Studies in Computational Intelligence 473

Roumen Kountchev
Barna Iantovics (Eds.)

Advances in Intelligent Analysis of Medical Data and Decision Support Systems

Springer

Studies in Computational Intelligence 473

Editor-in-Chief

Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw
Poland
E-mail: kacprzyk@ibspan.waw.pl

For further volumes:
http://www.springer.com/series/7092

Roumen Kountchev and Barna Iantovics (Eds.)

Advances in Intelligent Analysis of Medical Data and Decision Support Systems

Springer

Editors
Prof. Roumen Kountchev
Technical University of Sofia
Sofia
Bulgaria

Dr. Barna Iantovics
Petru Maior University
Targu Mures
Romania

ISSN 1860-949X
e-ISSN 1860-9503
ISBN 978-3-319-00028-2
e-ISBN 978-3-319-00029-9
DOI 10.1007/978-3-319-00029-9
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013930426

Printed on acid-free paper

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

Preface

The volume is a result of the very fruitful and vivid discussions during the MedDecSup'2012 International Workshop at which the participants decided to put together this volume to collect relevant body of knowledge, and new developments in this increasingly important field of medical informatics. A natural decision was made to base this volume on relevant papers presented at the Workshop, predominantly in their extended versions which have been prepared by taking into account the discussions and newest developments. The basic trend in these publications is the creation of new ideas aimed at the development of intelligent processing of various kinds of medical information and the perfection of the contemporary computer systems for medical decision support. The advance of the medical information systems for intelligent archiving, processing, analysis and search-by-content will improve without question the quality of the medical services for every patient and of the global healthcare system in all countries. This goal will certainly be facilitated by a general attitude adopted in the volume that combines in a synergistic way theoretical developments with implementability of the approaches developed.

The basic purpose of this book is to present the last developments and achievements in the research activity of the authors in the area of the medical informatics to wide range of readers: mathematicians, engineers, physicians, PhD students and other specialists.

In the book there are included 21 papers, covering topics in the areas mentioned above. The basic trends could be grouped as follows:

- Novelties in the application of the machine learning and data interpretation methods in the medical informatics;
- New approaches for improvement the quality of ultrasound, CT, X-ray, etc. medical images;
- New methods for decorrelation of groups of CT images, and for compression and filtration of US images;
- New descriptors for enhancement the efficiency of the search-by-content in medical image databases;

- New algorithms for early diagnostics through processing and analysis of various kinds of patients' medical images and the related text documents, which describe the history of the sickness;
- New applications of the theory of graphs used for solving some problems of the medical services;
- New generation of bio-medical equipment for signal processing, based on the FPGA, and on special software architecture for efficient processing of medical images.

The presented papers do not cover all aspects of the intelligent processing of the medical information and of the systems for medical decision support, but shall undoubtedly contribute for their further successful development.

We express our sincere thanks to all plenary speakers, section chairs, reviewers, and authors for their significant contribution for the preparation of this book. We thank also the Organizing committee of the MedDecSup'2012 workshop, the Technical University of Sofia, and our sponsor Smartcom-AD, who provided us with everything we needed for the successful performance of the workshop. We also express special thanks to the publishing board of Springer, who made this book possible.

We hope that this book will stimulate the development of new ideas and efficient solutions in the area of the medical systems for intelligent data processing and medical decision support.

December 2012 The Editors

Contents

Machine Learning Applications in Cancer Informatics

Abdel-Badeeh M. Salem

BioMedical Informatics and Knowledge Engineering Research Lab.,
Faculty of Computer and Information Sciences,
Ain Shams University, Abbasia, Cairo-Egypt
abmsalem@yahoo.com, asalem@eun.eg
http://elearn.shams.edu.eg/absalem/27.htm

Abstract. Cancer informatics is a multidisciplinary field of research. It includes oncology, pathology, radiology, computational biology, physical chemistry, computer science, information systems, biostatistics, machine learning, artificial intelligence (AI), data mining and many others. Machine learning (ML) offers potentially powerful tools, intelligent methods, and algorithms that can help in solving many medical and biological problems. The variety of ML algorithms enable the design of a robust techniques and new methodologies for managing, representing, accumulating, changing ,discovering ,and updating knowledge in cancer-based systems. Moreover it supports learning and understanding the mechanisms that will help oncologists, radiologists and pathologists to induce knowledge from cancer information databases. This paper presents the research results of the author and his colleagues that have been carried out in recent years on using machine learning in cancer informatics. In addition the talk discusses several directions for future research.

Keywords: Machine Learning, Cancer Informatics, Case-Based Reasoning (CBR), Ontological Engineering, Genetic Algorithms, Medical Knowledge Management.

1 Introduction

Cancer is a group of more than 200 different diseases; it occurs when cells become abnormal and keep dividing and forming either benign or malignant tumors. Cancer has initial signs or symptoms if any is observed, the patient should perform complete blood count and other clinical examinations. Then to specify cancer type, patient need to perform special lab-tests. Benign tumors can usually be removed and do not spread to other parts of the body. Malignant tumors, on the other hand, grow aggressively and invade other tissues of the body, allowing entry of tumor cells into the bloodstream or lymphatic system which spread the tumor to other sites in the body. This process of spread is termed metastasis; the areas of tumor growth at these distant sites are called metastasis.

From the informatics point of view, breast cancer classification, diagnosis and prediction techniques have been a widely researched area in the past decade in the world of medical informatics. Several articles have been published which tries to classify breast cancer data sets using various techniques such as fuzzy logic, support vector

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 1–14.
DOI: 10.1007/978-3-319-00029-9_1

machines, Bayesian classifiers, decision trees, neural networks, and case-based reasoning[1,2,3,4]. CBR researchers agree that the best way to satisfy the increasing demand of developing CBR applications is by development of CBR-based frameworks. CBR frameworks are provided by research groups to overcome the problem of disturbing the concentration of the researchers in different domains with programming of AI applications, and help researches focusing on building components that directly address key concepts on a provided infrastructure that facilitates application to large scale projects [4,5,6].

This paper discusses the application of machine learning techniques in cancer informatics. The paper describes the following applications; (a) Case-Based Reasoning for diagnosis of cancer diseases, (b) Ontological engineering for lung and breast cancer knowledge management, (c) data mining for assessing diagnosis of breast cancer, and(d) genetic algorithms based classifier for breast cancer disease.

2 Case-Based Reasoning Approach for Diagnosis of Cancer Diseases

This section presents a CBR-based expert system prototype for diagnosis of cancer diseases developed at Medical Informatics Group at Ain Shams University, Cairo [7, 8]. The main purpose of the system is to serve as doctor diagnostic assistant. The system provides recommendation for controlling pain and providing symptom relief in advanced cancer. It can be used as a tool to aid and hopefully improve the quality of care given for those suffering intractable pain. The system is very useful in the management of the problem, and its task to aid the young physicians to check their diagnosis.

2.1 Case-Based Reasoning Methodology

CBR receives increasing attention within the AI community [4]. CBR is an analogical reasoning method provides both a methodology for problem solving and a cognitive model of people. CBR means reasoning from experiences or "old cases" in an effort to solve problems, critique solutions, and explain anomalous situations. It is consistent with much that psychologist have observed in the natural problem solving that people do. People tend to be comfortable using CBR methodology for decision making, in dynamically changing situations and other situations were much is unknown and when solutions are not clear. CBR refers to a number of concepts and techniques that can be used to record and index cases and then search them to identify the ones that might be useful in solving new cases when they are presented. In addition, there are techniques that can be used to modify earlier cases to better match new cases and other techniques to synthesize new cases when they are needed.
Following to the CBR methodology, the algorithm of interpreting and assimilating a new case can be summarized in the following processes [4]:

1. Assign Indexes: the features of the new case are assigned as indexes characterizing the event.

2. Retrieve: the indexes are used to retrieve a similar past case from the case memory.
3. Modify: the old solution is modified to conform to the new situation, resulting in a proposed solution.
4. Test: the proposed solution is tried out. It either succeeds or fails.
5. Assign and Store: If the solution succeeds, then assign indexes and stores a working solution. The successful plan is then incorporated into the case memory.
6. Explain, Repair and Test: If the solution fails, then explain the failure, repair the working solution, and test again. The explanation process identifies the source of the problem. The predictive features of the problem are incorporated into the indexing rules knowledge structure to anticipate this problem in the future. The failed plan is repaired to fix the problem, and the revised solution is then tested.

To perform the CBR process, the following knowledge structures (KSs) are very essential:

1. Indexing Rules KS: These rules identify the predictive features in the input that provides appropriate indexes into the case memory.
2. Case Memory KS: Case memory is the episodic memory, which comprises of database of experience.
3. Similarity Rules KS: If more than one case is retrieved form episodic memory, the similarity rules (or metrics) can be used to decide which case is more like the current situation.
4. Modification Rules KS: If no old case is going to be an exact match for a new situation, the old case must be modified to fit. We require knowledge about what kinds of factors can be changed and how to change them.
5. Repair Rules KS: Once we identify and explain an expectation failure, we must try to alter our plan to fit the new situation. Again we have rules for what kinds of changes are permissible.

2.2 Technical Aspects of the Case-Based Expert System

The system consists of three main modules; user interface, case base reasoning module and computational module all are interacted with the main environment of cancer diseases [7, 8]. The user is cancer expert doctor, the interaction is through menus and dialogues that simulate the patient text sheet contain symptoms and lab examinations. Computational module uses rule-based inference to give diagnostic decision and new case is stored in case library. Patient cases are retrieved in dialogue with similarity matches using the nearest neighbor matching technique. Frame knowledge representation technique is used for patient case indexing, storage and retrieval.

Fig. 1 shows the architecture of the CBR-based system. The system's knowledge base is diverse and linked through a number of indices, frames and relationships. The bulk of this knowledge consists of actual case histories and includes 70 cancer patient cases; some are real Egyptian cases and some from virtual hospitals on the internet. Fig.2 shows example of an Egyptian liver cancer case description of old women.

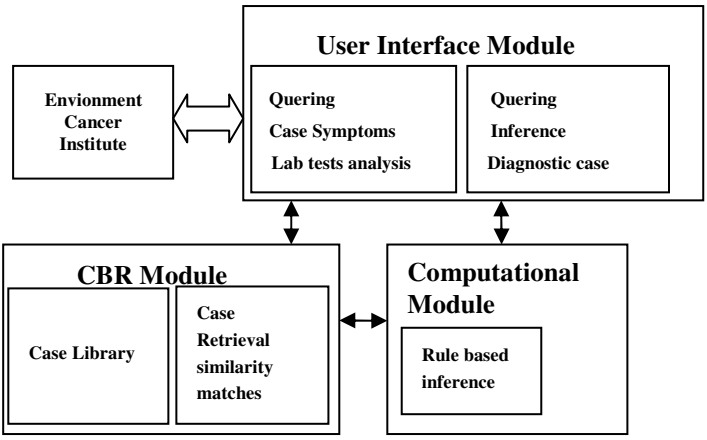


Fig. 1. Architecture of the CBR-based system for cancer diagnosis

Patient: 65-years old female not working, with nausea and vomiting.
Medical History: cancer head of pancreas
Physical Exam: tender hepatomgaly liver, large amount of inflammatory about 3 liters, multiple liver pyogenic abscesses and large pancreatic head mass.
Laboratory Findings: total bilrubin 1.3 mg/dl, direct bilrubin 0.4 mg/dl, sgot (ast) 28 IU/L, sgpt (alt) 26 IU/L.

Fig. 2. Egyptian liver cancer case

3 Ontological Engineering Approach for Cancer Knowledge Management

The term "ontology" is inherited from philosophy, in which it is a branch of metaphysics concerned with the nature of being. It began being used in Artificial Intelligence (AI) in the 1980s, and is now frequently used by computing and information science audiences. In AI ontology can be defined simply as a common vocabulary for describing a domain that can be used by humans as well as computer applications. Ontology is a formal explicit description of concepts in a domain of discourse (classes (sometimes called concepts)), properties of each concept describing various features and attributes of the concept (slots (sometimes called roles or properties)), and restrictions on slots (facets (sometimes called role restrictions)). Ontology together with a set of individual instances of classes constitutes a knowledge base. In reality, there is a fine line where the ontology ends and the knowledge base begins. Ontologies have a range of potential benefits and applications in medical domain including the sharing of medical information across medical systems, enabling reuse of medical knowledge and providing intelligent and personalized researcher support [9, 10].

This section presents the developed two web-based ontologies for lung and breast cancers [11, 12]. The two ontologies; were built using the Protégé-OWL editing environment. The main goals behind building these ontologies are to allow finding and locating information about lung and breast cancers needed for interested users and domain experts, integrating information about lung and breast cancers to be accessed in an easy manner and providing the availability and accessibility of lung and breast cancers knowledge over the web.

3.1 Web-Based Lung Cancer Ontology

The lung cancer ontology was encoded in OWL-DL format using the Protégé-OWL editing environment [11]. The knowledge concerning breast cancer is collected from many sources including: Cancerbackup , American Cancer Society, American Lung Association. In this ontology (Figure 3) we have the following four main super classes ;

- **People**, which has the sub classes; male and female.
- **Medical_Interventions** which has sub classes; *Treatment, Staging* and *Diagnosis.*
- **Disease** which has sub class *cancer* which has sub class; *lung_cancers.*
- **Disease_attributes which has sub classes; Causes, Disease_stage, Pathological_category, Staging_systems and Symptoms**

The lung cancers are described in terms of its symptoms, causes, stages, pathological category, diagnosis and treatment. In this context, we described causes, stages, pathological category and symptoms as disease attributes. While diagnosis (including determining the stage of the lung cancer) and treatment are described as medical interventions.

3.2 Web-Based Breast Cancer Ontology

Breast cancer ontology was encoded in OWL-DL format using the Protégé-OWL editing environment [12]. The knowledge concerning breast cancer is collected from many sources including: MedicineNet ,The World Health Organization (WHO) , The breastcancer.org , The ehealthMD and The National Comprehensive Cancer Network (NCCN) .

In this ontology we have two main super classes

- MedicalThings which has sub classes Diseases, Medical_Interventions, Pathological_Category, References.
- People which has the sub classes; men and women.

The class Diseases has a subclass Cancers which has a subclass Breast_Cancer. The class Medical_Interventions has subclasses Diagnostic and Therapeutic. The class References has subclasses Causes, Disease_Stage, Staging, Symptoms and TNM_Stage. Some of the subclasses motioned above may has its own sub classes as

shown in figure 4. These entire sub classes are related with is-a link. The breast cancers are described in terms of its symptoms, causes, stages, pathological category, diagnosis and treatment. In this context, we described causes, stages, and symptoms as references. While diagnosis and treatment are described as medical interventions as shown on figure 4.

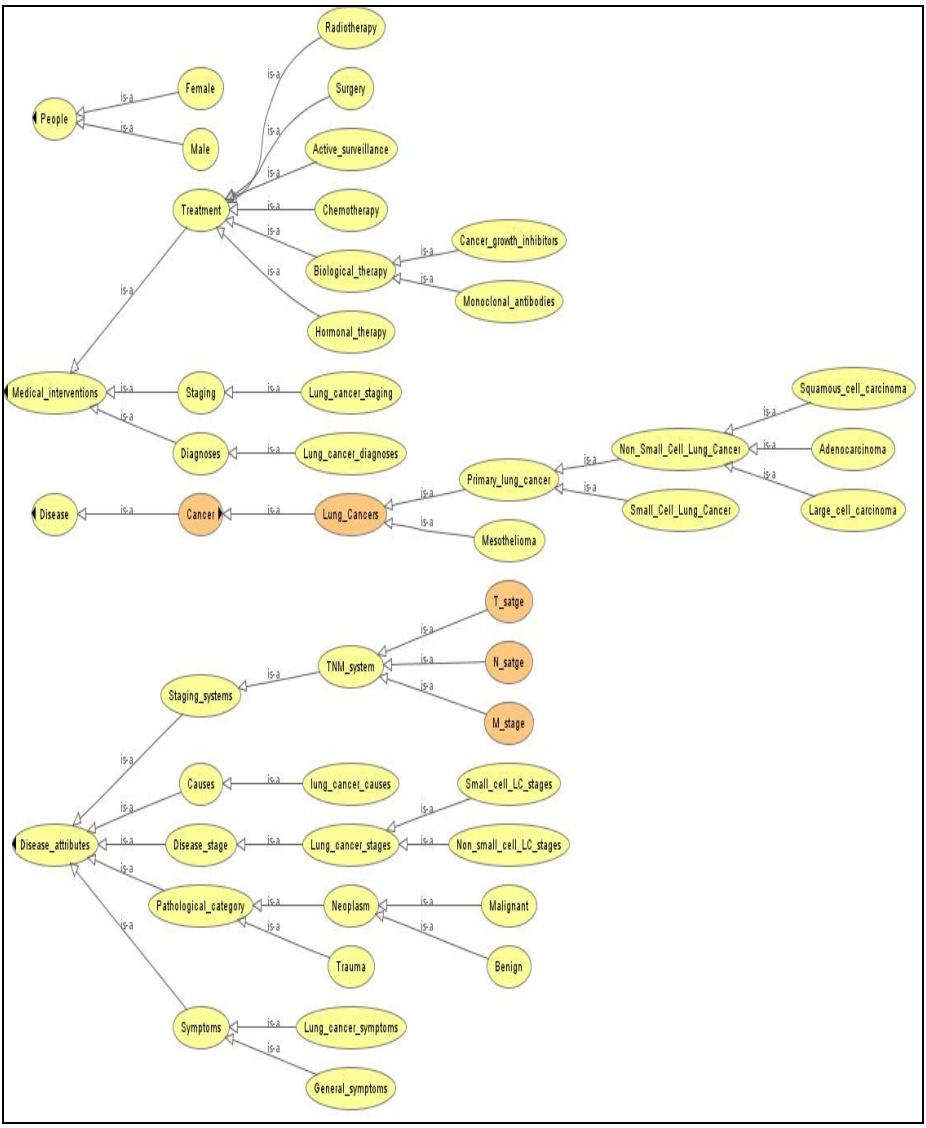


Fig. 3. The Lung Cancer Class Hierarchy

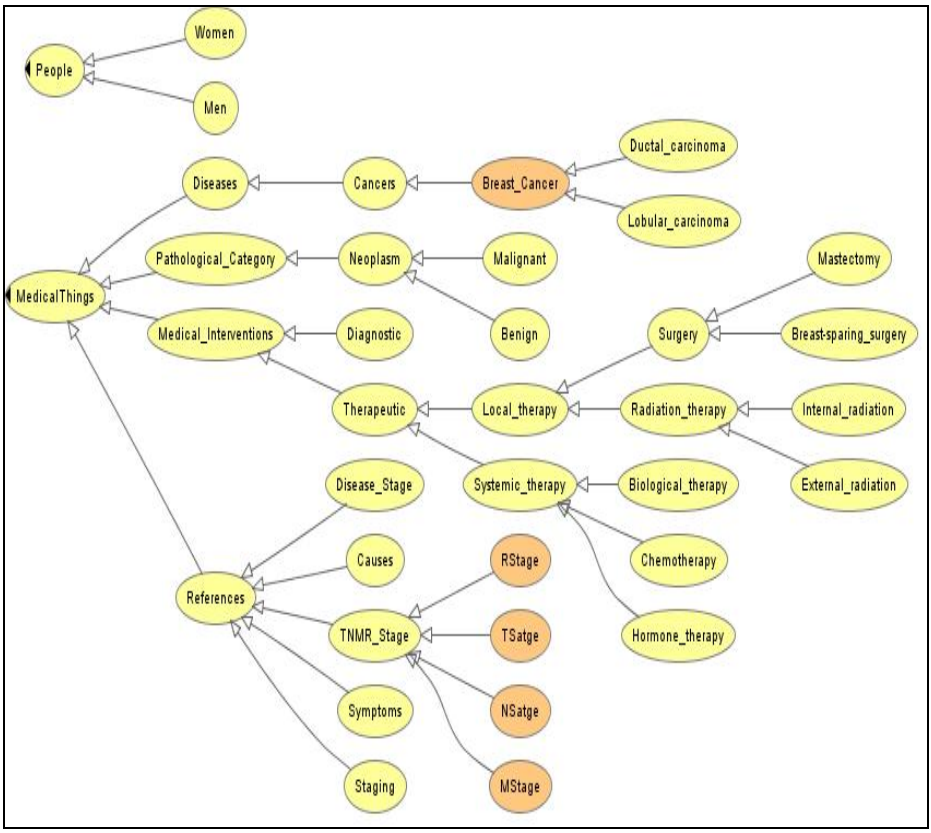


Fig. 4. The Developed Breast Cancer Ontology

In the breast cancer ontology, we described the diagnosis of the breast cancer as instances of the class Diagnostic. Also causes, stages, staging (how to determine the stage of the cancer) and symptoms of the breast cancer are described as instances. In the breast cancer ontology, the classes M_stage, N_stage and T_stage are defined as enumerated classes. Each one of these classes is described in terms of its instances.

4 Breast Cancer Classification Using a Combination of Ontology and Case-Based Reasoning

In our research [13, 14], we developed a new technique in the field of breast cancer classification. It uses a combination of ontology and case-based reasoning methodologies. Two recent frameworks are examined building the classifier. One is the open source jCOLIBRI system developed by GAIA group and provides a framework for building CBR systems. The other is the novel open source CBR tool, myCBR, developed at the German Research Center for Artificial Intelligence (DFKI). The objective

of this classifier is to classify the patient based on his/her electronic record whether he/she is benign or malignant.

4.1 Breast Cancer Domain

Breast cancer is the form of cancer that either originates in the breast or is primarily present in the breast cells. The disease occurs mostly in women but a small population of men is also affected by it. Early detection of breast cancer saves many thousands of lives each year. Many more could be saved if the patients are offered accurate, timely analysis of their particular type of cancer and the available treatment options. Since the breast tumors whether malignant or benign share structural similarities, it becomes an extremely tedious and time consuming task to manually differentiate them. As seen in Figure 5 there is no visually significant difference between the fine needle biopsy image of the malignant and benign tumor for an untrained eye.

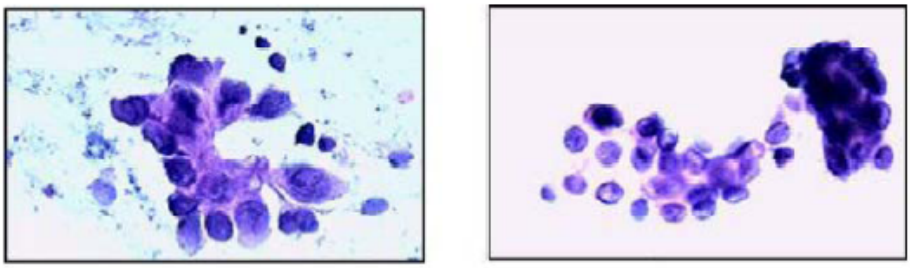


Fig. 5. Fine needle biopsies of breast. Malignant (left) and Benign (right) [15].

Accurate classification is very important as the potency of the cytotoxic drugs administered during the treatment can be life threatening or may develop into another cancer. Laboratory analysis or biopsies of the tumor is a manual, time consuming yet accurate system of prediction. It is however prone to human errors, creating a need for an automated system to provide a faster and more reliable method of diagnosis and prediction for the patients.

4.2 The Open Source jCOLIBRI Framework

jCOLIBRI framework is an evolution of the COLIBRI architecture [16], that consisted of a library of problem solving methods (PSMs) for solving the tasks of a knowledge-intensive CBR system along with ontology, CBROnto [8], with common CBR terminology. COLIBRI was prototyped in LISP using LOOM as knowledge representation technology. The design of the *jCOLIBRI* framework comprises a hierarchy of Java classes plus a number of XML files. The framework is organized around the following elements:

***Tasks and Methods*:** The tasks supported by the framework and the methods that solve them are all stored in a set of XML files.

***Case Base*:** Different connectors are defined to support several types of case determination, from the file system to a database.

***Cases*:** A number of interfaces and classes are included in the framework to provide an abstract representation of cases that support any type of actual case structure.

***Problem Solving Methods*:** The actual code that supports the methods included in the framework.

4.3 The Open Source myCBR Framework

myCBR is an open-source plug-in for the open-source ontology editor *Protégé*. *Protégé* is based on Java, is extensible, and provides a plug-and-play environment that makes it a flexible base for rapid prototyping and application development [17]. *Protégé* allows defining classes and attributes in an object-oriented way. Furthermore, it manages instances of these classes, which *myCBR* interprets as cases. So the handling of vocabulary and case base is already provided by *Protégé*. The *myCBR* plug-in provides several editors to define similarity measures for an ontology and a retrieval interface for testing. As the main goal of *myCBR* is to minimize the effort for building CBR applications that require knowledge-intensive similarity measures, *myCBR* provides comfortable GUIs for modeling various kinds of attribute specific similarity measures and for evaluating the resulting retrieval quality. In order to reduce also the effort of the preceding step of defining an appropriate case representation, it includes tools for generating the case representation automatically from existing raw data. The novice as well as the expert knowledge engineer are supported during the development of a *myCBR* project through intelligent support approaches and advanced GUI functionality. Knowledge engineer can go through the following four steps to develop a CBR System; (a) Generation of case representations, (b) Modeling similarity measures, (c) Testing of retrieval functionality, and (d) Implementation of a stand-alone application.

4.4 Experimental Results of Breast Cancer Classifications

In our study, the two mentioned CBR frameworks are tested by developing a CBR application that classifies the condition of the breast cancer tumor whether it is benign or malignant. Wisconsin breast cancer data set was used for building the case-bases. It is obtained from the University of Wisconsin Hospitals, Madison from Dr. William H. Wolberg [14]. Samples inside the data set arrive periodically as Dr. Wolberg reports his clinical cases. The number of instances inside the dataset is 699 (as of 15 July 1992). Each record contains ten attributes plus the class attribute.

During the implantation of the breast cancer diagnostic application using jCOLIBRI we found that jCOLIBRI is user-friendly and efficient to develop a quick application. The classifier was successful in classification of the selected data set. Results given by the testing of the developed application are quite satisfactory. The system can precisely retrieve most similar cases although the input query has a missing attribute value. Deployment of a standalone application is completely failed.

During the implantation of the breast cancer classifier using myCBR we noticed that myCBR is a really a tool for rapid prototyping of a new CBR application. In seconds, users may have a running standalone CBR application by using the CSV importing feature. myCBR is intelligent enough to build the case structure and the case base by parsing the provided CSV file. myCBR avoids reinventing the wheel by making the development of a new CBR application done inside Protégé. The classifier was successful in classification of the selected data set.

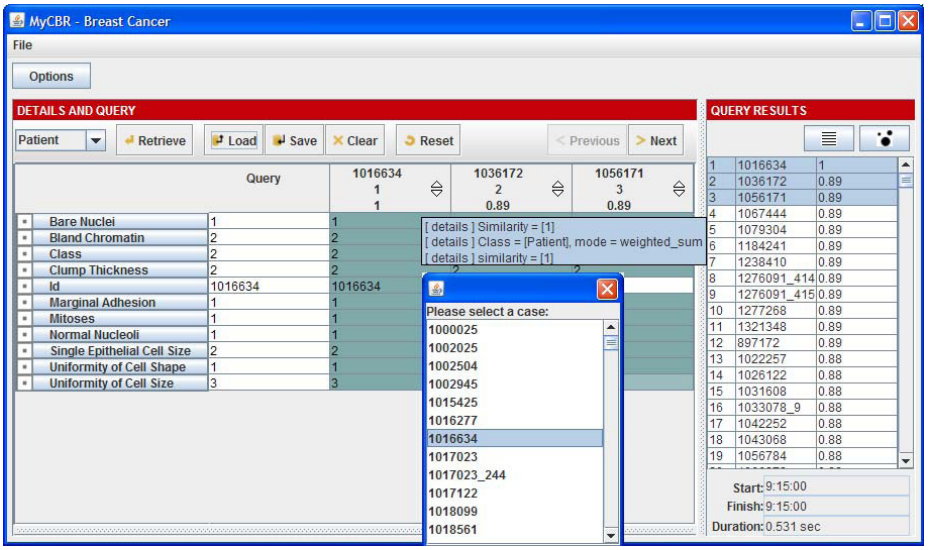


Fig. 6. Breast Cancer as a Stand-Alone Application

5 Genetic Algorithms Approach for Data Mining Breast Cancer Classification

5.1 Genetic Algorithms Approach

Genetic Algorithms (GA) provide an approach to learning that based loosely on simulated evolution. The GA methodology hinges on a population of potential solutions, and as such exploits the mechanisms of natural selection well known in evolution. Rather than searching from general to specific hypothesis or from simple to complex GA generates successive hypotheses by repeatedly mutating and recombining parts of the best currently known hypotheses. The GA algorithm operates by iteratively updating a poll of hypotheses (population). One each iteration, old members of the population are evaluated according a fitness function. A new generation is then generated by probabilistically selecting the fittest individuals form the current population. Some of these selected individuals are carried forward into the next generation population others are used as the bases for creating new off springs individuals by applying genetic operations such as crossover and mutation.

This section presents an application of genetic algorithms approach for data mining classification task [18,19]. Actually two C4.5 based classifiers were developed, the first classifier; RFC4.5 uses the RainForest framework approach while the second; GARFC4.5 is a hybrid classifier uses Genetic Algorithm. The role of C4.5 classifier is to construct a simple decision tree. The role of RainForest is to keep the scalability aspects in constructing the classifier. The role of Genetic algorithms is working as online or dynamic training. In what follow a brief description of both classifiers is given [17].

5.2 Breast Cancer Database

The Breast cancer database contains special laboratory examinations for breast cancer diseases. The problem is to predict whether a tissue sample taken from a patient's breast is malignant or benign. The original data tables have been imported and converted into access format. The database includes a number of instances (about 699 patients). The 16 instances with missing attribute values are removed from the database, leaving 683 instances. The main goal was to discover some sensitive and specific patterns as well as achieving a high prediction rate for breast cancer disease. The data record has 11 numerical attributes, attributes namely Clump Thickness, Uniformity of Cell Size, Uniformity of Cell Shape, Marginal Adhesion, Single Epithelial Cell Size, Bare Nuclei, Bland Chromatin, Normal Nucleoli, Mitoses and Class. These attributes are used to represent patient instances. The target attribute is the class attribute that has two possible values: benign or malignant. Class distribution: Benign: 458 (65.5%) and Malignant: 241 (34.5%).

5.3 Experimental Results of GAFRC4.5 Classifier

The GARFC4.5 hybrid classifier has two main components. The first one, RFC4.5 classifier, combines the previous described advantage of RainForest with the traditional C4.5 algorithm. RFC4.5 takes a set of training records from the database and produces a decision tree according to the randomly selected set of records. The second component, GA, performs the following operations: (a) taking, as input, the generated trees by RFC4.5, and (b) performs the GA operations. Crossover takes two generated trees from two different training data sets by RFC4.5, exchange two or more sub-tree, and produce another two different trees with different fitness function value (classification accuracy). Mutation operation takes one generated tree by first component RFC4.5, exchange two or more sub-tree of the same given tree or convert one or more subtree to a leaf then produce a new tree with different fitness function value. The GA operation proceed until the best tree is obtained from all the decision trees generated from randomly selected set of training data by RFC4.5 classifier component [20].

In our GARFC4.5 experiments, the probability of crossover is set to 0.9 and mutation rate to 0.02 after trying some other values and discovering that the previous values are the best (as known this value are problem specific). Twenty generation as the stopping criterion as noticed that the fitness of the individuals doesn't improve

anymore. The sampling percentages vary from 5% to 40% of the original dataset and the population size from 10 to 50. The database was divided randomly into 70% for training and 30% for testing.

Figure 7 shows the computational results for our genetic algorithms based classifier from different combination of population size and the sampling percentages comparative to RFC4.5 classifier (classifier without the genetic algorithms component). The result curves at figure 7 show that the genetic algorithms based classifier vibrates between high and low classification accuracy and this is due to the numerical nature of the breast cancer database attributes. The results show that GARFC4.5 was not good with numerical nature databases. The results showed that genetic algorithms approach improves the classification accuracy over the traditional decision tree approach. The classification rates are 94% and 81% for our genetic classifier and decision tree algorithm respectively. Moreover the genetic classifier has the privilege of online learning over traditional decision tree C4.5 algorithm.

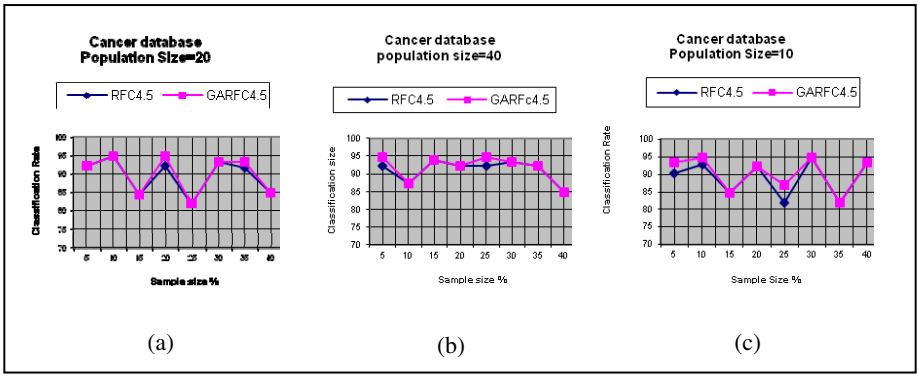


Fig. 7. Computational results of the classifiers GARFC4.5 and RFC4.5 for different population size

6 Conclusions

1. Machine Learning algorithms offer intelligent computational methods for accumulating, changing and updating knowledge in intelligent systems, and in particular learning mechanisms that will help us to induce knowledge from information or data.
2. CBR processes and techniques are very useful to develop CBR-based applications in medical domains but extensive effort is required to enhance their learning curve, usability and understandability.
3. Ontology-based CBR frameworks are very useful to develop robust CBR-based breast cancer classifiers that can play a very important role to help for early detecting the disease and hence right medications can be used to save lives
4. Ontological engineering approach is an effective methodology to manage and represent cancer knowledge.

5. Cancer ontologies are very useful in medical knowledge-based systems to facilitate knowledge sharing, refine, search and reuse.
6. The application of genetic algorithms approach improves the cancer classification accuracy over the traditional decision tree approach.

