Also the low standard of Kappa statistic (determining the extent to which a suggested model matches reality) for Filtered Classifier (0,29) and a small field under the ROC curve for KStar indicate a large probability of randomness.

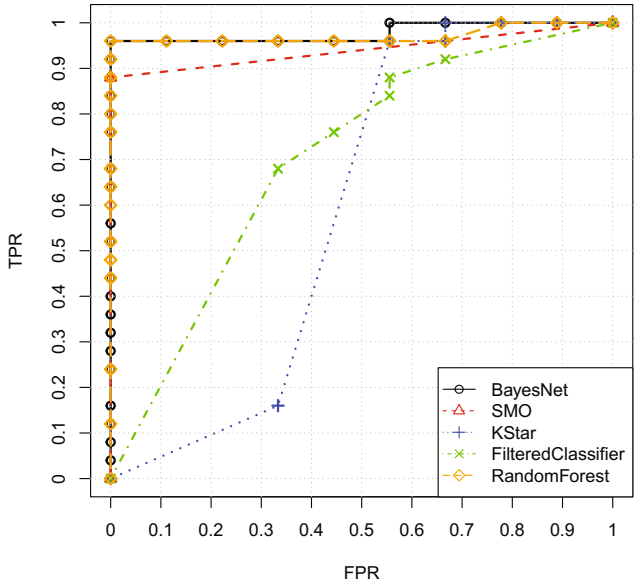


Fig. 5. ROC curve comparing different classifiers performance

5.2 Training Strategy Optimisation

For this evaluation scenario, we have used publicly available data obtained from Endomondo system. We have identified 9 similarly advanced runners participating in the same running contest/challenge, which was one year long. We have divided the data into two parts each 6 months long. First, the dataset was used to train our system while the second part was used for evaluation purposes.

The training dataset was aggregated in 12 evaluation evaluations periods (each 2 weeks the progress was measured). In each evaluation period, we have calculated the number of skipped trainings, distance covered, average distance in the evaluation period, average speed, a number of burnt calories, and amount of calories burnt per one meter.

The statistics calculated for each evaluation period (and each user) have been clustered with k-means algorithm, where $k = 6$. This allowed us to roughly estimate k different user training states.

Using the methodology described in Sect. 4.1, we have estimated the training policy and evaluated with the evaluation dataset (remaining 6 months of the competition).

In Fig. 4 there are presented cumulative achievements of 9 different users practising jogging during the period of 6 months. Also in the Table 1 final achievements of selected users are shown.

In Fig. 4 and in the Table 1, there are also characteristics showing achievements for the hypothetical runners (pol_1, pol_2, and pol_3). They follow different (suggested by the system) training policies.

The runner pol_1 follows the policy, which proposes the most probable training path – the one with highest transition probability for the estimated transition-states diagram. As expected, the results for this policy is average (in the sense of the covered distance). On the training data, the pol_1 runner took the fifth place and the seventh place on the evaluation dataset.

The runner pol_2 follows the policy, which proposes the most effective training path (the one which will allow the runner to cover the longest distance). As it is shown in the Table 1 the policy works also on the evaluation dataset. However, the amount of estimated burnt calories is quite high.

On the other hand, the runner pol_3 follows the policy, which balances the effectiveness with a number of burnt calories. It can be noticed that this policy will also work for the evaluation dataset. Nonetheless, a number of burnt calories is significantly lower.

6 Conclusions

In this paper, we have proposed innovative system supporting people practising running sports. There two main contributions of our research, namely: (i) supervised learning approach for training performance evaluation and (ii) reinforcement learning technique for finding optimal running strategies. In the proposed approach we have used two datasets to prove correctness and evaluate the performance of the system. The obtained results are promising and show that the proposed approach can be effectively used to evaluate runners training performance and to support them with suggestions for optimising the running strategies.

References

1. Endomondo training system (2017). https://www.endomondo.com/
2. Donciu, M., Ionita, M., Dascalu, M., Trausan-Matu, S.: The runner-recommender system of workout and nutrition for runners. In: 2011 13th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), pp. 230–238. IEEE (2011)

3. López-Matencio, P., Alonso, J.V., González-Castano, F., Sieiro, J., Alcaraz, J.: Ambient intelligence assistant for running sports based on k-nn classifiers. In: 2010 3rd Conference on Human System Interactions (HSI), pp. 605–611. IEEE (2010)
4. Schmidt, B., Wozniak, P., Knaving, K., Chen, C., Obaid, M.: ParkPal: towards ad-hoc route planning for runners. In: CoPDA@AVI, pp. 23–27 (2014)
5. Sidhu, G., Caffo, B., et al.: MONEYBaRL: exploiting pitcher decision-making using reinforcement learning. Ann. Appl. Stat. **8**(2), 926–955 (2014)
6. Vales-Alonso, J., López-Matencio, P., Alcaraz, J., Sieiro-Lomba, J., Costa-Montenegro, E., González-Castaño, F.: A dynamic programming approach for ambient intelligence platforms in running sports based on Markov decision processes. In: Human-Computer Systems Interaction: Backgrounds and Applications, vol. 2, pp. 165–181. Springer (2012)

Framework for Evaluating QoE for Remote Control of Autonomous Cars in Mobile Wireless Networks

Adam Flizikowski[1(✉)] and Mateusz Płócienniczak[2]

[1] University of Technology and Life Sciences, 85-791 Bydgoszcz, Poland
adamfli@utp.edu.pl
[2] ITTI Sp. z o.o., Rubież 464, 61-612 Poznań, Poland
mplocienniczak@itti.com.pl

Abstract. This paper introduces framework for designing and evaluating the video adaptation controllers - i.e. the enablers for the growing market of user generated traffic. The main targets are the use-cases connected with emerging market of the autonomous cars or drones that require remote assistance of the operator upon on-board controller failure. To validate the proposed architecture authors have designed test cases that validate usability of the proposed solution for the uplink surveillance and monitoring traffic adaptation. We propose a systematic approach to combine real network traces, network modelling and emulation, video transcoder as a service and the use of no-reference QoE metrics in order to assure effective means for video controllers design and tuning. Authors have described architecture as well as results of comprehensive sensitivity tests of QoE in uplink for mobile WiMAX networks. The results show that the proposed framework provides valuable potential for development and evaluation answering the demand of emerging scenarios for mobile surveillance.

Keywords: 4G · Video · Uplink · Autonomous car · QoE

1 Introduction

Video streams are dominating today's Internet traffic share especially in the down-link direction. They are one of the main contributors to the growing data volumes exchanged over today's networks (mobile video traffic accounted for 60% of total mobile data traffic in 2016 and is estimated exceed 69% of mobile data by 2018 [4]). In the advent of 5G the anticipated growth of video traffic will largely be motivated by the technology push from the vertical markets, that are tempted to generate great amounts of data (e.g. 4–40 Gb/hour/car) in order to deliver monitoring of vital operational parameters as well as value added services and eventually enhanced actuation capabilities [1]. Especially as the *the "real-time information monitoring helps companies deploy new video-based security systems, [...] bandwidth-intensive M2M connections are becoming more prevalent. Globally, M2M connections will grow from 780 million in 2016 to 3.3 billion*

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*,
Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_20

by 2021, a 34-percent CAGR - a fourfold growth". The standardization of 5G is gaining its momentum as autonomous car industry is. Although the 5G promises to bring "zero delay" and "gigabit throughput" to the wireless channel we are still well ahead of the deployed and operational 5G networks. Thus in order to adjust existing radio systems (e.g. 4G) and adapt it towards already growing needs of various verticals we need to focus on existing capabilities to improve 4G networks and/or the multimedia systems to especially support the user generated content. This paper addresses the challenges connected with delivering enablers for the growing market and demands of the video traffic in uplink direction. The most interesting version of so called "user generated content" deals with time-sensitive domains of remote control of autonomous cars, UAVs, UGS', etc. There is growing need for remote mobile surveillance solutions (area reconnaissance, crisis management operations, area crowd mapping, threat detection and mitigation for trucks on parking lots) and the emerging and disruptive market of autonomous cars strives to capitalize on that feature as well [5] as soon as around 2020. Remote operation of cars and especially trucks is the big promise for the logistic sector [10]. Our goal in this paper is to (i) propose a systematic approach to video controller development and evaluation for the 4G networks, but with potential to be also utilized for future use-cases, (ii) perform baseline validation of a QoE of relevant video feeds, by replaying variability of mobile channel (at low speeds) based on traces collected from real Wi-MAX networks across Poland and (iii) focus on selected set of QoE metrics (freezing, blockiness, blockloss) and evaluate quality for multiple settings.