References

1. Cortes, C., Vapnik, V.: Support Vector Networks. Machine Learning 20, 273–297 (1995)
2. Quinlan, J.R.: C4.5: Programming for Machine Learning. Morgan Kaufman Publishers (1993)
3. Goldberg, D.E.: Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Welsey Publishing Company, Reading (1989)
4. Kolonder, J.: Case-Based Reasoning. Morgan Kaufmann (1993)
5. Abdrabou, E.A., Salem, M.A.B.: Case-Based Reasoning Tools from Shells to Object-Oriented Frameworks. Advanced Studies in Software and Knowledge Engineering-Supplement to the Intern. J. "Information Technologies & Knowledge", 37–44 (2008); Markov, K., Ivanova, K., Mitov, I. (eds.). Institute of Information Theories and Applications FOI ITHEA, Sofia
6. Abdrabou, E.A.M., Salem, A.B.: eZ-CBR: A Generic Tool for Case-Based Reasoning. In: Proc. of 2nd International Conf. on e-Health and Bioengineering - EHB 2009, Iaşi-Constanţa, Romania (2009)
7. Salem, A.B.M., Roushdy, M., El-Bagoury, B.M.: An Expert System for Diagnosis of Cancer Diseases. In: Proc. of the 7th Intern. Conf. on Soft Computing, MENDEL, pp. 300–305 (2001)
8. Salem, A.-B.M., Nagaty, K.A., El-Bagoury, B.M.: A Hybrid Case-Based Adaptation Model For Thyroid Cancer Diagnosis. In: Proc. of 5th Int. Conf. on Enterprise Information Systems, ICEIS 2003, Angres, France, pp. 58–65 (2003)
9. Fernández-López, M., Gómez-Pérez, A.: Deliverable 1.4: A survey on methodologies for developing, maintaining, evaluating and reengineering ontologies. Part of a Research Project Funded by the IST Programme of the Commission of the European Communities as project number IST-2000-29243 (2002)
10. Salem, M.A.B., Alfonse, M.: Ontology versus Semantic Networks for Medical Knowledge Representation. In: Proceedings of 12th WSEAS CSCC Multiconference (Computers), Heraklion, Crete Island, Greece, pp. 769–774 (2008)
11. Salem, M.A.B., Alfonse, M.: Building Web-Based Lung Cancer Ontology. In: 1st National Symposium on e-Health and Bioengineering, EHB 2007, Iaşi, Romania, pp. 177–182 (2007)
12. Salem, M.A.B., Alfonse, M.: Ontological Engineering Approach for Breast Cancer Knowledge Management. In: Proc. of Med-e-Tel the International eHealth, Telemedicine and Health ICT for Education, Networking and Business, Luxembourg, pp. 320–324 (2009)
13. Abdrabou, E.A., Salem, M.A.B.: Application of Case-Based Reasoning Frameworks in a Medical Classifier. In: Proc. of Fourth International Conference on Intelligent Computing and Information Systems (ICICIS-4), Cairo, EG, pp. 253–259 (2009)
14. Abdrabou, E.A., Salem, M.A.B.: A Breast Cancer Classifier based on a Combination of Case-Based Reasoning and Ontology Approach. In: Proc. of 2nd International Multi-Conference on Computer Science and Information Technology, IMCSIT 2010, Wisła, Poland (2010)

15. Sewak, M., Vaidya, P., Chan, C.C., Duan, Z.H.: SVM Approach to Breast Cancer Classification. IMSCCS 2, 32–37 (2007)
16. Recio-García, J.A., Díaz-Agudo, B., González-Calero, P.A.: jCOLIBRI2 Tutorial, Group of Artificial Intelligence Application (GAIA). University Complutense of Madrid. Document Version 1.2 (2008)
17. Stahl, A., Roth-Berghofer, T.R.: Rapid Prototyping of CBR Applications with the Open Source Tool myCBR. In: Althoff, K.-D., Bergmann, R., Minor, M., Hanft, A. (eds.) ECCBR 2008. LNCS (LNAI), vol. 5239, pp. 615–629. Springer, Heidelberg (2008)
18. Cios, K.J., Pedrycz, W., Swiniarski, R.W.: Data Mining Methods for Knowledge Discovery. Kluwer (1998)
19. Jain, A.K., Murty, M.N., Flynn, P.J.: Data Clustering: A Review. ACM Computing Surveys 31(3), 264–323 (1999)
20. Salem, M.A.B., Abeer, M.M.: A Hybrid Genetic Algorithm-Decision Tree Classifier. In: Proc. of the 3rd Intern. Conf. on New Trends in Intelligent Information Processing and Web Mining, Zakopane, Poland, pp. 221–232 (2003)

Methods for Interpretation of Data in Medical Informatics

Boris Mirkin

Department of Data Analysis and Machine Intelligence, National Research University Higher School of Economics, 11 Pokrowski Boulevard, 109028, Moscow RF
Department of Computer Science, Birkbeck University of London, Malet Street, WC1E 7HX, London UK
mirkin@dcs.bbk.ac.uk

Abstract. An outline of a few methods in an emerging field of data analysis, "data interpretation", is given as pertaining to medical informatics and being parts of a general interpretation issue. Specifically, the following subjects are covered: measuring correlation between categories, conceptual clustering, and generalization and interpretation of empirically derived concepts in taxonomies. It will be shown that all of these can be put as parts of the same inquiry.

Keywords: data analysis, association between categories, clustering, hierarchical ontology, taxonomy, computational interpretation.

1 Introduction

In spite of the fact that medical informatics is one of the fastest growing areas both in research and in practice, as of this moment, there is no well developed system for the medical informatics domain. However, a number of focus areas are of interest to medical informatics:

- patient-centered systems: medical records and images;
- patient safety: error prevention and handling;
- clinical research informatics including new drugs and treatment methods;
- healthcare organization and administration;
- knowledge organization, updating and use.

So far most efforts and results have been related to the personal health support systems. However, each of the subjects is important in the health related efforts and can benefit significantly of informatics tools. Moreover, one cannot help but see the medical informatics as a pathfinder, a leader, in such computer-intensive areas of current interest as knowledge organization, updating, and use (see SNOMED CT ontologies development [12] and related efforts).

Currently, the issues of organization and maintenance of e-records are of urgent priority in medical informatics. Possibly, even more urgent are matters of reorganization of health services such as developing classifications of diseases and disorders

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 15–22.
DOI: 10.1007/978-3-319-00029-9_2

matching the common treatment practices. Yet there is a permanent need in automating of all aspects of data interpretation, which will become much apparent after the organizational issues have been addressed.

These are the subject of this presentation. Data of a set of patients may comprise tables, texts, and images. This paper refers mostly to the tabular data format corresponding to results of various tests over a set of patients, and, further down, to the author's attempts at developing methods for data interpretation. The current level of digitalization leads to growing popularity of exploratory data analysis and data mining approaches oriented towards finding patterns in data rather than testing hypotheses; the latter are prevailing in classical statistics frameworks still dominating many areas of the medical informatics discourse. Yet finding patterns is just an intermediate goal, the real challenge lies in developing data analysis methods in such a way that the result can be formulated in a way that a medical practitioner may find acceptable, understandable and reasonable. This is the niche I like to focus at.

I am going to present here a few data analysis methods oriented towards data interpretation issues:

(a) measuring correlation between categories,
(b) conceptual clustering and
(c) generalization and interpretation of empirically derived concepts in taxonomies.

It will be apparent in the end that these three are not as diverse as they seem to be. In fact, they are parts of the same inquiry.

2 Exploring Correlation between Categories: Interpretation Versus Statistics

This subject is of finding those sets of categories that are most correlated with each other. A simplest would be finding just pairs of correlated categories. First of all, I'd like to bring in an example showing the difference between the mathematical statistics and data interpretation approaches. A case for mathematical statistics: a lung cancer sufferer making a claim against an industrial company plant that they are responsible for the condition. To support their claim, the lung cancer sufferer's team refer to a statistical table (in the left part of Table 1).This table brings forward statistical testing of the plant company claim that the proportions of the condition near the plant and faraway from it, 0.05 and 0.03 on the sample, differ only because of the sampling bias and are, in fact, equal in the population. A two-sided z-test, like that in [4], would show that, under the assumption that the sampling has been random and independent, the hypothesis that the proportions are equal should be rejected, at 95% confidence level. Data analysis relates to a very different data and problem setting. The data comes from a database which has been collected from various sources, not necessarily independent or similar. There are many features in the database of which those mentioned in Table 1 could be just a couple. Moreover, the data may be much less balanced than in a goal-oriented sample. This is the case of data on the right in Table 1: only 60 cases from near the plant are in the dataset while the number of far-away-from-plant cases remains a thousand. Because of such a bias in the sample, the very

same z-test now decidedly supports the idea that the hypothesis that the proportions of the condition on the sample are the same cannot be rejected anymore, even in spite of the fact that they remain very much the same: 0.05 and 0.03. This is because the near-plant dwellers sample size is greatly reduced here so that the uncertainty of the situation increases.

Table 1. An illustrative example of contingency data in health statistics: (a) testing proportions, on the left side; (b) as happens in a data base, on the right side

	Classical statistics case			Data interpretation case		
Residence	No LC	LC	Total	No LC	LC	Total
Near plant	950	50	1000	57	3	60
Far from plant	970	30	1000	970	30	1000
Total	1920	80	2000	1027	33	1060

In contrast, the data interpretation view pays no attention to the classical mathematical statistics cause (except sometimes for the lip service only). The goal here is to capture the extent of correlation on the sample, rather than to see how the sample differs from the population – the latter is of no concern at all. The conditional probabilities, like those mentioned, 0.05 and 0.03 could be a good choice sometimes. Yet they can be used only in the case at which one subsample is compared to the other. A more universal measure has been proposed by the founding father of statistics A. Quetelet almost 200 years ago. Quetelet index compares the conditional probability of the event l at a given category k, with that on the entire set, not at a different subsample [8,9]:

$$q(l/k) = \frac{P(l/k) - P(l)}{P(l)} \quad (1)$$

That is, Quetelet index expresses correlation between categories k and l as the relative change in the probability of l when k is taken into account. In our case, q(LC/Near_Plant) = 3*1060/ (33*60) -1 = 0.606. That means that living near the plant increases the chances of acquiring LC by 60.6% - this should be taken into account in the court whatever considerations of the statistical significance are! (As one can notice, both parts of Table 1 are subject to this interpretation, not just the part on the left!)

It appears, the average Quetelet index, that is, the sum of $q(l/k)$ weighted by their probabilities, $P(k,l)$, coincides with the value of the well-known Pearson's chi-square coefficient which is widely used for assessing statistical independence, but not association, between categorical features (e.g., Daniel 1998 [1]). This sheds a different light over the Pearson's coefficient – that is an association measure, after all – and this is exactly the criterion for deriving a decision classification tree in some packages such as SPSS. A similar meaning can be assigned to other popular association measures such as Gini index.

3 Hierarchical Grouping: Conceptual Clustering

Hierarchical grouping with conceptual clustering was developed in 80es and recently has enjoyed some revival due to the emergence of ontologies and other conceptual structures (see, for example, Fanizzi et al. 2009 [2]). Our experience is based on the original developments in Russia in 80es for the analysis of data of large-scale sociological and health related surveys [7].

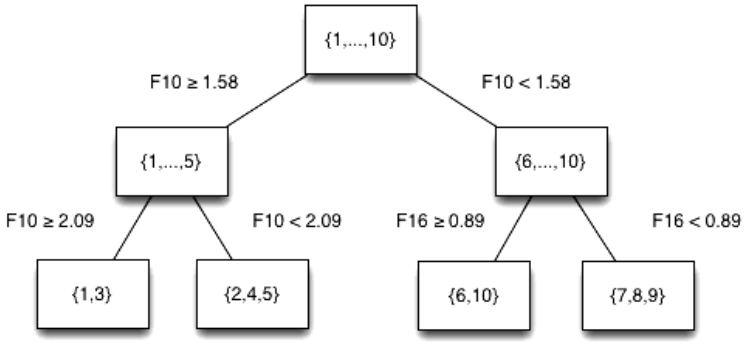


Fig. 1. An illustrative example of a conceptual grouping hierarchy

The result of application of a hierarchical grouping algorithm can be represented by a hierarchy resembling that of a decision tree [6]. Yet it is built automatically by sequential divisions of clusters, starting from the entire dataset, over features from a specified subset according to a criterion that is much similar to those used in clustering. Yet, in contrast to the classical cluster analysis, the clusters are split not over a multidimensional distance between entities but rather over just one of the features. If the feature x is quantitative, then the two split parts correspond to predicates "x>a" and "x≤a" for some feature value a. For a categorical feature, the split parts correspond to predicates "x=a" and "x≠a" for a category a. The algorithm tests all the candidate clusters and all the candidate features and chooses the split maximizing the summary association of that with all the features or, equivalently, the Ward's distance between the split parts' centroids [Mirkin 2011]. The obtained conceptual tree is much intuitive and, also, serves as an informative features selector (those actually used in the splits). The association of the hierarchic partition with the features is measured with the so-called correlation ratio, for quantitative features, or the Pearson chi square association coefficient, for categorical features [9]. The latter is to be modified to Gini coefficient depending on the data normalization, to keep the mathematical equivalence of the criterion to the so-called quadratic error criterion of k-means and similar clustering approaches. This is based on representation of the categories by the corresponding dummy variables with a follow-up standardization of them [8, 9].

In a large-scale survey conducted at Novosibirsk area (Russia) in early 80-es with regard to pneumonia, tuberculosis and other respiratory diseases, more than a dozen altogether, P. Rostovtsev and I built a hierarchical classification of the sample of more than 50000 individuals over the respiratory diseases and related features to find a final

conceptual clustering respiratory disease partition of about 20 clusters/disease types [11]. This partition was further used to find those features of the individuals' conditions that have been most correlated with it. The medical researchers were thinking of alcohol consumption and smoking as the two most important risk factors. In fact, the found partition of the individuals over respiratory conditions had no correlation with these whatsoever, which was very unfortunate because our findings could not be published at that time as being at odds with the dominating paradigm. Instead, we found two other features: "bad housing" and "the same disease in the family", as the real risk factors. The Quetelet coefficient for the former was about 600%.

4 Interpretation of Clusters over a Hierarchical Ontology of the Domain

Hierarchical ontologies, or taxonomies, are currently becoming a major format for computationally handling, maintaining and updating knowledge. A very recent international effort is being resulted in a set of hierarchical ontologies for the medicine SNOMED CT [12]. In fact, this is the very first example of the concept of ontology being developed as a device for practical purposes.

A hierarchical ontology is a set of concepts related by a tree-like hierarchical relation such as "A is a B" or "A is part of B". Of course, ontology of a domain may contain a rather small number of the domain concepts while many others, especially those new ones, remain out of the tree. The concept of ontology is much relevant to the medicine domains because it can encompass the mechanism of a disease and related disorders. The medical diagnostics process can frequently be put in terms of a decision tree related to a hierarchical ontology. The Manual [6] is an example of such an approach applied in the mental health domain.

Therefore, a problem of interpretation of concepts -"outsiders" in terms of the "insider" concepts emerges. Take, for instance, the International Association for Computing Machinery (ACM) classification of computing subjects – a hierarchical four-layer taxonomy of the computing world ACM-CCS. I realize that this may be considered as somewhat far from the medicine, but at this moment I have no application of the approach to be presented in the medical domain.

A recently emerged concept P, say "intuitionist programming", does not belong to the current ACM_CCS. To interpret that in terms of ACM_CCS take a look through a search engine like Yahoo! (because Yahoo was so much research-friendly) to find a profile of P.

Fuzzy profile of P (illustrative):

F.1 Computation by abstract devices - 0.60
F.3 Logics and meaning of programs - 0.60
F.4 Mathematical logic and formal languages - 0.50
D.1 Programming languages - 0.17.

(A fuzzy set, unconventionally normed in Euclidean metric so that the squares of the membership values sum to unity, because of another development by the author and S. Nascimento [10].)

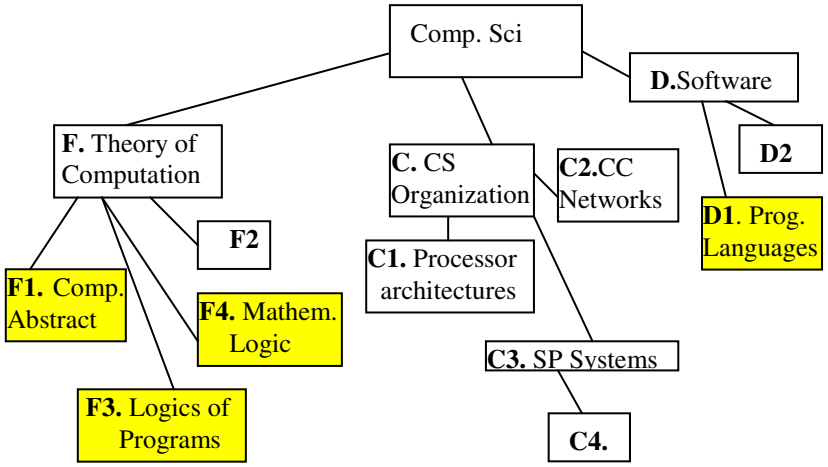


Fig. 2. A fragment of ACM-CCS taxonomy, along with the contents of a fuzzy topic set mapped to it (gray)

Mapped to ACM-CCS taxonomy as is (see Fig. 2), the contents of the fuzzy profile can be looked at through the taxonomy structure. Yet, when the contents counts a dozen or more topics well dispersed through the taxonomy tree, the mapping has some obvious drawbacks as being: (a) fragmentary, (b) not scalable, and (c) not quite cognition-friendly.

This is why we propose to interpret such a profile (fuzzy topic set) by lifting it to higher ranks of the hierarchy to minimize the number of subjects it embraces (Fig. 3). However, the lifting may make apparent some discrepancies, namely, gaps and offshoots. Therefore, the lifting penalty function should involve three types of elements: the "head subjects", the "gaps" and the "offshoots" so that their total, appropriately weighted, should be minimized at the interpretable result.

An algorithm, PARL, has been developed for optimally lifting a fuzzy topic set over a hierarchical ontology by recursively moving from the leaves to the root [10]. At each tree node, the algorithm specifies parsimonious events according to each of the two different scenarios: (a) the head subject has been inherited from the node's parent; (b) the head subject has not been inherited from the node's parent. The parsimony criterion is but an operational expression of the celebrated Occam's Razor principle of simplicity. To make the choice of the weights of different elements of the optimal scenario meaningful in a substantive domain, as many as possible concepts should be interpreted via the lifting process so that the probabilities of "gain" and "loss" of head subjects could be derived for the nodes. Then the "maximum parsimony" criterion can be changed for a "maximum likelihood" criterion at which the weights are defined by the maximum likelihood principle.

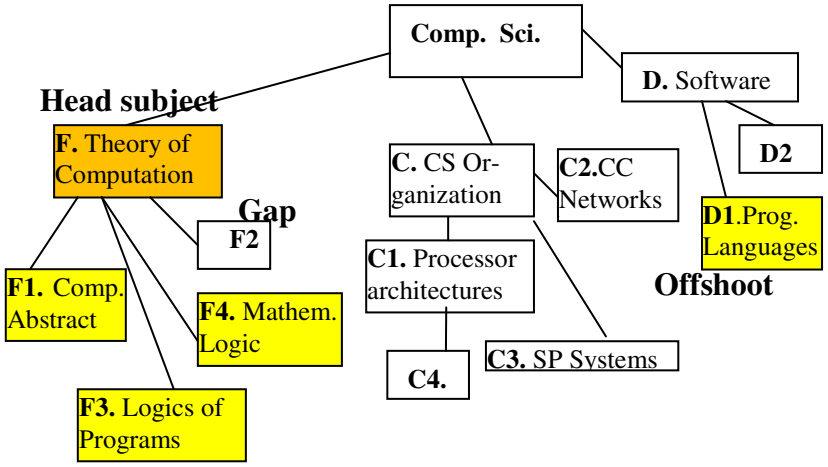


Fig. 3. Interpretation of the topic set by lifting it to "Head subject" F. Theory of computation (highlighted by a darker filling), with the price of having a "gap", F2, and an "offshoot", D1.

5 Conclusion

The discipline of computationally handling both data and knowledge is emerging as driven, to a large extent, by the medical informatics needs. The models and methods for interpretation of various patterns and facts will be an integral part to it. A few topics I just outlined are related quite closely: a topic set to be interpreted by lifting in a hierarchical ontology (section 4) can be derived with a conceptual clustering approach (section 3) which itself heavily relies on the ways for scoring category-to-category correlation (section 2).

References

1. Daniel, L.G.: Statistical Significance Testing: A Historical Overview of Misuse and Misinterpretation with Implications for the Editorial Policies of Educational Journals. Research in the Schools 5(2), 23–32 (1998)
2. Fanizzi, N., d'Amato, C., Esposito, F.: Metric-based stochastic conceptual clustering for ontologies. Information Systems 34(8), 792–806 (2009)
3. García, M.M., Allones, J.L.I., Hernández, D.M., Taboada Iglesias, M.J.: Semantic similarity-based alignment between clinical archetypes and SNOMED CT: An application to observations. International Journal of Medical Informatics (Available online March 13 (2012)
4. Joosse, S.A.: Two-proportion Z-test calculator (2011), `http://in-silico.net/statistics/ztest`
5. Ludwick, D.A., Doucette, J.: Adopting electronic medical records in primary care: Lessons learned from health information systems implementation experience in seven countries. International Journal of Medical Informatics 78(1), 22–31 (2009)
6. Manual of Mental Disorders (DSM-IV-TR), American Psychiatric Association (2000)

7. Mirkin, B.: Grouping in Socio-Economic Research. Finansy I Statistika Publishers, Moscow (1985) (in Russian)
8. Mirkin, B.: Eleven ways to look at the chi-squared coefficient for contingency tables. The American Statistician 55(2), 111–120 (2001)
9. Mirkin, B.: Core Concepts in Data Analysis: Summarization, Correlation, Visualization. Springer, London (2011)
10. Mirkin, B., Nascimento, S., Fenner, T., Felizardo, R.: How to Visualize a Crisp or Fuzzy Topic Set over a Taxonomy. In: Kuznetsov, S.O., Mandal, D.P., Kundu, M.K., Pal, S.K. (eds.) PReMI 2011. LNCS, vol. 6744, pp. 3–12. Springer, Heidelberg (2011)
11. Rostovtsev, P.S., Mirkin, B.G.: Hierarchical grouping in socio-economic research. In: Mirkin [7], Section 5.4, pp. 126–133 (1985)
12. SNOMED CT (Systematized Nomenclature of Medicine-Clinical Terms) (2012), `http://www.nlm.nih.gov/research/umls/Snomed/snomed_main.html` (visited May 27, 2012)

Visual System of Sign Alphabet Learning for Poorly-Hearing Children

Margarita Favorskaya

Siberian State Aerospace University, 31 Krasnoyarsky Rabochy, Krasnoyarsk, 660014 Russia
favorskaya@sibsau.ru

Abstract. Training visual systems have significant role for people with limited physical abilities. In this paper, the task of sign alphabet learning by poorly-hearing children was discussed using advanced recognition methods. Such intelligent system is an additional instrument for cultural development of children who can not learn alphabet in the usual way. The novelty of the method consists in proposed technique of features extraction and building vector models of outer contours for following identification of gestures which are associated with letters. The high variability of gestures in 3D space causes ambiguous segmentation, which makes the visual normalization necessary. The corresponding software has two modes: a learning mode (building of etalon models) and a testing mode (recognition of a current gesture). The Visual system of Russian sign alphabet learning is a real-time application and does not need high computer resources.

Keywords: Sign alphabet, gesture recognition, features extraction, spatio-temporal segmentation, skin classifiers.

1 Introduction

Human gestures are the universal instrument for non-verbal interaction, communication, and information exchange among people. At the latest years we see another type of communication – the so called interactive communication between man and computer. In this aspect, we need special devices, automated algorithms and procedures, and perhaps, another performance of interactive information [1-3]. In this paper, the main discussion will be related to visual system having a single video camera which can process and recognize hand gestures (signs) associating with letters of national alphabet. The limited number of letters in the alphabet determines the limited set of signs that simplifies the assigned task. In this case, the task of recognition becomes the task of identification with a pre-determined set of typical 3D objects.

Dactylology is a specific form of speech or communication that reproduces a spelling form of a word by fingers by using a sign alphabet. Dactylology includes the features of oral speech for operative communication and the specialties of writing speech having a form of sequential symbols according to spelling rules. Sometimes hearing people consider a sign alphabet as a symbol alphabet of deaf-mute people.

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 23–39.
DOI: 10.1007/978-3-319-00029-9_3

However, there are two different things: sounds of speech (letter by letter) are propagated by a sign alphabet and gesture symbols perform the words. At the initial stage of computer processing there is no difference between performances of sounds of speech and gesture symbols. The great difference will be on the last stage – a stage of a result interpretation.

An ideal gesture recognition system must be robust to handle and palm variations over a wide range. In common case, we have three main issues: (1) the start and the end of gesture definition in temporal domain (in sequence of continuous gestures); (2) 2D segmentation of 3D palms localization in spatial domain (overlapping images of fingers, warping transformations in 3D space); (3) feature extraction and recognition. A gesture that any human can quickly recognize may be the complex problem for machine-vision application. On the basic of existing intelligent technologies, some constraints and simplifications were introduced that permitted to solve the task and to design the corresponding software. Classification of a limited set of gestures permits to find a set of essential geometrical features according to which we may simultaneously identify most gestures; similar gestures are classified according to their vector models in selected frames.

The rest of the paper is organized as follows: related work will be discussed in Section 2; Section 3 describes the proposed classification of gesture associating with the letters of alphabet; in Section 4 the algorithms of segmentation, vectorization and identification will be discussed; Section 5 presents the experimental software "Russian Sign Language" and some experimental results, and the conclusions are given in Section 6.

2 Related Work

All methods which have been proposed for gesture segmentation and recognition in the literature can be classified into two major categories: separate segmentation/recognition [4, 5] and simultaneous segmentation/recognition [6, 7]. The methods from the first category detect the gesture boundaries for segmentation and extract various features for recognition, so the segmentation procedure precedes the recognition procedure. The methods from the second category execute segmentation and recognition procedures simultaneously. Most methods from both categories are based on various modifications of Hidden Markov Model (HMM), a Dynamic Programming technique (DP), Dynamic Time Warping (DTW), Continuous Dynamic Programming (CDP), Dynamic Bayesian Network (DBN), Artificial Neuron Network (ANN), etc.

2.1 Overview of Gesture Segmentation and Recognition Methods

Let's discuss briefly the basic methods of visual gesture segmentation and recognition using video sequences. Kim et al. [6] proposed the simultaneous gestures segmentation and recognition based on forward spotting accumulative HMMs. Firstly the start

and end points of gesture are determined as zero-crossed points of proposed probability function; secondly they are specified by using local maximums of gesture and non-gesture. The output scores are accumulated, and a current gesture is associated with the model which has a majority vote of all intermediate recognition results.

The HMM-based approach for dynamic hand gestures in video sequences was proposed in [9]. The output scores of every HMM are continuously observed at every time step using an improved normalized Viterbi algorithm which increased the output score of an HMM: peaks in the output scores of the respective models indicated the presence of gestures. Lee et al. [10] proposed a HMM-based threshold model for end points of gesture recognition using matched gesture patterns. The likelihood of all Gesture Models and the threshold models is calculated for candidate end points consideration. The start points of a gesture can be traced back using the Viterbi algorithm.

The DP technique combines motion detection and an explicit multi-scale search to find the start and end times of a gesture [8]. Kang et al. [11] proposed to detect possible start and end points of a gesture according to three criteria: abnormal velocity, a static gesture, and severe curvature. Then the segments between those candidate cuts were evaluated using DTW. Alon et al. [2] proposed an unified framework for spatio-temporal gesture segmentation and recognition. Multiple candidate hand locations are firstly detected; then by using CDP the feature vectors are matched to the model. During this stage a large number of hand gesture hypotheses are eliminated according to classifiers received from a training data set. The end point of a gesture is detected when the model gives the lowest matching cost in the interval.

In [12], Suk et al. proposed DBN model for hand gesture recognition that can be used to control media players or slides presentation. The suggested DBN model is preceded by steps of skin extraction, modeling, and motion tracking. Such gesture model is developed for one-hand or two-hands gestures and is based on a cyclic gesture network for modeling continuous gesture stream. Also authors had developed a DP-based real-time decoding algorithm for continuous gesture recognition. Some authors researched architecture on neural networks with different numbers of hidden layer neurons and different numbers of layers and built ANN [13]. They found the relationship between the interior parameters and the performance of ANNs having various numbers of hidden layers, neurons, model training parameters, and epoch size affects. If a large learning rate and small momentum of the model differ are achieved, the result is more precise.

2.2 Overview of Sign Languages Methods

Some recognition systems for national sign language are already well-known. In [14], Liang et al. proposed a gesture recognition system for Taiwanese sign language by using hand movements. Four parameters are used: posture, position, orientation and motion. A 3D tracker was also used to find the hand and fingers position. The end point of a gesture is detected by a time-varying parameter. When

the value of a time-varying parameter drops below a threshold, the motion becomes quasi-stationary, and the corresponding frame is considered as end point. Then HMM is used for the recognition of Taiwanese Sign Language at the syntax level.

In [5], a 48-dimensional vector is used to describe a Chinese sign language, including 36 hand shapes, six positions, and six orientations. To obtain these data, two cyber-gloves are used to collect the variation information of hand shape and fingers position. Three 3D-position trackers, where two of them are situated on the wrist of each hand and the third tracker is mounted at the signer's back, are used to collect the information of hands orientation and position. Segmentation and recognition of large-vocabulary continuous sign language are based on transition-movement models (TMMs). The transition movements between two adjacent signs are temporary clustered using the k-means algorithm. The iterative TMMs training algorithm is used for automatic segmentation of signs in a video sequence. The training algorithm of continuous sign language recognition is enhanced by new TMMs of sign models using the Viterbi algorithm.

Hand motion in German sign language has been described by hand shape, orientation and location in work [15]. To obtain these data, the signer is required to wear colored gloves on each hand, with dominant and non-dominant hands marked by different colors. While the glove for the non-dominant hand is one uniform color, the glove for the dominant hand has seven different colors, marking each finger, the palm and the back of the hand.

Bobick et al. [16] proposed to use a set of motion-energy image (MEI) (binary cumulative motion images) and motion-history image (MHI) (scalar-valued images, where more recently moving pixels are brighter) to represent a human movement for each view/movement combination. Seven Hu moments are calculated to describe the shapes of MEI and MHI under affine transitions (translation and scale). In a real-time American Sign Language (ASL) recognition system, hand blobs had been extracted and tracked based on skin color instead of a hand shape detailed description [17]. A sixteen-element feature vector contains information about following parameters: hand's position, change in position between frames, area, angle of axis of least inertia (founded from the first eigenvector of the blob), length of this eigenvector, and eccentricity of bounding ellipse. The effectiveness of the hand gesture descriptor provides high recognition accuracy of ASL at the sentence level.

In work [18], skin color blobs are extracted from each frame and then a relational distribution of edge pixels in the skin blobs is obtained for automatic signs extraction from continuous sign language. The sign is represented as a trajectory in a low-dimensional space called Space of Relation Distribution (SoRD), which implicitly captures the shape and motion of the sign.

Various national sign alphabets have different signs associating with the same letter. Examples can be seen on Fig. 1, where three letters A, F, and T are shown according to four classifications: American, German, International, and Spanish sign alphabets.

Fig. 1. Examples of some signs: 1st column – American, 2nd column – German, 3rd column – International, and 4th column – Spanish sign alphabets

3 Signs Classification

Dactylology statements influence significantly the segmentation and recognition processes. The rules of signs demonstration are identical in many sign alphabets. Russian sign alphabet underlies in the basic of this research. In Section 3.1, dactylology statements will be represented. In Section 3.2, the process of sign features extraction for Russian sign alphabet will be commented.

3.1 Dactylology Statements

The gesture communication of poorly-hearing people has a complex structure and comprises speaking, tracing (gestures and signs), and combined gesture speech. In our case, motions of hands are a special kinetic system in which gestures indicate the letters of national alphabets but not events of visual environment. We may call such speech as a speech by a sign notation of letters. Using a set of signs, a signer follows to a verbal language grammar (English, Russian, etc.). So motions of hands are a singular kinetic form of the verbal communication. The functions of hands motions are quite a few: from communication between hearing (teachers, parents) and non-hearing people to interpersonal communication of non-hearing people by a gesture speech.

The designers of the first national sign alphabets chose such positions of fingers that makes signs looking like letters. The main criterion was the similarity paradigm. Later, one can find other criteria suggested by I. Geilman in 1981: a configuration (one-hand, two-hands, and combined), a forming mode (tracing and variant), and a notation (alphabetic, syllabic, and integrated). According to this classification, Russian, Spain and some other sign alphabets are one-handed, tracing and alphabetic; English sign alphabet is a two-hands, tracing and alphabetic (apart from the letter "C"). Chinese sign alphabet is determined as combined, variant, and integrated, and is presented as two sets of signs: the first set is for the initial syllables, and the second set is for the last syllables. The elements of sign alphabet are similar to letters of written language, and they distantly recall the letters of type font.

As writing, reading and spelling, this process requires a verbal speech observance and is executed by following rules:

- Writing sign process is realized according to spelling rules.
- Sign speech is mandatory tracing by articulation (sentences spelling).
- Signs are shown exactly and clearly.
- The learning process is a smooth and unified process. Sharp and break learning process is a rough mistake.
- Sign words are separated by pause; phrases are separated by stopping.
- The hand is not close to face.
- The learning process is realized by the right hand.
- During the learning process the hand is bended, situated forward on hand level and tended by palm to a signer.
- The hand is shifted left to avoid of mirror writing.
- In the case of mistake or non-understanding, all words are multiply repeated.
- The signer is seeing on his pupil.

During communication, the signer shows the letters as signs and a pupil follows visually the hand motion. If a pupil can not see vision symbols, than he senses them by touch (so called sign-contact speech). If during communication with blind deaf-mute man, the character outline is represented on the palm then it will be not dactylology, but dermography.

3.2 Signs Features Extraction

Let's discuss the process of gesture features extraction using the Russian sign alphabet, shown on Fig. 2. For robust recognition of signs from Russian sign alphabet, the advanced technique based on fingers position estimations was used. We analyze such features as fingers position relatively each other, fingers orientation relatively the horizontal axis of the image, the hand skewing position, the presence of inner closed contour (circle or ellipse), the shifting hand motion, and the number of separated fingers. A set of proposed features is determined by possibilities of 2D video sequence received from a single video camera. The determined parameters of signs classification are situated in Table 1.

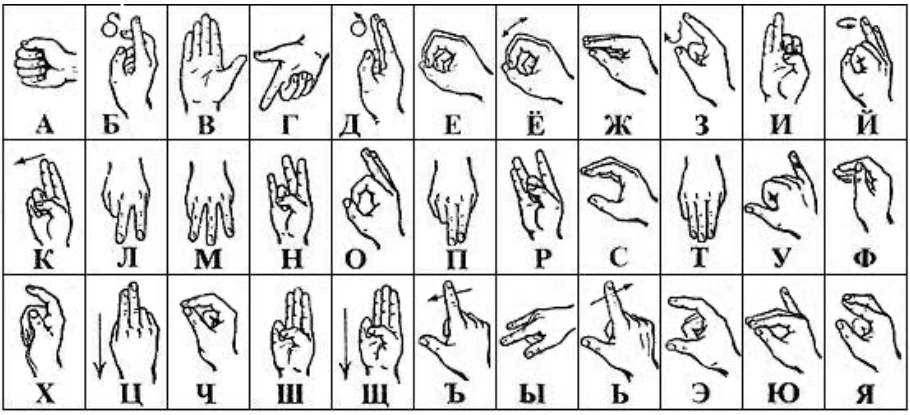


Fig. 2. The Russian sign alphabet (arrows are pointed the directions of motion)

The main features such as fingers position and orientation, position of hand skewing, inner closed contour, shifting hand motion, and number of separated fingers were chosen for building of classification rules. Each feature is coded as normalized parameter into previously determined values interval. As one can see, some letters have

Table 1. Extracted features for signs recognition

Letter	Fingers position	Fingers orientation	Hand skewing position, °	Inner closed contour	Shifting motion	Number of separated fingers
А	closed	left	$\varphi = 0$	–	–	non-defined
Б	closed	up	$\varphi \in [90, 120]$	+	+	non-defined
В	closed	up	$\varphi \in [75, 105]$	–	–	1
Г	open-ended	down	$\varphi \in [210, 240]$	–	–	2
Д	closed	up	$\varphi \in [90, 120]$	–	+	non-defined
Е	closed	left	$\varphi \in [120, 150]$	+	–	non-defined
Ё	closed	left	$\varphi \in [120, 150]$	+	+	non-defined
Ж	closed	left	$\varphi \in [120, 150]$	+	–	non-defined
З	open-ended	up	$\varphi \in [120, 150]$	–	+	1
И	closed	up	$\varphi \in [75, 105]$	–	–	non-defined
Й	closed	up	$\varphi \in [90, 120]$	–	+	non-defined
К	closed	up	$\varphi \in [90, 120]$	–	+	non-defined
Л	open-ended	down	$\varphi \in [255, 285]$	–	–	2
М	open-ended	down	$\varphi \in [255, 285]$	–	–	3
Н	open-ended	up	$\varphi \in [90, 120]$	–	–	3
О	closed	up	$\varphi \in [120, 150]$	+	–	non-defined
П	closed	down	$\varphi \in [255, 285]$	–	–	non-defined
Р	open-ended	up	$\varphi \in [90, 120]$	–	–	1
С	open-ended	left	$\varphi \in [120, 150]$	–	–	1
Т	closed	down	$\varphi \in [255, 285]$	–	–	non-defined
У	closed	up	$\varphi \in [120, 150]$	–	–	2
Ф	closed	left	$\varphi \in [90, 120]$	–	–	non-defined
Х	open-ended	up	$\varphi \in [90, 120]$	–	–	1
Ц	closed	up	$\varphi \in [75, 105]$	–	+	non-defined
Ч	closed	left	$\varphi \in [120, 150]$	+	–	non-defined
Ш	closed	up	$\varphi \in [75, 105]$	–	–	non-defined
Щ	closed	up	$\varphi \in [75, 105]$	–	+	non-defined
Ъ	open-ended	up	$\varphi \in [90, 120]$	–	+	2
Ы	open-ended	left	$\varphi \in [210, 240]$	–	–	2
Ь	open-ended	up	$\varphi \in [90, 120]$	–	+	2
Э	open-ended	left	$\varphi \in [120, 150]$	–	–	2
Ю	open-ended	up	$\varphi \in [120, 150]$	+	–	1
Я	open-ended	up	$\varphi \in [120, 150]$	+	–	1

identical features according to Table 1 (for example, two last letters); that's why we need in following vector models based on outer contours of signs. Because of the continuity of video sequences, the mentioned features would not be extracted without some additional criteria. Let's lay down three criteria for definition of concrete time moment when a gesture is taking place, as follows:

Criterion 1. The hand position ought to be rigid, at least it has to be fixed during the interval t_g. In practice, the minimal duration t_g = 0.5 s. If frequency of frames equals 25 frame/s then it means that the hand position ought to be rigid during 12 sequential frames and more.

Criterion 2. The hand which shows a gesture ought to be in a stationary position. Such requirement is provided by additional constraint on the center of gravity position. This center ought to be in limited threshold radius r_g at least during the interval t_g. Experiments show that value r_g = 30 (pixels) is enough for carrying out the Criterion 2 when a square of hand image is 10–14 % of frame sizes.

Criterion 3. The speed of hand motion ought to have the minimal value. For minimal value definition, let's introduce delay t_l. In experiments this value was equal t_l = 0.3 s.

Based on extracted features of signs and determined main and additional criteria, algorithms of segmentation, vectorization, and identification were designed. The algorithm for sign segmentation uses the spatial and temporal domains: in the spatial domain, methods of skin segmentation are applied, and in the temporal domain, the active contour method is tracking hand shifts.

4 Proposed Algorithms of Segmentation, Vectorization and Identification

The aim of sign alphabet recognition comprises the interpretation of extracted features for semantic matching of a current hand shape from video sequence and a sign shape from a knowledge database. The main methods for hand shape detection are the pattern and model matching. In the first case, the prototype is matched on a pixel level; an enumerative technique is usually used with all corresponding disadvantages connecting with a non-invariance to shifts, rotation and scale, and high computational cost. The second case is more acceptable, but the complexity of the task is determined by possible warping transformations which are not described by any transformation model. That's why we simplified the assigned task and used a representation of complex 3D motion by affine motion model.

The success of hand segmentation algorithm determines the efficiency of following object identification. The hand segmentation algorithm is not a trivial sub-task because of a continuous hand movement in the 3D space and its mapping in the 2D video stream. The complex segmentation algorithms were not used because of the requirement for real-time application. Let's discuss the proposed algorithms for segmentation, vectorization and identification of signs which are represented by a set of sequential frames in a video sequence.

4.1 Spatio-Temporal Segmentation of Gestures

The spatial segmentation algorithm contains following steps: (1) receiving of current frame, (2) detection of skin region, (3) morphological processing, and (4) hand segmentation. The temporal segmentation algorithm executes (1) the tracking of detected hand during sequential frames and (2) the detection of moments (start and end points) when a current sign is shown to a pupil. Let's consider primarily the spatial segmentation of gestures.

Detection of skin region is based on its characteristic hue that permits to segment successfully a skin in color images. Such characteristic hue does not depend on brightness and race color, and has low computational cost. This approach proposes the choice of a color space and application of some classifiers, i.e. rules according to which the current pixel will or will not be considered as a pixel of skin image. Skin classifiers for most popular color spaces are presented in Table 2 [19]. Experiments show that the use of the skin color as a feature of gesture in the image is the effective method for automatic localization in real-time applications. Sometimes "broken" skin regions and "holes" into skin region appear. The following morphological processing removes this disadvantage. Also one can mention another restriction: the background color must differ from the skin color; otherwise the system will detect all regions looking as skin image in the color space.

Table 2. Skin classifiers for color spaces

Color space	Skin classifier
RGB-color space	(R, G, B) is classified as skin if: $R > 95$ and $G > 40$ and $B > 20$ and $\max\{R, G, B\} - \min\{R, G, B\} > 15$ and (10)
	$\lvert R - G\rvert > 15$ and $R > G$ and $R > B$
YUV-color space	(Y, U, V) is classified as skin if:
	$-10 < U < 8$ and $15 < V < 32$, where $U, V = [0, 255]$
YCbCr -color space	(Y, Cb, Cr) is classified as skin if:
	$Y > 80$ and $85 < Cb < 135$ and $135 < Cr < 180$,
	where $Y, Cb, Cr = [0, 255]$
HSV-color space	(H, S, V) is classified as skin if:
	$0 < H < 50$ and $0.23 < S < 0.68$,
	where $H = [0, 360]$ and $S, V = [0, 1]$
Normalized RGB-color space	(Normalized R, G, B) is classified as skin if:
	$g < g_u$ and $g > g_d$ and $W > 0.0004$,
	where u is a top boundary, d is a bottom boundary, $g_u = J_u r^2 + K_u r + L_u$, $g_d = J_d r^2 + K_d r + L_d$, $W = (r - 0{,}33)^2 + (g - 0{,}33)^2$, $J_u = -1{,}377$, $K_u = 1{,}074$, $L_u = 0{,}145$, $J_d = -0{,}776$, $K_d = 0{,}560$, $L_d = 0{,}177$
Log opponent IRgBy-color space	$(hue, saturation)$ is classified as skin if:
	$110 \le hue \le 180$ and $0 \le saturation \le 130$,
	$hue = \arctan^2(Rg/By)$, $saturation = \sqrt{Rg^2 + By^2}$.

The morphological processing is a sequential series of erosion and dilatation operators which permits to joint "broken" skin regions and to destroy "holes" into skin region. For both operators (erosion and dilatation), the structured element (so called the mask of morphological filter) is determined. Non-zero values in the mask show which one from adjacent pixels will be considered during the morphological processing. Erosion and dilatation operators have a view:

$$\mathrm{A} \oplus \mathrm{B} = \left\{ a \in \mathrm{A}, b \in \mathrm{B}, (a+b) \in \mathrm{R}^2 \right\}, \quad \mathrm{A} \theta \mathrm{B} = \left(\mathrm{A}^C \oplus \mathrm{B} \right)^C, \tag{1}$$

where $\boldsymbol{A}$ is an initial image; $\boldsymbol{B}$ is a structured element; a and b are elements of $\boldsymbol{A}$ and $\boldsymbol{B}$ sets correspondingly; $\boldsymbol{A}^C$ is a complement of a set $\boldsymbol{A}$; $\oplus$ and θ are the symbols of erosion and dilatation operators correspondingly; $\boldsymbol{R}^2$ is a *2D* space of an initial image.

Such rules are applied non-recursively for all pixels. During erosion, the structured element is one of the masks of size *3×3* or *5×5* pixels; as a result "broken" skin regions and "holes" into skin region disappear. During following dilatation, the mask of size *3×3* pixels is used for outer edges smoothing of connected regions. The morphological operators may be used repeatedly for more effective skin regions merginf.

The hand segmentation is based on a term – pixels "connectivity". It is well known in two variants: 4-connectivity and 8-connectivity (permits to find the contour more accurately). The algorithm of "immersion" firstly creates the connection table for labeled pixels; secondly after the algorithm pass, common regions are relabeled and merged according to 4-connectivity. The proposed algorithm of "immersion" has following steps:

Step 1. The beginning from the dark pixels – the points on initial "basins" (regions from which streams from all points drain to a single common point).

Step 2. For each brightness level k: for each connected component, (a) if a current pixel belongs to one existing basin only, then adds it to basin; (b) if a current pixel belongs to more than one existing basin, then labels it as a watershed; (c) in other case, create a new basin.

Step 3. The transition to next pixel.

Step 4. The recycling of Steps 2 and 3 until all pixels are processed.

However we may receive "false" skin regions. The segmentation algorithm calculates such geometrical features as a square of a hand image S (sum of skin region, near 0.3 – 2.5 % pixels of whole image) and coordinates of a mass center R_c (average of coordinates of the skin region). Also we may use the assumption that a gesture will be in predetermined screen area, so coordinates of a mass center R_c can not be less then 1/3 of the image height. Such algorithms and conditions permit to realize the spatial segmentation of a hand image. The following stage is the temporal segmentation of gestures.

In temporal domain, there are two categories of segmentation methods: motion segmentation and contour detection. Motion segmentation methods include block-matching method and its modifications, method of phase correlation, feature points tracking, and method of optical flow [20]. Contour detection methods contain gradient methods and active contour methods. Active contours (shakes) are widely used in the tasks connecting with image segmentation and boundary extraction. For boundary detection on images, curves with minimal energy are applied. It is supposed that a

target boundary is a smooth closed line [21]. Such assumptions well fit with hand segmentation process.

The initial contour is accepted as a simple figure (for example, a circle). Then the contour is distorted, and contour points tend to object boundary by minimization of contour energy. For each point p_i in neighborhood, energy is determined as

$$E_i(p_i) = \alpha \cdot E_{int}(p_i) + \beta \cdot E_{ext}(p_i), \tag{2}$$

where $E_{int}(p_i)$ is an energy function which is determined by contour shape; $E_{ext}(p_i)$ is an energy function which is determined by image properties, in particular a gradient in the neighborhood of point p_i; α and β are the constants which correct the energy value. For each point p_i, $p_i \to p_i'$ that corresponds the minimal value $E(p_i)$.

Thereby the active contour method includes three main stages: the forming of initial contour, anchor points detection, and tracking of anchor points. The algorithm of initial contour forming is as follows:

Step 1. The region of interest is checked up to the first transition between background and object (skin region).

Step 2. When the transition is detected, the algorithm finds a following neighboring pixel with the determined connectivity. Only pixels on boundary are analyzed, which prevents contour propagation into the object.

Step 3. The total contour is outlined in the same manner clockwise or anticlockwise (the direction is remained persistent).

Step 4. The edge-point linking is finished when algorithm achieves its initial point from Step 1.

The second stage of active contour method is the anchor points' detection on the received contour. The anchor point is a point, where a knee curve is watched. For this purpose, the Douglas-Peucker algorithm may be used, as given below:

Step 1. The single rib is considered connecting the first and the last contour points (for closed contour it will be a single point).

Step 2. We search a point which is at the maximal distance from the considered segment.

Step 3. If the distance from this point to the segment is less then the threshold value ε then all points on this segment are rejected; otherwise we remember this point, and the segment from the previous step is divided into two parts.

Step 4. We recursively repeat Step 3 for a new polyline building. As a result, the curve containing only the saved points will be received.

The last stage of tracking of anchor points is executed for each frame:

Step 1. The initial object image $f_0(x, y)$ is received.

Step 2. The active contour is located near to a moving object $f_0(x, y)$, $f_{act}(x, y) = f_0(x, y)$.

Step 3. Until shake does not lose the object, the external energy E_{ext} is calculated from the function $f_{act}(x, y)$. Then we find the energy minimum $E(p)$ (Eq. 2) until a shake is converged. Function $f_{act}(x, y)$ ought to be such that the shake center of gravity is situated in the center of image region.

The detection of moments (start and end points) of a current sign is searched as pause in hand motions according to dactylology rules from Section 3.1.

4.2 Vector Model of Gesture

The gesture vector model is based on the anchor points of the outer contour which were detected during hand tracking (Section 4.1). The vector model **VM** is a set of sequentially connected vectors *Vi*, *i=1,…, n*:

$$\mathrm{VM} = \{\mathrm{V}_1, \mathrm{V}_2, \ldots, \mathrm{V}_i, \mathrm{V}_{i+1}, \ldots, \mathrm{V}_n\}. \tag{3}$$

Here the beginning of each following vector $\mathbf{V}_{i+1}$ is the ending of the previous vector $\mathbf{V}_i$; so we form a closed contour. Each vector $\mathbf{V}_i$ includes two components: a length L_i and an angle γ_i (between a vector direction and the axis *OX*) in relative coordinates:

$$L_i = \sqrt{x_i^2 + y_i^2}\ , \qquad \gamma_i = arctan\left(y_i / x_i\right), \tag{4}$$

where x_i and y_i are the relative coordinates of the vector $\mathbf{V}_i$: $x_i = x_{j+1} - x_j$, $y_j = y_{j+1} - y_j$; (x_j, y_j) and (x_{j+1}, y_{j+1}) are the coordinates of current and following anchor points correspondingly.

Then a vector model **VM** (Eq. 3) is reconstructed by procedures of compression and normalization for the purpose of invariance to affine transforms (shifts, rotation, and scale). The compression procedure permits to reject non-essential vectors which connect the adjacent anchor points by a sequential comparison of the vectors length with experimentally predetermined threshold value. The normalization procedure executes a normalization of vectors lengths (sum of all vectors lengths becomes equal to 1) according to a following expression:

$$L_i^{nr} = L_i \Big/ \sum_{i=1}^{n} L_i\ , \tag{5}$$

and a normalization of vectors directions when the current vector direction is replaced by one of normalized directions from the unit vectors set. Such unit vector has a view:

$$\mathrm{U}_i = \left\{1, \gamma_i^{nr}\right\}, \quad \gamma_i^{nr} = i \cdot 2\pi / Z\ , \tag{6}$$

where *Z* is a number of directions (usually $Z = 8$ or $Z = 16$).

Thereby we receive a set of normalized vector models according to the starting point of the contour:

$$\begin{aligned}
\mathrm{VM}_1^{nr} &= \left\{\mathrm{V}_1^{nr}, \mathrm{V}_2^{nr}, \ldots, \mathrm{V}_i^{nr}, \ldots, \mathrm{V}_{n-1}^{nr}, \mathrm{V}_n^{nr}\right\}, \\
\mathrm{VM}_2^{nr} &= \left\{\mathrm{V}_2^{nr}, \mathrm{V}_3^{nr}, \ldots, \mathrm{V}_i^{nr}, \ldots, \mathrm{V}_n^{nr}, \mathrm{V}_1^{nr}\right\}, \\
&\ldots \qquad\qquad \ldots \qquad\qquad \ldots \\
\mathrm{VM}_N^{nr} &= \left\{\mathrm{V}_n^{nr}, \mathrm{V}_1^{nr}, \ldots, \mathrm{V}_i^{nr}, \ldots, \mathrm{V}_{n-2}^{nr}, \mathrm{V}_{n-1}^{nr}\right\},
\end{aligned} \tag{7}$$

where *N* is the set power.

However the sequential search of all received vector models $\mathbf{VM}_i^{nr}$ requires additional computer resources. It is more rational to design a vector model with invariance to the choice of a starting vector. If a vector $\mathbf{V}_k^{nr}$ with a minimal length will be a single element then this vector model is accepted; otherwise the search of the "best" vector is realized according to a following procedure:

Step 1. The first determined vector of a minimal length is chosen.
Step 2. The sum of lengths of all following vectors of minimal length is calculated.
Step 3. The following determined vector of a minimal length is chosen.
Step 4. If not all vectors of minimal lengths are processed, then jump to Step 2.
Step 5. Finally such vector is chosen (the "best" vector) the total length of which has the minimal value.

During the training stage, algorithms and procedures from Sections 4.1 and 4.2 are applied to all gestures represented in Table 1. As a result, we have a designed vector models for each gesture (etalon vector models). During working stage, the current gesture is processed in the same manner (current vector model). Additionally other essential features are calculated for the identification procedure.

4.3 Identification and Decision Marking

All parameters, presented in Table 1, are coded in a numerical form; additionally we built the "best" vector model of the outer gesture contour. These data are the input features for the identification procedure. There are various well known approaches for recognition such as the PCA-based (Principal Component Analysis), the HMM-based (Hidden Markov Models), and the NN-based (Neural Networks) recognition. Each approach has many modifications aimed at various tasks. Their main disadvantage is the high computational cost. That's why a method with fewer resources was chosen, and it is the maximum likelihood method.

By using blob-detectors [22] we may determine fingers position and orientation, and the number of separated fingers. Also we can calculate the hand skewing position, inner closed contours and the shifting motion during gesture demonstration applying well known image processing methods. For these six parameters we built a set of decision rules based on Table 1 as a production knowledge model. Mostly the application of such decision rules is enough for sign identification. In these cases, a comparative analysis of vector models (etalon and current) is an additional possibility of decision marking confirmation. In other cases, we use the maximum likelihood method for the final decision which calculates a maximum of similarity *DM* between etalon description **E** and the current description $\mathbf{VM}^{nr}$:

$$DM = arg \max_j \max_l P\left(\mathrm{E}_l, \mathrm{VM}_j^{nr}\right), \tag{8}$$

where P is a probability function between the etalon and current descriptions (calculated during a training stage), l is a number of signs, $l = 33$ for Russian sign alphabet.

The similarity measure D_j between the etalon j and current descriptions is determined in the Euclidian metric as a sum of vectors differences:

$$D_j = \sum_{i=1}^{N} \left(\left| L_i^{nr}, \gamma_i^{nr} \right| - \left| L_i^j, \gamma_i^j \right| \right)^2 , \tag{9}$$

where L_i^j is one of normalized vector length; γ_j^i is one of normalized vector directions.

5 Software "Russian Sign Language" and Experimental Results

The software "Russian Sign Language" includes five main modules: the module "Frame Extraction" contains functions for frames extraction from video sequences in various formats; the module "Hand Detection" includes functions for hand detection; the module "Hand Tracking" executes hand tracking considering the actions when a hand appears or disappears in the frame; the module "Recognition" has functions and procedures for gesture identification; the module "Interface Control" provides interaction between user and application. The software has some additional functions such as embedded video sequences player, speed control, and editor of signs vocabulary.

User may create, replace, and remove a sign pushing the correspondence button. All signs are listed in a list "Value" where we can replace an old value by a new value. If a sign image was loaded from the main screen of application then its contour will be automatically calculated and displayed. If a sign image was loaded from a file then a manual mode is used as lasso with pointing anchor points of hand contour. Additional image processing includes some filters: correction filter "Grey world", median filter, inverse filter, grey-scale transformation, and binary filter.

The work begins if a user initiated the process of recognition; otherwise the application will be in a waiting mode. If a hand image is successfully detected in a sequence of frames, then these data pass to the tracking module. If a hand disappears in a frame, then the tracking process will be restarted. During recognition, a module "Recognition" holds the identification process and writes the results in a journal file. The main screen of the experimental software "Russian Sign Language" is represented on Fig. 3. The software is realized in the environment of Rapid Application Design "Borland Delphi 7" with using the utility "DirectShow" for working with video stream.

During experiments own test video sequences with frame sizes 360×238 pixels were used. Segmentation of skin regions was accomplished in the YUV-color space. The probability of sign identification had been calculated by using four own test video sequences on which several signers show the same sign. Such test demonstrates the invariance of gesture recognition for sizes and shapes of different hands. One can see results on Fig. 4. The non-skin similar background color is the main condition. Also it is desirable that the background intensity will be homogeneous. Conditions of hand motion are situated in Section 3.2.

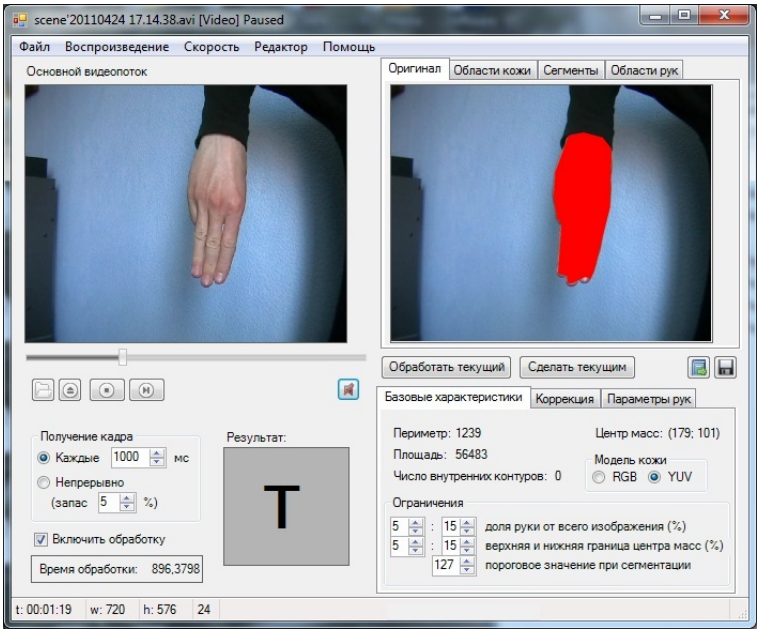


Fig. 3. The main screen of system "Russian Sign Alphabet" in learning mode

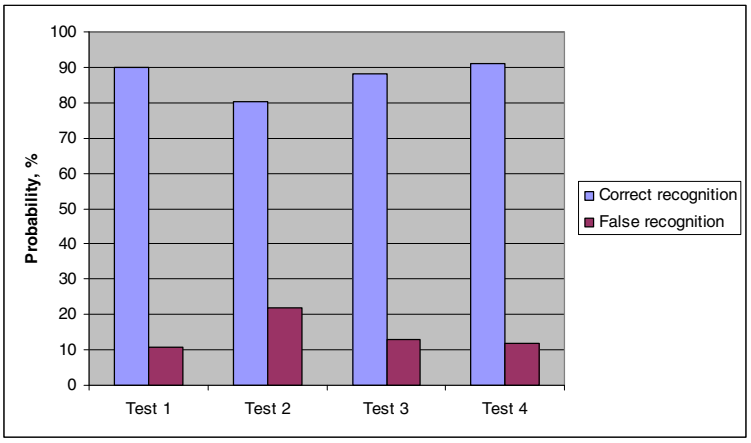


Fig. 4. Results of correct and false recognition of sign "F" by four signers

Average probability value of the correct recognition is near 88-90 % and this of the false recognition is less than 13-15 %. Sometimes during a hand displacement, software makes mistakes and tries to recognize a sign, that's why the average value of the false recognition is increased.

Also some signs from different video sequences were tested for checking vector models. The probability of their correct recognition is presented on Fig. 5. Average probability value equals 82-87 %. The worse results were achieved for dynamic signs and static signs had the better results.

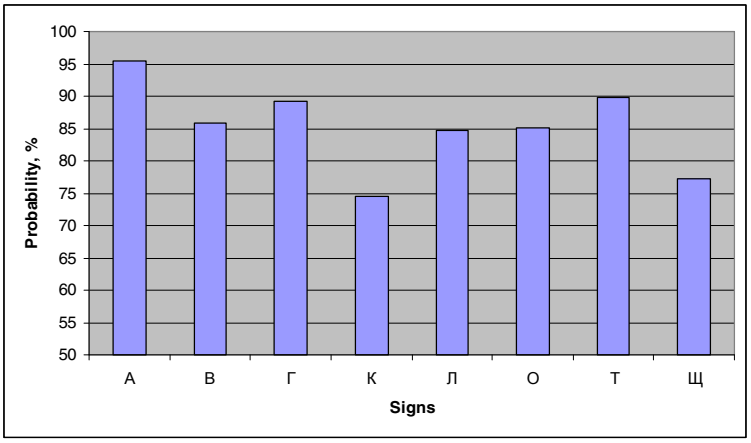


Fig. 5. Results of recognition probability of signs vector models

6 Conclusions

The intelligent system for sign alphabet learning includes advanced technique for segmentation and recognition of 3D visual objects. The Russian sign alphabet was analyzed in detail that permits to extract more essential gesture features. Based on dactylology statements, the main and additional criteria were formulated and used in segmentation and recognition procedures. The hand segmentation algorithm is not a trivial task and requires the design of complex segmentation in a spatio-temporal domain. Vectorization algorithm increases the efficiency of following object identification from the limited set of signs.

Software "Russian Sign Language" includes modules which realize frame extraction from video sequences, hand detection and tracking, recognition, and interface control. During experiments, it was find out that the average probability value for the correct recognition is near 88-90 % and this of the false recognition less than 13-15 %. Such test demonstrates the invariance of gesture recognition for sizes and shapes of signers' hands. Also some signs were tested for checking vector models. Average probability value achieved was 82-87 %; worse results were achieved for dynamic signs and static signs had better results.

References

1. Mitra, S., Acharya, T.: Gesture Recognition: a Survey. IEEE Trans. Syst. Man. Cybern. Part C Appl. Rev. 37(3), 311–324 (2007)
2. Alon, J., Athitsos, V., Yuan, Q., Sclaroff, S.: A Unified Framework for Gesture Recognition and Spatio-temporal Gesture Segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 31(9), 1685–1699 (2009)
3. Li, H., Greenspan, M.: Model-based Segmentation and Recognition of Dynamic Gestures in Continuous video streams. Pattern Recognition 44, 1614–1628 (2011)

4. Li, H., Greenspan, M.: Segmentation and Recognition of Continuous Gestures. In: Int. Conf. on Image Proc., vol. 1, pp. I365–I368 (2007)
5. Fang, G., Gao, W., Zhao, D.: Large-vocabulary Continuous Sign Language Recognition Based on Transition-movement Models. IEEE Trans. Syst. Man. Cybern. Part A: Syst. Humans 37(1), 1–9 (2007)
6. Kim, D., Song, J., Kim, D.: Simultaneous Gesture Segmentation and Recognition Based on Forward Spotting Accumulative HMMs. Pattern Recognition 40, 3012–3026 (2007)
7. Roh, M., Christmas, B., Kittler, J., Lee, S.: Gesture Spotting for Low-resolution Sports for Video Annotation. Pattern Recognition 41, 1124–1137 (2008)
8. Yang, R., Sarkar, S., Loeding, B.: Handling Movement Epenthesis and Hand Segmentation Ambiguities in Continuous Sign Language Recognition Using Nested Dynamic Programming. IEEE Trans. Pattern Anal. Mach. Intell. 32(3), 462–477 (2010)
9. Morguet, P., Lang, M.: Spotting Dynamic Hand Gestures in Video Image Ssequences Using Hidden Markov Models. In: Int. Conf. on Image Processing, vol. 3, pp. 193–197 (1998)
10. Lee, H.-K., Kim, J.H.: An HMM-based Threshold Model Approach for Gesture Recognition. IEEE Trans. Pattern Anal. Mach. Intell. 21(10), 961–973 (1999)
11. Kang, H., Lee, C.W., Jung, K.: Recognition-based Gesture Spotting in Video Games. Pattern Recognition 25, 1701–1714 (2004)
12. Suk, H.-I., Sin, B.-K., Lee, S.-W.: Hand Gesture Recognition Based on Dynamic Bayesian Network Framework. Pattern Recognition 43, 3059–3072 (2010)
13. Tsai, C.-Y., Lee, Y.-H.: The Parameters Effect on Performance in ANN for Hand Gesture Recognition System. Expert Systems with Applications 38, 7980–7983 (2011)
14. Liang, R.-H., Ouhyoung, M.: A Real-time Continuous Gesture Recognition System for Sign Language. In: IEEE Int. Conf. Automatic Face and Gesture Recognition, pp. 558–563 (1998)
15. Bauer, B., Hienz, H.: Relevant Features for Video-based Continuous Sign Language Recognition. In: IEEE Int. Conf. on Automatic Face and Gesture Recognition, pp. 440–445 (2000)
16. Bobick, A.F., Davis, J.W.: The Recognition of Human Movement Using Temporal Templates. IEEE Trans. Pattern Anal. Mach. Intell. 23, 257–267 (2001)
17. Starner, T., Weaver, J., Pentland, A.: Real-time American Sign Language Recognition Using Desk and Wearable Computer Based Video. IEEE Trans. Pattern Anal. Mach. Intell. 20(12), 1371–1375 (1998)
18. Nayak, S., Sarkar, S., Loeding, B.: Automated Extraction of Signs from Continuous Sign Language Sentences Using Iterated Conditional Modes. In: Proceedings of IEEE Conf. on Computer Vision and Pattern Recognition, pp. 2583–2590 (2009)
19. Favorskaya, M.N., Pahirka, A.I.: Models of Face Localization on Images. Control Systems and Information Technologies 3(33), 404–408 (2008) (in Russian)
20. Favorskaya, M.: Motion Estimation for Object Analysis and Detection in Videos. In: Kountchev, R., Nakamatsu, K. (eds.) Advances in Reasoning-based Image Processing, Analysis and Intelligent Systems: Conventional and Intelligent Paradigms, Springer, Heidelberg (2012)
21. Zinbi, Y., Chahir, Y., Elmoataz, A.: Moving Object Segmentation Using Optical Flow with Active Contour Model. In: Intern. Conf. on Information and Communication Technologies: From Theory to Applications (2008), doi:10.1109/ICTTA.2008.4530112
22. Tuytelaars, T., Mikolajczyk, K.: Local Invariant Feature Detectors: A Survey. Foundations and Trends in Computer Graphics and Vision 3(3), 177–280 (2007)

Decorrelation of Sequences of Medical CT Images Based on the Hierarchical Adaptive KLT

Roumen Kountchev and Peter Ivanov

Technical University of Sofia, Department of Radio Communications and Video Technologies, Boul. Kl. Ohridsky 8, Sofia 1000, Bulgaria
rkountch@tu-sofia.bg, peter.n.ivanov@gmail.com

Abstract. In this work is presented one new approach for processing of sequences of medical CT images, called Hierarchical Adaptive Karhunen-Loeve Transform (HAKLT). The aim is to achieve high decorrelation for each group of 9 consecutive CT images, obtained from the original larger sequence. In result, the main part of the energy of all images in one group is concentrated in a relatively small number of eigen images. This result could be obtained using the well-known Karhunen-Loeve Transform (KLT) with transformation matrix of size 9×9. However, for the implementation of the 2-levels HAKLT in each level are used 3 transform matrices of size 3×3, in result of which the computational complexity of the new algorithm is reduced in average 2 times, when compared to that of KLT with 9×9 matrix. One more advantage is that the algorithm permits parallel processing for each group of 3 images in every hierarchical level. In this work are also included the results of the algorithm modeling for sequences of real CT images, which confirm its ability to carry out efficient decorrelation. The HAKLT algorithm could be farther used as a basis for the creation of algorithms for efficient compression of sequences of CT images and for features space minimization in the regions of interest, which contain various classes of searched objects.

Keywords: Decorrelation of medical CT image sequences, Hierarchical Adaptive Karhunen-Loeve Transform (HAKLT), Group of Medical Images.

1 Introduction

In the last years, large number of new technologies and systems for digital processing of medical images had been created [1, 2, 3], such as: the standard Digital Imaging and Communication in Medicine (DICOM) [4] used for storage, transfer and visualization of images, obtained from computer tomographic systems, MRI and ultrasound scanners; Picture Archiving and Communication System (PACS); systems for digital radiography, teleradiology, etc. Medical images could be still or moving, such as Magnetic Resonance Image (MRI), Nuclear Magnetic Resonance Image (NMRI), Magnetic Resonance Tomography Image (MRTI), etc. Moving images are represented by sequences of still images, obtained in consecutive time moments, or

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 41–55.
DOI: 10.1007/978-3-319-00029-9_4

spatial positions of the Computer Tomography (CT) scanner. Because of the large volume of the visual medical information, various algorithms are used for its compression. For still MRI are usually used algorithms based on the DCT, wavelet decomposition for prediction or zero-tree/block coding [5, 6, 7], etc. For compression of CT images sequences are used: interframe decorrelation based on hierarchical interpolation (HINT) [8, 9], spatial active appearance model [10], JPEG-LS and JPEG2000 with interframe motion compensated prediction [11, 12] and distributed representation of image sets based on Slepian-Wolf coding [13].

One of the most efficient methods for decorrelation and compression of groups of images is based on the KLT, also known as transform of Hotelling, or Principal Component Analysis (PCA) [14] - [23]. For its implementation the pixels with the same spatial position in a group of N images compose the corresponding N-dimensional vector. The basic difficulty of the KLT implementation is related to the large size of the covariance matrix. For the calculation of its eigenvectors is necessary to calculate the roots of a polynomial of n^{th} degree (characteristic equation) and to solve a linear system of N equations. For large values of N, the computational complexity of the algorithm for calculation of the transform matrix is significantly increased.

One of the possible approaches for reduction of the computational complexity of KLT for N-dimensional group of medical images is based on the "Hierarchical Adaptive KLT" (HAKLT), offered in this work. Unlike the famous hierarchical KLT (HKLT) [18], this transform is not related to the image sub-blocks, but to the whole image from one group. For this, the HAKLT is implemented through dividing the image sequence into sub-groups of 3 images each, on which is applied Adaptive KLT (AKLT), of size 3×3. This transform is performed using equations, which are not based on iterative calculations, and as a result, they have lower computational complexity. To decorrelate the whole group of medical images is necessary to use AKLT of size 3×3, which to be applied in several consecutive stages (hierarchical levels), with rearranging of the obtained intermediate eigen images after each stage. In result is obtained a decorrelated group of 9 eigen medical images.

The paper comprises the following: the principle for decorrelation of CT images group through HAKLT, the calculation of eigen images through AKLT with 3×3 matrix, experimental results, evaluation of the computational complexity and conclusions.

2 Principle for Decorrelation of a Group of CT Images through Hierarchical AKLT

The sequence of medical images is divided into Groups of 9 Images (GMI), for which is supposed that they are highly correlated. On the other hand, each GMI is further divided into 3 sub-groups.