2 Related Work

In order to deal with remote monitoring of the moving objects (e.g. car, drone) environment and provide capabilities for its control by a remote operator one need to consider few crucial substrates: (i) particular scenario requirements in order to assure smooth operation of the car (e.g. being able to follow the view of the street and the guiding lines, detection of obstacles), (ii) network coverage variability (e.g. holes in the coverage) and (iii) channel dynamics while a car is moving (e.g. slow/fast-fading mitigation). On the other hand it is important to understand that the typical surveillance systems are usually identified as "target recognition videos". It means that the crucial performance indication is whether an operator can properly detect (recognize) an object or a situation which a camera captures. In the market there are solutions that can be used for real-time monitoring of mobile assets (e.g. police cars, busses). For example the multipath video streaming solution are offered by companies like [7] or [11] dealing with security monitoring. Still those solutions do not seem to assume any means of context-awareness in their architecture. Moreover there is no information that they would be using QoE metrics which are tailored to the surveillance applications. Most of the existing video QoE metrics are derived from VoD like technologies [2] and thus mainly related to entertainment multimedia. Recently the ITU-T is keeping track of this need in amendments to its P.912 specification for "video recognition tasks". The methods used so far in improving QoE

of video streaming focus mostly on the pixel-level QoE evaluation and can be divided into: Full-Reference, Reduced-Reference, No-Reference categories. These categories differ in the availability of source videos during metric calculation. Major focus (for multimedia scenarios) is put into enhancing and improving image quality frame by frame by evaluating specific, video-oriented metrics such as blurriness or jerkiness. However such approach does not take into consideration contextual requirements of the operators nor variations caused by variable network conditions [15]. From the current paper's point of view, most important QoE metrics are: Blockiness, Blockloss and Freezing. They are important as they directly relate to the critical aspects of video fluency when experiencing wireless network misbehaviour. In order to build and tune video controllers it is necessary to deliver relevant emulation of the target network scenarios, so that the controllers can be tuned based on realistic settings of the target environment. Several standalone network emulation tools available on the market e.g. [3,9]. For the purpose of this research we have performed dedicated examination of IXIA ANUI [6] which simulates radio conditions and impairments including signal delays, jitter or packet drops in WAN networks. Unfortunately test we have performed proved it does not support dynamically changing bit rate values (on the order of hundreds of changes per second), which is none of the main disadvantage of the solution given this papers aims. Considering this limitation authors have decided to use proprietary solution based on the NETEM framework [8]. From the perspective of the video adaptation for wireless systems authors in [24] address approaches to improve the delivery of data such as video over disadvantaged networks. This work focuses on utilizing reliable multicast protocol's hooks i.e. in the NORM protocol to provide a network information service with access to path bandwidth, delay, and lost packets. Resulting network characteristics are derived to drive the video transcoder at the server which chooses from set of profiles that include settings for: video resolution, frame-rate, and encoding bit-rate to allow the server attempt to fit the video stream into the available bandwidth. The work of [23] relies on context-aware services to adapt system behaviors based on the retrieved context data (context is represented in a way of ontology). Besides choosing appropriate video content based on user profile the dynamic media adaption is performed to improve the video quality perceived by the end user in response to changes in: varying wireless channel quality, available energy of the end equipment, network congestion and application Quality of Services (QoS). The results show that utilizing context may help improving video quality (PSNR) by 2–3 dB. Whereas [25] focuses on the optimal control of traffic variability in the downlink direction of LTE radio access by the use of smoothing buffer. Authors provide algorithms and analytical solution to find optimal state switching strategy. Control of the smoothing buffer should be linked to the channel state variability. The proposed algorithm allows for reduction of the variation rate for the video stream (ca. 20%–40%). In [16] authors introduce the methodology that aims to provide a frame-work that addresses, in a combined way, the energy consumption and the end-user perception of the video traffic. This methodology was validated on two distinct technologies i.e. WLAN and

WIMAX. Authors use the SSIM metric to assess image quality according to the full reference approach. Videos of four data rates (2M bps, 4M bps, 6 Mbps and 8 Mbps) used in tests were compressed with the MPEG-4 codec using the *ffmpeg* application. On the opposite [20] suggests that in order to adapt videos sent from UAVs *"instead of reacting to packet loss [he] uses an increase in queueing delay at the router [or CPE] to detect phases of throughput degradation"*. Kacianka in [19] takes this research further and proves that there is correlation between high delays (higher than 1 s) in receiving ACK for a RTP packet sent from an UAV using WiFi and bad video quality. Still none of the authors focus on the QoE adaptation concerning the needs of the remote monitoring and operation of car or drone. Complementary approaches based on utilization of a cross-layering design can be seen in number of proposals found in recent literature: admission control schemes studied together with adaptive modulation and coding on the performance of elastic traffic [17], fair and efficient QoS guaranteed resource allocation for a mixture of real-time and non-real-time service flows is achieved with a joint packet scheduling and slot allocation scheme interaction in both MAC and PHY layers [14,22], cross-layering optimization mechanism for multimedia traffic has been proposed in [26,27] while a cross-layering scheduler that employs AMC scheme at the PHY layer, according to the SNR on wireless fading channels, has been described in [21]. Majority of works employs AMC scheme at the PHY layer over wireless fading channels. Therefore to adjust to the user expectations, the codecs need to be adjusted according to the information from the PHY/MAC layers. For this reason authors propose solution that will enable delivery of architecture suitable for working in laboratory environment but successfully validated for number of representative cases collected during traces.

3 Congestion Control Framework Design

The overall, high-level description of the proposed WiMAX network emulation framework is presented in the figure (Fig. 1). The system allows the user to execute customizable emulation scenarios related to video streaming in WiMAX network. Such scenarios comprise many variables including network properties and conditions as well as various users' activities. The proposed software includes rtPS scheduler implementation compliant with requirements of the WiMAX class of service. To run it, it is required to specify and customize the overall network properties, which include the following: overall bandwidth available within a network in the uplink direction based on Downlink/Uplink ratio and number of symbols, TDD frame duration, over-head behavior for rtPS terminals, scheduler's efficiency (variable deciding how efficiently scheduler allocates bits in symbols) and network delay distribution scheme (for efficient delay emulation). The above emulator architecture enables definition of real-time (rtPS) flows, together with its key parameters (MBR, MBA, prio) whether CBR or VBR/adaptive traffic. Based on the configured rtPS parameters the built-in scheduler is able to perform scheduling compliant with the WiMAX rtPS scheduler. The PC1 plays a role of video source with transcoder in order to mimic the camera mounted inside a car

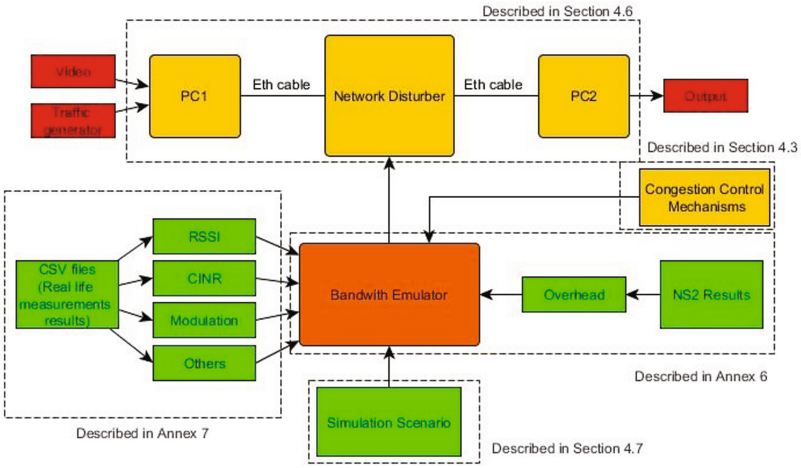


Fig. 1. Network emulation framework for WiMAX

or on a drone (whether front, rear or sideways). On the side of PC1 it is possible to utilize (i) real camera attached via network interface, (ii) prere-corded video from files or (iii) MGEN traffic generator. The PC2 mimics a traffic receiver *"inside the control room"* - where e.g. Uber employees are triggered to provide remote guidance to a member of the fleet of autonomous cars upon request from the particular car's controller [13]. Technically during any tests and evaluations with the proposed architecture the PC2 requests video stream from the "car node" (PC1) via RTSP request to initiate the transmission. For this purpose, a video client like e.g. ffplay with dedicated shell scripts that initiate video delivery from PC1 are used at the side of PC2.

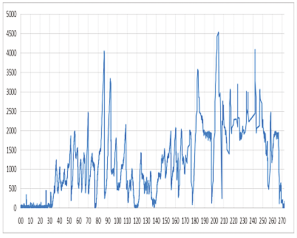
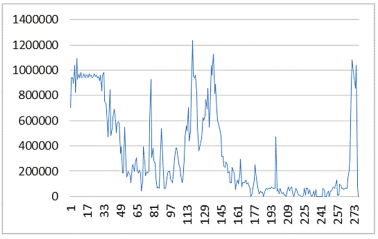
The *"Network Disturber"* block is realized by a Linux router which plays a role of:

- Base station - with scheduling of traffic in the uplink, modulation adjustment (AMC), as well as introducing signaling overhead (rtPS signalling, BE signalling).
- Uplink radio channel - by introducing modulation driven rate adaptations caused by mobility (slow-fading, multi-path), NLOS, packet drops and delays.

This solution uses real values of modulation, rate and delay collected from the trace file. The *Bandwidth Emulator* is an external application that based on traces from drive tests performed by the authors in multiple (selected as representative) locations produces two configuration scripts that have to be deployed on the *Network disturber*. The first output script (TC_script) contains a set of commands for the traffic control framework installed on the Linux router. It creates hierarchy of traffic control nodes such as qdisc, class and filters, and attaches such structure to the outgoing network interface of the *Network Disturber*. The NETEM qdisc is used for defining delay and loss for the user's flow. Additionally,

for isolating any other traffic from the simulated transmission the script defines qdisc and filter for all remaining traffic flowing through the interface and passes further it without any disturbing. It allows defining all crucial parameters such as OFDM symbols count, frame duration and over-head. For target scenarios different users relate to different cameras. To prepare a test one needs to specify the overall number of cameras in a scenario, provide their operation status updates within given timespans, and their respective class parameters including: priority, minimum bandwidth reserved and maximum bandwidth available. For each camera, the user also defines the radio quality behavior by assigning CSV trace from a particular drive/field test in real WiMAX network. Each transmitter that is sending video footage to the receiver (i.e. each data flow) is experiencing network conditions according to the CSV file assigned for this specific transmitter in the process of pre-paring tests. In order to provide representative samples of channel variability authors have performed drive tests (or walking tests) in different networks. The following network behavior sample (Table 1) was collected while driving the car in Choszczno in Poland along four streets in circle (at a speed of 20 km/h).

Table 1. Sample network traces from drive tests (driving a car)

Case	Delay [ms]	Throughput [bit/s]
Driving a car		

The traffic rate sent from the transmitter was 1 Mb/s and it was sent to the node representing security operator premises (PC2). The case above represents situation where there were both degradation causes present in parallel: intense NLOS and mobility. That is why the throughput level is largely degraded throughout the plot. Results of above mentioned tests are reported in more details in the paper [18]. Our remaining work aimed to define sample video controller with the following features: traffic shaping - introduced by properly adjusting transcoder rate based on information from WiMAX modem at the transmitter (e.g. car), overhead injection - the implemented scheduler subtracts fixed number of symbols every frame for a given user to mimic various congestion control algorithms. The implemented scheduler delivers treatment compliant with the rtPS WiMAX class with its crucial parameters (maximum sustained

traffic rate and minimum reserved traffic rate) - as it needs to be considered when dealing with congestion control mechanisms. The next chapter presents results of the framework validation with video from external camera as well as from files.

4 Measurements

The emulation testbed architecture has been developed and in turn we have designed and then performed a set of tests in order to validate the framework and then focus on a systematic evaluation of the QoE of various video feeds. Drive tests were performed in various locations however it is worth highlighting the observed gap between fidelity of modelling coverage of the measured base station utilizing dedicated software [12] and utilizing the in-house developed signal analyser (RaspberryPi based) with GPS coordinates readings (see Fig. 2). It is worth noticing that the circular areas of equal signal coverage (the right column) mismatch the instantaneous modulation readings from the same area.

Fig. 2. Drive test (Choszczno mobile) - modulations

The coverage analysis based on the Splat! model gives only a rough estimate of the coverage as compared to the detailed modulation traces.

4.1 Validating *Network Disturber*

At the stage of framework development and configuration authors have been connecting the proposed *Network Disturber* and the transmitting and receiving endpoints using wireless (WiFi) but eventually switched to wired network (Ethernet) due to identified unstable behavior of the wireless card driver in connection with Netem, manifested in anomalous delays. After removing this problem we have been especially interested in validating the fidelity of effectively replaying the radio conditions based on baseline traces. The tests have confirmed that both instantaneous bandwidth and modulations are properly recovered. However already at this stage we have found that the delays resulting from sequentially modifying channel (i.e. available rate) can exhibit large variations. On one hand

the delays simply indicate that the channel capabilities of a user (car, UAV) have decreased on the other this is the clear reason for adaptations to be applied. In order to minimize those extra delays from the Netem queues it is essential that the source traffic is *properly adapted* to accommodate to the instant values of the artificial channel. The excessive delays appear when arrival rate on the ingress of the emulator queues exceeds the emulated service rate emulated on the egress of the *Network Disturber*. In order to mitigate such mismatch we have created an option in the *Bandwidth emulator* that configures traffic generation script timing in exact synchronization with the channel changes of the emulator. The resulting data flow (UDP packets) mimics the video source with "ideal feed-back". However if this latter option is used just for controlling QoS of the data stream under particular channel variability emulation (i.e. when it is enabled no video is sent so one cannot evaluate video QoE). The next figure (Fig. 3) shows the two plots which represent the effective instantaneous rate of "adaptive flow" at the transmitter (orange plot) and the rate after the packets have been received at the receiver (PC2). It is clearly seen that there is a mismatch between the timing of various "spikes" at both plots. This is caused by "desynchronization" between the flow that starts to grow whenever emulated bandwidth experiences "rate drops". Such drops cause packets in the delay queues of Netem at the *Network Disturber* to face the head of line blocking due to slower channel.

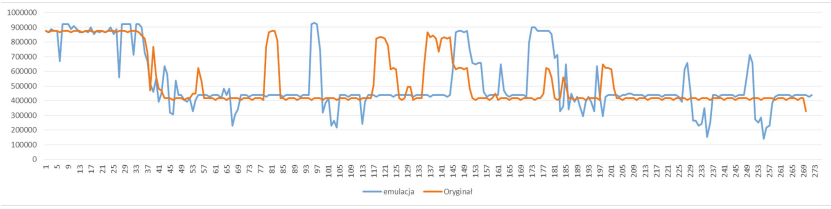


Fig. 3. Emulated rate at the receiver (channel - Choszczno mobile)

Following plots demonstrate the delay of the "adaptive traffic" represented by the orange plot on the Fig. 3. The first of them represents exactly the case presented above. The Fig. 5 has been plotted after shifting the "orange peaks" to the right, to cover respective "blue peaks".

We can see that delay in the Fig. 5 seems to follow the shape of delay from Fig. 4 but the "spikes" seen in the latter in the intervals of "80–98 s" and "120–170 s" have been reasonably decreased. It shows the influence of properly adapting the sending rate at the sender (e.g. car, UAV). Only if rate adaptation is tuned to the emulated channel the Netem-based emulation can deliver more realistic delays. However it can be seen that even though the macro adaptation is applied some smaller variations (mismatches) still cause delay spikes in the Fig. 5.

This is caused by the temporary behaviour of emulator where "packets delayed by X seconds" are being processed in parallel with "packets delayed

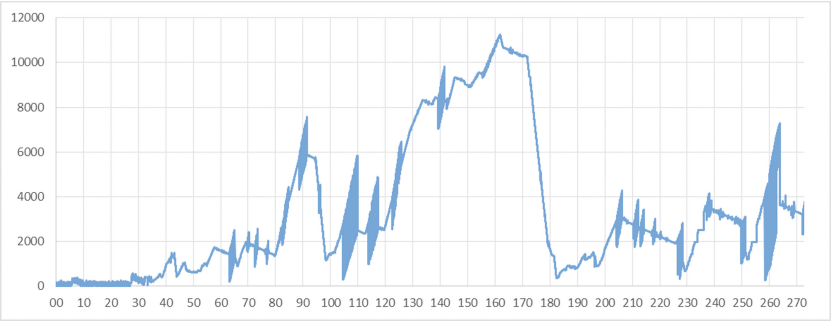


Fig. 4. Delay at the receiver - desynced case (trace: Choszczno mobile)

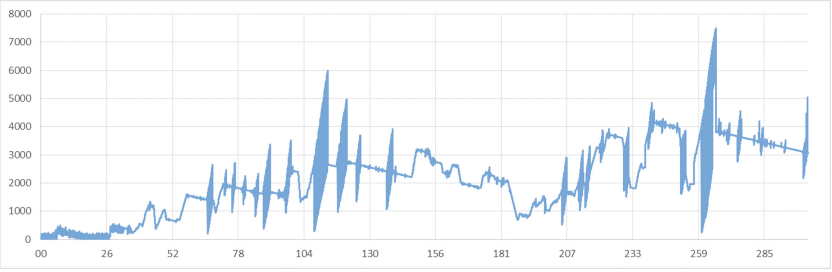


Fig. 5. Delay at the receiver - after synchronization was made between source and emulated channel (trace: Choszczno mobile)

by Y seconds", as the delay manipulations are introduced sequentially by the TC scripts in the Network Emulator (see Fig. 1).

4.2 QoE Measurements

After validating the emulation framework with artificial traffic we decided to evaluate video from real camera connected to the PC1 (so there was no transcoder at the transmitting node). The receiver used the VLC for playing the video. The video was 1 Mb/s H.264 encoded scene captured from behind the window (small patio with one tree, no people, building seen from the window). First we have tested the influence of increased delay on the quality metrics (values in parentheses should be observed in a video without any distortion): Blockiness (0.9–1.01), Blockloss (0–5) and Freezing (0). At the first stage delay ramps-up through values 50–100–200–500 ms and back to 50 ms after the maximum of 500 ms was reached. The full cycle takes 6 min with a single delay step of 60 s. Such cycle is repeated four times - but each time the "delay steps" are shortened by half. The results show no visible influence of introducing such delay for the QoE metrics. Further we have tested the influence of emulating delays across the stable periods (i.e. when channel is stable) as normal distribution (configuration

parameter in Netem) but this time the camera traffic was 2 Mb/s. Our rationale was that owing to such approach we should decrease the number of delay adjustments within the test. We performed these tests with UDP and TCP as transport protocols for comparison, each repeated with no emulation (denoted by "C") and with delay emulated utilizing the normal distribution (denoted by "D") for particular average value and variation according to the settings: no emulation > 50 ms/30 > 50 ms/50 > 50 ms/80. Each setting took one minute. The results showed that emulation utilizing normal distribution of the delay introduced freezing for both transport protocols, even if the standard deviation was negligible (1–5–10 ms). But tests with delay distribution switched off shown no difference between QoE for UDP or TCP transport - see Table 2.