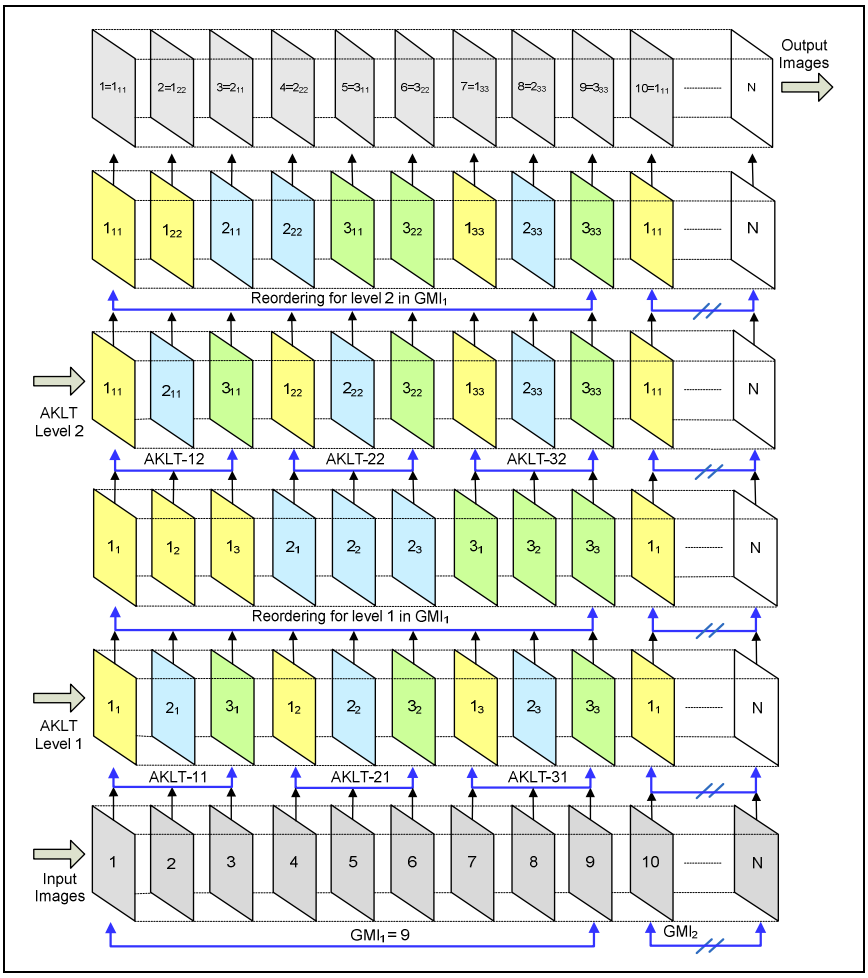


Fig. 1. Algorithm for 2-levels Hierarchical Adaptive KLT for Group of 9 Medical Images

The algorithm for 2-levels HAKLT for one GMI is shown on Fig. 1. As it is easily seen there, on each sub-group of 3 images from the first hierarchical level of HAKLT is applied AKLT with matrix of size 3×3. In result are obtained 3 eigen images, colored in yellow, blue and green correspondingly. After that, the eigen images are rearranged so that the first sub-group of 3 eigen images to comprise the first images from each group, the second group of 3 eigen images – the second images from each group, etc. For each GMI of 9 intermediate eigen images in the first hierarchical level is applied in similar way the next AKLT, with a 3×3 matrix, on each sub-group of 3 eigen values. In result are obtained 3 new eigen images (i.e. the eigen images of the group of 3 intermediate eigen images), colored in yellow, blue, and green correspondingly in the second hierarchical level. Then the eigen images are rearranged again so, that the first group of 3 eigen images to contain the first images from each group

before the rearrangement; the second group of 3 eigen images - the second image before the rearrangement, etc. At the end of the processing is obtained a decorrelated sequence of eigen images, using which and through inverse HAKLT could be restored the original sequence.

3 Calculation of Eigen Images through AKLT with 3×3 Matrix

For the calculation of eigen images through AKLT with 3×3 matrix for GMI sub-group is used the approach, given in [24] for the representation of the 3D color vector in the KLT space. From each sub-group with 3 medical images of S pixels each, shown on Fig. 2, are calculated the vectors $\vec{C}_s = [C_{1s}, C_{2s}, C_{3s}]^t$ for $s=1, 2, \ldots, S$ (on the figure are shown the vectors for the first 4 pixels only, resp. $\vec{C}_1 = [C_{11}, C_{21}, C_{31}]^t, \vec{C}_2 = [C_{12}, C_{22}, C_{32}]^t$, $\vec{C}_3 = [C_{13}, C_{23}, C_{33}]^t, \vec{C}_4 = [C_{14}, C_{24}, C_{34}]^t$). Each vector is then transformed into corresponding vectors $\vec{L}_s = [L_{1s}, L_{2s}, L_{3s}]^t$ through APCA with the matrix $[\Phi]$ of size 3×3. Its elements Φ_{ij} are defined below:

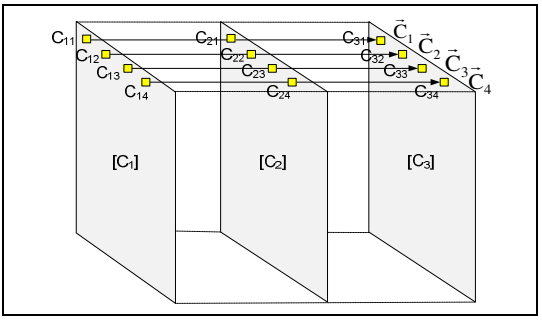


Fig. 2. Sub-group of 3 images from the GMI

❖ The covariance matrix $[K_C]$ of size 3×3 for vectors $\vec{C}_s$ is calculated:

$$[K_C] = \left[\frac{1}{S}\sum_{s=1}^{S}\vec{C}_s\vec{C}_s^t\right] - \vec{m}_c\vec{m}_c^t = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix}, \quad (1)$$

where $\vec{m}_C = [\overline{C}_1, \overline{C}_2, \overline{C}_3]^t$ is the mean vector. Here $\bar{x} = E(x_s) = \frac{1}{S}\sum_{s=1}^{S} x_s$; $E(.)$ – operator for calculation of the mean value of x_s for $s = 1, 2, \ldots, S$.

❖ The elements of the mean vector $\vec{m}_c$ and of the matrix $[K_C]$ are defined in accordance with the relations:

$$\overline{C}_1 = E(C_{1s}),\ \overline{C}_2 = E(C_{2s}),\ \overline{C}_3 = E(C_{3s}), \quad (2)$$

$$k_{11}= k_1=E(C_{1s}^2)-(\overline{C}_1)^2, k_{22}= k_2=E(C_{2s}^2)-(\overline{C}_2)^2, k_{33}= k_3=E(C_{3s}^2)-(\overline{C}_3)^2, \quad (3)$$

$$k_{12}= k_{21}= k_4=E(C_{1s}C_{2s})-(\overline{C}_1)(\overline{C}_2), k_{23}= k_{32}= k_6=E(C_{2s}C_{3s})-(\overline{C}_2)(\overline{C}_3), \quad (4)$$

$$k_{13}= k_{31}= k_5=E(C_{1s}C_{3s})-(\overline{C}_1)(\overline{C}_3). \quad (5)$$

❖ The eigen values $\lambda_1, \lambda_2, \lambda_3$ of the matrix $[K_C]$ are defined in accordance to the solution of the characteristic equation:

$$det| k_{ij}-\lambda\delta_{ij} |=\lambda^3+a\lambda^2+b\lambda+c=0, \quad (6)$$

where: $\delta_{ij}=\begin{cases} 1, & i=j, \\ 0, & i\neq j. \end{cases}$

$$\begin{aligned} &a=-(k_1+k_2+k_3), \quad b=k_1k_2+k_1k_3+k_2k_3-(k_4^2+k_5^2+k_6^2), \\ &c=k_1k_6^2+k_2k_5^2+k_3k_4^2-(k_1k_2k_3+2k_4k_5k_6), \end{aligned} \quad (7)$$

Since the matrix $[K_C]$ is symmetric, its eigen values are real numbers. For their calculation could be used the equations of Cardano for "casus irreducibilis" (i.e., the so-called "trigonometric solution"):

$$\begin{aligned} &\lambda_1=2\sqrt{\frac{|p|}{3}}\cos\left(\frac{\varphi}{3}\right)-\frac{a}{3}; \lambda_2=-2\sqrt{\frac{|p|}{3}}\cos\left(\frac{\varphi+\pi}{3}\right)-\frac{a}{3}; \\ &\lambda_3=-2\sqrt{\frac{|p|}{3}}\cos\left(\frac{\varphi-\pi}{3}\right)-\frac{a}{3} \quad for \ \lambda_1\geq\lambda_2\geq\lambda_3\geq 0, \end{aligned} \quad (8)$$

$$\begin{aligned} &q=2(a/3)^3-(ab)/3+c, \ p=-(a^2/3)+b<0, \\ &\varphi=\arccos\left[-q/2\Big/\sqrt{(|p|/3)^3}\right], \end{aligned} \quad (9)$$

❖ The eigen vectors $\vec{\Phi}_1, \vec{\Phi}_2, \vec{\Phi}_3$ of the covariance matrix $[K_C]$ are the solution of the system of equations below:

$$[K_C]\vec{\Phi}_m=\lambda_m\vec{\Phi}_m \text{ and } \left|\vec{\Phi}_m\right|^2=\sum_{i=1}^{3}\Phi_{mi}^2=1, \text{ for } m=1,2,3. \quad (10)$$

Eq. 10 follows from the condition for orthogonality and normalization of all 3 eigenvectors:

$$\vec{\Phi}_s^t\vec{\Phi}_k=\sum_{i=1}^{3}\Phi_{is}\Phi_{ik}=\begin{cases} 1 & for \ s=k; \\ 0 & for \ s\neq k. \end{cases} \text{ for } s, k = 1,2,3. \quad (11)$$

The solution of the system of equations (10) is used to calculate components of m[th] eigenvector $\vec{\Phi}_m = [\Phi_{1m}, \Phi_{2m}, \Phi_{3m}]^t$, which corresponds to the eigen value λ_m:

$$\Phi_{1m} = A_m / P_m;\ \Phi_{2m} = B_m / P_m;\ \Phi_{3m} = D_m / P_m \text{, for } m = 1,2,3; \tag{12}$$

$$A_m = (k_3 - \lambda_m)[k_5(k_2 - \lambda_m) - k_4 k_6],\ B_m = (k_3 - \lambda_m)[k_6(k_1 - \lambda_m) - k_4 k_5], \tag{13}$$

$$D_m = k_6[2k_4 k_5 - k_6(k_1 - \lambda_m)] - k_5^2(k_2 - \lambda_m),\ P_m = \sqrt{A_m^2 + B_m^2 + D_m^2} \neq 0. \tag{14}$$

The KLT matrix $[\Phi]$ comprises the eigenvectors $\vec{\Phi}_m = [\Phi_{1m}, \Phi_{2m}, \Phi_{3m}]^t$:

$$[\Phi] = \begin{bmatrix} \vec{\Phi}_1^t \\ \vec{\Phi}_2^t \\ \vec{\Phi}_3^t \end{bmatrix} = \begin{bmatrix} \Phi_{11} & \Phi_{21} & \Phi_{31} \\ \Phi_{12} & \Phi_{22} & \Phi_{32} \\ \Phi_{13} & \Phi_{23} & \Phi_{33} \end{bmatrix} \text{, for } m = 1,2,3. \tag{15}$$

The direct AKLT for vectors $\vec{C}_s = [C_{1s}, C_{2s}, C_{3s}]^t$, from which are obtained vectors $\vec{L}_s = [L_{1s}, L_{2s}, L_{3s}]^t$, is:

$$\begin{bmatrix} L_{1s} \\ L_{2s} \\ L_{3s} \end{bmatrix} = \begin{bmatrix} \Phi_{11} & \Phi_{21} & \Phi_{31} \\ \Phi_{12} & \Phi_{22} & \Phi_{32} \\ \Phi_{13} & \Phi_{23} & \Phi_{33} \end{bmatrix} \begin{bmatrix} (C_{1s} - \overline{C}_1) \\ (C_{2s} - \overline{C}_2) \\ (C_{3s} - \overline{C}_3) \end{bmatrix} \text{ for } s = 1,2,...,S. \tag{16}$$

The components of vectors $\vec{L}_s = [L_{1s}, L_{2s}, L_{3s}]^t$ could be processed in various way (such as for example: orthogonal transforms, quantization, decimation and interpolation, etc.). In result are obtained the corresponding vectors $\vec{L}_s^q = \psi(\vec{L}_s) = [\psi_1(L_{1s}), \psi_2(L_{2s}), \psi_3(L_{3s})]^t$ with components $L_{1s}^q = \psi_1(L_{1s})$, $L_{2s}^q = \psi_2(L_{2s})$, $L_{3s}^q = \psi_3(L_{3s})$, where $\psi_1(.), \psi_2(.), \psi_3(.)$ are the functions of the used transform. For the restoration of the vectors $\vec{L}_s^q$ are used the functions for inverse transform of the components $\hat{L}_{1s} = \psi_1^{-1}(L_{1s}^q)$, $\hat{L}_{2s} = \psi_2^{-1}(L_{2s}^q)$, $\hat{L}_{3s} = \psi_3^{-1}(L_{3s}^q)$ and in result are obtained the decoded vectors $\vec{\hat{L}}_s = [\hat{L}_{1s}, \hat{L}_{2s}, \hat{L}_{3s}]^t$. Using the inverse AKLT, the vectors $\vec{\hat{L}}_s$ are transformed into vectors $\vec{\hat{C}}_s = [\hat{C}_{1s}, \hat{C}_{2s}, \hat{C}_{3s}]^t$:

$$\begin{bmatrix} \hat{C}_{1s} \\ \hat{C}_{2s} \\ \hat{C}_{3s} \end{bmatrix} = \begin{bmatrix} \Phi_{11} & \Phi_{12} & \Phi_{13} \\ \Phi_{21} & \Phi_{22} & \Phi_{23} \\ \Phi_{31} & \Phi_{32} & \Phi_{33} \end{bmatrix} \begin{bmatrix} \hat{L}_{1s} \\ \hat{L}_{2s} \\ \hat{L}_{3s} \end{bmatrix} + \begin{bmatrix} \hat{\overline{C}}_1 \\ \hat{\overline{C}}_2 \\ \hat{\overline{C}}_3 \end{bmatrix} \text{ for } s = 1,2,...,S. \tag{17}$$

Here the matrix of the inverse APCA is:

$$\begin{bmatrix} \Phi_{11} & \Phi_{12} & \Phi_{13} \\ \Phi_{21} & \Phi_{22} & \Phi_{23} \\ \Phi_{31} & \Phi_{32} & \Phi_{33} \end{bmatrix} = [\Phi]^{-1} = [\Phi]^{t} = [\vec{\Phi}_1, \vec{\Phi}_2, \vec{\Phi}_3]. \tag{18}$$

For the restoration of vectors $\vec{\hat{C}}_s = [\hat{C}_{1s}, \hat{C}_{2s}, \hat{C}_{3s}]^t$ through inverse AKLT are needed not only the vectors $\vec{\hat{L}}_s = [\hat{L}_{1s}, \hat{L}_{2s}, \hat{L}_{3s}]^t$, but also the elements Φ_{ij} of the matrix $[\Phi]$, and the values of $\overline{C}_1, \overline{C}_2, \overline{C}_3$ as well. The total number of these elements could be reduced representing the matrix $[\Phi]$ as the product of matrices $[\Phi_1(\alpha)], [\Phi_1(\beta)], [\Phi_1(\gamma)]$, and rotation around coordinate axes for each transformed vector in Euler angles α, β and γ correspondingly:

$$[\Phi] = \begin{bmatrix} \Phi_{11} & \Phi_{21} & \Phi_{31} \\ \Phi_{12} & \Phi_{22} & \Phi_{32} \\ \Phi_{13} & \Phi_{23} & \Phi_{33} \end{bmatrix} = [\Phi_1(\alpha)][\Phi_2(\beta)][\Phi_3(\gamma)] = [\Phi(\alpha, \beta, \gamma)], \tag{19}$$

where

$$\begin{aligned} [\Phi_1(\alpha)] &= \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}; [\Phi_2(\beta)] = \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix}; \\ [\Phi_3(\gamma)] &= \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned} \tag{20}$$

In this case the elements of the matrix $[\Phi]$ are represented by the relations:

$$\begin{aligned} &\Phi_{11} = \cos\alpha\cos\beta\cos\gamma - \sin\alpha\sin\gamma; \\ &\Phi_{21} = -(\cos\alpha\cos\beta\sin\gamma + \sin\alpha\cos\gamma); \Phi_{31} = -\cos\alpha\sin\beta; \\ &\Phi_{12} = \sin\alpha\cos\beta\cos\gamma + \cos\alpha\sin\gamma; \Phi_{22} = -\sin\alpha\cos\beta\sin\gamma + \cos\alpha\cos\gamma; \\ &\Phi_{32} = -\sin\alpha\sin\beta; \Phi_{13} = \sin\beta\cos\gamma; \Phi_{23} = -\sin\beta\sin\gamma; \Phi_{33} = \cos\beta. \end{aligned} \tag{21}$$

The matrix of the inverse AKLT is defined by the relation:

$$[\Phi]^{-1} = [\Phi_3(-\gamma)][\Phi_2(-\beta)][\Phi_1(-\alpha)] \tag{22}$$

Then, for the calculation of the elements of the inverse matrix $[\Phi]^{-1}$ is enough to know the values of the 3 rotation angles α, β and γ, defined by the relations:

$$\alpha = -\arcsin\left(\Phi_{32} \Big/ \sqrt{1-\Phi_{33}^2}\right); \ \beta = \arccos(\Phi_{33}); \ \gamma = \arccos\left(\Phi_{13} \Big/ \sqrt{1-\Phi_{33}^2}\right). \tag{23}$$

In result, the number of the needed values for the calculation of the matrix $[\Phi]^{-1}$ is reduced from 9 down to 3, i.e. 3 times reduction. The elements L_{1s}, L_{2s}, L_{3s} for $s=1,2,...,S$ comprise the pixels of the first, second and third eigen image in the sub-group of medical images C_{1s}, C_{2s}, C_{3s}.

4 Experimental Results

On the basis of the 2-levels HAKLT algorithm, shown on Fig. 1, were done experiments with sequences of CT images of size 512×512 pixels, 8 bpp. The sequence was divided into groups (Set 1,..,Set R), each containing 9 consecutive CT images. As an example, on Fig. 3 is shown one of the groups - Set 3, which contains CT Image 1,.., Image 9.

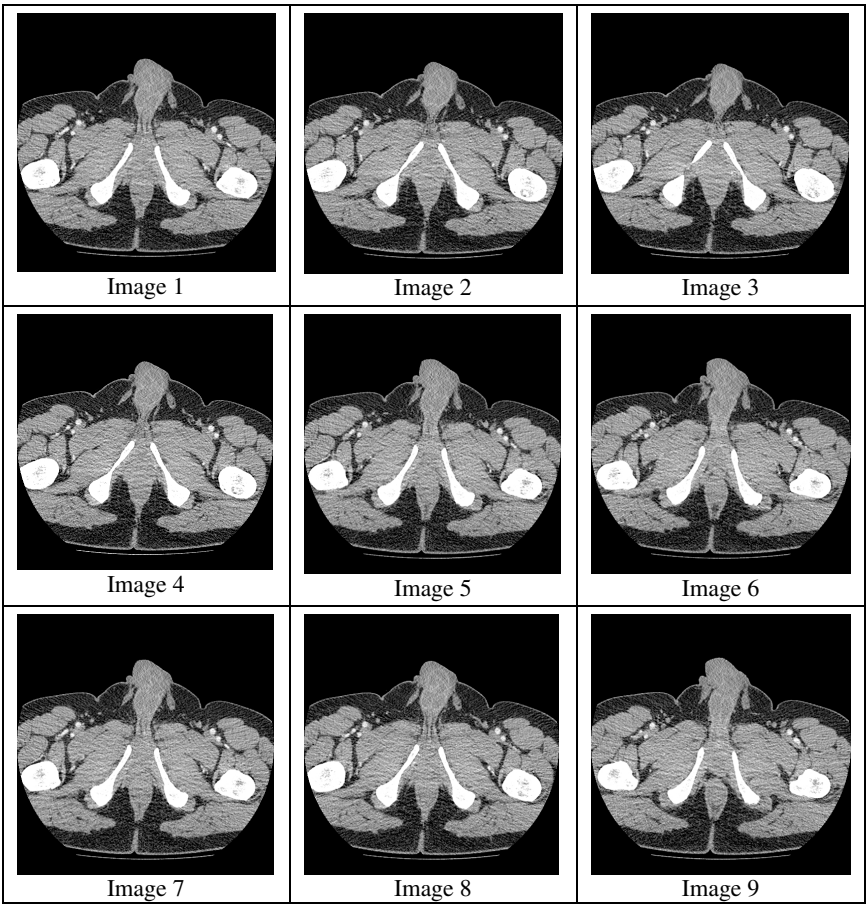


Fig. 3. Group of 9 consecutive CT images in Set 3

On Fig. 4 are shown the corresponding eigen images, obtained in result of applying the 2-levels HAKLT algorithm on the group of images (Set 3).

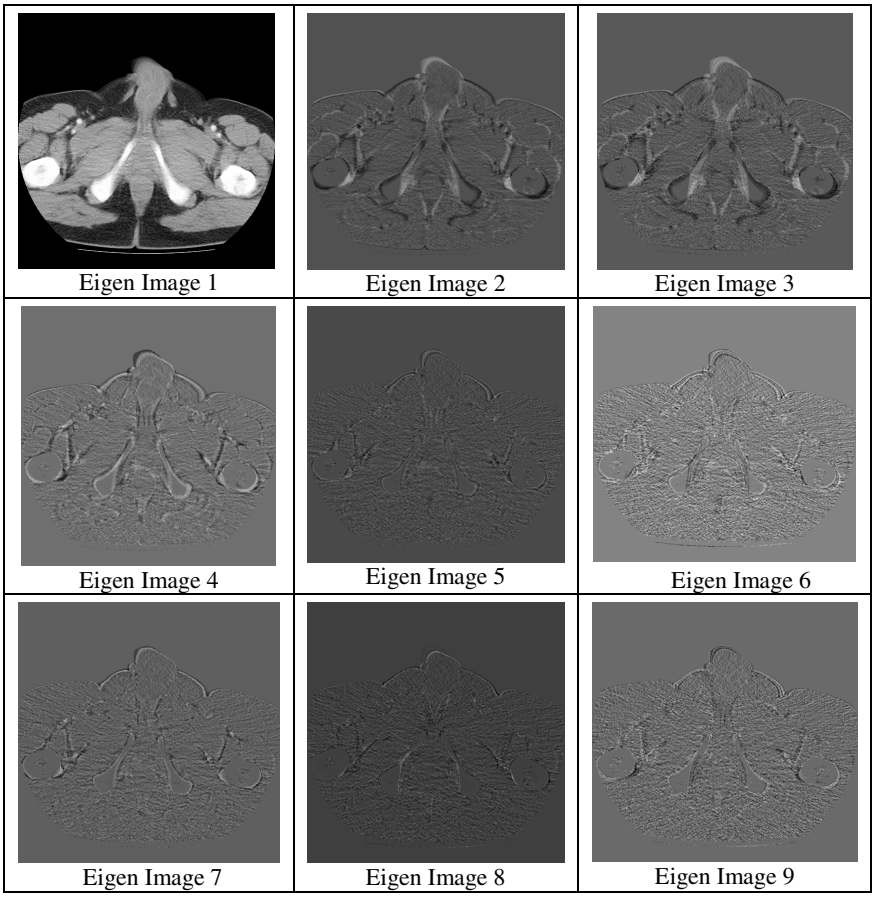
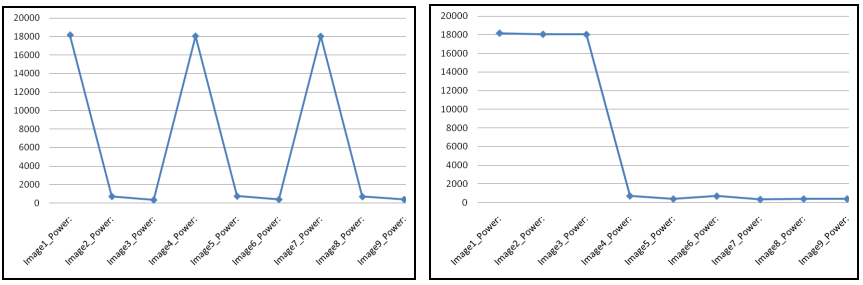


Fig. 4. Eigen images, obtained for Set 3 after performing 2-levels HAKLT

As it could be seen from the results on Fig. 4, on the first eigen image is concentrated the main part of the energy of all 9 images, and the energy of each next eigen image decreases quickly. This conclusion is confirmed by the data given in Table 2, where is given the power distribution of pixels of eigen images from Set 3 after first and second level of HAKLT, before and after their rearrangement in correspondence to Fig. 1. In Table 1 is given the power distribution of all eigen images in Set 3 before and after each operation and the relative mean power distribution. On the basis of data given in Table 1 are build the corresponding graphics, representing the power distribution of all 9 eigen images, shown correspondingly on Figs. 5 - 7.

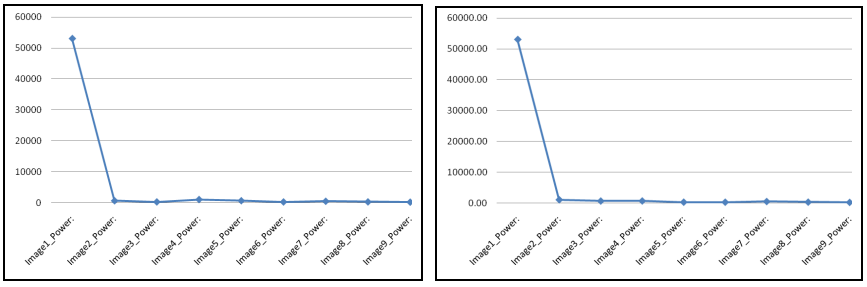
Table 1. Power distribution of all eigen images in Set 3 before and after each operation and relative mean power distribution

Name	Level 1 (not arranged)	Level 1 (arranged)	Level 2 (not arranged)	Level 2 (arranged)	Relative mean
Eigen Im. 1	18170	18170	53041	53041	219
Eigen Im. 2	715	18056	686	1100	5
Eigen Im. 3	341	18029	316	686	3
Eigen Im. 4	18056	715	1100	710	3
Eigen Im. 5	748	389	710	316	1
Eigen Im. 6	389	694	305	305	1
Eigen Im. 7	18029	341	523	523	2
Eigen Im. 8	694	389	326	326	1
Eigen Im. 9	394	394	242	242	1



a b

Fig. 5. Power distribution for Set 3, level 1: a - not arranged, b - arranged



a b

Fig. 6. Power distibution for Set 3, level 2: a - not arranged, b - arranged

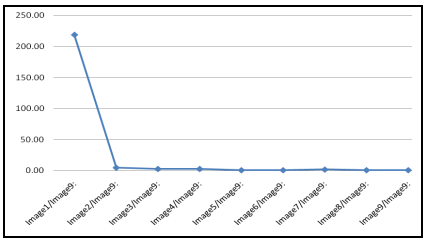
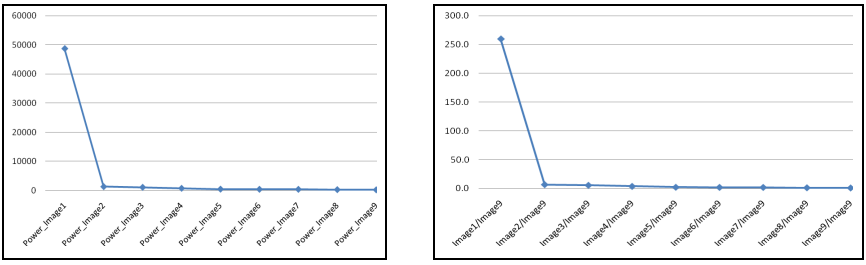


Fig. 7. Relative mean power distribution for Set 3, level 2 (arranged)

In Table 2 are given the mean and relative mean power distribution of pixels of all 9 eigen images in Set 1,.., Set 7 (R=7), and on Fig. 8 a, b - their corresponding graphic distributions. The data in the last column of Table 2 show, that in the first 3 eigen images are concentrated 95,7 % of the total mean power of all 9 images in GMI.

Table 2. Power Distribution, Mean Power Distribution, Relative Mean Power Distribution and Relative Mean % of Power Distribution for all eigen images in Set 1,.., Set 7

Name	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Mean	Relative mean	Reltive mean %
Eigen Im.1	49992	49749	53041	53547	53774	43272	37701	48725	259.6	91.4
Eigen Im.2	949	811	1100	875	2331	1770	1094	1276	6.8	93.8
Eigen Im.3	683	2325	686	1062	625	834	1144	1051	5.6	95.7
Eigen Im.4	808	710	710	512	460	811	950	709	3.8	97.1
Eigen Im.5	522	566	316	425	300	442	364	419	2.2	97.8
Eigen Im.6	350	529	305	306	317	402	435	378	2.0	98.6
Eigen Im.7	206	222	523	317	554	306	430	365	1.9	99.2
Eigen Im.8	172	198	326	261	312	251	218	248	1.3	99.6
Eigen Im.9	130	171	242	173	254	167	177	188	1.0	100.0



a. b.

Fig. 8. a. Mean Power Distribution; b. Relative Mean Power Distribution for all sets of eigen images

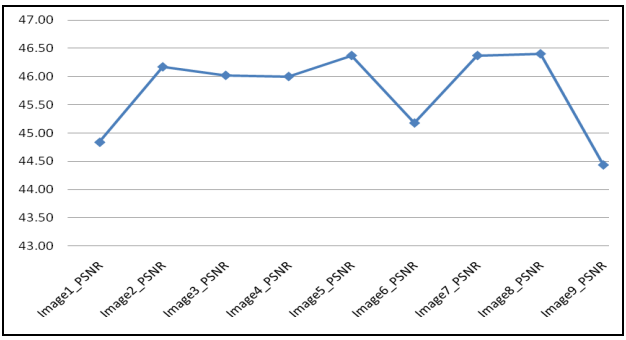


Fig. 9. Quality evaluation (through PSNR in dB) of the restored images from Set 3 after Inverse 2-level HAKLT on the eigen images from Fig. 4

From Fig. 8 b follows that the mean power of the first eigen image for all sets is more than 250 times larger than that of each of the next 8 eigen images.

The values for pixels of the eigen images, obtained in result of the direct 2-level HAKLT, were calculated with full accuracy, and after corresponding rounding could be transformed into 8-bit numbers. Then, if on the 8 bpp eigen images is applied the inverse 2-level HAKLT, the quality of corresponding restored images in GMI, evaluated by their peak signal-to-noise ratio (PSNR), is ≥ 45 dB. This was confirmed by the results from Fig. 9, obtained for the eigen images in Set 3 of Fig. 4 after inverse HAKLT in correspondence with the algorithm, shown on Fig.1. Hence, the sequence of 9 images could be restored with retained visual quality. This result illustrates the ability for efficient compression of a sequence of CT images, when HAKLT is used.

The experimental results were obtained with the software implementation of HAKLT, in Visual C.

5 Evaluation of the Computational Complexity

The computational complexity of the 2-level HAKLT algorithm, based on 3×3 matrices will be compared with that of the KLT algorithm with a matrix of size 9×9, because 2-level HAKLT is equivalent of the KLT for 9-component vector. For this, both algorithms are compared in respect to the performed number of operations S (additions and multiplications) [25] needed for the calculation of the following components:

- covariance matrices $[K_C]$ – in total 6 for the first algorithm, each of size 3×3, and one matrix $[K_C]$ of size 9×9 – for the second algorithm;
- eigen values and eigen vectors of the corresponding matrices $[K_C]$;
- eigen images of each GMI, obtained using both algorithms.

On the basis of the computational complexity analysis given in [24] for AKLT with matrix of size 3×3 and for KLT with a matrix of size $N \times N$ follows, that for the 2-level HAKLT with 3×3 matrices and for the KLT with a 9×9 matrix we have:

- The number of operations needed for the calculation of all elements k_{ij} for all 6 matrices $[K_C]$ of size 3×3 (for the 2-level HAKLT) and for one matrix $[K_C]$ of size 9×9 (for the KLT) is:

$$S_k(N)\big|_{N=3} = 3N(N+1)[N(N-1)+2(N+2)] = 576. \tag{24}$$

$$S_k(N)\big|_{N=9} = (1/2)N(N+1)[N(N-1)+2(N+2)] = 4230. \tag{25}$$

- The number of operations needed for the calculation of the eigenvalues of matrices $[K_C]$ for the 2-level HAKLT and of the $[K_C]$ matrix for KLT, when the QR decomposition and the Householder transform of *(N-1)* steps [24] were used, is:

$$S_{val}(N)\big|_{N=3} = 282. \tag{26}$$

$$S_{val}(N)\big|_{N=9} = (N-1)\left(\frac{4}{3}N^2 + \frac{17}{6}N + 7\right) = 1124. \tag{27}$$

- The number of operations needed for the calculation of the eigen vectors of matrices $[K_C]$ for the 2-level HAKLT and for the matrix $[K_C]$ of KLT, in case that iterative algorithm with 4 iterations is used, is correspondingly:

$$S_{vec}(N)\big|_{N=3} = 275. \tag{28}$$

$$S_{vec}(N)\big|_{N=9} = N[2N(4N+5) - 1] = 6633. \tag{29}$$

- The number of operations needed for the calculation of a group of 9 eigen images (each of P pixels), obtained in result of the direct 2-level HAKLT and of KLT for zero mean vectors, is correspondingly:

$$S_{HAKLT}(N)\big|_{N=3} = 6PN(2N-1) = 90P. \tag{30}$$

$$S_{KLT}(N)\big|_{N=9} = PN(2N-1) = 153P. \tag{31}$$

Then the total number of operations *SS* for the 2-level HAKLT and for KLT is correspondingly:

$$\begin{aligned} SS_1(3) &= [S_k(3) + S_{val}(3) + S_{vec}(3) + S_{HAKLT}(3)] = \\ &= 576 + 282 + 275 + 90P = 1133 + 90P, \end{aligned} \tag{32}$$

$$\begin{aligned} SS_2(9) &= [S_k(9) + S_{val}(9) + S_{vec}(9) + S_{KLT}(9)] = \\ &= 4239 + 1124 + 6633 + 153P = 11996 + 153P. \end{aligned} \tag{33}$$

The reduction of the total number of operations needed for the 2-level HAKLT, compared to that of the KLT could be evaluated using the coefficient η:

$$\eta(P) = \frac{SS_2(9)}{SS_1(3)} = \frac{11996 + 153P}{1133 + 90P}. \quad (34)$$

For example, for $P=100$ $\eta(100) = 2.96;$ for P=1000 correspondingly $\eta(1000) = 1.81$ and $\eta(\infty) \to 1.7.$ Hence, $SS_1(P)$ is at least 1.7 times smaller than $SS_2(P)$ for each value of P (in average, about 2 times).

6 Conclusions

The basic qualities of the offered HAKLT for processing a group of sequential medical images are:

1. Lower computational complexity than KLT for the whole GMI, due to the lower complexity of AKLT compared to the case, for which for the calculation of the KLT matrix are used numerical methods [15, 16];
2. Ability for efficient lossy compression of GMI with retained visual quality of the restored images and for lossless compression also;
3. Ability for minimization of features space in the regions of interest in a group of medical images, which contain searched objects of various kinds;
4. There is also a possibility for further development of the HAKLT algorithm, through: use of Integer KLT for lossless coding of medical images by analogy approach with [23]; compression of video sequences from stationary TV camera; sequences of multispectral and multi-view images, etc.

Acknowledgement. This paper was supported by the Joint Research Project Bulgaria-Romania (2010-2012): "Electronic Health Records for the Next Generation Medical Decision Support in Romanian and Bulgarian National Healthcare Systems" DNTS 02/19 and the Research Project No 122PD0061-07, supported by RD sector of TU-Sofia.

References

1. Gonzales, R., Woods, R.: Digital Image Processing, 3rd edn. Prentice Hall (2008)
2. Rangayyan, R.: Biomedical Image Analysis. CRC Press, Boca Raton (2005)
3. Dougherty, G.: Digital Image Processing for Medical Applications. Cambridge University Press (2009)
4. Graham, R., Perris, R., Scarsbrook, A.: DICOM demystified: A review of digital file formats and their use in radiological practice. Clinical Radiology 60, 1133–1140 (2005)
5. Wu, Y.: Medical image compression by sampling DCT coefficients. IEEE Trans. on Information Technology in Biomedicine 6(1), 86–94 (2002)

6. Ramesh, S., Shanmugam, D.: Medical image compression using wavelet decomposition for prediction method. International Journal of Computer Science and Information Security (IJCSIS) 7(1), 262–265 (2010)
7. Lalitha, Y., Latte, M.: Image compression of MRI image using planar coding. Intern. J. of Advanced Computer Science and Applications (IJACSA) 2(7), 23–33 (2011)
8. Roos, P., Viergever, M.: Reversible interframe compression of medical images: a comparison of decorrelation methods. IEEE Trans. Medical Imaging 10(4), 538–547 (1991)
9. Reed, T. (ed.): Digital Image Sequence Processing, Compression, and Analysis. CRC Press (2004)
10. Szilágyi, S.M., Szilágyi, L., Benyó, Z.: Echocardiographic Image Sequence Compression Based on Spatial Active Appearance Model. In: Rueda, L., Mery, D., Kittler, J. (eds.) CIARP 2007. LNCS, vol. 4756, pp. 841–850. Springer, Heidelberg (2007)
11. Miaou, S., Ke, F., Chen, S.: A Lossless Compression Method for Medical Image Sequences Using JPEG-LS and Interframe Coding. IEEE Trans. on Inform. Techn. in Biomedicine 13(5), 818–821 (2009)
12. Bitaa, I., Barretb, M., Phamc, D.: On Optimal Transforms in Lossy Compression of Multicomponent Images with JPEG2000. SP 90(3), 759–773 (2010)
13. Thirumalai, V.: Distributed Compressed Representation of Correlated Image Sets, Thesis No 5264, Lausanne, EPFL (2012)
14. Jain, A.: A fast Karhunen-Loeve transform for a class of random processes. IEEE Trans. Commun. COM-24, 1023–1029 (1976)
15. Dony, R.: Karhunen-Loeve Transform. In: Rao, K., Yip, P. (eds.) The Transform and Data Compression Handbook, ch. 1. CRC Press, Boca Raton (2001)
16. Jolliffe, I.: Principal Component Analysis, 2nd edn. Springer, NY (2002)
17. Ujwala, P., Uma, M.: Image Fusion using Hierarchical PCA. In: Intern. Conf. on Image Information Processing (ICIIP), pp. 1–6 (2011)
18. Hanafi, M., Kohler, A., Qannari, E.: Shedding new light on Hierarchical Principal Component Analysis. J. of Chemometrics 24(11-12), 703–709 (2010)
19. Langs, G., Bischof, H., Kropatsch, W.G.: Hierarchical Top Down Enhancement of Robust PCA. In: Caelli, T.M., Amin, A., Duin, R.P.W., Kamel, M.S., de Ridder, D. (eds.) SSPR&SPR 2002. LNCS, vol. 2396, pp. 234–243. Springer, Heidelberg (2002)
20. Grasedyck, L.: Hierarchical Singular Value Decomposition of Tensors, Preprint 20, AG Numerik/Optimierung, Philipps-Universitat Marburg, pp. 1–29 (2009)
21. Diamantaras, K., Kung, S.: Principal Component Neural Networks: Theory and Applications. John Wiley & Sons, NY (1996)
22. Solo, V., Kong, X.: Performance analysis of adaptive eigen analysis algorithms. IEEE Trans. Signal Processing 46(3), 636–645 (1998)
23. Hao, P., Shi, Q.: Reversible Integer KLT for Progressive-to-Lossless Compression of Multiple Component Images. IEEE ICIP, Barcelona (2003)
24. Kountchev, R., Kountcheva, R.: Image Color Space Transform with Enhanced KLT. In: Nakamatsu, K., Phillips-Wren, G., Jain, L.C., Howlett, R.J. (eds.) New Advances in Intelligent Decision Technologies. SCI, vol. 199, pp. 171–182. Springer, Heidelberg (2009)
25. Arora, S., Barak, B.: Computational Complexity: A Modern Approach. Cambridge University Press (2009)

Compression with Adaptive Speckle Suppression for Ultrasound Medical Images

Roumen Kountchev[1], Vladimir Todorov[2], and Roumiana Kountcheva[2]

[1] Technical University of Sofia, Department of Radio Communications and Video Technologies, Boul. Kl. Ohridsky 8, Sofia 1000, Bulgaria, rkountch@tu-sofia.bg
[2] T&K Engineering Co., Mladost 3, POB12, Sofia 1712, Bulgaria, {todorov_vl,kountcheva_r}@yahoo.com

Abstract. In the paper is presented one new approach for efficient processing of ultrasound medical images. The application of the algorithm for image compression based on the inverse difference pyramid (IDP) permits together with considerable compression ratio to achieve suppression of the specific (speckle) noise in ultrasound medical images. The paper describes the principle of image decomposition and its modification, designed for this medical application. Special attention is paid to achieve relatively low computational complexity of the used algorithms. Besides, an adaptive filtration aimed at the visual quality improvement of the restored image is also included. At the end of the paper are given experimental results and comparison with other contemporary methods for image archiving based on the JPEG and JPEG 2000 standards.

Keywords: Image compression, Speckle noise suppression, Digital adaptive fuzzy filter.

1 Introduction

Contemporary medical information comprises different kinds of documents: texts, images, biomedical data, etc. Significant attention require the visual information, obtained through various technologies, such as radiography, magnetic resonance imaging, fiduciary markers, nuclear medicine, photo acoustic imaging, breast thermography, tomography, and ultrasound, because it needs large volumes of computer memory and in together with this – special tools for image quality improvement. For medical images archiving is used DICOM, which is based on the JPEG standard [1,2]. This approach offers significant image data compression and retained visual quality. Together with these advantages, there also exist some difficulties. For example, there are no special tools for image content protection and for large images (for example, X-ray chest images) – no tools for selection and fast access to pre-selected regions of interest. In [3, 4] are introduced methods based on the Laplacian and wavelet decompositions. They all have very high computational complexity because of the use of the wavelets, which require longer time for processing of the visual information.

Most medical images are of low contrast and possess some specific noises. Ultrasound images are relatively large part of this information. They give significant

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 57–68.
DOI: 10.1007/978-3-319-00029-9_5

information about patients' body, but togetherr with this, they have many problems concerning the image quality and processing. Speckle noise is an inherent property of medical ultrasound imaging. Ultrasound beams propagating in biological tissues undergo distortions due to local inhomogeneities of the acoustic parameters and the nonlinearity of the medium. This usually results in reduction of the image resolution and contrast and as a consequence – in reduction of the diagnostic value of this imaging modality. So, the speckle noise reduction is an important prerequisite, whose significance is as high as that of the image compression. Various methods had already been developed aimed at the speckle noise removal. In [5] is introduced a speckle-reduction method, based on soft thresholding of the wavelet coefficients of a logarithmically transformed medical ultrasound image. The method is based on the generalised Gaussian distributed (GGD) modelling of sub-band coefficients. In [6], instead of using the multiplicative model of speckled image formation is introduced a simple preprocessing procedure, which modifies the acquired radio-frequency images, so that the noise in the log-transformation domain becomes very close in its behavior to a white Gaussian noise. As a result, the preprocessing allows filtering methods based on assuming the noise to be white and Gaussian, to perform in nearly optimal conditions. One more approach is presented in [7], where the logarithmic transform of the original image is first analyzed into the multiscale wavelet domain and then the authors design a Bayesian estimator to exploit the specific image statistics. In [8] is proposed a robust wavelet domain method for noise filtering in medical images. All these methods have significant computational complexity.

In this paper is introduced a method based on the Inverse Pyramid Decomposition (IPD), which permits to perform also adaptive suppression of the speckle noise in ultrasound images. The paper is arranged as follows: in section 2 is given a brief description of the IPD; in Section 3 is presented the approach used for the speckle noise suppression; Section 4 presents an adaptive fuzzy filter for image quality enhancement; in Section 5 are given some experimental results and Section 6 contains the Conclusions.

2 Basic Principles of IPD

The essence of the IPD is explained here in brief for 8-bit grayscale images. The image matrix is processed with two-dimensional, orthogonal transform using a limited number of coefficients only. The values of the transform coefficients build the first (lowest) pyramid level. Then, using these values, the image is restored with inverse orthogonal transform and subtracted pixel by pixel from the original. The difference image, which is of same size as the original, is divided into four equal parts (sub-images) and each is then processed with the two-dimensional, orthogonal transform again. Then each sub-image is restored and subtracted from the preceding difference image. Afterwards, it is divided into four new sub-images, which are processed in similar way as the already described preceding levels. In this way, all pyramid levels, consisting of coefficients values only, are calculated. The set of the selected coefficients of the orthogonal transform for every pyramid level could be different. The last

pyramid level is reached when the size of constituting sub-images remains only 2×2 pixels. In practice, the image decomposition stops when the required image quality is reached - usually earlier than the last possible pyramid level. The coefficients resulting from the orthogonal transform from all pyramid levels are sorted in accordance with their frequency, and losslessly compressed. The lossless compression technique, used here, is developed by taking into account the specific features of the data, obtained in result of the IPD.

High image quality is obtained using large number of spectrum coefficients and pyramid levels. The image quality deteriorates together with the reduction of the number of the used coefficients, but the compression ratio is higher. To restore the image, all the operations have to be performed in reverse order.

The algorithm is suitable for processing 24-bit color bmp images, as well. In this case, for each of the color components (for example, *Y*, *U* and *V*) is built individual pyramid.

The processing could be explained as follows. Let the input image be represented by the matrix [*B (i, j)*] of size $2^n x 2^n$ for *i, j = 0, 1, 2,...,* 2^n *-1*, and the luminance of every pixel is *B (i, j)*$\in$ [*0,1, 2, ..., 255*]. The image element *B(i, j)* is:

$$B(i,j) = \tilde{B}_0(i,j) + \sum_{p=1}^{n-1} \tilde{E}_{p-1}^{k_p}(i,j) \ , \tag{1}$$

where $k_p = 1,2,...,4^p$ is the number of the corresponding sub-image W_{k_p} consisting of $2^{n-p} x 2^{n-p}$ pixels in pyramid level p, divided into 4^p equal blocks (sub-images). The term $\tilde{B}_0(i,j)$ is the "zero" IPD component, which corresponds to the top of the inverse pyramid and is represented as:

$$\tilde{B}_0(i,j) = 4^{-n} \sum_{k=0}^{(d/2)-1} \sum_{l=0}^{(d/2)-1} s(k,l)\, t(k,l,i,j) \ \text{ for } \quad p=0, \tag{2}$$

where

$$s(k,l) = \sum_{i=0}^{2^n-1} \sum_{j=0}^{2^n-1} B(i,j)\, t(k,l,i,j) \quad \text{for} \quad k, l = 0, 1, \ldots, (d/2)-1. \tag{3}$$

Here *t(k, l, i, j)* is the two-dimensional function with spatial frequency *(k, l),* defined by the selected orthogonal transform; *s(k, l)* is the spectrum coefficient obtained after the orthogonal transform; *d* is the number of the spectrum coefficients used for the inverse orthogonal transform.

The term $\tilde{E}_{p-1}^{k_p}(i,j)$ in Eq. (1) is the IPD component obtained after low-frequency filtration of each block $E_{p-1}^{k_p}(i,j)$ in the difference $E_{p-1}(i,j)$ from pyramid level *p*, when *p*≥*1*. The difference is defined as follows:

$$E_0(i,j) = B(i,j) - \tilde{B}_0(i,j) \qquad \text{for } p=1, \tag{4}$$

$$E_{p-1}(i,j) = E_{p-2}(i,j) - \tilde{E}_{p-2}(i,j) \text{ , for } p=2,3, ..., n-1 \tag{5}$$

The term $\tilde{E}^{k_p}_{p-1}(i,j)$ is defined in the same way as the "zero" component, using the low-frequency coefficients in the spectrum of sub-images $E^{k_p}_{p-1}(i,j)$, obtained in result of dividing differences $E_{p-1}(i,j)$ into 4^p blocks W_{k_p}, i. e.

$$\tilde{E}^{k_p}_{p-1}(i,j) = 4^{n-p} \sum_{k=0}^{(d_p/2)-1} \sum_{l=0}^{(d_p/2)-1} s_{k_p}(k,l)\, t_{k_p}(k,l,i,j),$$

where (6)

$$s_{k_p}(k,l) = \sum_{(i,j)\in W_{k_p}} \sum E^{k_p}_{p-1}(i,j)\, t_{k_p}(k,l,i,j) \quad \text{for} \quad k,l = 0,1,\ldots,(d_p/2)-1.$$

In order to decrease the information excess in the decomposition, statistic features of the image and especially – the correlation between the neighbor pixels is used. In the case of same resolution in horizontal and vertical direction, the correlation interval for natural images in each direction is about 30-40 pixels. The fact that the human eye is less sensitive to luminance transitions with diagonal orientation ensures additional information excess. This permits decomposition reduction without noticeable deteriorations in the restored image, performing the following:

- The pyramid is "truncated" from the bottom. For this, the number of levels is limited.
- The number of coefficients in the spectrum for the sub-images W_{k_p} in the next pyramid levels is decreased removing these with small amplitudes.

After pyramid levels truncation retaining levels from *4* up to *(n-1),* for Eq. (1) is obtained:

$$B_{4,n-1}(i,j) = \tilde{B}^{k_4}_3(i,j) + \sum_{p=5}^{n-1} \tilde{E}^{k_p}_{p-1}(i,j). \tag{7}$$

The compression ratio for image, represented with a reduced pyramid with levels *4, 5,...,(n-1),* using *4* coefficients for every sub-image up to the level *(n-2)* and *2* coefficients for the last level *(n-1),* is defined in correspondence with the equation:

$$K(p) = 4^n / N_p \ , \tag{8}$$

where N_p is the total number of used coefficients, and for N_p is obtained:

$$N_{4,n-1} = 4\sum_{i=4}^{n-2} 4^i + 2.4^{n-1} = \frac{1}{3}4^5\left(\frac{5}{2}4^{n-5} - 1\right) \approx \frac{5}{6}4^n \qquad \text{for } n \geq 9. \tag{9}$$

3 Modified IPD for Speckle Noise Suppression

The application of the IDP method for US medical images requires adaptation to their statistical properties. In research works [9,10] is investigated that speckles mean statistical size for most frequently used ultrasound images, is smaller than 8 pixels. On the other hand, this size (8 pixels) is big enough to ensure successful boundary restoration in the speckled US medical images. For these reasons, in this work is offered a modification of the IDP decomposition, which consists of two levels only. The initial level ("0") is calculated for sub-images of size 8×8, and the next level ("1") - for sub-images of size 4×4. This selection permits speckle suppression, together with retaining the image boundaries. The choice of pyramid of two levels was made taking into consideration the ability to achieve maximum compression ratio. Specific for this approach is that the two pyramid levels use different orthogonal transforms: in the lower is used DCT and in the higher level – Walsh-Hadamard transform (WHT). The DCT transform ensures the highest energy concentration in the low-frequency coefficients of the spectrum compared with other determinated transforms; the WHT transform has lower computational complexity and for small sizes *(4×4),* it is practically equivalent to DCT in respect to approximation efficiency [11].

The description of the modified method is given below. Let us assume that the image matrix [*B*] of size M×N pixels is divided into blocks of *8×8* pixels. Then the *B(i,j)* pixel of each block is approximated with the pixel *B'(i,j)* from the restored image, in accordance with the relation:

$$B'(i,j) = B'_0(i,j) + E'_0(i,j) \quad \text{for } i,j=0,1,\ldots,7. \tag{10}$$

The two components in Eq. (10) correspond to decomposition levels 0 and 1. The first component is calculated using inverse DCT:

$$B'_0(i,j) = \frac{1}{4}\sum_{k=0}^{7}\sum_{l=0}^{7} C(k)C(l)s'_f(k,l)t_c(k,l,i,j), \tag{11}$$

where: $C(n) = \begin{cases} 1/\sqrt{2} & \text{for} \quad n=0; \\ 1 & \text{for} \quad n=1,2,..,7. \end{cases}$

The *2D* transform functions in Eq. (11) are represented using two-dimensional cosine functions [11]:

$$t_c(i,j,k,l) = \cos\left[\frac{(2i+1)\pi k}{16}\right]; \qquad \cos\left[\frac{(2j+1)\pi l}{16}\right]. \tag{12}$$

Here $s'_f(k,l)$ is the filtered coefficient of the block transform, obtained after quantization/dequantization of $s_f(k,l)$, i.e.:

$$s'_f(k,l) = s_{fq}(k,l)q_0(k,l), \tag{13}$$

$$s_{fq}(k,l) = [s_f(k,l)/q_0(k,l)]_{integer}. \tag{14}$$

The coefficient $q_0(k,l)$ is from the quantization matrix $[Q_0]$ of size 8×8, applied at level *p=0*, and used to reduce the number of "meaning" coefficients, i.e. – to enhance the compression. The quantization matrix was selected in correspondence with results obtained from experiments with medical ultrasound images. On the other hand, the coefficient $s'_f(k,l)$ in Eq. (14) is defined as:

$$s_f(k,l) = s(k,l)m_0(k,l) = \begin{cases} s(k,l) & if \quad m_0(k,l) = 1; \\ 0 & if \quad m_0(k,l) = 0, \end{cases} \tag{15}$$

where $q_0(k,l)$ is one element from the binary filtering matrix $[Q_0]$ with dimensions 8×8, used for IDP level *p=0*. This matrix defines the place of the retained coefficients $s_f(k,l)$ in the transform of each block. The matrix was chosen in result of investigating the compression ratio and filtration efficiency for the specific ultrasound image noise (speckles). The coefficient $s_f(k,l)$ in Eq. (15) is defined using direct DCT, applied on the original image *B (i, j):*

$$s(k,l) = \frac{1}{4}\sum_{i=0}^{7}\sum_{j=0}^{7} C(k)C(l)B(i,j)t_c(i,j,k,l) \quad \text{for} \quad k,l=0,1,\ldots,7. \tag{16}$$

The second component in Eq. (10), corresponding to pyramid level *p=1*, is:

$$E'_0(i,j) = B(i,j) - B'_0(i,j) \qquad \text{for} \quad i,j=0,1,\ldots,7. \tag{17}$$

Each block is divided into 4 sub-blocks of *4×4* pixels and is processed with WHT. The transform coefficients are:

$$s(k,l) = \frac{1}{16}\sum_{i=0}^{3}\sum_{j=0}^{3} E'_0(i,j)t_w(i,j,k,l) \qquad \text{for} \quad k,l=0,1,2,3, \tag{18}$$

where the basic image is a two-dimensional Walsh-Hadamard function:

$$t_w(i,j,k,l) = (-1)^{\sum_{r=0}^{2}[q_r(k)i_r \oplus q_r(l)j_r]}, \text{ for } k = \sum_{r=0}^{1} k_r 2^r \text{ and } l = \sum_{r=0}^{1} l_r 2^r, \tag{19}$$

$q_1(k) = k_1 \oplus k_0$, $q_0(k) = k_1$, $q_1(l) = l_1 \oplus l_0$, $q_0(l) = l_1$.

The WHT coefficients are filtered in similar way as in the preceding level:

$$s_f(k,l) = s(k,l)m_1(k,l) = \begin{cases} s(k,l) & if \quad m_1(k,l) = 1; \\ 0 & if \quad m_1(k,l) = 0, \end{cases} \tag{20}$$

where $m_1(k,l)$ is one element of the binary matrix $[M_1]$, which defines the places of retained coefficients for each sub-block. The matrix was chosen in a way, similar

to $[M_0]$. The coefficients of the blocks (sub-blocks) from each pyramid level are arranged in accordance to their spatial frequencies, quantizated and losslessly coded.

4 Fuzzy Adaptive Digital Filter

Here is given one approach for post-processing of decompressed still images, based on the use of fuzzy digital filters [12, 13], which became very popular recently.

The algorithm, which describes the performance of the presented two-dimensional adaptive fuzzy filter (2DAFF), specially developed for the visual quality enhancement of restored images after IPD, is given by Eq. (21). The filter uses a sliding window of size $M \times N$ pixels ($M=2R+1$ and $N=2S+1$):

$$x_F(i,j)=\begin{cases} \left\lfloor \dfrac{\sum_{r=-R}^{R}\sum_{s=-S}^{S}\mu(i+r,j+s)x(i+r,j+s)}{\sum_{r=-R}^{R}\sum_{s=-S}^{S}\mu(i+r,j+s)} \right\rfloor & \text{for } \sum_{r=-R}^{R}\sum_{s=-S}^{S}\mu(i+r,j+s)\geq T, \\ \left\lfloor (1/MN)\sum_{r=-R}^{R}\sum_{s=-S}^{S}x(i+r,j+s) \right\rfloor & \text{- in all other cases} \end{cases} \tag{21}$$

where $\lfloor \circ \rfloor$ is the rounding operator; T - a threshold; and $x(i,j)$ and $x_F(i,j)$ are pixels belonging correspondingly to the input and to the filtered image;

$$\mu[(i+r,j+s)] = \begin{cases} 1 & \text{for} \quad \Delta(i+r,j+s) \leq \alpha; \\ \dfrac{\Delta(i+r,j+s)-\beta}{\alpha-\beta} & \text{for } \alpha \leq \Delta(i+r,j+s) \leq \beta; \\ 0 & \text{for} \quad \Delta(i+r,j+s) \geq \beta, \end{cases} \tag{22}$$

is the selected membership function with parameters α and β ($\beta > \alpha$), which define the fuzziness area. The argument Δ is the module of the difference between the central pixel $x(i, j)$ in the filter window and the pixel $x(i+r, j+s)$, which is at a distance (r, s) from this pixel:

$$\Delta(i+r,j+s)=\left| x(i,j)-x(i+r,j+s) \right| \text{ for } r=\overline{-R,R} \text{ and } s=\overline{-S,S}. \tag{23}$$

The values of parameters α and β are defined in accordance to image contents and to distortions, which should be corrected. In the case, when they are block artifacts, α and β are defined in accordance with the compression strength and the kind of the used compression algorithm. Here the 2DAFF filter is applied on images, with block

artifacts, due to IPD decomposition. The image quality improvement starts from CR higher than 10. As a rule, the filter parameters should not be too small so that the noisy edges can be sufficiently smoothed. The basic filter parameters were set as follows: filter fuzziness center equal to ½ of the maximum difference and filter width was 5 (experimentally the center of the filter fuzziness was set to be 65, i.e. approximately, equal to half of the maximum error in the approximating image).

5 Experimental Results

The compression ratio for an image, coded with the modified IDP without using lossless coding for coefficients' values, was calculated in accordance with the equation:

$$K = 64 /[\sum_{k=0}^{7}\sum_{l=0}^{7} m_0(k,l) + \frac{1}{4}\sum_{k=0}^{3}\sum_{l=0}^{3} m_0(k,l)] \quad . \tag{24}$$

As an example, for $M=N=512$, and $\sum_{k=0}^{7}\sum_{l=0}^{7} m_0(i,j) = 8, \sum_{k=0}^{3}\sum_{l=0}^{3} m_1(i,j) = 3$, from Eq. (24) was obtained $K=7.31$. After lossless coding of $s_{fq}(k,l)$ coefficients, this ratio is increased 4-6 times.

Advantage of the described method is that it permits easy removal of certain spectral coefficients, eliminating this way the corresponding most significant noise components in the restored image. The investigations show that the predominant part of the specific noises existing in ultrasound images (speckles), are of size 2×1, 1×2, 2×2, 2×3 or 2×0, 0×2 pixels or their enlarged (scaled) equivalents. When the values of the corresponding spectral components are made equal to zero in conformity with Eqs. (15) and (20), the visual quality of the restored medical image is noticeably improved. In this case, the best evaluation of the image quality is the visual one - when the peak mean square error is calculated (or if some other evaluation technique is used) the calculated error will be large because the noise removal will be assumed as quality deterioration.

The noise suppression depends on the choice of the filtrating matrices $[M_0]$ and $[M_1]$, of size 8×8 and 4×4 respectively. On Fig. 1a,b are shown several matrices, which were selected in result of experiments with big number of ultrasound medical images.

On Fig. 2 are shown two-dimensional orthogonal functions with highest correlation with the speckle noise structure, suppressed in accordance with Matrix $[M_0]$-*1* selection (Fig.1a).

The ability to select different combinations of two-dimensional coefficients for the filtration permits to suppress speckles with different size and shape. In case when the original image contains bigger speckles, it is possible to use pyramid with three or even more levels in order to increase the size of the filtered patterns and to suppress bigger forms. Usually this does not decrease significantly the compression ratio.

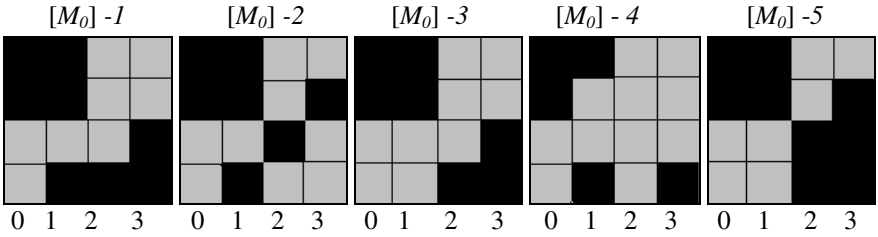


Fig. 1.a. Five matrices [M_0]-*1,2,..,5* of size *8×8* elements (here are shown only the 16 low-frequency coefficients of these matrices; the remaining 48 are assumed equal to zero). Black square - retained coefficient; white square - suppressed coefficient.

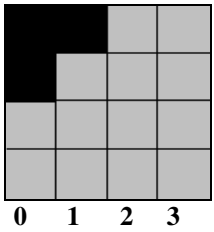


Fig. 1.b. Matrix [M_1] of size *4×4* elements

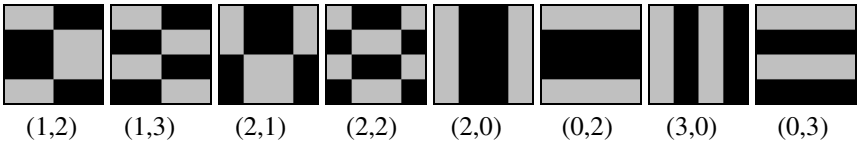


Fig. 2. Examples of basic two-dimensional Walsh-Hadamard functions *(k,l)* of size *8×8* pixels, which best correspond to speckle noise characteristics and are suppressed by the matrix [M_0] *-1*.

The modified IPD method for compression and speckle noise suppression of US medical images was compared with JPEG-based algorithm (version of Microsoft Photo Editor in Microsoft Office XP Professional). The results obtained show that for compression ratios up to 50:1 both methods are practically equivalent. The main advantages of the modified IPD method are that it offers the ability for speckle noise suppression together with the compression and the lower computational complexity, because the WHT transform needs smaller number of calculations. Besides, for higher compression ratios (above 50) the visual quality of restored images after IPD compression is much better than that for JPEG or even for JPEG 2000. For illustration only on Fig. 3 is shown the restored test image "Axial" after compression of 137:1 with IDP and JPEG, and on Fig.4 – after compression 24:1 (for this compression all images are visually same).

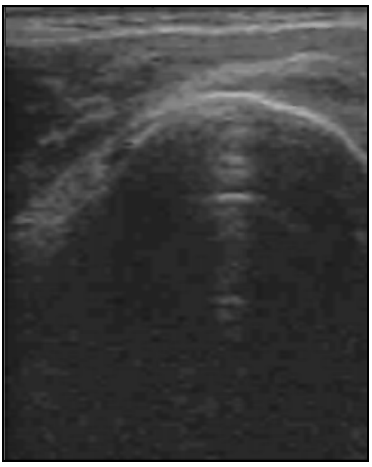


Fig. 3.a. Test image "axial" after IDP compression 137:1

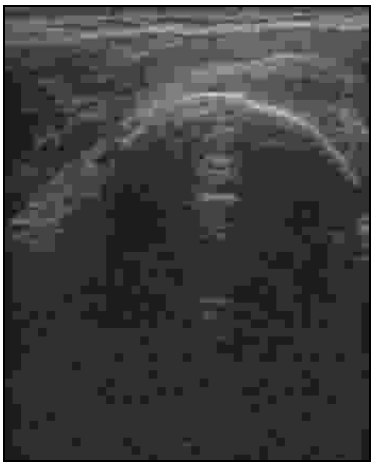
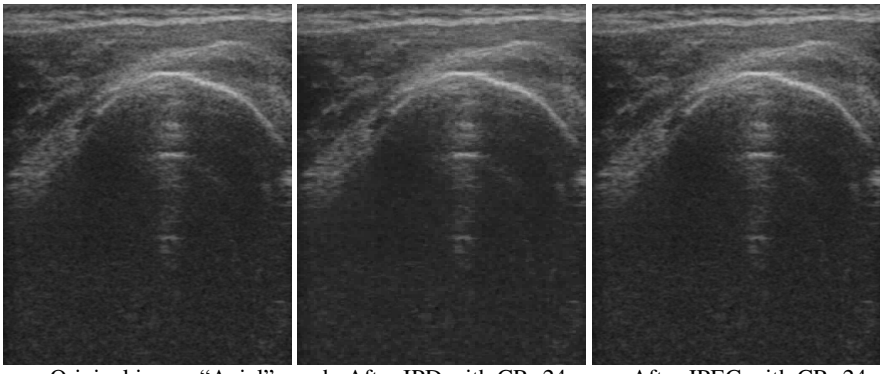


Fig. 3.b. Test image "axial" after JPEG compression 137:1



a. Original image "Axial" b. After IPD with CR=24 c. After JPEG with CR=24

Fig. 4. Comparison of IPD and JPEG obtained for compression ratio (CR) = 24

For the example image on Fig. 4 the PSNR is not calculated, because the suppression of the speckle noise will result in lower value. The visual quality of the image is retained. The detailed description of the influence of the digital filter parameters on the restored image quality is given in earlier publication of the authors [15].

6 Conclusions

A new method for compression and speckle suppression for medical ultrasound imaging is presented. The main advantages of the method are:

- The method permits ultrasound medical images to be compressed with high compression ratio, together with significant speckle noise suppression;

- A set of filter matrices is selected which offers high flexibility for various medical applications.
- The method is suitable for processing and archiving of medical images used in image databases.
- The method could be used successfully for other kinds of medical images (roentgen, etc.) as well.
- The computational complexity of the new method is much lower than that of JPEG 2000. Corresponding comparison is given in earlier publications of the authors [14].

The future development of the method could be in the following directions:

- Enhancement of the processing efficiency for color images, using the KL transform;
- Enhancement of the noise suppression with pre- and post- image processing.

Acknowledgement. This work was supported by the Joint Research Project Bulgaria-Romania (2010-2012): "Electronic Health Records for the Next Generation Medical Decision Support in Romanian and Bulgarian National Healthcare Systems".

References

1. Taubman, D., Marcellin, M.: JPEG2000: Image Compression Fundamentals, Standards and Practice. Kluwer Academic Publishers, Boston (2002)
2. Information Technology–JPEG 2000 Image Coding System: Part 9–Interactivity Tools, APIs and Protocols, no.15444-9, ISO/IECJTC1/SC29/WG1IS, Rev. 3 (2003)
3. Cincotti, G., Loi, G., Pappalardo, M.: Frequency Decomposition and Compounding of Ultrasound Medical Images with Wavelet Packets. IEEE Trans. on Medical Imaging 20(8), 764–771 (2001)
4. Saad, A.: Simultaneous Speckle Reduction and Contrast Enhancement for Ultrasound Images: Wavelet Versus Laplacian Pyramid. Pattern Recognition and Image Analysis 18(1), 63–70 (2006)
5. Gupta, S., Chauhan, R., Sexana, S.: Wavelet-based Statistical Approach for Speckle Reduction in Medical Ultrasound Images. Medical and Biological Engineering and Computing 42(2), 189–192 (2004)
6. Michailovich, V., Tannenbaum, A.: Despeckling of Medical US Images. IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control 53(1), 64–78 (2006)
7. Achim, A., Bezerianos, A., Tsakalides, P.: Novel Bayesian Multiscale Method for Speckle Removal in Medical Ultrasound Images. IEEE Trans. on Medical Imaging 20(8), 772–783 (2001)
8. Pizurica, A., Philips, W., Lemahieu, I., Acheroy, M.: A Versatile Wavelet Domain Noise Filtration Technique for Medical Imaging. IEEE Trans. on Medical Imaging 22(3), 323–331 (2003)
9. Dutt, V., Greenleaf, J.: Speckle Analysis Using Signal to Noise Ratios Based on Fractional Order Moments. In: Ultrasonic Imaging, vol. 17, pp. 251–268. Academic Press (1995)
10. Czerwinski, R., Jones, D., O'Brien, W.: Detection of Lines and Boundaries in Speckle Images–Application to Medical Ultrasound. IEEE Trans. on Medical Imaging 18(2), 126–136 (1999)

11. Rao, K., Yip, P.: Discrete Cosine Transform: Algorithms, Advantages, Applications. Academic Press (1990)
12. Nachtegael, M., et al.: Fuzzy Filters for Image Processing. Springer (2003)
13. Nie, Y., Barner, K.: Optimized Fuzzy Transformation for Image Deblocking. IEEE ICME 1, 541–544 (2003)
14. Kountchev, R., Milanova, M., Ford, C., Kountcheva, R.: Multi-layer Image Transmission with Inverse Pyramidal Decomposition. In: Halgamuge, S., Wang, L. (eds.) Computational Intelligence for Modeling and Predictions, vol. 2(13), pp. 179–196. Springer (2005)
15. Kountchev, R., Milanova, M., Todorov, V., Kountcheva, R.: Adaptive Fuzzy Filter for Reduction of Blocking Artifacts in Images Compressed with IDP Decomposition and JPEG. WSEAS Trans. on Signal Processing 7(2), 941–948 (2006)

Adaptive Approach for Enhancement the Visual Quality of Low-Contrast Medical Images

Vladimir Todorov and Roumiana Kountcheva

T&K Engineering Co., Mladost 3, POB12, Sofia 1712, Bulgaria
{todorov_vl,kountcheva_r}@yahoo.com

Abstract. In the paper is presented one specific approach aimed at improvement of the visual quality of underexposed or low-contrast medical images. For this are developed adaptive contrast-enhancement algorithms, based on the segmentation of the image area with relatively high density of dark elements. The problem is solved changing the brightness intervals of the selected segments followed by equalization (in particular – linear stretch and skew) of the corresponding parts of the histogram. The implementation is relatively simple and permits easy adaptation of the contrasting algorithms to image contents, requiring setting of small number of parameters only. The corresponding software tools permit to change the image with consecutive steps and to evaluate the visual quality of the processed images. The original image is also available, which permits easy comparison and evaluation. The obtained results prove the efficiency of the new methods for image quality enhancement.

Keywords: Image contrast enhancement, Adaptive contrast enhancement, Image segmentation.

1 Introduction

Contemporary medical diagnostic is highly related to image analysis. However, most of these images are of relatively bad quality, which complicates the diagnostic process. In most cases these images have low contrast and this makes their understanding too difficult. The problem with their quality enhancement is of high importance and had been explored by many researchers. The usual approach is based on histogram equalization [1-6]. This method increases the global contrast of the processed images, especially when the usable data of the image is represented by close contrast values. In result, the intensities are better distributed on the histogram. This allows the dark (low-contrast) areas to gain a higher contrast. Histogram equalization accomplishes this by effectively spreading out the most frequent intensity values. In particular, the method leads to better views of bone structure in x-ray images. In theory, if the histogram equalization function is known, then the original histogram can be recovered and the calculation is not computationally intensive. A disadvantage of the method is that it is indiscriminate. It may increase the contrast of background noise, while decreasing the usable signal. However the method is very useful for scientific images like thermal, satellite or x-ray images. Also histogram equalization

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 69–78.
DOI: 10.1007/978-3-319-00029-9_6

can produce undesirable effects (like visible image gradient) when applied to images with low color depth. For example, if applied to 8-bit image displayed with 8-bit gray-scale palette it will further reduce color depth (number of unique shades of gray) of the image. Histogram equalization will work the best when applied to images with much higher color depth than palette size, for example, for 12- or 16-bit gray-scale images.