Table 2. QoE metrics for varying delays and transport protocols (video from camera 2 Mb/s)

	Blockiness	Blockloss	Freeze
TCP-C	0,88	2,48	0,00
TCP-D	0,88	2,64	0,04
UDP-C	0,88	2,47	0,00
UDP-D	0,88	2,55	0,04

Table 3. Influence of introducing delay in different modes (normal distribution, direct value)

Delay option	Blockiness	Blockloss	Freeze
Normal distr	0,83	1,41	0,10
Direct value	0,83	1,22	0,00

We have zoomed-in on the problem of the "delay from distribution" and it was evident that activating this feature for channel trace which has reasonably long periods of stable delay introduced problems in video playout. We have investigated that the reason was the introduction of the RTP packet reordering. So the quality degradation was even worse for theoretically better traces (Table 3). We have also checked influence of setting various camera rates: 1 Mb/s and 2 Mb/s. The results show that blockloss and blockiness are improved as we increase video rate, which is intuitively logical (Table 4). Eventually we have tested the influence of multiple architectural settings at the source node: Test 1 (Src>QoE), Test 2 (Src>VLC>QoE), Test 3 (Src>FFmpeg>Transcoder>VLC>QoE). All the tests used the video captured by camera pointed outside the building on a quiet street. The baseline test (Test 1) is the video simply evaluated without playing on a player, the next test (Test 2) showed the influence of player and the last test (Test 3) shows influence of ffmpeg conveying video to the transcoder,

Table 4. Evaluating influence of the video rate

Rate	Blockiness	Blockloss	Freeze
1M	0,74	2,10	0,00
2M	0,82	1,31	0,00

Table 5. Validation of the reason for freezing (1 Mb/s video)

QoE metric	Test 1	Test 2	Test 3
	1 MB		
Blockiness	0,88	0,86	0,80*
Blockloss	4,89	4,8	6,87*
Freezing	0	0	0,04*
Spatial activity	87	87	81,48
Letterbox	0	0	0
Pillarbox	0	0	0
Blur	2,66	2,66	2,74
Temporal activity	2,19	2,19	2,32
Blackout	0	0	0
Exposure	113,29	113,29	113,42
Contrast	71,61	71,61	72,47
Interlace	0,039	0,039	0,034
Noise	0,27	0,27	0,29
Slice	4,44	4,44	5,57
Flickering	$-0,78$	$-0,78$	$-0,79$

that transcodes and sends video to the VLC. Eventually the video in Test 3 is stored into a file and then evaluated for the QoE metrics (Table 5).

The results show that when transcoder is used freezing is more likely to happen and the blockloss and blockiness are worsened. This is caused by the fact of not enough processing power of the laptop used for the tests (the Intell i5 laptop). We have performed additional tests with higher rate videos (2 Mb/s, 4 Mb/s) and it became clear that at 4 Mb/s the freezing happened for each of all the tests performed.

5 Summary

In this paper we have shown architecture and application of trace based OFDM network emulator for WiMAX networks. The emulator has been thoroughly validated considering real and artificial traffic (camera, emulator), real channel behaviour (from traces), multiple transport protocols, with adaptive traffic (idealistic feedback) etc. The framework can be utilized to ease the costs and effort

required to perform real tests. Thanks to the proposed architecture we also gain repeatability of results when testing new versions of video control algorithms. Such algorithms can be placed on the transmitter node (PC1) and connected with the Bandwidth emulation block directly so that it replays channel behaviour in real-time. This way controller installed on PC1 can be tuned as it would react also in real-time to the perceived channel (modulation) changes. As future work we plan to deal with video controllers that learn from experience (i.e. history based) or can adjust to the location (also based on some previously observed network parameters in a location).

References

1. 5G-PPP whitepaper "5G automotive vision". https://5g-ppp.eu/. Accessed 01 July 2017
2. AGH telecommunications department home page. http://vq.kt.agh.edu.pl/. Accessed 01 July 2017
3. Candelatech. http://www.candelatech.com/wiser50_product.php. Accessed 01 July 2017
4. Cisco visual networking index. http://www.cisco.com/. Accessed 01 July 2017
5. Financial times home page. https://www.ft.com/. Accessed 01 July 2017
6. Ixia home page. http://www.ixiacom.com/products/ixia-network-emulators. Accessed 01 July 2017
7. Mushroom networks home page. https://www.mushroomnetworks.com/. Accessed 01 July 2017
8. Netem home page. https://wiki.linuxfoundation.org/networking/netem. Accessed 01 July 2017
9. Nsdie emulator. http://www.cs.cmu.edu/~glennj/JuddNSDIEmulator.pdf. Accessed 01 July 2017
10. Science digest home page. http://www.scdigest.com/. Accessed 01 July 2017
11. Servision home page. http://www.servision.net/solutions/mobile/. Accessed 01 July 2017
12. Splat! home page. http://www.qsl.net/kd2bd/splat.html. Accessed 01 July 2017
13. Uber home page. https://newsroom.uber.com/. Accessed 01 July 2017
14. Ali-Yahiya, T., Beylot, A.L., Pujolle, G.: An adaptive cross-layer design for multiservice scheduling in OFDMA based mobile WiMAX systems. **32**, 531–539 (2009). Adaptive Multicarrier Communications and Networks. http://www.sciencedirect.com/science/article/pii/S014036640800488X
15. Bernardo, M.V., Pinheiro, A.M., Fiadeiro, P.T., Pereira, M.: Quality assessment of chromatic variations: a study of full-reference and no-reference metrics. In: 2014 Proceedings of the 22nd European Signal Processing Conference (EUSIPCO), pp. 216–220. IEEE (2014)
16. Bernardo, V., Curado, M.: A methodology for assessing video transmission energy consumption and quality. In: 2012 IEEE International Conference on Communications (ICC), pp. 6308–6313, June 2012
17. Elayoubi, S.E., Fourestie, B.: Performance evaluation of admission control and adaptive modulation in OFDMA wimax systems. IEEE/ACM Trans. Netw. (TON) **16**, 1200–1211 (2008)

18. Flizikowski, A., Płócienniczak, M., Rybicki, M., Hołubowicz, W., Podyma, M.: Aassessing capabilities of commercial wimax networks for deliverling real-time surveillance video traffic in uplink. Image Process. Commun. **20**(1), 33–44 (2015). ISSN 1425–140X 33–44 (2015). ISSN 1425–140X
19. Kacianka, S., Hellwagner, H.: Adaptive video streaming for UAV networks. In: Proceedings of the 7th ACM International Workshop on Mobile Video, MoVid 2015, pp. 25–30. ACM, New York (2015). http://doi.acm.org/10.1145/2727040.2727043
20. Kofler, I.: In-network adaptation of scalable video content, vol. 2, pp. 7–8. ACM, New York, December 2010. http://doi.acm.org/10.1145/2039331.2039335
21. Li, X., Wu, X., Li, W., Wang, X.: An adaptive cross-layer scheduling algorithm for multimedia networks. In: 2008 International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 52–55, August 2008
22. Liu, Q., Wang, X., Giannakis, G.B.: A cross-layer scheduling algorithm with QoS support in wireless networks. IEEE Trans. Veh. Technol. **55**, 839–847 (2006)
23. Luo, H., Ci, S., Wu, D., Tang, H.: Adaptive wireless multimedia communications with context-awareness using ontology-based models. In: 2010 IEEE Global Telecommunications Conference, GLOBECOM 2010, pp. 1–5, December 2010
24. Rickenbach, B., Griffin, P., Rush, J., Flanagan, J., Adamson, B., Macker, J.P.: Adaptive data delivery over disadvantaged, dynamic networks. In: 2011 IEEE Military Communications Conference, MILCOM 2011, Baltimore, MD, USA, 7–10 November 2011, pp. 1628–1632 (2011). https://doi.org/10.1109/MILCOM.2011.6127542
25. Shuaib, K., Sallabi, F.: Smoothing of video transmission rates for an LTE network. In: 2010 IEEE 6th International Conference on Wireless and Mobile Computing, Networking and Communications, pp. 713–719, October 2010
26. Triantafyllopoulou, D., Passas, N., Salkintzis, A.K., Kaloxylos, A.: A heuristic cross-layer mechanism for real-time traffic in IEEE 802.16 networks. In: 2007 IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 1–5, September 2007
27. Triantafyllopoulou, D.K., Passas, N., Kaloxylos, A.: A cross-layer optimization mechanism for multimedia traffic over IEEE 802.16 networks. In: European Wireless (2007)

Moodle: Practical Advices for University Teachers

Jens Myrup Pedersen[1(✉)] and Mehmet Şükrü Kuran[2]

[1] Department of Electronic Systems, Aalborg University, Aalborg, Denmark
jens@es.aau.dk
[2] Abdullah Gul University, Kayseri, Turkey
sukru.kuran@agu.edu.tr

Abstract. Moodle is a widely used Learning Management System, with a market share of 20% in the US/Canada and 65% in Europe. However, it is our experience that the system is too often used just as a website or repository for classical teaching material such as literature references, slides and problems for students to solve after the lectures, and that the fully potential of the platform is not exploited. In this paper we demonstrate some of the functionalities that university teachers can make use of to increase the learning experience of the students. For each of the features we demonstrate, we both show how it can be used, and give some didactic considerations. We have tested all of the presented features ourself in a blended learning course carried out as part of an Erasmus+ Strategic Partnership.

1 Introduction

Moodle [4] is a widely used Learning Management System, with a market share of 20% in the US/Canada and 65% in Europe [2]. It offers a wide range of tools for organizing university courses. However, it is our experience that many university teachers only use the platform as a repository for the course material, without being aware of the powerful options that Moodle delivers. While the tools offered by Moodle should not dictate how courses are presented and conducted, at least teachers and instructors should be aware of them and consider them as additional tools they can use where it fits.

The tools described in this paper have been used and tested during a blended learning course given by seven different universities, described in further details in [7]. While they only represent a small fraction of the capabilities in Moodle, it is our experience that they are generally very useful: Easy to use for teachers, well received by students, and supporting active learning strategies. Teachers experimenting with flipped classrooms [6] and blended learning strategies will probably find them particularly helpful.