The local contrast enhancement is aimed at modification of the local characteristics of the processed image. Specific for this approach is that it increases "local" contrast in smaller regions, while at the same time preventing an increase in "global" contrast — thereby protecting large-scale shadow/highlight detail. Unfortunately, it achieves this by making some pixels in the histogram cross over each other, which is not possible when enhancing contrast using levels or curves. Local contrast enhancement works similarly to sharpening with an unsharp mask; however the mask is instead created using an image with a greater blur distance. This results in a local contrast mask which maps larger-scale transitions than the small-scale edges which are mapped when sharpening an image. The quality improvement of low-contrast images is usually performed transforming the brightness of every pixel in accordance with a function of some kind [1-5]. These functions could be in general classified as follows:

- Linear, piecewise linear and non-linear transforms: logarithmic, hyperbolic, exponential, power-low transformations, etc.
- Statistical (adaptive) transforms: histogram equalization and modification, bi-histogram transform [6] and transforms, based on local statistics information (mean value, standard deviation, etc.);
- Membership functions for contrast adaptation by: minimization of image fuzziness; fuzzy histogram;
- Transforms of the kinds, pointed above, based on the global image histogram data or on the local histogram, framed by a window, placed around every pixel [8];
- Transforms, based on the local information analysis of the color position in the color space [9].

The image contents analysis is of high importance when the objects boundaries should be detected – for example, to define a tumor, or its real size [10-12]. In these cases the image contrast enhancement is an additional tool for medical decision support.

The main disadvantage of these methods is their high computational complexity.

Each of the methods for image contrast enhancement offers very good results for image quality improvement, when certain image classes are concerned, without being a universal tool for this aim.

In this work is presented one new approach for contrast enhancement in images with low-contrast areas, modifying their dynamic range. In section 2 are presented the investigated methods, section 3 contains part of the obtained results and section 4 presents the conclusions. The comparison of the new method with the traditional approaches for contrast enhancement proves its efficiency for image quality improvement and medical decision support.

2 Adaptive Contrast Enhancement

In this section are given 2 algorithms for adaptive contrast enhancement. Each algorithm has its advantages and disadvantages, but bound together in special software, they turn into an intelligent tool for visual quality enhancement and medical decision support. All algorithms are based on the image histogram analysis.

2.1 Two-Part Histogram Enhancement

The low-contrast medical ultrasound images usually contain relatively small part of the possible brightness values - usually, about 30%. The test image "axial" and its histogram are shown on Fig. 1.a,b.

Fig.1.a. Test ultrasound image "axial"

b. The histogram of the test image

The algorithm for two-part histogram enhancement is presented below:

- The image histogram is calculated in accordance with the relation:

$$h(k) = n(k)/n \quad \text{for} \quad k=0,1,2,\ldots,k_{max}, \tag{1}$$

where: $n(k)$ is the number of the pixels in the discrete brightness level k;
n is the total number of pixels in the image;
k_{max} is the maximum number of brightness levels.

- The maximum of the histogram is detected.

$$h_{max} = max\{h(k)\} \text{ for } k=0,1,2,\ldots,k_{max}. \tag{2}$$

- The ends of the "meaning" histogram values are detected. These are the brightness values, which represent number of points, larger than 0.1% of the pixels in

the image. The corresponding part of the histogram is defined by the blue lines on Fig. 2.

- Two points are calculated, which expand the histogram with approximately 30% in each direction, i.e. - higher or lower than the end "meaning" values (Fig. 2). The new end points are placed at δ_1 and δ_2 outside the "meaning" area, placed correspondingly. All other brightness values are excluded from the processing (in fact, usually there are no pixels of such brightness values in ultrasound images).

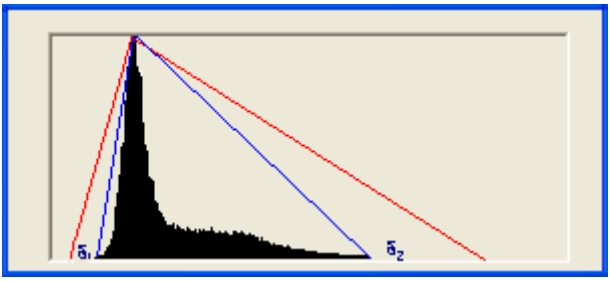


Fig. 2. The boundaries of the expanded image histogram

- The brightness of each pixel is then recalculated, using special tables following the corresponding linear relations.

2.2 Three-Part Histogram Enhancement Based on Histogram-Adaptive Segmentation

In order to achieve easier processing, the image contrast is first evaluated. For this is used the contrast coefficient K, calculated in accordance with the relation below [3]:

$$K = \frac{k_{max} - k_{min}}{k_{max} + k_{min}} \tag{3}$$

where k_{max} and k_{min} are correspondingly the maximum and the minimum brightness levels in the processed image.

In case, that $K<K_0$ (where K_0 is a threshold value), defining the image as one of low contrast, then its brightness histogram $h(k)$ is strongly distorted and it has an outlined maximum, h_{max}. The image histogram is calculated in accordance with Eq. 1.

The method for image contrast enhancement comprises two consecutive steps: brightness image segmentation (respectively - gray level segmentation) analyzing its histogram $h(k)$, and brightness transformation for the pixels in the defined segments in accordance with table, corresponding to the segment histogram.

In the first step for the case, when $K<K_0$, is performed image segmentation, using the thresholds k_1 and k_2, in result of which the brightness range is divided into three segments (A, B, C). The assumption is that in the second segment B is clustered too big number of pixels with close brightness values, what distorts the image contrast. In order to improve the image quality, the dynamic range of this segment should be

"stretched", and the dynamic ranges of the two remaining segments – correspondingly "skewed". The algorithm for the calculation of the thresholds is presented as follows:

- The image histogram is calculated
- The maximum of the histogram is defined (Eq.2):
- The value of the segmentation boundary is calculated as $t=\alpha h_{max}$, where $\alpha<1$;
- The values of the thresholds k_1 and k_2 are defined in accordance to relations below:

$$h(k)\leq t \text{ for } k=0,1,2,\ldots,k_1-1,$$
$$h(k)\geq t \text{ for } k=k_2+1, k_2+2,\ldots,k_{max}, \quad (4)$$

In the second step, the gray-level k for every pixel in the three segments A, B and C is transformed in accordance to individual table, as follows:

$$g(k)=\begin{cases} g_A(k) & if \quad 0\leq k<k_1; \\ g_B(k) & if \quad k_1\leq k\leq k_2; \\ g_C(k) & if \quad k_2<k\leq k_{max}. \end{cases} \quad (5)$$

Here $g_A(k)$, $g_B(k)$ and $g_C(k)$ are the corresponding tables for brightness transformation for segments A, B and C. In order to improve the image quality of the low-contrast image areas the thresholds k_1, k_2 of the segment B are widened (stretched) up to $(k_1-\delta_1)\geq 0$ and $(k_2+\delta_2)\leq 255$ skewing the segments A and C. In this case δ_1 and δ_2 are parameters, which define the contrast enhancement for the segment B and correspondingly – the change of the contrast range for segments A and C. Each brightness transformation table is defined in accordance to the condition for histogram equalization for the corresponding image segment A, B or C with modified (stretched or skewed) dynamic range:

$$g_A(k)=(k_1-\delta_1)\sum_{l=0}^{k}h_A(l),$$
$$g_B(k)=(k_2-k_1+\delta_1+\delta_2)\sum_{l=k_1-\delta_1}^{k}h_B(l)+(k_1-\delta_1), \quad (6)$$
$$g_C(k)=(k_{max}-k_2-\delta_2)\sum_{l=k_2+\delta_2}^{k}h_C(l)+(k_2+\delta_2).$$

In particular, in case that the histogram of the corresponding segment is constant:

$$h_A(k)=\frac{1}{k_1}; \text{ for } k=0,1,..,k_1-1;$$
$$h_B(k)=\frac{1}{k_2-k_1}; \text{ for } k=k_1,k_1+1,..,k_2; \quad (7)$$
$$h_C(k)=\frac{1}{k_{max}-k_2}, \text{ for } k=k_2+1,k_2+2,..,k_{max},$$

The table for the brightness transform for each segment is linear and the relations are correspondingly defined as:

$$g_A(k)=\left(\frac{k_1-\delta_1}{k_1}\right)k\ ;$$

$$g_B(k)=\left(\frac{k_1-k_2-\delta_1-\delta_2}{k_1-k_2}\right)k+\left(\frac{\delta_2 k_1+\delta_1 k_2}{k_1-k_2}\right); \tag{8}$$

$$g_C(k)=\left(\frac{k_{max}-k_2-\delta_2}{k_{max}-k_2}\right)k+\left(\frac{\delta_2 k_{max}}{k_{max}-k_2}\right).$$

In this case the brightness levels in the range (k_1, k_2) are stretched in accordance with a linear relation and corresponding inverse operations are performed for the two remaining histogram segments $(0, k_1\text{-}1)$ and (k_2+1, k_{max}).

In the case, when $k=1$ or $k_2=k_{max}$ the image histogram is divided into two segments (*A* and *B*) only and the corresponding brightness levels are processed in similar way, as it was already explained for three-segment (*A*, *B* and *C*) case. For this the dynamic range of the segment *A* (or *B*), which contains the low-contrast objects is stretched, and the range of the other segment – correspondingly skewed.

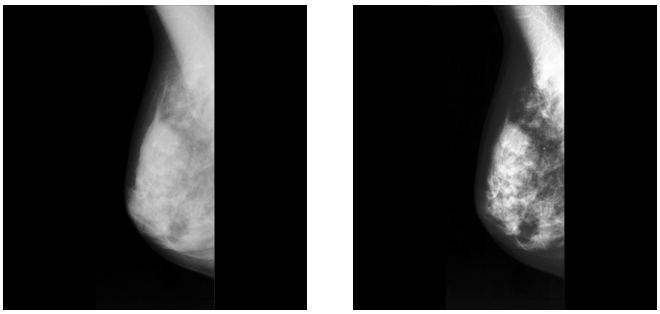
The same method is also suitable for processing of color *R, G, B* images. In this case, the contrast enhancement is performed after transformation into the *Y, Cr, Cb* format, after which the brightness component *(Y)* is processed in the already described way. Then the three components *Y, Cr, Cb* are transformed back into *R, G, B* format.

3 Experimental Results

For the experiments was used the still image database of the Laboratory for Image and Sound Processing of the Technical University of Sofia, Bulgaria. For the experiments were used more than 800 medical images of various sizes (.bmp files) from the image database of the Laboratory. The software implementation of the method was in C++, Windows environment. The values of the most important parameters permit changes, which result in corresponding changes in the restored images. Besides, the original image is retained and this is a good basis for comparison. Part of the experimental results is shown below.

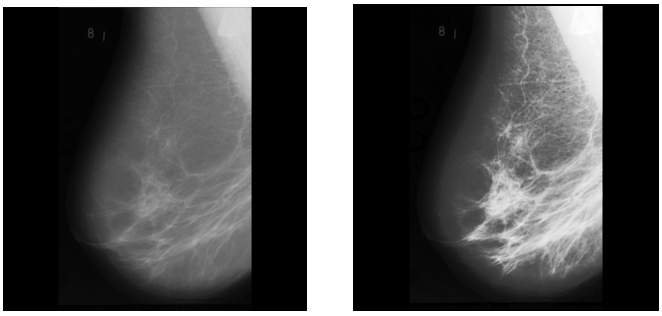
3.1 Experiments for the Two-Part Histogram Enhancement

The method is suitable for contrast enhancement of low contrast images of any kind, but it best suits mammographic images (examples shown on Figs. 3 and 4).



a. Original test image b. After histogram enhancement

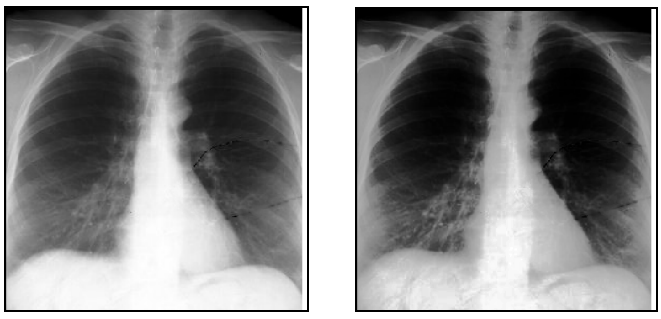
Fig. 3. Result obtained after Two-part histogram enhancement



a. Original test image b. After histogram enhancement

Fig. 4. Result obtained after Two-part histogram enhancement

The results obtained for X-ray images are promising also (Fig. 5)



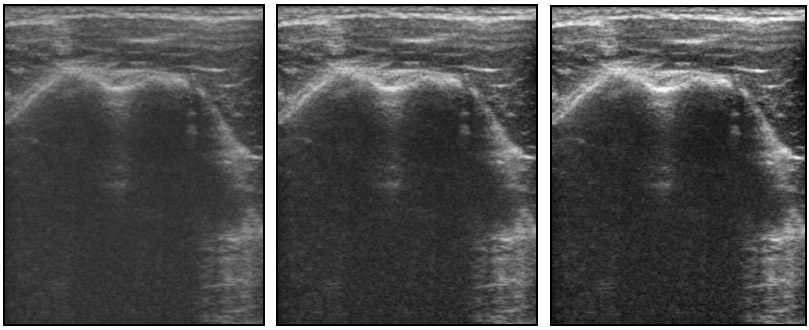
a. Original test image b. After histogram enhancement

Fig. 5. Results for X-ray images obtained after Two-part histogram enhancement

3.2 Experiments for the Three-Part Histogram Enhancement

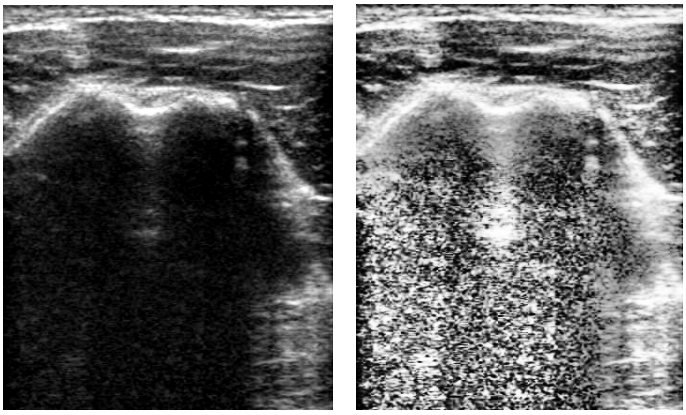
The test image "axial 2" is of size 350×432 pixels, 8 bpp (Fig.6). The histogram comprises brightness values from 33 up to 180. The histogram enhancement was done

stretching the "meaning" part up to 70% (Fig. 6b); and 80% (Fig. 6c). Each of the presented results shows up different parts of the object, depending on the depth, which corresponds to selected threshold brightness values.



a b c

Fig. 6. a. Original image; b. After 3-part histogram enhancement up to 70%; c. up to 80%



a b

Fig. 7. a. The image from Fig. 6.a, after global contrast enhancement; b. After global histogram equalization

For comparison, on Fig. 7 are shown results obtained after the global contrast enhancement (the proposed method) and after global histogram equalization (using Corel Photo Paint).

The same method is suitable for processing of low-contrast images of any kind. Example is shown on Fig. 8. The original image "Cave" is of size 800 x 600 pixels, 24 bpp. The image is of low contrast (the brightness values are in the range from 0, up to 110). The histogram is widened up to 70% of the global range and the result is shown on Fig. 6.b. It is easy to notice, that some objects become visible, retaining the natural view of the picture.

Advantage of the method implementation is that the original image is always retained and available for visual comparison. Similar results were obtained for mammographic images.

a b

Fig. 8. a. Original test image "Cave"; b. After 3-part histogram enhancement up to 70%.

4 Conclusions

The methods for contrast enhancement of medical images, presented in this work, have the following advantages compared to other similar methods:

The contrast enhancement permits to visualize various depths of the image, giving new visual information.

The computational complexity of the method is relatively low: Compared to methods based on image processing, which comprise wavelet transforms, the offered method is much faster and permits to perform real-time processing even for remote diagnostic purposes.

The use of the proposed methods is user-friendly, because the processing depends on low number of parameters, which could be easily changed by the operator.

The obtained results are of high importance for the efficient compression and archiving of medical information based on the special format, developed at TUS, because the contrast enhancement also enhances the compression efficiency.

The computational complexity of the compression method and its efficiency were compared to famous contemporary standards, JPEG and JPEG 2000 (the medical DICOM standard is based on JPEG). The advantages of the new format and the comparison with JPEG and JPEG 2000 standards were given in detail in earlier publications of the authors [13-16].

Acknowledgement. This work was supported by the Joint Research Project Bulgaria-Romania (2010-2012): "Electronic Health Records for the Next Generation Medical Decision Support in Romanian and Bulgarian National Healthcare Systems".

References

1. Bovik, A. (ed.): Handbook of Image and Video Processing. Academic Press (2000)
2. Bow, S.T.: Pattern Recognition and Image Preprocessing. Marcel Dekker (2002)
3. Pratt, W.: Digital Image Processing. John Wiley and Sons (2001)
4. Gonzalez, R., Woods, R.: Digital Image Processing. Prentice-Hall (2002)
5. Sharma, G. (ed.): Digital Imaging Color Handbook. CRC Press, NY (2003)
6. Kim, Y.: Contrast Enhancement Using Brightness Preserving Bi-Histogram Equalization. IEEE Trans. on Consumer Electronics 43(1), 1–8 (1997)
7. Jahne, B.: Computer Vision and Applications. Academic Press (2000)
8. Rowberg, A., Malcolm, B.: Distortion-free Image Contrast Enhancement. In: Inchingolo, P., Pozzi-Mucelli, R. (eds.) EuroPACS-MIR in the Enlarged Europe, pp. 357–360 (2004)
9. Laihanen, P.: Hue-dependent Contrast Enhancement. Graphic Arts in Finland 30(2), 1–4 (2001)
10. Chalana, V., Kim, Y.: A Methodology for Evaluation of Boundary Detection Algorithms on Medical Images. IEEE Trans. on Medical Imaging 16(5), 642–652 (1997)
11. Czerwinski, R., Jones, D., O'Brien, W.: Detection of Lines and Boundaries in Speckle Images -Application to Medical Ultrasound. IEEE Trans. on Medical Imaging 18(2), 126–136 (1999)
12. Ladak, H., Mao, F., Wang, Y., Steinman, D., Fenster, A., Downey, B.: Prostate Boundary Segmentation from 2D Ultrasound Images. Med. Phys. 17(8), 1–12 (2000)
13. Milanova, M., Kountchev, R., Todorov, V., Kountcheva, R.: New Method for Lossless Compression of Medical Records. In: Proc. of 8th IEEE Intern. Symp. on Signal Processing and Information Technology (ISSPIT 2008), Bosnia, Herzegovina, pp. 23–28 (2008)
14. Milanova, M., Kountchev, R., Kountcheva, R., Todorov, V.: Efficient Compression of Sequences of Medical and Multispectral Images. In: SPIE Multiconference, Airborne ISR Systems and Applications, USA, pp. 8020–8031 (2011)
15. Kountchev, R., Mironov, R., Kountcheva, R.: Efficient Compression of Medical Images Based on Adaptive Histogram Modification. In: Proc. of 46th Int. Conf. on Information, Communication and Energy Systems and Technologies, Serbia, vol. 1(1), pp. 13–16 (2011)
16. Kountchev, R., Todorov, V., Kountcheva, R.: New Method for Adaptive Lossless Compression of Still Images Based on the Histogram Statistics. In: Tsihrintzis, G.A., Virvou, M., Jain, L.C., Howlett, R.J. (eds.) IIMSS 2011. SIST, vol. 11, pp. 61–70. Springer, Heidelberg (2011)

An Adaptive Enhancement of X-Ray Images

Veska Georgieva, Roumen Kountchev, and Ivo Draganov

Department of Radio Communications and Video Technologies, Technical University - Sofia, Bul. Kl. Ohridsky 8, Sofia 1797, Bulgaria
{vesg,rkountch,idraganov}@tu-sofia.bg

Abstract. Most of the X-ray images are no truly isotropic and its quality varies depending on penetration of X-rays in different anatomical structures and on the technologies of their obtaining. The noise problem arises from the fundamentally statistical nature of photon production. This paper presents an approach for X-ray image enhancement based on contrast limited adaptive histogram equalization (CLAHE), following by morphological processing and noise reduction, based on the Wavelet Packet Decomposition and adaptive threshold of wavelet coefficients in the high frequency sub-bands of the shrinkage decomposition. Implementation results are given to demonstrate the visual quality and to analyze some objective estimation parameters in the perspective of clinical diagnosis.

Keywords: X-ray images, CLAHE, Noise reduction, Wavelet Transformations.

1 Introduction

X-rays are especially useful in the detection of pathology of bony structures, as well as detecting some disease processes in soft tissue. An "artifact" on a diagnostic X-ray image may appear as light or dark spots, lines, fogging, specks, etc. They can be caused by motion, poor contact between the film and the cassette that holds the film, and so on. The quantum noise is dominant and comes from the quantization of energy into photons. It is Poisson distributed and independent of measurement noise. The measurement noise is additive Gaussian noise and usually negligible relative to the quantum noise. It comes from the motion of patient [1].

Many techniques are available to enhance the quality of X-ray images; one of the most well known image enhancement techniques is the histogram equalization technique. There are still situations where image enhancement on a uniform histogram may not be the best approach. So other histogram techniques may need to be used, such as adaptive histogram equalization. Adaptive histogram equalization has the disadvantage to enhance not only the image, but also it enhances the noise in the image [2]. The wavelet thresholding scheme [3], which recognizes that by performing a wavelet transform of a noisy image, random noise will be represented principally as small coefficients in the high frequency sub-bands. So by setting these small coefficients to zero, will be eliminated much of the noise in the image.

In this work is presented an integrated approach for increasing contrast of the X-ray image or of its selected regions of interest (ROI), based on contrast limited

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 79–88.
DOI: 10.1007/978-3-319-00029-9_7

adaptive histogram equalization (CLAHE) and noise suppression and detail preservation abilities of the image, based on morphological processing and wavelet transformations. By properly choosing of opening, closing and top & bottom hat filtration and suitable form of structuring element, local structures can be eliminated or local geometry of the investigated object can be modified [4]. The reduction of noise components is made on the base of 2D wavelet packet transformations. To improve the diagnostic quality of the medical objects some parameters of the wavelet transforms are optimized such as: determination of best shrinkage decomposition, threshold of the wavelet coefficients and value of the penalized parameter of the threshold. This can be made adaptively for which image on the base of calculation and estimation of some objective parameters.

The paper is arranged as follows: In Section 2 is given the basics stages of the proposed approach; in Section 3 are presented some experimental results, obtained by computer simulation and their interpretation and Section 4 - the Conclusion.

2 Basic Stages for X-Ray Image Enhancement

Image enhancement techniques are applied to grayscale X-ray images which exhibited disease processes. We purpose to define a region of interest (ROI) in the image in the case of very big images, too. It can be selected in interactive procedure from the operator. The result of ROI image is written in a file format that can be used in next processing.

The general algorithm consists of three basic stages, used to improve the image quality.

- Contrast limited adaptive histogram equalization (CLAHE) for contrast enhancement ;
- Morphological processing for detail preservation capabilities;
- Noise reduction based on wavelet packet decomposition and adaptive threshold.

2.1 Contrast Limited Adaptive Histogram Equalization (CLAHE)

Contrast limited adaptive histogram is a technique utilized for improving the local contrast of images. It is a generalization of ordinary histogram equalization and adaptive histogram equalization. CLAHE does not operate on the whole image like ordinary Histogram Equalization (HE), but it works on small areas in images, named tiles. Each tile's contrast is enhanced, so that the histogram of the output area roughly matches the histogram determined by the 'Distribution' parameter. This parameter can be selected depending on the type of the input image. The adjacent tiles are then combined using bilinear interpolation to eliminate artificially induced boundaries.The contrast, particularly in homogeneous regions, can be limited to avoid amplifying any unwanted information like noise which could be existing in images. The algorithm

CLAHE limits the slope associated with the gray level assignment scheme to prevent saturation. This process is accomplished by allowing only a maximum number of pixels in each of the bins associated with the local histograms. After "clipping" the histogram, the clipped pixels are equally redistributed over the whole histogram to keep the total histogram count identical. The CLAHE method can be divided into steps to achieve as following [5]:

- The X-ray image is divided into contextual regions which are continuous and non-overlapping. Each contextual region size is M×N pixels;
- The histograms of each contextual regions are calculated;
- The histograms of each contextual regions are clipped.

For limiting the maximum slope is to use a clip limit β to clip all histograms. This is a contrast factor that prevents over-saturation of the image specifically in homogeneous areas. These areas are characterized by a high peak in the histogram of a particular image tile due to many pixels falling inside the same gray level range. This clip limit can be related to what is referred to as clip factor, α in percent, as follows [6]:

$$\beta = \frac{MN}{L}\left[1 + \frac{\alpha}{100}\left(S_{max} - 1\right)\right], \tag{1}$$

where $M \times N$ are numbers of pixels of each region and L is the number of grayscales.

If the clip factor is equal to zero the clip limit becomes exactly equal to (MN/L), moreover if clip limit is equal to 100 the maximum allowable slope is S_{max}. Normally S_{max} is set to four for still X-ray images. However, for any other application, it is recommended to obtain a good choice for S_{max} by experiment. As clip factor is changing between zero to hundred, the maximum slope, in each mapping, is changing between 1 to S_{max}.

Finally, cumulative distribution functions (CDF) of the resultant contrast limited histograms are determined for grayscale mapping. The result mapping at any pixel is interpolated from the sample mappings at the four surrounding samplegrid pixels. Pixels in the borders of the image outside of the sample pixels need to be processed specially. The neighboring tiles were combined using bilinear interpolation and the gray scale values were altered according to the modified histograms [5].

The procedure of CLAHE can be applied to Y component of the selected image that is processing in YUV system as more effectiveness.

2.2 Morphological Processing

Mathematical morphology is a powerful tool for the representation and description of region shape like boundaries, skeletons and the convex hull. Besides there are many morphological techniques for pre- or post-processing such as erosion, dilatation, opening, closing, thinning and top & bottom hat filtering. Morphology is based on set theory.

Two sets, A and B, are combined by one or more operations such as translation, reflection, complement or difference. Set A is usually referred to as the image element, while set B is referred to as the structuring element.

The next step in the purposed approach included morphological processing by the operators: opening, closing and top & bottom hat filtering, which are used to detail preservation capabilities [4].

Opening (A○ B) and closing (A • B), two useful morphological operations, can be built by combining dilation and erosion as follows:

$$A \circ B = (A \ominus B) \oplus B \tag{2}$$

$$A \bullet B = (A \oplus B) \ominus B \tag{3}$$

Both, opening and closing, tend to smooth the contour of an image, but whereas opening breaks narrow isthmuses and eliminates thin protrusions, closing fuses narrow breaks, eliminates small holes and fills gaps in the contour. The top & bottom hat filtering extracts the original image from the morphologically closed version of the image.

$$T_{hat}(A) = A - (A \circ B) \tag{4}$$

$$B_{hat}(A) = (A \bullet B) - A \tag{5}$$

One principal application of these transforms is in removing objects from an image by using a structuring element in the opening and closing that does not fit the objects to be removed. The difference then yields an image with only the removed objects. The top-hat is used for light objects on a dark background and the bottom-hat – for dark objects on a light background. An important use of top-hat transformation is in correcting the effects of non-uniform illumination.

Local structures can be eradicated or local geometry of the inspected object can be customized by appropriate selection of opening and closing filtration, top and bottom hat filtration and proper form of structuring elements and its parameters. In addition each of morphological operation can be used many times. All these elements of the procedure of morphological processing can be determined on the base of the calculated estimation parameters. Peak signal to noise ratio (PSNR) and Effectiveness of filtration (E_{FF}) values are higher for better quality of the X-ray image, where the value of Noise reduction rate (NRR) is lower.

2.3 Noise Reduction Based on Wavelet Packet Decomposition and Adaptive Threshold

Noise in X-ray images is a multi-source problem and arises mostly from the fundamentally statistical nature of photon production. In an image contained Poisson noise can be presented as an additive noise model for each pixel is as follows (6):

$$s(x, y) = f(x, y) + n(x, y) \tag{6}$$

where $f(x, y)$ is the desired image without noise; $n(x, y)$ is the noise.

The algorithm for noise reduction based on wavelet packet transform contains the following basic stages:

1) Decomposition of the X-ray image

The wavelet packet methods for noise reduction give a richer presentation of the image, based on functions with wavelet forms, which consist of 3 parameters: position, scale and frequency of the fluctuations around a given position. They propose numerous decompositions of the image, that allows estimate the noise reduction of different levels of its decomposition. For the given orthogonal wavelet functions exists library of bases, called wavelet packet bases. Each of these bases offers a particular way of coding images, preserving global energy, and reconstructing exact features. Based on the organization of the wavelet packet library, the decomposition can be determined from a given orthogonal wavelets. An optimal decomposition is used with respect to a conventional criterion. In case of denoising the 2D joint entropy of the wavelet co-occurrence matrix is used as the cost function to determine the optimal threshold. In this case 2D Discrete Wavelet Transform (DWT) is used to compose the noisy image into wavelet coefficients [3].

We propose in the paper another adaptive approach. The criterion is a minimum of three different entropy criteria: the energy of the transformed in wavelet domain image, Shannon entropy and the logarithm of energy [7].

Looking for best adaptive shrinkage decomposition to noise reduction, two important conditions must be realized together [8]. The conditions are as follows:

$$E_k=min, \text{ for } k=1, 2, 3 \ldots n \tag{7}$$

where E_k is the entropy in the level $\boldsymbol{K}$ for the best tree decomposition of image.

$$s_{ij} \geq T \tag{8}$$

2) Determination of the threshold and thresholding of detail coefficients

By determination of the global threshold it is used the strategy of Birge-Massart [9]. It uses spatial-adapted threshold, which allows to determinate the thresholds in three directions: horizontal, vertical and diagonally. The threshold can be hard or soft. The soft-thresholding function takes the argument and shrinks it toward zero by the threshold. The soft-thresholding method is chosen over hard-thresholding, because it yields more visually pleasant images over hard-thresholding. To become more precisely determination of the threshold for noise reduction in the image we can penalize adaptively the sparsity parameter α. Choosing the threshold too high may lead to visible loss of image structures, but if the threshold is too low the effect of noise reduction may be insufficient.

3) Restoration of the image

The restoration of the image is on the base on 2D Inverse Wavelet Packet Transform. The reconstructions level of the denoised image is dependent on the level of its best shrinkage decomposition.

4) Estimation of filtration.

The procedure for noise reduction can be determined on the base of the calculated estimation parameters. All adaptive procedures in the proposed algorithm are made

automatically, based on calculated estimation parameters. PSNR and E_{FF} values are higher for better denoised X-ray image where the value of NRR is lower.

3 Experimental Results

The formulated stages of processing are presented by computer simulation in MATLAB, version 7.14 environment with using the IMAGE PROCESSING and WAVELET TOOLBOXES. In analysis are used 20 real grayscale X-ray images with different sizes in cases of bone structures from the head. The original images have been done in jpeg file format. By post processing they are converted in bmp format. The results from the experiments are illustrated with X-ray image of the head with size 2869x3851 pixels, presented in Figure1.

Fig. 1. The original X-ray image

The obtained averaging results by processing of ROI images with size 1220x3260 pixels are shown in Table 1. It presents the values of the objective quantitative estimation parameters such as PSNR, Signal to noise ratio in the noised image (SNR_Y), Signal to noise ratio in the filtered image (SNR_F), Effectiveness of filtration (E_{FF}) Noise reduction ratio (NRR) in the different stages of the algorithm.

The best results by contrast enhancement using CLAHE are obtained by bell-shaped form of histogram (Rayleigh distribution) and clip limit 0.05. In MATLAB environment the clip limit must be between 0 and 1 and is the normalized value, for example 0.5 means 127.5 for a uint8 image and 1.0 is 255. Higher clip limit values will clip fewer values and thus they will be spread out more, hence more contrast.

The best results of morphological processing (MP) are obtained by line structuring element (LEN=2; DEG=60) using top& bottom hat operator. The values of PSNR and Effectiveness of filtration (E_{FF}) are maximum and the value of NRR is minimum (0.5), and shows that the noise is about two times reduced. By using of disk-shaped

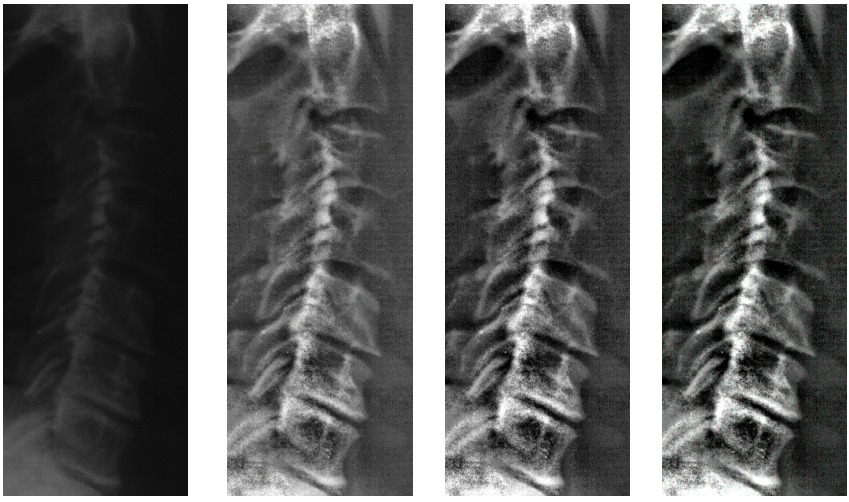
structuring element and especially a diamond-shaped structuring element can be obtained well-defined outlines of some tissues, but the values of the noised components in the image are greater.

The best results by noise reduction of Poisson noise are obtaining by Coiflet wavelet packet functions, adaptive shrinkage decomposition (best tree) on the base of the second level and minimum of the Shannon entropy criteria, by using of hard penalized threshold. By using of the log energy and energy criteria the effectiveness of the filtration is smaller. In order to quantify how much noise is suppressed by the proposed noise reduction approach, the noise reduction rate is computed. The obtained average results for NRR are around 0.3 and shows that the noise is three times reduced. The values of PSNR and Effectiveness of filtration (E_{FF}) are sufficient.

Table 1. Simulation results for stages of processing of X-ray ROI images

Stage of processing	*PSNR [dB]*	*SNR_Y [dB]*	*SNR_F [dB]*	*E_{FF} [dB]*
CLAHE	25.421	-	-	-
Morph. Processing (MP)	28.483	10.518	12.301	1.783
Noise Reduction	29.315	12.301	14.233	1.932

A visual presentation of X-ray ROI image and its modifications as a result of different stages of processing can be seen in Figure 2.



a b c d

Fig. 2. The selected ROI from the original X-ray ROI image and its modifications as a result of processing: a) original; b) after CLAHE; c) after CLAHE and MP; d) after CLAHE and MP and noise reduction on the base of WPT

The so obtained results can be compared with other methods for X-ray image enhancement and noise reduction, based on histogram equalization (HE), CLAHE and wavelet discrete transform (DWT) [2,10,11]. Table 2 contains the averaging results from the simulation obtained by comparison of the methods: histogram equalization (HE); histogram equalization following by noise reduction, based on wavelet packet transformation (HE+WPT); CLAHE, following by noise reduction, based on wavelet transformation (HE+DWT); Median filtration, following by CLAHE and noise reduction, based on wavelet transformation (MF+CLAHE+DWT). The best results are obtained by the proposed approach.

Table 2. Simulation results obtained by different methods for X-ray image enhancement

Method of processing	*PSNR [dB]*	SNR_Y *[dB]*	SNR_F *[dB]*	E_{FF} *[dB]*
HE	20.321	-	-	-
HE+WPT	28.678	9.372	12.188	2.816
CLAHE	25.421	-	-	-
CLAHE+DWT	28.923	9.372	13.646	3.914
MF+CLAHE+DWT	27.514	9.372	13.110	3.738
CLAHE+MP+WPT	29.815	9.372	14.233	4.861

A visual presentation of the X-ray ROI image and its modifications as a result of processing by different methods can be seen in Figure 3.

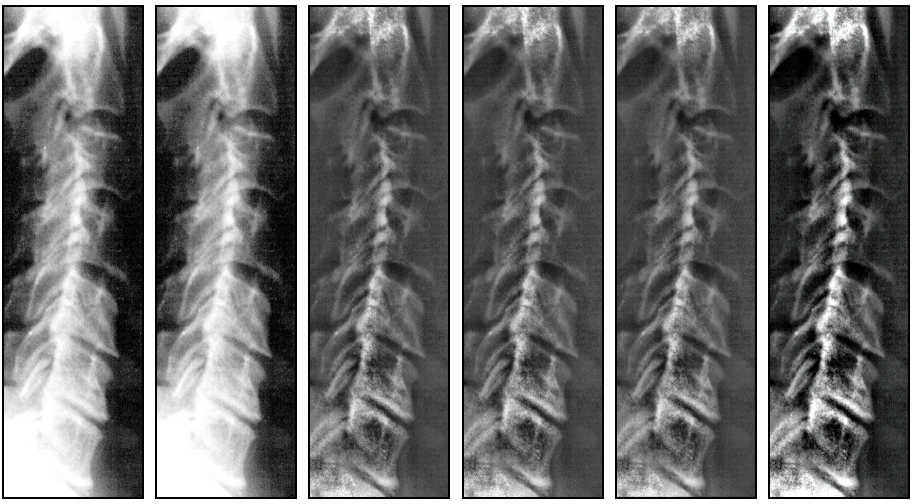
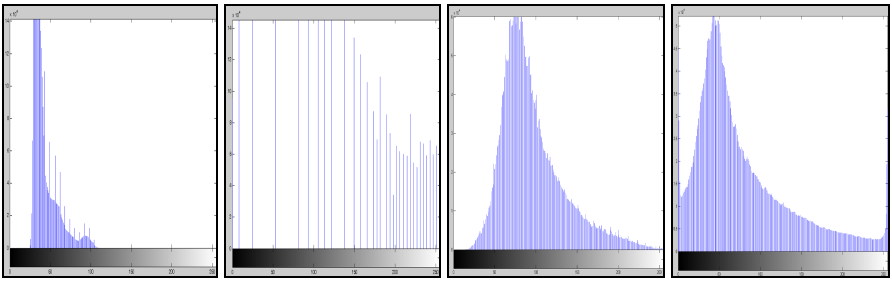


Fig. 3. The modifications of a X-ray ROI image as a result of processing: a) HE; b) HE+WPT; c) CLAHE; d) CLAHE+DWT; e) MF+CLAHE+DWT; f) proposed approach

In Figure 4 are presented the calculated original histogram of the X-ray ROI image and its modifications after histogram equalization (HE), CLAHE and after processing by proposed approach.



a b c d

Fig. 4. The histograms of the X-ray ROI image: a) original; b) equalized (HE); c) equalized (CLAHE); d) after processing

The experimental results showed enhancement of the image and increasing of image information in consequence of brightness level restoration.

The implemented studying with X-ray images and the obtained experimental results has shown that:

- The proposed approach, based on CLAHE, shows good results when used to enhance X-ray images. However artifacts are considerably amplified when tiles are more than 16×16;
- CLAHE can enhance not only the contrast of the image but it can also reduce the noise in the homogenous areas;
- The proposed effective approach for contrast increasing and noise reduction based on WPD can be adaptive applied for every stage of image processing;
- A complementary adjustment can be made in the case of the level of wavelet shrinkage decomposition and the sparsity parameter α of the penalized threshold.

4 Conclusions

In the paper is proposed a new and effective adaptive approach for X-ray image enhancement. It's based on CLAHE, morphological processing and wavelet packet transformation. The implemented algorithm provides a basis for further investigations in several directions:

- The enhanced X-ray images can be post processed with methods for segmentation, based on adaptive WPD for detection of specifically regions and edges with more diagnostic information;
- The enhanced X-ray images can be used for better visualization in 3D reconstruction;

- Some statistical characteristics such as histogram of the images and the functions of noise distribution can be analyzed for full automatically noise reduction.

The obtained image database can be easily implicated for classification of different diseases.

Acknowledgements. This paper was supported by the Joint Research Project Bulgaria-Romania (2010-2012): "Electronic Health Records for the Next Generation Medical Decision Support in Romanian and Bulgarian National Healthcare Systems", DNTS 02/19.

References

1. Smith, M., Docef, A.: Transforms in telemedicine applications. Kluwer Academic Publishers (1999)
2. Parven, N.R., Sathik, M.: Enhancement of bone fracture by equalization methods. In: Proc. of the Conference on Computer Technology and Development, pp. 391–394. IEEE Xplore Press, Kota Kinabalu (2009)
3. Zeyong, Aviyente, S.: Image denoising based on the wavelet co-occurrence matrix. IEEE Trans. on Image Processing 9(9), 1522–1531 (2000)
4. Georgieva, V.: Computed Tomography Images Enhancement with Morphology- based Filters. Proceedings of TU-Sofia 58(1), 128–134 (2008)
5. Xu, Z., Lin, X., Chen, X.: For removal from video sequences using contrast limited histogram equalization. In: Proceedings of the International Conference on Computational Intelligence and Software Engineering, pp. 1–4. IEEE Xplore Press, Wuhan (2009)
6. Ali, M.: Realization of CLAHE for Real-time Image Enhancement. Journal of VLSI Signal Processing 38, 35–44 (2004)
7. Coifmann, R., Wickerhauser, M.: Entropy based Algorithms for best basis selection. IEEE Transaction on information theory 38, 713–718 (1992)
8. Georgieva, V., Kountchev, R.: An influence of the wavelet packet decomposition on noise reduction in ultrasound images. In: Proceedings of International Scientific Conference on Information, Communication and Energy Systems and Technology, Sofia, Bulgaria, pp. 185–188 (2006)
9. MATLAB User's Guide, `http://www.mathworks.com`
10. Fayad, L., Jin, Y., Laine, A., Berkmen, Y., Pearson, G., Feedman, B., Heerturn, R.: Chest CT with multiscale Adaptive Histogram Equalization. Radiology 233(3), 845–852 (2002)
11. Saleh, H., Ahmed, S., Nordin, M.: Improving Diagnostic Viewing of medical Images using Enhancement Algorithms. Journal of Computer Science 7(12), 1831–1838 (2011)

Medical Images Transform by Multistage PCA-Based Algorithm

Ivo Draganov, Roumen Kountchev, and Veska Georgieva

Department of Radio Communications and Video Technologies,
Technical University - Sofia, Bul. Kl. Ohridsky 8, Sofia 1797, Bulgaria
{idraganov,rkountch,vesg}@tu-sofia.bg

Abstract. In this paper a novel approach for medical images transform by the Multistage Principal Component Analysis (MPCA) algorithm is presented. It consists of applying PCA over series of pixels grouped two by two in multiple stages. The process is extremely straightforward and the computation complexity is considerably reduced in comparison to the full PCA performed over the whole image. Promising results are achieved experimentally over a multitude of test images and the proposed approach is considered very perspective for both lossy and lossless compression of medical visual data.

Keywords: Medical image transform, Compression, 2D Multistage PCA.

1 Introduction

The Principal Component Analysis (PCA) [1-6] is one very useful statistical data transform technique, optimal in relation to energy compacting and mean-squared error minimization after truncation in the transform domain. By applying PCA to an image, it creates the image basis vectors, which are orthogonal, and thus, a full decorrelation of their components is achieved. By assuming the row and column-independent statistics, the generation and implementation of the 2D-PCA are simplified by using two consecutive 1D-PCA.

The main problems arising when applying the PCA for the image transform are:

- the considerably higher computational complexity compared to those of the deterministic orthogonal transforms, such as the Discrete Fourier Transform (DFT), the Discrete Cosine Transform (DCT), the Hadamard Transform, etc.
- in the general case, there is no "fast" algorithm for PCA (there is such only for a class of images which can be represented as First Order Markov Process [1]);
- the criterion for the minimal mean squared error on which the PCA is based, is not always acceptable for subjective evaluation of the vast majority of real images.

In spite of this, the PCA has been used as a benchmark in evaluating the performance of other transforms. It has also provided an incentive for the researchers to develop signal independent (fixed) transforms that not only have fast algorithms but also approach PCA in terms of performance. In the recent years a number of algorithms for lossless image compression were proposed, such as the integer reverse PCA [5], the

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 89–100.
DOI: 10.1007/978-3-319-00029-9_8

stream and parallel faster implementations of PCA [6], the recursive implementation of PCA [7, 8], the PCA transform with the use of neural networks [11, 12]. For the last two approaches the calculation of the PCA matrix $2^n \times 2^n$ ($n \geq 2$) needs from several hundred up to several thousand iterations. These algorithms are practically hard to use in image processing systems with higher demands for execution time. There is a large amount of practical applications far beyond image compression with such requirements, e.g. numerous GIS [13], text analysis systems [14], etc.

In this paper, a new block-based algorithm is proposed for 2D Multistage PCA (2D-MPCA) with reduced computational complexity and easy implementation in streamed fashion. This approach makes use of several stages including PCA with a transform matrix of size 2×2. The paper continues with the following parts: presentation of the algebraic method for image transform by PCA with a matrix of size 2×2, the description of the block-based 2D-MPCA algorithm, the evaluation of its computational complexity, experimental results for the transformation of test medical images by 2D-MPCA, and conclusions.

2 Image Transform by PCA with a Matrix of Size 2×2

For each grayscale image represented in a matrix form [*C*], consisting of $P=M \times N$ pixels, can be defined $S=P/2$ vectors $\vec{C}_s = [C_{1s}, C_{2s}]^T$ ($s=1,2,..,S$). As an example in Fig. 1 is given the image [*C*] of 4×4 pixels, forming a group of $S=8$ vectors: $\vec{C}_1=[C_{11}, C_{21}]^T$, $\vec{C}_2=[C_{12}, C_{22}]^T$, ... , $\vec{C}_8=[C_{18}, C_{28}]^T$ (the direction of the vectors was chosen to be horizontal).

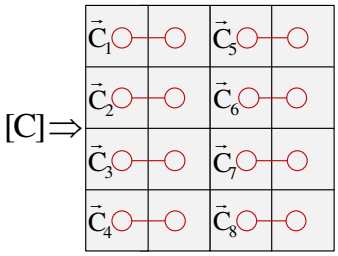


Fig. 1. Digital image with P = 16 pixels, consisting of S = 8 vectors

Each vector is then transformed into the corresponding vectors $\vec{L}_s = [L_{1s}, L_{2s}]^T$ through adaptive PCA, and using the matrix [Φ], of size 2×2 (1D-PCA 2×2). Its elements Φ_{jt} are calculated following the general algorithm for KLT [2, 3], as given below:

1. The covariance matrix $[K_C]$ of size 2×2 for the vectors $\vec{C}_s$ is calculated:

$$[K_C] = \frac{1}{S}\sum_{s=1}^{S} \vec{C}_s \vec{C}_s^t - \vec{\mu}_c \vec{\mu}_c^T = \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix}, \quad (1)$$

where $\vec{\mu}_c = [\overline{C}_1, \overline{C}_2]^T$ is the mean vector ($\bar{x} = E(x_s) = \frac{1}{S}\sum_{s=1}^{S} x_s$ - mean operator).

2. The elements of the mean vector $\vec{\mu}_c$ and the covariance matrix $[K_C]$ are defined by the relations:

$$\overline{C}_1 = E(C_{1s}),\ \overline{C}_2 = E(C_{2s}), \tag{2}$$

$$k_{11} = k_1 = E(C_{1s}^2) - (\overline{C}_1)^2,\ k_{22} = k_2 = E(C_{2s}^2) - (\overline{C}_2)^2, \tag{3}$$

$$k_{12} = k_{21} = k_3 = E(C_{1s}C_{2s}) - (\overline{C}_1)(\overline{C}_2). \tag{4}$$

3. The eigenvalues λ_1, λ_2 of the matrix $[K_C]$ are the solution of the characteristic equation:

$$det|\, k_{ij} - \lambda\delta_{ij}\,| = \lambda^2 + (k_1 + k_2)\lambda + (k_1k_2 - k_3^2) = 0. \tag{5}$$

Since the matrix $[K_C]$ is symmetric, its eigenvalues are real numbers:

$$\lambda_1 = \frac{1}{2}\left[(k_1 + k_2) + \sqrt{(k_1 - k_2)^2 + 4k_3^2}\right],\ \lambda_2 = \frac{1}{2}\left[(k_1 + k_2) - \sqrt{(k_1 - k_2)^2 + 4k_3^2}\right] \tag{6}$$

4. The eigenvectors $\vec{\Phi}_1$ and $\vec{\Phi}_2$ of the matrix $[K_C]$ are the solutions of the systems of equations below:

$$\sum_{j=1}^{2} k_{ij}\Phi_{jt} = \lambda_t\Phi_{jt} \text{ and } \left|\vec{\Phi}_t\right|^2 = \sum_{j=1}^{2}\Phi_{jt}^2 = 1, \text{ for } i,\ t\ =\ 1,\ 2. \tag{7}$$

The solution of each system of equations (7) is used to calculate the components of t^{th} eigenvector $\vec{\Phi}_t = [\Phi_{1t}, \Phi_{2t}]^T$, which corresponds to the eigenvalue λ_t:

$$\Phi_{11} = \frac{\alpha + \gamma}{\sqrt{2(\gamma^2 + \alpha\gamma)}},\ \Phi_{21} = \frac{\beta}{\sqrt{2(\gamma^2 + \alpha\gamma)}} \text{ for } t = 1, \tag{8}$$

$$\Phi_{12} = \frac{\alpha - \gamma}{\sqrt{2(\gamma^2 - \alpha\gamma)}},\ \Phi_{22} = \frac{\beta}{\sqrt{2(\gamma^2 - \alpha\gamma)}} \text{ for } t = 2, \tag{9}$$

where $\alpha = k_1 - k_2$, $\beta = 2k_3$ and $\gamma^2 = \alpha^2 + \beta^2$.

In the Eq. (9), in order to satisfy the relation $\sqrt{2(\gamma^2 - \alpha\gamma)} \neq 0$ it's necessary to have satisfied the requirement $k_3 \neq 0$ (i.e. the mutual covariation between the vectors $\vec{C}_s$ should be positive or negative). If the opposite is true (i.e., for $k_3 = 0$), the 1D-PCA 2×2 is not applied because the vectors $\vec{C}_s$ are decorrelated.

The matrix $[\Phi]$ for the 1D-PCA2×2 comprises the eigenvectors $\vec{\Phi}_1 = [\Phi_{11}, \Phi_{21}]^T$ and $\vec{\Phi}_2 = [\Phi_{12}, \Phi_{22}]^T$:

$$[\Phi] = \begin{bmatrix} \vec{\Phi}_1^t \\ \vec{\Phi}_2^t \end{bmatrix} = \begin{bmatrix} \Phi_{11} & \Phi_{21} \\ \Phi_{12} & \Phi_{22} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{\alpha+\gamma}{\sqrt{\gamma^2+\alpha\gamma}} & \frac{\beta}{\sqrt{\gamma^2+\alpha\gamma}} \\ \frac{\alpha-\gamma}{\sqrt{\gamma^2-\alpha\gamma}} & \frac{\beta}{\sqrt{\gamma^2-\alpha\gamma}} \end{bmatrix}. \tag{10}$$

In particular, if $\alpha = 0$ (for $k_1 = k_2$) from Eqs. (8,9) follows that $\vec{\Phi}_1 = \frac{1}{\sqrt{2}}[1, 1]^t$ and $\vec{\Phi}_2 = \frac{1}{\sqrt{2}}[-1, 1]^t$. Then the matrix for the 1D-PCA 2×2: $[\Phi] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ coincides with the rotation matrix of size 2×2, by which the right-oriented coordinate system (C_1, C_2) is rotated at an angle (π/4) in the counter-clockwise direction, relative to the right-oriented coordinate system (L_1, L_2).

5. The direct form of the 1D-PCA 2×2 for the vectors $\vec{C}_s = [C_{1s}, C_{2s}]^T$, from which are obtained the vectors $\vec{L}_s = [L_{1s}, L_{2s}]^T$, is:

$$\begin{bmatrix} L_{1s} \\ L_{2s} \end{bmatrix} = \begin{bmatrix} \Phi_{11} & \Phi_{21} \\ \Phi_{12} & \Phi_{22} \end{bmatrix} \begin{bmatrix} (C_{1s} - \overline{C}_1) \\ (C_{2s} - \overline{C}_2) \end{bmatrix} \text{ for } s = 1, 2, \ldots, S. \tag{11}$$

6. The inverse form of the 1D-PCA 2×2 for the vectors $\vec{L}_s$, that are transformed into the vectors $\vec{C}_s = [C_{1s}, C_{2s}]^T$, is:

$$\begin{bmatrix} C_{1s} \\ C_{2s} \end{bmatrix} = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} L_{1s} \\ L_{2s} \end{bmatrix} + \begin{bmatrix} \overline{C}_1 \\ \overline{C}_2 \end{bmatrix} \text{ for } s = 1, 2, \ldots, S. \tag{12}$$

Here, the matrix of the inverse 1D-PCA 2×2 is:

$$\begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} = [\Phi]^{-1} = [\Phi]^t = \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{\alpha+\gamma}{\sqrt{\gamma^2+\alpha\gamma}} & \frac{\alpha-\gamma}{\sqrt{\gamma^2-\alpha\gamma}} \\ \frac{\beta}{\sqrt{\gamma^2+\alpha\gamma}} & \frac{\beta}{\sqrt{\gamma^2-\alpha\gamma}} \end{bmatrix}. \tag{13}$$

7. The elements Φ_{ij} of $[\Phi]$ are functions of θ to which the coordinate system (L_1, L_2) is rotated in relation to the initial system (C_1, C_2), as a result of the 1D-PCA 2×2. If the coordinate systems (C_1, C_2) and (L_1, L_2) are selected to be right-oriented, then

the matrix $[\Phi(\theta)]$, used for the rotation at the angle θ in the counter-clockwise direction, is given by the expression:

$$[\Phi(\theta)]=\begin{bmatrix}\Phi_{11}(\theta) & \Phi_{21}(\theta)\\ \Phi_{12}(\theta) & \Phi_{22}(\theta)\end{bmatrix}=\begin{bmatrix}cos\theta & sin\theta\\ -sin\theta & cos\theta\end{bmatrix}, \tag{14}$$

where $\theta = arctg\left(\dfrac{\Phi_{21}(\theta)}{\Phi_{11}(\theta)}\right)=arctg\left(\dfrac{\beta}{\alpha+\gamma}\right)=arctg\left(\dfrac{2k_3}{k_2-k_1+\sqrt{(k_1-k_2)^2+4k_3^2}}\right)$.

The elements of the rotation matrix $[\Phi(\theta)]$ are:

$$cos\theta = cos[arctg\left(\frac{\beta}{\alpha+\gamma}\right)]=\frac{\alpha+\gamma}{\sqrt{2\gamma(\alpha+\gamma)}}, \tag{15}$$

$$sin\theta = sin[arctg\left(\frac{\beta}{\alpha+\gamma}\right)]=\frac{\beta}{\sqrt{2\gamma(\alpha+\gamma)}}. \tag{16}$$

If $tg2\theta = 2tg\theta/(1-tg^2\theta)$, the rotation angle θ can be estimated in the following way:

$$\theta=\frac{1}{2}arctg\left(\frac{\beta}{\alpha}\right)=\frac{1}{2}arctg\left(\frac{2k_3}{k_1-k_2}\right). \tag{17}$$

Then from Eq. (14) follows that the rotation matrix for the 1D-PCA 2×2 is:

$$[\Phi(\theta)]=\begin{bmatrix}cos\theta & sin\theta\\ -sin\theta & cos\theta\end{bmatrix}=\frac{\alpha+\gamma}{\sqrt{2\gamma(\alpha+\gamma)}}\begin{bmatrix}1 & \dfrac{\beta}{\alpha+\gamma}\\ -\dfrac{\beta}{\alpha+\gamma} & 1\end{bmatrix}. \tag{18}$$

In case that $\alpha=0$, the rotation angle is $\theta=\pi/4$ and $\cos\theta=\sin\theta=1/\sqrt{2}$. In correspondence to Eq. (14) the matrix for the 1D-PCA 2×2 is: $[\Phi(\pi/4)]=\dfrac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}$.

When the initial coordinate system is right-oriented, and the system rotated at an angle θ is left-oriented (the case of rotation with reflection), the rotation matrix is:

$$[\Phi(\theta)]=\begin{bmatrix}cos\theta & sin\theta\\ sin\theta & -cos\theta\end{bmatrix}=\frac{\alpha+\gamma}{\sqrt{2\gamma(\alpha+\gamma)}}\begin{bmatrix}1 & \dfrac{\beta}{\alpha+\gamma}\\ \dfrac{\beta}{\alpha+\gamma} & -1\end{bmatrix}. \tag{19}$$

When $\alpha=0$, the matrix for 1D-PCA 2×2 coincides with that of Hadamard.

From Eqs. (8), (9), (11) follows:

$$L_{1s} = \frac{(\alpha+\gamma)(C_{1s}-\overline{C}_1)+\beta(C_{2s}-\overline{C}_2)}{\sqrt{2\gamma(\alpha+\gamma)}}, \tag{20}$$

$$L_{2s} = \frac{-\beta(C_{1s}-\overline{C}_1)+(\alpha+\gamma)(C_{2s}-\overline{C}_2)}{\sqrt{2\gamma(\alpha+\gamma)}}, \tag{21}$$

where L_{1s} *and* L_{2s} are the decorrelated components of the transformed vector $\vec{L}_s$.

In accordance with Eqs. (8), (9), (12) for $s = 1,2,...,S$, the components C_{1s} and C_{2s} of the recovered vectors $\vec{C}_s$ are:

$$C_{1s} = \frac{(\alpha+\gamma)L_{1s}}{\sqrt{2\gamma(\alpha+\gamma)}} - \frac{\beta L_{2s}}{\sqrt{2\gamma(\alpha+\gamma)}} + \overline{C}_1, \tag{22}$$

$$C_{2s} = \frac{\beta L_{1s}}{\sqrt{2\gamma(\alpha+\gamma)}} + \frac{(\alpha+\gamma)L_{2s}}{\sqrt{2\gamma(\alpha+\gamma)}} + \overline{C}_2. \tag{23}$$

From Eqs. (11)-(12) and (18)-(19) follows that for the direct or inverse 1D-PCA 2×2 is necessary to know 3 parameters only (these are θ, $\overline{C}_1$ and $\overline{C}_2$) to execute the transform successfully.

3 The 2D Multistage PCA-Based Algorithm

In order to apply the 2D-PCA over a grayscale image, its matrix should be divided into blocks and over each one of them the 1D-PCA has to be implemented – at first horizontally over the rows and then - vertically over the columns. The main difficulty as suggested in the first part of this paper appears, when rows and columns have significant length leading to high computational complexity of the transform. Here a new approach is revealed in the form of the 2D-MPCA, taking in use the 1D-PCA already described above. For each block of size $2^n \times 2^n$ (n = 2, 3,...) the 1D-PCA with a matrix 2×2 is performed sequentially *n* times over the consecutive rows at first and then over the consecutive columns, where the vectors being processed are two-dimensional. After each 1D-PCA 2×2, the components of the transformed vectors are rearranged, and after that - used for the calculation of new covariance and transform matrices. The goal of the rearrangement is to obtain a new vector with all of its components arranged in decreasing order - the rearrangement in Fig. 2 is an exemplary. Prior to performing each 1D-PCA 2×2, the condition $k_3 \neq 0$ is checked and if it is not satisfied, the rearrangements of the vectors in the selected direction are ended. After finishing, all stages of the 2D-PCA decorrelated elements for each block of size $2^n \times 2^n$ are stored.

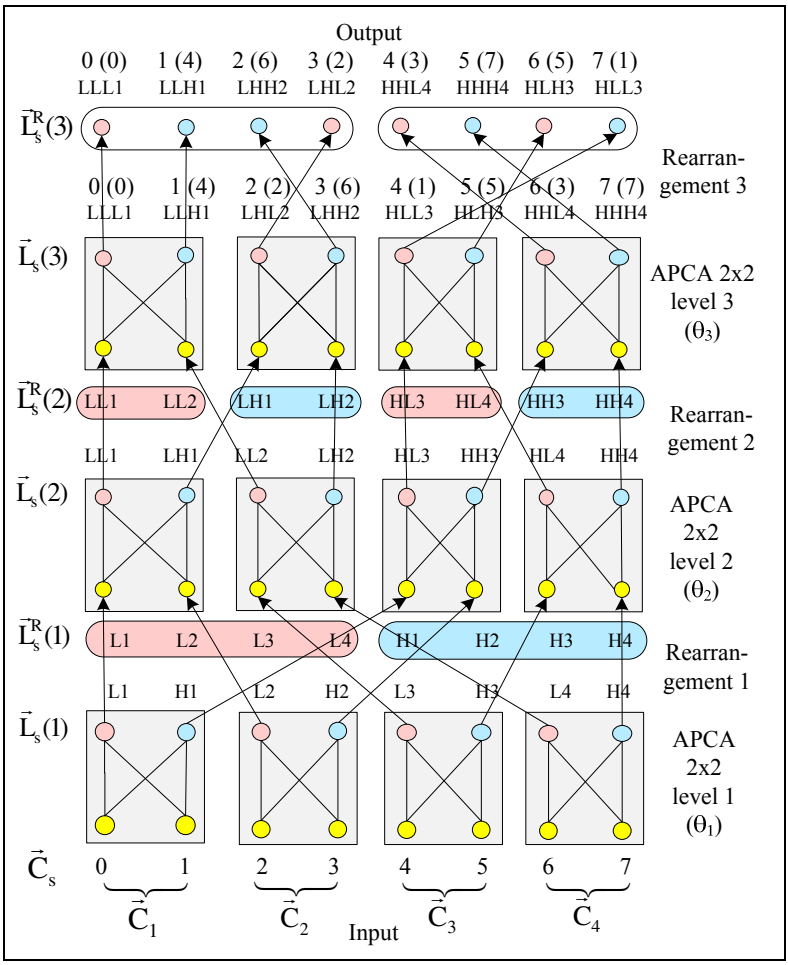


Fig. 2. Direct 3-level 1D-MPCA for the vectors $\vec{C}_s=[C_{1s},C_{2s}]^T$, s = 1,2,3,4

The whole process is illustrated in Fig. 2. There are 4 vectors with two components $\vec{C}_1=[C_{11},C_{21}]^T,..,\vec{C}_4=[C_{14},C_{24}]^T$. In light shades (pink) are shown the L components of the transformed vectors and in the dark shades (blue) are shown the H components after each stage. There are 3 stages shown here. When the rotation matrix $[\Phi(\theta)]$ with a reflection is used, the components L and H of the transformed vectors are estimated by the relations:

$$\begin{bmatrix} L_{i+1} \\ H_{i+1} \end{bmatrix} = \begin{bmatrix} cos\theta_i & sin\theta_i \\ sin\theta_i & -cos\theta_i \end{bmatrix} \begin{bmatrix} L_i - \overline{L}_i \\ H_i - \overline{H}_i \end{bmatrix} \tag{24}$$

where $$\theta_i = \frac{1}{2} arctg\left(\frac{2k_{3i}}{k_{1i} - k_{2i}}\right) \text{ for } i = 1,2,3. \quad (25)$$

Here the parameters L_i, H_i, θ_i are calculated from the vectors $\vec{L}_s(i) = [L_s(i), H_s(i)]^T$ in the i^{th} stage of the processing, in accordance to Eqs. (2)-(4) for $\vec{L}_s(0) = \vec{C}_s = [C_{1s}, C_{2s}]^T$.

4 Evaluation of the Computational Complexity of 2D-MPCA

The computational complexity of the 2D-MPCA, described above, could be evaluated considering the consecutive steps, and then - compared to the general 2D-PCA algorithm complexity for a matrix of size $N \times N$, when $N=2^n$. For this purpose, a notion is introduces for the total number of operations S (additions and multiplications) when calculating the covariance matrix $[K_C]$ also $N \times N$, the eigenvalues and the eigenvectors and finally the transformed image itself. The total number of operations, needed for the calculation of all elements of $[K_C]$, is:

$$S_k(N) = \frac{1}{2} N(N+1)[N(N-1) + 2(N+2)]. \quad (26)$$

As a case of $[K_C]$ with dimensions 2×2 ($N=2$) the result is $S_k(2) = 30$.

The eigenvalues of $[K_C]$ for $N \times N$ are calculated using the algorithm for QR decomposition, based on the Householder transformation [16] of $(N-1)$ steps. In this case the number of operations S_{val} is given by:

$$S_{val}(N) = \sum_{m=1}^{N-1} [3(N-m+1)^2 + (N-m+1)(N-m) + 4] = (N-1)\left(\frac{4}{3}N^2 + \frac{17}{6}N + 7\right). \quad (27)$$

In case, that the components Φ_{jt} of the t^{th} eigenvector $\vec{\Phi}_t$ of $[K_C]$ are calculated by using the iterative algorithm from [16], then:

$$\Phi_{jt}^{(l+1)} = \Phi_{jt}^{(l)} - \frac{1}{k_{jj}} [\sum_{p=1}^{N} (k_{jp} - \lambda_t \delta_{jp}) \Phi_{pt}^{(l)}] \text{ for } j = 1,2,..,N \text{ and } l = 0,1,2,\ldots \quad (28)$$

where $\Phi_{jt}^{(l)}$ and $\Phi_{jt}^{(l+1)}$ are the values of the (j,t) component in iterations (l) and $(l+1)$.

From Eq. (22) follows that if for the calculation of Φ_{jt} are needed L iterations, then the respective number of operations for a single eigenvector $\vec{\Phi}_t$ is $2LN(N+1)$. The cumulative number of operations for the calculation of N eigenvectors is:

$$S_{vec}(N) = 2LN^2(N+1) + N(2N-1) = N[2N(LN+L+1) - 1]. \quad (29)$$

For the direct PCA, applied over *N*-component vectors, the number of needed operations is *N(2N-1).* This transform needs to be done *N* times over all rows and in result, the total number of operations is:

$$S_{PCA}(N) = N[N^2 + N(N-1)] = N^2(2N-1). \tag{30}$$

For the direct separable 2D-PCA over an image of $N \times N$ pixels, the respective number is:

$$S_{2D\text{-}PCA}(N) = 2S_{PCA}(N) = 2N^2(2N-1). \tag{31}$$

For the 1D multistage PCA algorithm, applied on the *N*-component vector, the needed stages (levels) are $n = lg_2 N$. The number of operations here is:

$$S_{MPCA}(N) = 6n(N/2) = 3Nlg_2 N, \tag{32}$$

and for the 2D-MPCA - respectively $S_{2D\text{-}MPCA}(N) = 6Nlg_2 N$.

All operations $SS_1(N)$, needed for the separable 2D-PCA, are:

$$SS_1(N) = N^2(N^2-1) + 2N(N+1)(N+2) + \frac{1}{3}(N-1)(8N^2+17N+42) + 4N^2(4N+5) + \\ + 2N^2(2N-1) - 2N. \tag{33}$$

Hence, the computational complexity is $O_1(N^4)$. Besides the number of operations needed for the separable 2D-MPCA, including the complexity of the sorting operation at each level $N \times lg_2 N$ if quick sort is used, are:

$$SS_2(N) = 2n[S_k(2) + S_{vec}(2) + S_{MPCA}(N)] = 6(13 + N(lg_2 N + 1))lg_2 N, \tag{34}$$

and the computational complexity then is $O_2(n^2 N)$. The speeding up η of the decorrelation, which represents the relation between the corresponding computational complexities of both algorithms, is:

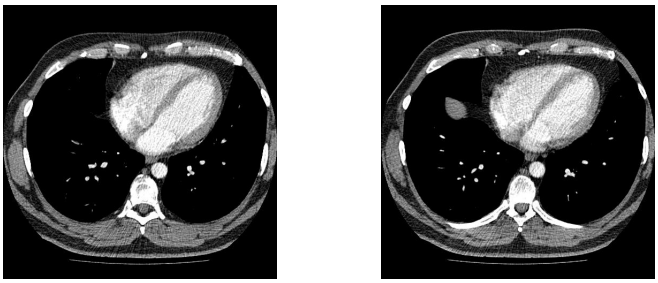
$$\eta(N) = \frac{SS_1(N)}{SS_2(N)} = \frac{3N(N^3-3) + 2N^2(37N+42) - 2}{18(13 + N(lg_2 N + 1))lg_2 N}. \tag{35}$$

From this follows that the value of the coefficient η is increasing together with *N*.

Besides, the computational complexity of the 2D-MPCA algorithm decreases in respect to that of the 2D-PCA, and for $N \to \infty$ the value of η approaches N^3/n^2.

5 Experimental Results

For the experiments were used CT test-images (10 grayscale slices, obtained by a computer tomography of abdominal areas). The size of all images is $M \times N = 512 \times 512$ pixels, 16 bpp. In Fig. 3.a,b are shown the first and the last image from the test sequence of 10.



a b

Fig. 3. CT test images from a group of 10: a) first; b) tenth

For the experiments all images were separated in blocks 8×8 pixels, and from each, 8 vectors, corresponding to a row of pixels from a block of 8 components, were calculated. All 8 vectors (block by block) were stored in a matrix of 32768 rows and 8 columns going through the blocks in a progressive scan through the end of the image.

Then over the resulting matrix, the proposed 2D-MPCA of 3 levels (from 0 to 2), was applied. For comparison, the general 2D-PCA was also performed. The execution times were calculated for each test image, being processed by both algorithms. The test environment was IBM® PC® compatible computer with P4 processor, running at 2 GHz with 2 GB of RAM under MS® Windows® XP® SP3 (32 bit) and Matlab R2009A workspace.

One more aspect of the experiments was to estimate the decorrelation for the proposed algorithm in each level of the 2D-MPCA. For this, a new parameter was introduced, called *Covariance Ratio (CovR)* which actually represents the ratio between the sum of the squared values of all diagonal elements in the covariance matrix and the sum of the squared values of the non-diagonal elements:

$$CovR = \sum_{i=1}^{8}\sum_{j=1}^{8} k^2_{ij(i=j)} \Big/ \sum_{i=1}^{8}\sum_{j=1}^{8} k^2_{ij(i\neq j)} \cdot . \tag{36}$$

For the general 2D-PCA algorithm this ratio is obviously infinity for any image. In Table 1 the experimental results are given for all 10 test images. The graphical representation of the numerical results from Table 1 is given in Fig. 4.

There is a significant difference among the decorrelation degrees from the level 0 up to the level 2 by a factor of a hundred units. It's obvious that if higher decorrelation is needed and thus - a higher compression within a complete compressing algorithm, all three levels should be used. In contrast to the general 2D-PCA, the execution times achieved are lower by a factor of 10, which is in consent with the theoretically derived reduction of the computational complexity by (N^3/n^2) times. Some deviation from the predicted values and the experimental results could be noticed due to the rearrangement of the vector components in each level, which were not included in the expressions.