The paper is based on the authors own experience with the different Moodle features, obtained from a number of courses given in different institutions including a course on Future Internet Opportunities given across seven European universities, a course on Communication Networks given at Aalborg University,

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*,
Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_21

and a course on Network Security given at Riga Technical University. With the paper, we aim for giving university teachers a quick and short introduction to some of the tools in Moodle that we have tested ourself successfully. Some of the tools would be implemented in most institutions already, whereas others require installation of plugins that will often be done by Moodle administrators. The paper is focused on the technical aspects of setting up resources and activities in Moodle rather than the didactical aspects of how the different tools are used, even if we also will discuss some didactical considerations.

The paper is organised as follows: In Sect. 2 we introduce a number of overall features related to the course organisation, in Sect. 3 we describe features related to interactive content (automated), and in Sect. 4 we present features related to interaction between students and teachers. Section 5 concludes the paper.

2 Course Organisation

2.1 Bootstrap Elements

Our experience is that the traditional Moodle layout can be quite messy and confusing for the students. Bootstrap elements allow for e.g. creating dropdown menus as the one illustrated in Fig. 1. Also, the use of different themes can help to improve the look and feel. When Bootstrap elements are enabled, it can simply be added to a course like any other resource. One consideration when changing the layout is to remember that the students are usually following a number of courses, and if all courses are organized in different ways it can be confusing even if the intention was to simplify. This is even more so the case when multiple teachers share a course.

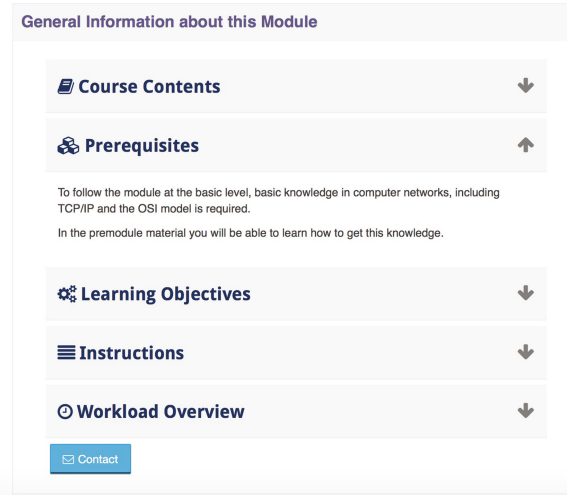


Fig. 1. Example of how a Bootstrap element can be used to make drop-down menus.

2.2 Completion Tracking

Another way to make it easier for students to keep track of the different activities in a course is to use completion tracking. We found it especially helpful to use when implementing some of the interactive activities described later in the paper, because it allows students to track their own progress. Moreover, it also allows teachers to get a good overview of the progress of the students in a class - both individual activities and performances as well as statistics for the whole class.

Figure 2 shows how it looks for a student, where the boxes to the right will be ticked whenever an activity has been registered as completed.

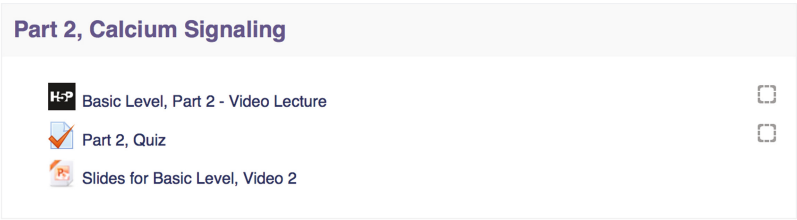


Fig. 2. Example of how completion tracking looks like for the student.

To enable completion tracking in a course, it has to be first enabled in the general course settings, and then added for each resource or activity where it is relevant. In most cases it is necessary to specify the conditions for when a resource or activity is included: It can be done either manually or automatically, based on e.g. the download of a resource, if a video has been seen, or if a quiz has been completed with sufficient grades. However, these options also illustrate the limitations of the tool, for example having downloaded a file or clicked on a video does not really determine whether the student has done the work required. We find that it is important to explain to the students how the completion tracking works, and what the intention is (e.g. to help the students in keeping an overview). This should also make it clear that a completed activity does not necessarily imply that the student has learned everything he should. This is also a reason why we would not recommend to use the tool for "checking" students behavior - the quizzes could probably be used this way, but in our implementation the focus was to provide the students with feedback, and we also did not implement any measures to prevent cheating.

2.3 Access Management

If completion tracking is used, it opens up for another useful tool, namely access management: For each resource or activity in a course access can be restricted in terms of e.g. completion of previous activities. For example, it can be used to ensure that one or more activities in a previous part of the course are completed

before the next part of the course becomes accessible to a student. The student view is illustrated in Fig. 3.

In our experience it is important to consider carefully how strict the access is implemented: For example, Some students might want to download course material in advance for offline use, or simply browse through the materials to get an overview of the course and topic - sometimes even before signing up for the course. Also, if access depends on manually graded activities, the student might need to wait unnecessarily long time before gaining access.

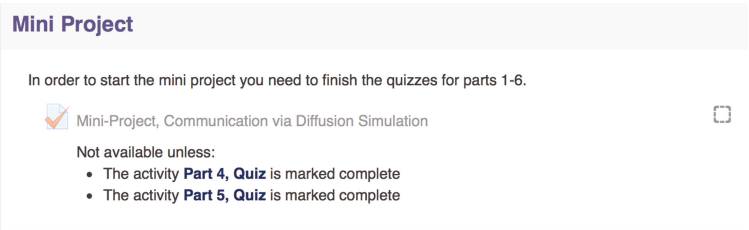


Fig. 3. Example of how access restrictions look like for the student.

2.4 Data and Statistics

Moodle offers some possibilities for teachers to access logs of student activity. In general we did not find it very useful for checking individual students behavior, since the logs available do not tell that much: For example, knowing that a student has accessed a document does not imply that he actually read it. On the other hand, "activity reports" can be useful for providing an overview of how much different course activities/materials have been used by the students (even if it is not too exact). All in all we found it more useful to use completion tracking as described above, and follow the activities in this way. When videos are made available through Youtube [5], the statistics provided here can also be useful (e.g. number of views and average view duration). However, here it is not possible to follow the activities of each student, and other viewers can influence the statistics.

3 Interactive Content

3.1 Videos and Quizzes

Videos are easy to embed in Moodle, either uploaded directly to Moodle or accessed through other websites such as Youtube. Both approaches have their pros and cons: Keeping everything in Moodle will often be the easiest solution with respect to content management and copyrights - on the other hand, Youtube makes the content accessible to the general public, and also allows for playing smoothly on a large range of devices.

The videos can be combined with a more interactive element, namely quizzes. In our experience, short videos of 5–10 min followed by quizzes with questions that cover the learning objectives of the videos are very useful, and also keep the student active. This approach with rather short videos also makes it easier to re-make videos with updated content later on, without the need for re-recordings of everything.

A common misconception is that quizzes can contain only multiple choice questions: In fact there are a wide range of question types, including some which requires manual grading. To set up a quiz, the questions are usually stored in a question bank, and questions can then be added to a quiz when it is set up (or at a later stage). It is also possible to define different questions categories, and then randomly select questions whenever a quiz is taken. There are many options for setting up the quizzes with respect to time limits, allowed number of attempts, timing, grading, passing criteria, etc. For the teacher, it is always possible to see the results achieved by each student.

Even with the many different question possibilities, we find it important to mention that not all learning objectives fits with the quiz format, and that the quiz format is not suitable for all learning activities, so quizzes should be used with consideration. It is also important to consider how the quizzes are used, for example if they are used as a way of providing feedback to students on their learning, or if they are used for test/examination purposes.

3.2 H5P Interactive Quizzes

H5P [3] is a plugin to Moodle that allows for a wide range of interactive activities. One such activity is interactive quizzes, which are embedded as a layer on top a video (which can be either shown directly in Moodle or through e.g. Youtube). For example, this makes it possible for students to view a video, and at a defined time the video is paused and a quiz pops up on top of the video. This makes it a much more integrated experience than when videos and quizzes are added as separate activities. On the other hand, the type of questions are more limited, and it is not possible to keep track of students results the same way as for the quizzes in Moodle.

If the H5P plugin is installed to Moodle, the elements are just added like any other activity in Moodle (or they can be created on the H5P website and imported, if this is preferred).

Figure 4 illustrates the integration of a simple quiz on top of a video on Youtube, using the H5P plugin.

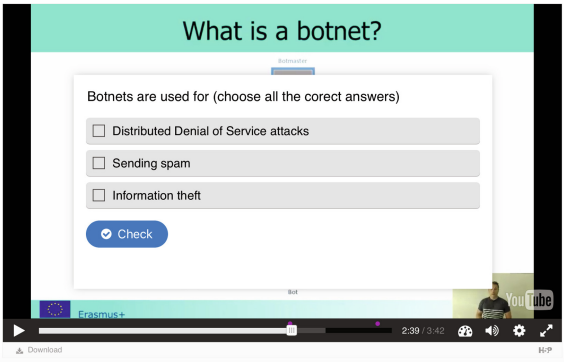


Fig. 4. Example of integrated video and quiz using the H5P plugin.

4 Interaction Between Students and Teachers

4.1 Forums

The Forums in Moodle are known by most teachers and course responsible, and will not receive much attention in this paper. However, we think it is important that teachers are aware of how they are used and integrated in the courses. Specifically, it is important to consider if the forums are used for one-way communication or for discussions - in the latter case it is important to ensure that the students have the rights to post questions. It is also important to be aware of the possibility to have several forums in a course, with different purposes.

In our experience it has been hard to get much activity in the forums, unless specific tasks or assignments are given.

4.2 Assignments

Assignments allow students to submit assignments, and teachers to assess and grade these. They are easy to set up, as they are added like any other activity in Moodle. There are several advantages of using the Moodle functionality for this instead of e.g. receiving submissions by email: It provides a single space for all submissions, the grading is recorded and stored, and it supports completion tracking and access management as described above.

Assignments can be added like any other activity, and as for other activities it is important to be aware of the timing settings (i.e. deadlines and cut-off times).

4.3 Workshops

Workshops is also a way of submitting assignments, but they allow for peer assessments, i.e. students reviewing the work of other students. In the end, the teacher can also assess each submission as well as the comments received by

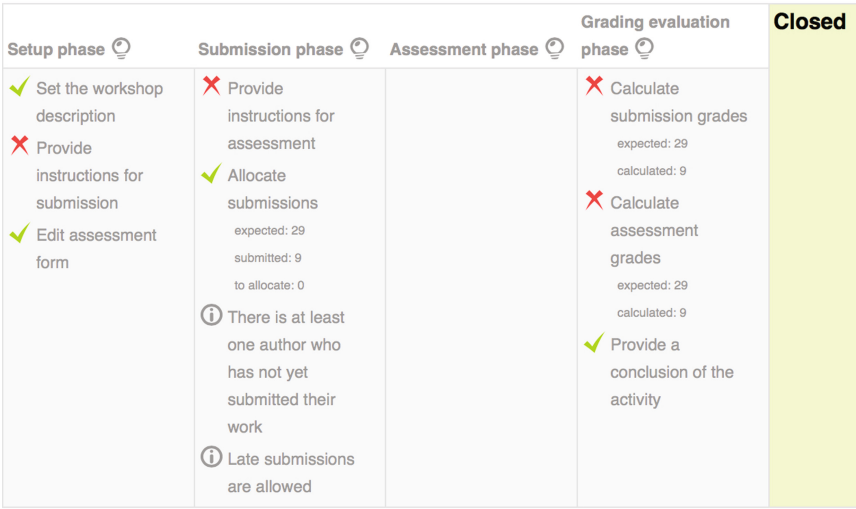


Fig. 5. Example of teacher view of a workshop.

the other students. Automatic grading can be done based on peer assessments, teacher assessments (of assignments and assessments), or a combination hereof.

We consider this a powerful tool, but the flexibility also complicates the use of it, and it is important to make sure that everything is set up correctly. Figure 5 illustrates the teachers view on a workshop that is set up. In particular, we found the following aspects important to be aware of:

A workshop is organized with four phases: A setup phase, a submission phase, an assessment phase, and a grading/evaluations phase. It is important to ensure either that the phases are shifting automatically by defining the timers accordingly, or that this is done automatically. In our experience, this is easy to forget. In particular, it is important to know that the students cannot see each others assessments until the final phase.

Also, it should be very clear what the students are expected to do, and how they are expected to assess each others work. Moodle provides the functionality of defining a set of criteria, so that each student should evaluate their peers work according to these with both grades and comments.

Another important thing to be aware of is how assignments are distributed among students. Moodle offers different ways to do this, either manually or automatically. For example, it is possible to set it up so that assignments are distributed randomly by the submission deadline, and the phase automatically switches to assessment phase. It is up to the teacher to define how many peers will review each assignment, and it is possible to exclude students who did not submit their own from being assigned the work of other students.

We found the workshop to be a powerful tool, and a fine way of implementing peer learning. However, a repeating challenge was students with late submissions, and how this should be handled: We attempted to distribute late submissions

to other students with late submissions, but partly because late submissions are not submitted at the same time it ended up with ad-hoc manual distributions.

5 Conclusion

It is our experience that many university teachers, even in institutions using Moodle, are not aware of the tools and functionalities offered that can ease their work and increase the learning experience of the students. In this short paper, we have introduced some of the most useful features and tools, and we have explained how they can be set up and used, and discussed points to be particularly aware of. We hope that this work can contribute to an easier take up of these tools, and inspire to explore the options in Moodle further.

While this paper has focused on the technical aspects of tools and features, we would like to stress that the activities implemented should of course make sense didactically, and not just be implemented for the sake of technology. We would encourage teachers interested also to study relevant paper on e.g. flipped classrooms and blended learning, but that is outside the scope of this paper.

Finally, the authors would like to thank the European Commission for supporting the work through the Erasmus+ Strategic Partnership Colibri [1].

References

1. Colibri project. http://erasmus-colibri.eu/. Accessed 2017
2. Edutechnica: LMS data - spring 2017 data. http://edutechnica.com/2017/03/12/lms-data-spring-2017-updates/. Accessed 2017
3. H5P. https://h5p.org/. Accessed 2017
4. Moodle. http://www.moodle.com/. Accessed 2017
5. Youtube. http://www.youtube.com/. Accessed 2017
6. Alvarez, B.: Flipping the classroom: homework in class, lessons at home. Educ. Dig. **77**(8), 18 (2012)
7. Pedersen, J.M., Lazaro, J.A., Mank, L., Eichhorn, V.: Blended learning and problem based learning in a multinational and multidisciplinary setting. In: IRSPBL 2017, Bogota, Columbia (2017, accepted)

Comparison of Visual Descriptors for 3D Reconstruction of Non-rigid Planar Surfaces

Michał Bednarek(✉)

Institute of Control and Information Engineering,
Poznan University of Technology, Poznań, Poland
michal.gr.bednarek@doctorate.put.poznan.pl

Abstract. In the deformable surface reconstruction, one relies on local descriptors to be able to perform matching between a reference frame and a deformed frame and afterwards perform optimisation to obtain the reconstruction. The quality of surface modelling highly depends on the matching process. Therefore, I am testing the performance of different detector-descriptor pairs in order to provide hints, which of these is best suitable for 3D reconstruction of non-rigid planar surfaces.

Keywords: Computer vision · Image processing · 3D reconstruction · Matching · Non-rigid surface

1 Introduction

The main goal for the computer vision is to find objects in the image and possibly reconstruct and track them. In the context of rigid objects, a lot has been done so far. However, the field of non-rigid objects to a large extent is left unexplored. In most of the cases, the tools from rigid object tracking are implemented to the non-rigid case. This is true for example when we talk about the process of feature detection and description. In my work, I am focusing my attention on an evaluation of different types of detectors and descriptors known in the literature applied to the task of 3D reconstruction of Non-Rigid Planar Surfaces.

The main problem in this scenario is local out-of-plane rotations of the object that are causing an improper description of such regions and therefore missing correct matching. The assessment of the descriptors in terms of their robustness against out-of-plane rotations is needed.

Therefore, in my work, I am concentrating on 6 different detectors and descriptors and provide a quantitative comparison of their performance in the task of non-rigid shape reconstruction. For the fair comparison, the results are

This work was supported by grant No. LIDER/3/0183/L-7/15/NCBR/2016 funded by The National Centre for Research and Development (Poland).
The original version of this chapter was revised: Frontmatter of the chapter has been updated with the provided text. The erratum to this chapter is available at https://doi.org/10.1007/978-3-319-68720-9_23

M. Chorać and R.S. Chorać (eds.), *Image Processing and Communications Challenges 9*,
Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9_22

obtained on a publicly available dataset and performed using algorithms implementation provided in the OpenCV library.

Firstly, I will describe the state of the art. Then, I will briefly present each of used detectors and descriptors. Following that, the results will be provided. Finally, some concluding remarks with future work plans will be given.

2 Related Work

One of the most prominent groups working on the problem of deformable shapes reconstruction is a group led by Ron Kimmel [2]. Their field of study is an identification of faces regardless of surface deformations caused by facial expressions. The Very important problem of deformable objects reconstruction was described in [4], where the variational methods were used. This work was extended by the same group in [20]. Yet another approach using the RGB-D camera was presented in [21]. The approach based on Laplacian mesh was described in [10]. Other work is focused on the evaluation of different methods of deformable objects search [13]. Latest work in computer vision which is focused on volume deformation is described in [5]. Another research field where the modelling of non-rigid objects is crucial is the reconstruction of tissues inside human body [15].

A lot of interest is also seen in the robotics community, as robots are able to manipulate deformable objects. Mainly it is done using point clouds [16]. One of the aspects of this field of study is clothes manipulation, authors are able to categorise them [19] and to obtain their pose [8]. Another project which named CloPeMa is focused on textiles. The authors are addressing the problem of clothes folding [3]. In this line of research, a paper on a strategic point of view selection [6] could be found.

3 Method

In this section I will present the approach that I selected. The algorithm that I am using for deformable surface reconstruction, relies on finding correspondences in *reference frame* and *input frame* given base 3D model of the object. If I have textured objects I can rely on local feature detector-descriptor.

The code used in the evaluation is the one provided by the authors of the paper [10]. In this work modelling of non-rigid surfaces is based on the extension of the Laplacian formalism, which is used for converting the 3D shape reconstruction from monocular image of a deformable surface given correspondences with a reference image into a better-posed problem. The benefit of this approach is quick and reliable method of elimination of outliers achieved by solving least squares problem. In this step one obtain initial 3D shape reconstruction, which 2D projections are accurate. The reconstruction is then refined using constrained optimisation problem.

In my work I have changed the front-end part which is dealing with the detection, description and matching process. I left unchanged the optimisation framework provided by the authors. Therefore, the only source of changes in the deformation results are based on the selection of the detector-descriptor pair.

In my work I have focused my attention on 6 different pairs of detectors and descriptors. The choice of these particular methods was driven by the fact that reference implementation were available in OpenCV. Below I am presenting a short review, which outlines main features of the selected detector-descriptor pairs.

SIFT is one of the most popular algorithms for finding and computing image features [9]. SIFT consists of 4 main steps. First of all, to achieve scale invariance, scale space extremes are localised using a difference-of-Gaussian function (over all scales and image locations). After that, low-contrast candidates and edges are rejected and the orientation, based on local image gradients directions, is assigned to each selected keypoint. At the end, descriptors for each point of interests are created – local image gradients computed from the neighbourhood of keypoint at the selected scale are transformed into a vector of floating point values.

SURF was designed to be a speeded-up version of SIFT [1]. The main difference is to approximate Laplacian of Gaussian with Box Filter. Wavelet responses are used for orientation assignment and final feature description. In theory, SURF supposed to be faster than SIFT in the task of finding keypoints as well as computing descriptors. Unfortunately, OpenCV implementation does not support this statement.

KAZE authors of KAZE detector and descriptor [11] noticed that building scale space by Gaussian blurring takes into account neither noise nor natural boundaries of objects in the scene. This leads to decrease of the accuracy in localisation and uniqueness of found keypoints. To eliminate mentioned difficulties, authors kindly proposed to build nonlinear scale space using Additive Operator Splitting technique [18] and variable conductance diffusion. Their idea makes object recognition system locally adaptive to the input image.

AKAZE was a step forward relative to significantly time-consuming KAZE algorithm [12]. Acceleration was possible by using Fast Explicit Diffusion to build nonlinear scale space. In addition, to describe detected keypoints the authors used Modified-Local Difference Binary descriptor (unlike descriptor KAZE, which operates on floating point numbers), which additionally speeded up the performance of the system and saved memory.

BRISK is a combination of FAST-based, scale-space detection method and description carried out through analysis of local intensity gradients [7].

ORB was algorithm designed to be alternative to SIFT and SURF in terms of computational cost and – what is more important – patents. ORB is free for its use [14]. For detection, it is using FAST algorithm with Harris corner measure in scale space. To achieve orientation invariance authors proposed to save some special measure known as intensity centroid. For description, ORB system uses slightly modified BRISK version, which deals with orientation.

In my work not only I use standardised methods, but additionally, all the results are performed on the same benchmark dataset provided in the same work [10] and recorded previously in [17].

4 Results

In this section, I will provide a comparison of quality of 3D reconstruction in terms of mean squared error (expressed in millimetres) using a variety of

detector-descriptor pairs. Based on these results, I will further relate the performance against the time required to perform the whole process. Establishing this relationship is crucial for the online solutions where we can sacrifice precision a little bit to obtain real time operation of the system.

4.1 Root Mean Square 3D Reconstruction Error

In my work, I have focused on the common benchmark datasets provided in work [17]. It contains two textured sequences: *paper* and *t-shirt*. The example images from the former are shown in Fig. 4a, b and for the latter are shown in Fig. 4c, d. I performed extensive tests for both sets of images. In Fig. 1a one can observe the performance of different detectors when using SIFT as a descriptor. In this test, I wanted to check the detection phase only. As it could be observed the best performance is achieved for KAZE descriptor. However, SIFT and AKAZE are close to this result.

In the case of T-shirt dataset, the results are slightly different from the one provided in the paper dataset. This may be due to the fact that paper is better textured and used detector has a simplified task in finding keypoints, hence the 3D reconstruction has better performance. In result for T-shirt dataset, there is a decreased number of correspondences and it has a negative influence on final 3D reconstruction result. In Fig. 2b, when detector-descriptor pairs of the same family are used, the best result was obtained for SURF. In Fig. 2a also there are changes in overall reconstruction error. Here the best performance was obtained for SIFT. In the less textured dataset, SIFT detector appears to find the biggest

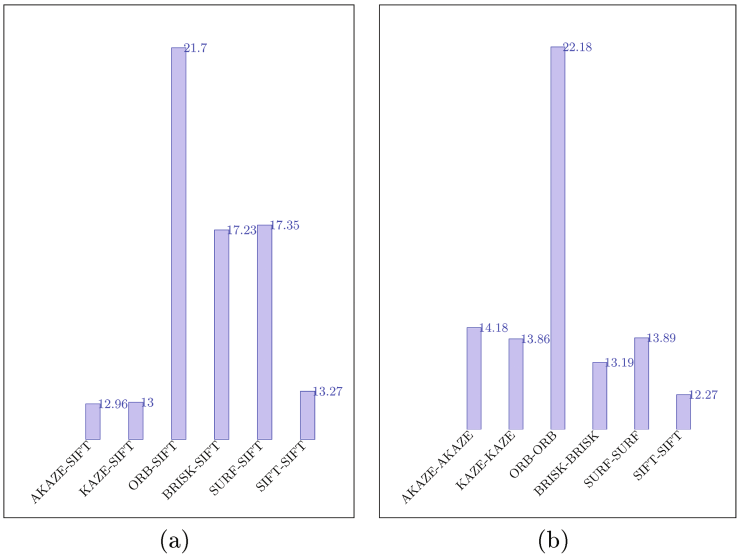


Fig. 1. The square root of mean square 3D reconstruction error in [mm] for paper dataset using (a) variety of detectors with SIFT descriptor and (b) descriptor and detector pairs of the same family for the paper dataset.

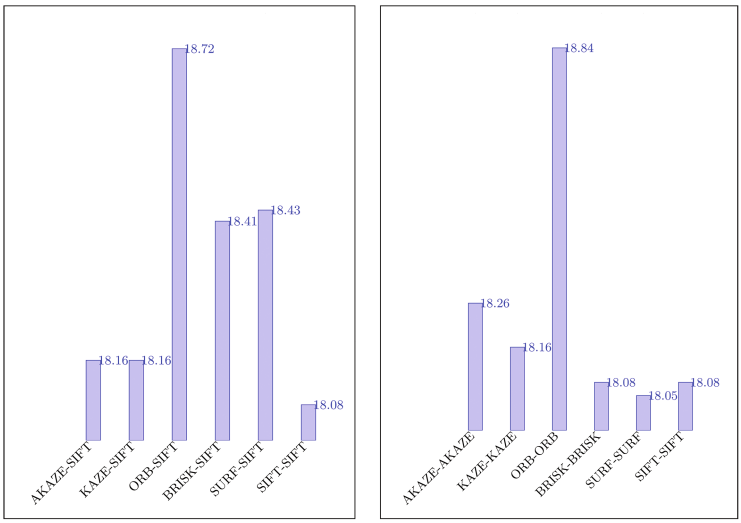


Fig. 2. The square root of mean square 3D reconstruction error [mm] for T-shirt dataset using (a) variety of detectors with SIFT descriptor and (b) descriptor and detector pairs of the same family for the shirt dataset.

number of correspondences between reference and input frame and – as a result – increases the quality of 3D reconstruction.

4.2 Time Performance

If we compare how much time is needed to obtain the results Fig. 3, the use of AKAZE is advisable, if we want to obtain a fast method and not to compromise the precision of the reconstruction.

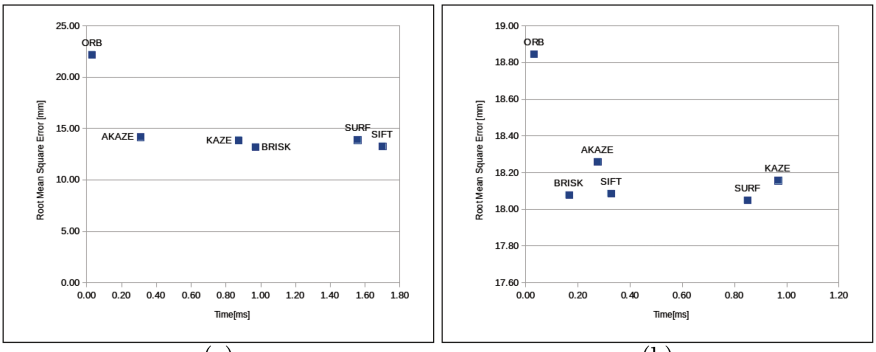


Fig. 3. Mean squared error depending on the time of detection and description in the same method performed on the paper dataset (a) and shirt dataset (b).

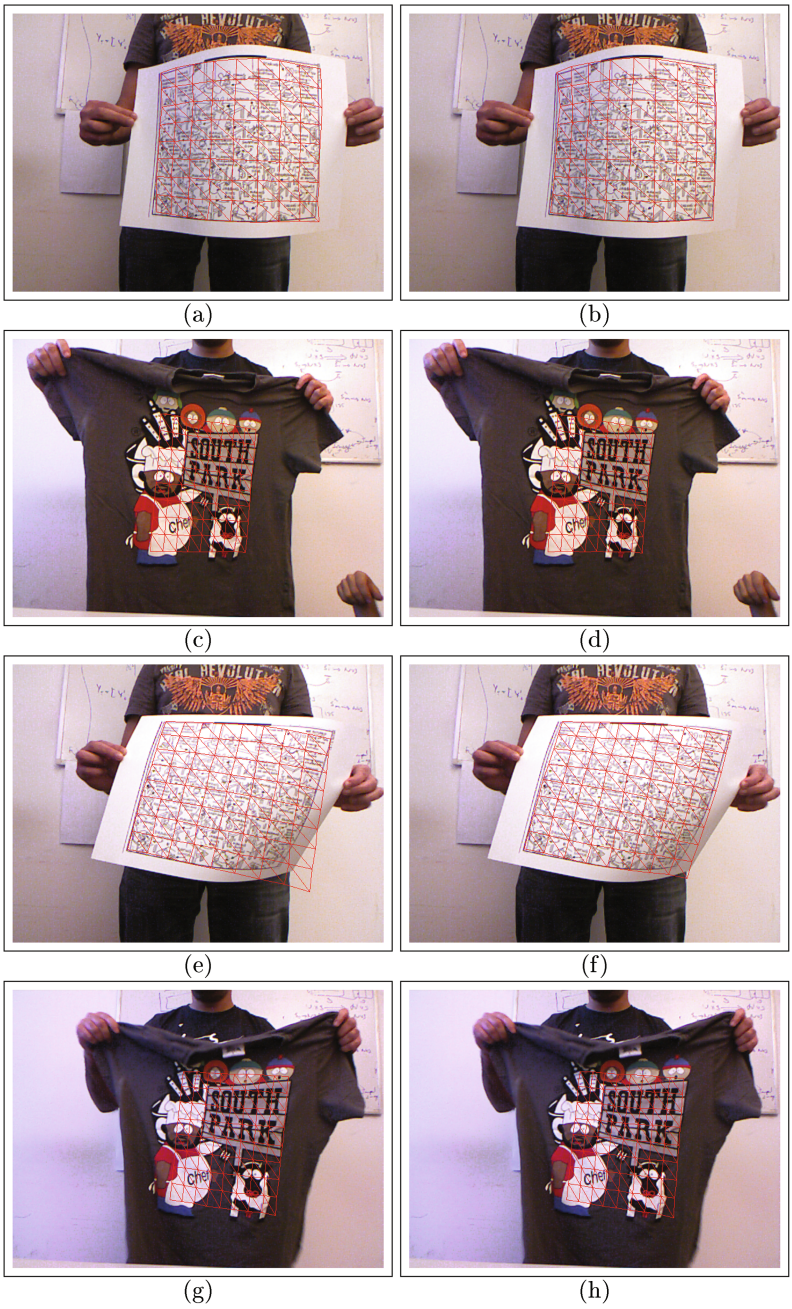


Fig. 4. The results of shape reconstruction using variety of detector-descriptor pairs: (a) BRISK-SIFT, (b) KAZE-SIFT, (c) BRISK-SIFT, (d) SIFT-SIFT, (e) AKAZE-AKAZE, (f) SIFT-SIFT, (g) KAZE-KAZE, (h) ORB-ORB

ORB is the fastest algorithm among all compared methods, but the quality of its final result is the worst. Also, as it can be seen – in better-textured paper dataset – SIFT has the slowest performance, but in T-Shirt dataset, it is significantly faster. It may be due to the fact, that this algorithm is able to find a large number of keypoints and when the number increases – time performance is decreasing, due to the costly matching process with floating point descriptors. It appears to be very well implemented in OpenCV, so when the number of keypoints is not that large – its performance is comparable even to binary descriptors as it can be observerd in Fig. 3b.

5 Conclusions

In this paper, I have compared the performance of 6 different detector-descriptor pairs in the task of 3D reconstruction of non-rigid planar surfaces. It turns out that binary descriptors are not losing that much performance when compared to the real-valued descriptors, but due to the faster computation of distances Hamming Distance, and L-2 norm respectively, binary descriptors are substantially faster. In the future work, I am planning to implement my own optimisation process for modelling 3D non-rigid surfaces. This paper is the first step towards achieving this goal.

Acknowledgments. This work was supported by grant No. LIDER/3/0183/L-7/15/NCBR/2016 funded by The National Centre for Research and Development (Poland).

References

1. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: SURF: Speeded Up Robust Features. Comput. Vis. Image Underst. (CVIU) **110**(3), 346–359 (2008)
2. Bronstein, A.M., Bronstein, M.M., Kimmel, R.: Numerical Geometry of Non-Rigid Shapes. Springer, Heidelberg (2009)
3. Doumanoglou, A., Kargakos, A., Kim, T.-K., Malassiotis, S.: Autonomous active recognition and unfolding of clothes using random decision forests and probabilistic planning. In: IEEE International Conference on Robotics and Automation (ICRA), pp. 987–993, May 2014
4. Garg, R., Roussos, A., Agapito, L.: Dense variational reconstruction of non-rigid surfaces from monocular video. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1272–1279, June 2013
5. Innmann, M., Zollhöfer, M., Nießner, M., Theobalt, C., Stamminger, M.: VolumeDeform: Real-Time Volumetric Non-rigid Reconstruction, pp. 362–379. Springer, Cham (2016)
6. Kita, Y., Kanehiro, F., Ueshiba, T., Kita, N.: Clothes handling based on recognition by strategic observation. In: 11th IEEE-RAS International Conference on Humanoid Robots (Humanoids), pp. 53–58, October 2011
7. Leutenegger, S., Chli, M., Siegwart, R.: BRISK: binary robust invariant scalable keypoints. In: IEEE International Conference on Computer Vision (ICCV), pp. 2548–2555 (2011)

8. Li, Y., Wang, Y., Case, M., Chang, S., Allen, P.K.: Real-time pose estimation of deformable objects using a volumetric approach. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 1046–1052, September 2014
9. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. (IJCV) **60**(2), 91–110 (2004)
10. Ngo, T.D., Östlund, J., Fua, P.: Template-based monocular 3D shape recovery using Laplacian meshes. IEEE Trans. Pattern Anal. Mach. Intell. **38**(1), 172–187 (2016)
11. Bartoli, A., Alcantarilla, P.F., Davison, A.J.: KAZE Features. Springer, Heidelberg (2013)
12. Bartoli, A., Alcantarilla, P.F., Nuevo, J.: Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces (2013)
13. Pickup, D., Sun, X., Rosin, P.L., Martin, R.R., Cheng, Z., Lian, Z., Aono, M., Ben Hamza, A., Bronstein, A., Bronstein, M., Bu, S., Castellani, U., Cheng, S., Garro, V., Giachetti, A., Godil, A., Han, J., Johan, H., Lai, L., Li, B., Li, C., Li, H., Litman, R., Liu, X., Liu, Z., Lu, Y., Tatsuma, A., Ye, J.: SHREC 2014 track: shape retrieval of non-rigid 3D human models. In: Proceedings of the 7th Eurographics Workshop on 3D Object Retrieval, EG 3DOR2014. Eurographics Association (2014)
14. Rublee, E., Rabaud, V., Konolige, K., Bradski, G.: ORB: an efficient alternative to SIFT or SURF. In: IEEE International Conference on Computer Vision (ICCV), pp. 2564–2571 (2011)
15. Samir, C., Kurtek, S., Srivastava, A., Canis, M.: Elastic shape analysis of cylindrical surfaces for 3D/2D registration in endometrial tissue characterization. IEEE Trans. Med. Imaging **33**(5), 1035–1043 (2014)
16. Schulman, J., Lee, A., Ho, J., Abbeel, P.: Tracking deformable objects with point clouds. In: IEEE International Conference on Robotics and Automation (ICRA), pp. 1130–1137, May 2013
17. Varol, A., Salzmann, M., Fua, P., Urtasun, R.: A constrained latent variable model. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012, pp. 2248–2255. IEEE Computer Society (2012)
18. Weickert, J., Romeny, B.M.T.H., Viergever, M.A.: Efficient and reliable schemes for nonlinear diffusion filtering. IEEE Trans. Image Process. **7**, 398–410 (1998)
19. Willimon, B., Walker, I., Birchfield, S.: A new approach to clothing classification using mid-level layers. IEEE International Conference on Robotics and Automation (ICRA), pp. 4271–4278, May 2013
20. Yu, R., Russell, C., Campbell, N.D.F., Agapito, L.: Direct, dense, and deformable: template-based non-rigid 3D reconstruction from RGB video. In: International Conference on Computer Vision (ICCV) (2015)
21. Zollhoefer, M., Niessner, M., Izadi, S.M., Rehmann, C., Zach, C., Fisher, M., Wu, C., Fitzgibbon, A., Loop, C., Theobalt, C., Stamminger, M.: Real-time non-rigid reconstruction using an RGB-D camera. ACM Trans. Graph. **33**(4), 156:1–156:12 (2014). Lipiec

Erratum to: Comparison of Visual Descriptors for 3D Reconstruction of Non-rigid Planar Surfaces

Michał Bednarek(✉)

Institute of Control and Information Engineering,
Poznan University of Technology, Poznań, Poland
michal.gr.bednarek@doctorate.put.poznan.pl

Erratum to:
Chapter "Comparison of Visual Descriptors for 3D Reconstruction of Non-rigid Planar Surfaces" in:
M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*, Advances in Intelligent Systems and Computing 681,
https://doi.org/10.1007/978-3-319-68720-9_22

In the original version of the book, Frontmatter of Chapter 22 has been updated by including the following text in the footnote:

This work was supported by grant No. LIDER/3/0183/L-7/15/NCBR/2016 funded by The National Centre for Research and Development (Poland).

The updated online version of this chapter can be found at https://doi.org/10.1007/978-3-319-68720-9_22

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*,
Advances in Intelligent Systems and Computing 681, https://doi.org/10.1007/978-3-319-68720-9_23

Author Index

M. Choraś and R.S. Choraś (eds.), *Image Processing and Communications Challenges 9*,
Advances in Intelligent Systems and Computing 681, DOI 10.1007/978-3-319-68720-9