Table 1. The Covariance Ratio and the Execution Times for the 2D-MPCA (levels 0 – 2), vs the full separable 2D-PCA

Image	*Level 0*		*Level 1*		*Level 2*		*2D-PCA*
	CovR	*Execution Time, sec*	*CovR*	*Execution Time, sec*	*CovR*	*Execution Time, sec*	*Execution Time, sec*
1	0.28	0.0076	16.25	0.0142	68.07	0.0195	0.1624
2	0.17	0.0066	9.94	0.0122	40.99	0.0177	0.1558
3	0.23	0.0057	12.94	0.0126	55.67	0.0176	0.1654
4	0.31	0.0051	17.42	0.0097	74.98	0.0149	0.1471
5	0.25	0.0047	13.96	0.0095	62.31	0.0140	0.1585
6	0.27	0.0055	15.73	0.0102	66.42	0.0152	0.1529
7	0.19	0.0048	10.97	0.0095	46.55	0.0151	0.1353
8	0.29	0.0057	17.10	0.0100	69.99	0.0150	0.1553
9	0.30	0.0060	17.41	0.0137	72.79	0.0194	0.1496
10	0.22	0.0049	13.01	0.0103	54.09	0.0157	0.1350

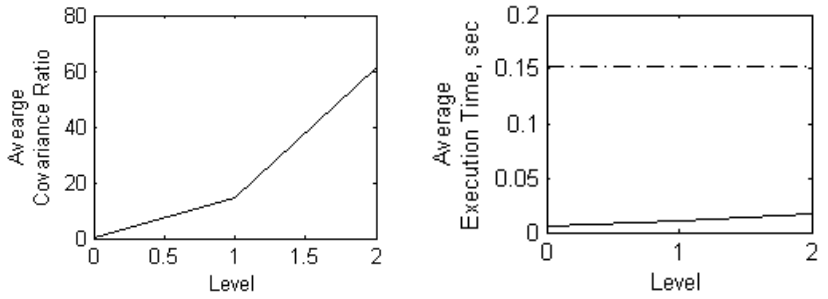


Fig. 4. a) Average Covariance Ratio of the MPCA for the levels 0,1,2; b) Comparison of the average execution times of 2D-MPCA and 2D-PCA

6 Conclusion

The proposed approach 2D-MPCA, which uses a set of low-dimensionality transforms, achieves high enough values for the decorrelation of the intensities of medical images which result could be easily extended for grayscale images in general. The reduction of the computational complexity compared to that of the general 2D-PCA is significant, which can be incorporated in real-time image or video processing systems. The applications which can make use of the suggested algorithm are: images compression (video frames) with no visual losses, integer adaptive multistage transforms with different basis, video motion compensation techniques where the intensities of the pixels are substituted by siftings as input data, fast lossless or lossy compression of series of medical or multispectral images considering the spatial correlation from slice (band) to slice (band). Lossy compression can be easily achieved by truncation of some of the low-energy components of the vectors after the last

transform level. The developed approach is also considered as very promising when used as a part of algorithms for remote sensing, distant control, pattern recognition, data mining, machine intelligence, etc.

Acknowledgements. This paper was supported by the Joint Research Project Bulgaria-Romania (2010-2012): "Electronic Health Records for the Next Generation Medical Decision Support in Romanian and Bulgarian National Healthcare Systems", DNTS 02/19.

References

1. Jain, A.: A fast Karhunen-Loeve Transform for a Class of Random Processes. IEEE Trans. Commun. COM-24, 1023–1029 (1976)
2. Dony, R.: Karhunen-Loeve Transform. In: Rao, K.R., Yip, P.C. (eds.) The Transform and Data Compression Handbook. CRC Press (2001)
3. Jolliffe, I.: Principal Component Analysis, 2nd edn. Springer, NY (2002)
4. Gonzales, R., Woods, R.: Digital Image Processing, 2nd edn. Prentice Hall (2002)
5. Hao, P., Shi, Q.: Reversible Integer KLT for Progressive-to-Lossless Compression of Multiple Component Images. In: IEEE ICIP, Barcelona, Spain, vol. 1, pp. 633–636 (2003)
6. Fleury, M., Dowton, A., Clark, A.: Karhunen-Loeve Transform: An Exercise in Simple Image Processing Parallel Pipelinnes. In: Euro-Par 1997, pp. 1–25 (1997)
7. Li, W., Yue, H., Cervantes, S., Qin, S.: Recursive PCA for adaptive process monitoring. Journal of Process Control 10, 471–486 (2000)
8. Erdogmus, D., Rao, Y., Peddaneni, H., Hegde, A., Principe, J.: Recursive Principal Components Analysis Using Eigenvector Matrix Perturbation. EURASIP Journal on Advances in Signal Processing 13, 2034–2041 (2004)
9. Hanafi, M., Kohler, A., Qannari, E.: Shedding New Light on Hierarchical Principal Component Analysis. Journal of Chemometrics 24(11-12), 703–709 (2010)
10. Grasedyck L.: Hierarchical Singular Value Decomposition of Tensors, Preprint 20, AG Numerik/Optimierung, Philipps-Universitat Marburg, pp. 1–29 (July 8, 2009)
11. Diamantaras, K., Kung, S.: Principal Component Neural Networks: Theory and Applications. John Wiley & Sons, New York (1996)
12. Solo, V., Kong, X.: Performance Analysis of Adaptive Eigen Analysis Algorithms. IEEE Trans. Signal Processing 46(3), 636–645 (1998)
13. Lazarova, M., Angelova, M.: GIS Web Services for Distributed Computing Systems. In: Proc. of Fourth International Bulgarian-Greek Conference - Computer Science 2008, Kavala, Greece, pp. 1010–1015 (2008)
14. Brodić, D., Milivojević, D.: An Algorithm for the Estimation of the Initial Text Skew. Information Technology and Control 41(3), 211–219 (2012)
15. Hoffman, J.: Numerical Methods for Engineers and Scientists. Marcel Dekker (2001)
16. Korn, G., Korn, T.: Mathematical Handbook for Scientists and Engineers. McGraw-Hill Book Company, NY (2000)

A New Histogram-Based Descriptor for Images Retrieval from Databases

Kidiyo Kpalma[1], Cong Bai[1], Miloud Chikr El Mezouar[1,2], Kamel Belloulata[2], Nasreddine Taleb[2], Lakhdar Belhallouche[2], and Djamal Boukerroui[3]

[1] UEB - INSA, IETR, UMR 6164, F-35708 Rennes, France
kidiyo.kpalma@insa-rennes.fr
[2] RCAM – Université de Sidi Bel Abbès, Algeria
[3] Heudiasyc UMR 7253 UTC, F-60205, Compiègne, France

Abstract. In this paper, we propose a new approach for designing histogram-based descriptors. For demonstration purpose, we generate a descriptor based on the histogram of differential-turning angle scale space (d-TASS) function and its derived data. We then compare the proposed histogram-based descriptor with the traditional histogram descriptors in terms of retrieval performance from image databases. Experiments on three shapes databases demonstrate the efficiency and the effectiveness of the new technique: the proposed technique of histogram-based descriptor outperforms the traditional one. These experiments showed also that the proposed histogram-based descriptor using d-TASS function and the derived features performs well compared with the state-of-the-art. When applied to texture images retrieval, the proposed approach yields higher performance than the traditional histogram-based descriptors. From these results, we believe that the proposed histogram-based descriptor should perform efficiently for medical images retrieval so we will focus on this aspect in the future work.

Keywords: pattern recognition, image description, image retrieval, texture image, differential-turning angle scale space, turning angle scale space, bull's eye performance.

1 Introduction

With the extraordinary growth of image databases due to the development of digital systems, automatic image retrieval within large databases becomes necessary. To make them more efficient and more easy to use, retrieval systems must be based preferably on images content.

In the present study, we use the d-TASS (differential-Turning Angle Scale Space) function and two related measurements to generate their corresponding histograms from which the shape descriptor is built. The d-TASS function introduced in [4] and [5] is known to yield characteristic properties and to be invariant under translation, rotation and scale change.

The rest of the paper is organized as follows: after a brief description of the d-TASS construction and the definition of the related features, in section 2, the

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 101–112.
DOI: 10.1007/978-3-319-00029-9_9

proposed descriptor is presented in section 3. Section 4 presents some experimental results and compares the proposed and the traditional histogram-based descriptor generation. We also present some comparison of the proposed approach and some existing techniques of the state of the art. Finally, section 5 gives some discussions and concludes the paper.

2 Description of the Features

As presented in [5], the input contour is defined by a set of *N* points ordered counter clockwise in the plane. Given a starting point P_0, the points are numbered from P_0 to P_{N-1}. The vector $\vec{V}_n$ originating at P_n and oriented towards P_{n+1} makes an angle θ_n with the x-axis. This angle is called the turning angle. Figure 1 illustrates the principle. From this angle, we define the differential turning angle (d-TA) function φ_n:

$$\varphi_n = \theta_n - \theta_{n-1} \tag{1}$$

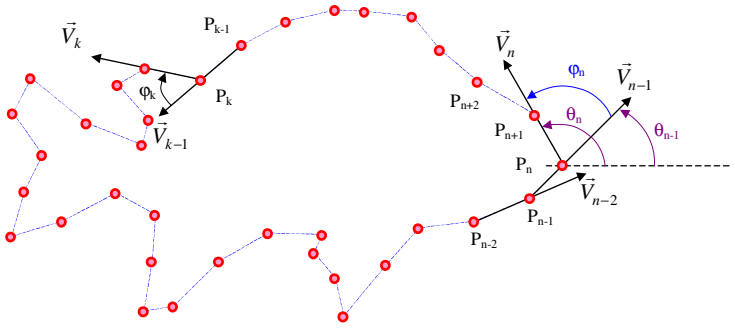


Fig. 1. Illustration of the differential-turning angle

The d-TA function yields an expression of the curvature of the contour thus its zero-crossing corresponds to that obtained from curvature analysis [12, 13].

2.1 Generation of the d-TASS Function

By definition, the d-TASS function is generated, by progressively smoothening the contour with a Gaussian filter [4, 5, 6]. Given a Gaussian kernel with a standard deviation σ_0, a progressive filtering is performed by iterating the operation so that the filtering scale $\sigma_s = \sigma_0 \sqrt{s}$ is reached after s iterations. In this study, a Gaussian kernel of size 3 (g=[0.25, 0.50, 0.25]) corresponding to $\sigma_0 = \sqrt{2/\pi}$ is used. After each iteration, the d-TA function is computed. Thus, given a range of scales (a range of iterations), one obtains the d-TASS function.

From the d-TA function, three essential points are derived to give the d-TASS map [5]. These essential points are:

- α-points corresponding to local minimums of the d-TA function,
- β-points corresponding to local maximums of the d-TA function,
- γ-points corresponding to zero-crossings of the d-TA function.

2.2 Derived Features

Generating the d-TASS function and the corresponding map is done according to the following procedure. The input contour is resampled to generate 360 equidistant points [4, 5] ordered counter clockwise. From this input sequence, the d-TA function is computed to generate corresponding essential points. In this study, we focus on the γ-points only because they will be used to set the stop condition of the iteration process. Like in the case of curvature scale space (CSS) [12], the stop condition is that there is no more γ-point at the current scale.

At each scale σ_s (iteration s), γ-points are detected and the following desired features are computed:

- an area function: area of the triangle defined by the current point P_n, the following one P_{n+1} and the mid-point P_M between two consecutive γ-points (see Fig.2.b).
- a distance function: distance between the current contour point P_n and the mid-point P_M (see Fig.2.c).

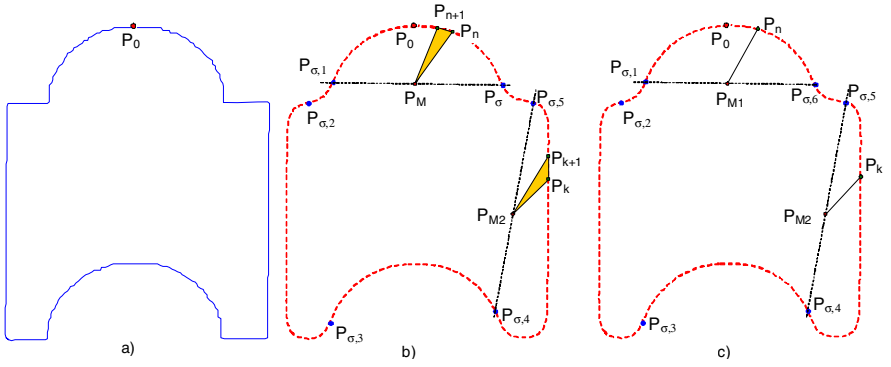


Fig. 2. a) original contour and the illustration of b) the area function and c) the distance function

These measures are characteristic of the contour segment. By iterating the process for a range of scales (until the stop condition), we obtain the area scale space function (ASSF) and the distance scale space function (DSSF), respectively, for the area function and the distance function. Figure 3 illustrates the three corresponding functions for the contour shown in Fig.2.a. The superimposed red curves represent the map of γ-points. On this figure, the x-axis represents the curvilinear abscissa and the y-axis the scale (or the iteration number).

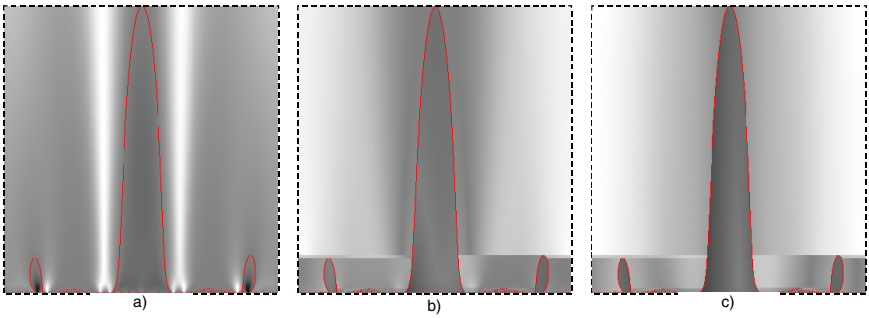


Fig. 3. Illustration of a) d-TASS function, b) ASSF function and c) DSSF function

3 Definition of the Proposed Descriptor

As said before, the proposed shape descriptor is based on the histograms of the d-TASS function and the related ASSF and DSSF functions. The idea of the proposed technique of histogram generation is based on the approach used in [1] and in [7]. The histogram of the whole database is generated to produce a reference histogram from which we calculate decision levels that enable us to generate descriptors. Commonly and particularly in [7], histogram descriptors are designed by simply uniformly quantizing the feature of interest (say the image grey scale) and then counting the number of elements having the same value and then taking the number of desired bins. This approach has proved to provide efficient results.

The proposed technique is borrowed from image coding or image restoration where histogram equalization is applied. In image processing, it is shown that this adaptive quantification can help to reduce the bit rate and to enhance signal-to-noise ratio. Starting from this observation, we propose to design the histogram-based descriptor by defining the histogram bins like in the case of adaptive quantization.

First, the reference histogram of the database is generated and normalized to sum to the unity. Then the adaptive quantization is designed such a way that the "probability" to get a value between any two consecutive decision levels is equal. Knowing the desired number K of bins, this probability equals $PK=1/K$. This enables to generate decision levels that are then used to define the descriptor for the query. Finally, for each individual histogram, use the generated decision levels and sum up the bins between two consecutive levels. This leads to the descriptor for the contour under study. Figure 4 illustrates the principle and indicates the decision levels; it also shows an example of generated histogram bins.

The proposed descriptor is thus defined by the concatenation of individual descriptors, corresponding to the three features, as follows:

$$V=[H_{dTASS}\ H_{ASSF}\ H_{DSSF}] \tag{2}$$

where H_{dTASS}, H_{ASSF} are H_{DSSF} are the respective generated histograms descriptors of d-TASS, ASSF and DSSF functions. Thus the length of the generated descriptor V is $D=3K$.

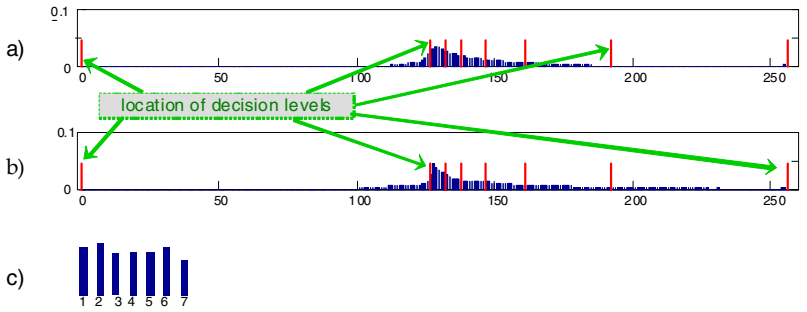


Fig. 4. a) Reference histogram of the whole database, b) histogram of a contour ("bonefishes"), and c) its 7-bins' proposed histogram

4 Evaluation of the Proposed Descriptor

To evaluate the ability of the proposed descriptor to discriminate objects, we have conducted experiments of image retrieval from database of planar objects. In order to compare our results with those from other methods, the experiments are done on three well-known databases that are the kimia99 dataset [8], the multi-view curve dataset (MCD) [10, 11] and that extracted from the SQUID marine animals [12]. The performance of the descriptors is evaluated using the Precision-Recall curve averaged over the whole database and particularly we present the bull's eye performance (BEP). The similarity measure is based on Manhattan distance. As proposed in [12], some global parameters like eccentricity, aspect ratio and circularity are taken into account. Based on this, a two-steps hierarchical procedure is performed: given a global parameter g, a threshold th_g is set so that all models giving $\frac{|g_Q - g_M|}{\max(g_Q, g_M)} \leq th_g < 1$ are discarded from the retrieval procedure, where g_Q and g_M are the parameters corresponding, respectively, to query Q and model M.

Then we define the similarity measure between the query and a model, represented by their respective descriptor vectors V_Q and V_M, as follows:

$$\mathrm{Sim}(Q,M) = \alpha \sum_{i=1}^{G} |g_Q(i) - g_M(i)| + \beta \sum_{i=1}^{D} |V_Q(i) - V_M(i)| \tag{3}$$

where G is the number of global parameters (7 in this case), α and β are the weights applied, respectively, to global parameters distance and to that of the histogram descriptor.

4.1 Performance Assessment

Precision and Recall are two parameters commonly used to assess the performance of a retrieval system. They are given by relations (4) and (5)

$$Precision = \frac{\#(Retrieved\ \ Relevant)}{\#\,Retrieved} \tag{4}$$

$$Recall = \frac{\#(Retrieved\ Relevant)}{\#\ Relevant} \tag{5}$$

The BEP indicates the percentage of the retrieved relevant contours after we have retrieved twice the number of relevant contours in the database: this corresponds to the Recall for the twice the number of relevant contours in the database. It provides an objective measurement of the efficiency of a retrieval system: the higher it is, the better is the system.

Results presented in the next sections correspond to the average values all over the database by using each contour of the database as a query.

4.2 Test Sets

The evaluation was done on three databases. The first one is proposed by Klein and Kimia [8]. It consists of shapes gathered from different sources to form 9 different categories. Each category contains 11 shapes leading to variations in form, occlusion, articulation, missing parts, etc. This gives a total of 99 shapes. Figure 5 illustrates this dataset. As our method utilizes contour-based analysis, an edge detector is applied to each of these elements to extract the corresponding contours.

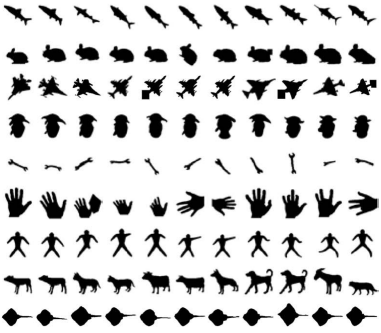


Fig. 5. Content of the kimia99 dataset

The second dataset is the multi-view curve dataset (MCD) [14]. It consists of 40 classes drawn from the MPEG-7 Core Experiment Shape-1. Each class contains 14 contours corresponding to different perspective distortions of the original one. Figure 6 shows a sample set taken from this database.

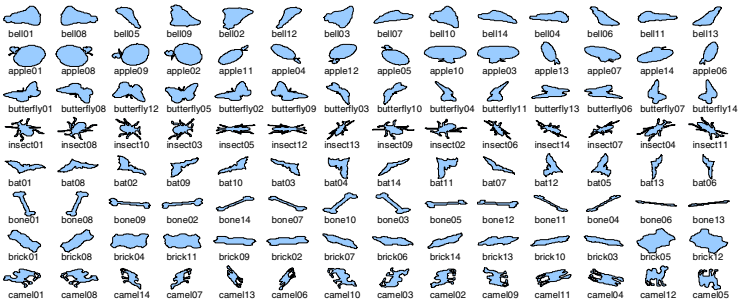


Fig. 6. A sample set of MCD database contours

The MCD dataset presents realistic perspective transformations that one can encounter while creating images of real objects. Indeed, in the construction process, the authors [10] have printed the 40 original contours on white paper and for each one, they took images from 7 different view angles using a digital camera and then contours are extracted. By adding random rotations and reflections to these samples, the number of samples in each class is doubled to 14, leading to a total of 560 shapes in the database.

The third dataset is the SQUID subset used by F. Mokhtarian and M. Bober in [12] and by F. Mokhtarian et al. in [13]. It consists of 17 classes containing from 6 to 8 shapes selected from the SQUID marine animals database [15] represented on Fig. 7, containing 131 shapes.

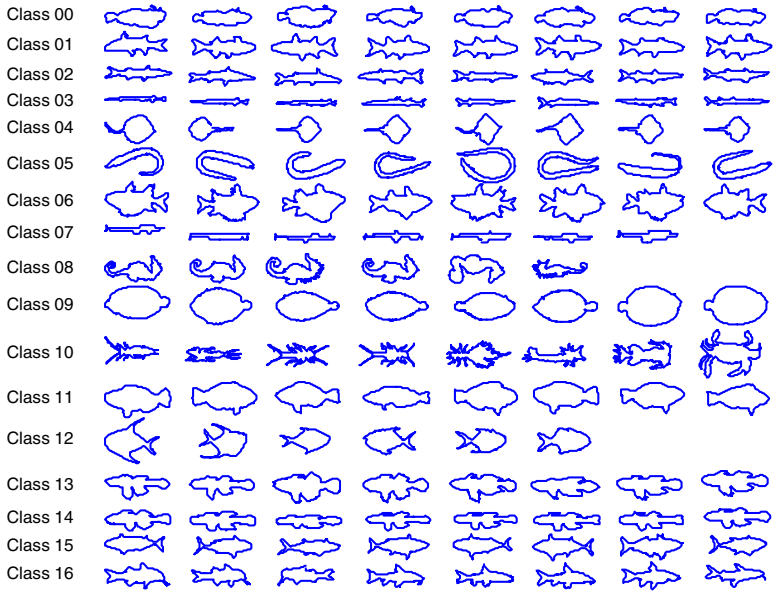


Fig. 7. Dataset-III extracted from the SQUID database

The evaluation on this database is done by using the same procedure as proposed in [12]. Similar to the concept of Recall, the procedure is as follows:

1. For a given class, use each of the contours as the query and determine the first 15 most similar outputs. Count the number of retrieved relevant contours and compute the ratio of this number to the number of relevant contours in the class. The mean value of this ratio all over the class is the performance measure for the concerned class (see individual columns in table 3).
2. Repeat 1 for all the classes in the dataset. The overall performance is then obtained by averaging the performance provided by the 17 classes. Column 'Overall' in table 3, gives that overall performance for the four involved methods.

4.3 Experimental Results

In the experiments we have evaluated the effect of the length of the descriptor. This led us to observe that, for both histogram generation techniques, the performance varies a lot at the beginning and then establishes from a limit length above which there is no gain in performance.

4.3.1 Comparison of the Two Histograms Techniques

Figure 8 demonstrates the BEP performance computed for a range of the number of bins. This figure shows a performance gap between both techniques of histogram-based descriptor generation. This gap is observed all over the three tested databases.

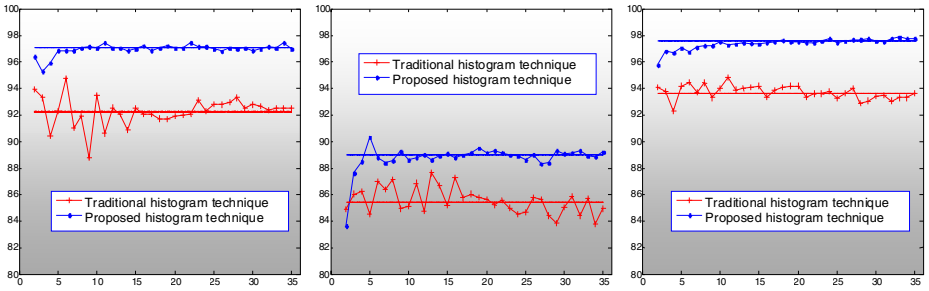


Fig. 8. BEP vs. number of bins for a) kimia99 b) SQUID and c) MCD databases

Table 1 summarizes the retrieval results from the three databases. We present the mean, the standard deviation and the maximal values of BEP over the number of bins used. On the right of the table, we give the differences of those respective values between the two techniques: the value obtained from the new technique minus that of the traditional technique. From this table, it appears clearly that the proposed technique gives the best performance with more than 3.58% mean BEP improvement.

Notice also that the proposed technique gives more stable descriptor since the BEP values vary the least: we obtain negative difference values indicating standard deviation reduction (of more than 0.30).

Table 1. Comparison of the techniques

	Traditional histogram generation (THi)			Proposed histogram generation (PHi)			Difference (PHi-THi)		
	kimia99	SQUID	MCD	kimia99	SQUID	MCD	kimia99	SQUID	MCD
Mean	92.82	85.41	93.68	96.70	88.99	97.56	3.88	3.58	3.88
σ	0.66	0.97	0.45	0.18	0.29	0.15	-0.48	-0.68	-0.30
Max	94.95	87.63	94.82	97.34	90.35	97.87	2.39	2.72	3.05

4.3.2 Comparison of the Proposed Approach with Other Approaches

In this subsection, we evaluate the proposed approach and compare it with some existing methods by applying it to appropriate databases.

A - Kimia99 Dataset: in this experiment, we retrieve the top 15 most similar for each query. Figure 9 shows an example of this retrieval. The left column represents the query contour. As each category (or class) is composed of 11 contours, there are 10 relevant contours for each query: hence columns 1~10 represent the first top 10. This figure shows that, in most cases, most of the contours from the query category are among the first 10 retrieved contours. In general, the whole class is retrieved after 15 retrievals except for the 8th category (row 8) that misses one shape. The BEP performance for this database is of 97.43% for 11 bins (leading to a 33-length descriptor). In a recent survey, Chandan Singh and Pooja [2] compared their methods with others and showed that their proposal outperforms the best on this dataset. Their proposed descriptor (ZM+HLTC) is hybrid and consists of two kinds of features: 1) local features extracted from the contour by using Hough transform and 2) global (region-based) features extracted from the region of the shape by using Zernike moments (ZM).

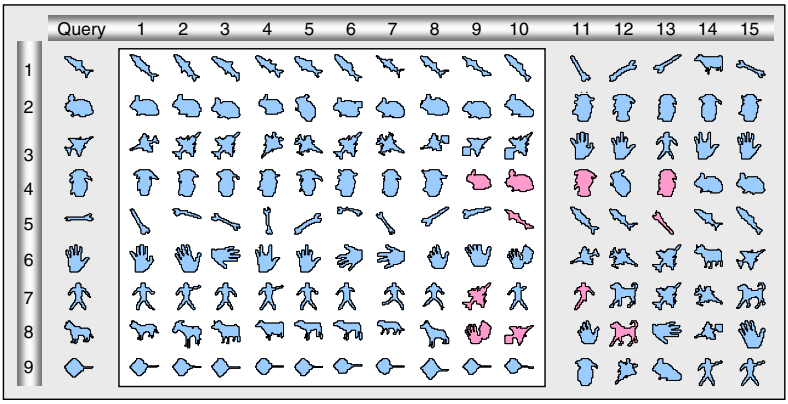


Fig. 9. Results of retrieval from kimia99 database

The BEP obtained is 99.55%: this performance is very challenging. The comparison reported in [2] is done with methods using local, global or hybrid descriptors. Table 2 shows the performance of involved methods: as can be observed, the proposed method ranks second just after the challenging ZM+HLTC method. As our descriptor uses only local features, we can reasonably believe that we could improve our performance to meet the latter if we integrate global features.

Table 2. Comparison of average BEP with other methods applied to kimia99

FD local	WLD local	CPDH local	MI global	GFD global	ZMD global	TCS hybrid	QLS hybrid	ZM+HLTC hybrid	**Proposed local**
84.29	74.39	90.05	8.74	87.93	92.68	84.29	82.81	99.55	**97.43**

B - SQUID Dataset: Table 3 shows the retrieval results from the SQUID database. For comparison purpose, we use results from the well-known curvature scale space (CSS) descriptor, the Fourier descriptor (FD) and the moment invariants (MI) reported in [9]. Columns 1~17 indicate the mean success retrieval rate for each class

and the column "Overall" indicates the mean success retrieval rate over the whole database. As can be seen on this table, the proposed approach outperforms the FD and MI descriptors and ranks first (ex aequo with CSS method).

Table 3. Comparison of retrieval rate with other methods applied to SQUID dataset

	Class N°	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	**Overall**
Descriptors	CSS	89	86	81	95	100	100	91	98	75	92	81	100	94	89	100	95	69	**91**
	Proposed	88	91	96	100	100	97	93	98	75	97	74	96	92	85	91	99	68	**91**
	FD	100	84	100	75	78	42	80	100	36	98	13	70	78	98	73	97	48	75
	MI	34	78	91	100	88	41	72	86	47	53	23	91	53	45	72	73	58	65

C - MCD Dataset: Figure 10 shows the Precision vs. Recall graphs obtained from the MCD database. Compared to the results reported in recent studies [3] and [11], the proposed approach performs very well. The Recall-Precision graph indicates more than 91.86% Recall for more than 92% Precision for the proposed approach. This outperforms the reported performance that are 86% Recall for 85% Precision for Helmholtz curve descriptor (HCD) [11] and 90% Recall for 90% Precision obtained by Ekombo et al. with the invariant Fourier descriptor [3].

5 Conclusions

The subject of this study is the description of planar objects based on the analysis of their contours. For this purpose, we introduced two contributions: 1) first, we proposed a new descriptor based on the histograms of three features derived from the turning angle scale space analysis of the contours and 2) we adopted another technique for histogram descriptor generation as opposed to traditional histogram-based descriptors.

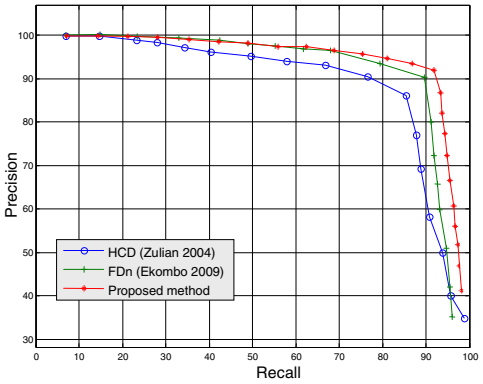


Fig. 10. Example of Precision vs. Recall graphs

The proposed technique of histogram generation jointly with the new features provides an efficient and effective descriptor that outperforms the traditional histogram descriptor. Applied to three databases, the proposed technique gave more than 3.58% improvement of the bull's eye performance regarding the traditional histogram descriptors.

Initially designed for d-TASS-based descriptor for shape retrieval, the novel technique of histogram-based descriptor construction is successfully applied for texture retrieval. These encouraging results suggest a potential application (using an appropriate set of features: DCT, ...) of the novel proposed histogram technique on medical images retrieval.

Finally, as demonstrated by experimental results, the proposed descriptor allows reaching state-of-the-art performance. Since this descriptor is based only on contour information, we believe that the descriptor can be improved by combining it with a region-based feature.

From the previous observations, the future developments of this study will be of two main types: the first will be to generate a hybrid descriptor by combining contour-based and region-based features in order to improve the shape descriptor and the second will be to extend the proposed histogram generation technique to medical images description for their retrieval from databases.

Acknowledgements. This work was partially supported by the Franco-Algerian cooperation program PHC TASSILI grants. Entitled "Système Conjoint de Compression et d'Indexation Basé-Objet pour la Vidéo (SCCIBOV)", the project gathers members of Institut National des Sciences Appliquées (INSA) de Rennes, Université de Technologie de Compiègne (UTC), the University of Sidi Bel Abbès and the University of Mascara.

References

1. Bai, C., Kpalma, K., Ronsin, J.: A New Descriptor Based on 2D DCT for Image Retrieval. In: International Conference on Computer Vision Theory and Applications (VISAPP), Roma, Italy, February 24-26, pp. 714–717 (2012)
2. Singh, C., Pooja: Improving Image Retrieval Using Combined Features of Hough Transform and Zernike Moments. Optics and Lasers in Engineering 49(12), 1384–1396 (2011)
3. Ekombo, P.L.E., Ennahnahi, N., Oumsis, M., Meknassi, M.: Application of Affine Invariant Fourier Descriptor to Shape-based Image Retrieval. International Journal of Computer Science and Network Security (IJCSNS) 9(7), 240–247 (2009)
4. Kpalma, K., Yang, M., Ronsin, J.: Planar Shapes Descriptors Based on the Turning Angle Scalogram. In: Campilho, A., Kamel, M.S. (eds.) ICIAR 2008. LNCS, vol. 5112, pp. 547–556. Springer, Heidelberg (2008)
5. Kpalma, K., Ronsin, J.: Turning Angle Based Representation for Planar Objects. Electronics Letters / IEE Electronics Letters 43(10), 561–563 (2007)
6. Kpalma, K., Ronsin, J.: Multiscale Contour Description for Pattern Recognition. Pattern Recognition Letters 27(13), 1545–1559 (2006)
7. Zhong, D., Defée, I.: DCT Histogram Optimization for Image Database Retrieval. Pattern Recognition Letters 26, 2272–2281 (2005)

8. Klein, S.T.B., Kimia, P.N.: B.B. Recognition of Shapes by Editing their Shock Graphs. IEEE Trans. Pattern Anal. Machine Intell. 26(5), 550–571 (2004)
9. Zhang, D., Lu, G.: Review of Shape Representation and Description Techniques. Pattern Recognition 37, 1–19 (2004)
10. Zuliani, M., Bhagavathy, S., Manjunath, B., Kenney, C.S.: Affine-invariant Curve Matching. In: IEEE International Conference on Image Processing, ICIP (2004)
11. Zuliani, M., Kenney, C., Bhagavathy, S., Manjunath, B.S.: Drums and Curve Descriptors, British Machine Vision Conference (BMVC) (September 2004)
12. Mokhtarian, F., Bober, M.: Curvature Scale Space Representation: Theory, Applications and MPEG-7 Standardization. Kluwer Academic Publishers (2003)
13. Mokhtarian, F.A., Mackworth, K.: A Theory of Multiscale, Curvature-Based Shape Representation for Planar Curves. IEEE Transactions on Pattern Analysis and Machine Intelligence 14(8), 789–805 (1992)
14. http://vision.ece.ucsb.edu/~zuliani/Research/MCD/MCD.shtml (visited November 25, 2010)
15. http://www.ee.surrey.ac.uk/CVSSP/demos/css/demo.html (visited November 25, 2010)

Combining Features Evaluation Approach in Content-Based Image Search for Medical Applications

Antoaneta A. Popova and Nikolay N. Neshov

Department of Radio Communications and Video Technologies
Technical University of Sofia, Bul. Kl. Ohridsky 8, Sofia 1797, Bulgaria
{antoaneta.popova,nneshov}@tu-sofia.bg

Abstract. In this paper we propose an approach for a feature combination helping to distinguish searched images from databases by retrieving relevant images. The retrieval effectiveness of 11 well known image features, commonly used in Content Based Image Retrieval (CBIR) systems, is investigated. We suggest a combined features approach including features' performance comparison of 57 various medical image categories from IRMA Database. The most informative 3 features, adaptive to image categories, are defined. Based on experiments and image similarity accuracy analysis we suggest a set of 3 low level features Color Layout, Edge Histogram and DCT Coefficients. The developed approach achieves better similar images retrieval results for more image classes. The results show an accuracy improvement of 14.49% on Mean Average Precision (MAP). The comparison is done to the same type performance measure of the best individual feature in different medical image categories.

Keywords: CBIR, image similarity search, feature selection, query by example, visual features, medical images.

1 Introduction

New digital technologies in recent years produce a huge amount of images in the areas of medicine, education, entertainment, Internet libraries, galleries, media services, life, etc. In the medical field a large number of images of various imaging modalities e.g., Computer Tomography (CT), Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), etc. are produced daily and used to support clinical decision making or to develop medical decision support systems.

Many CBIR systems were developed in the last decade for indexing and retrieving images automatically, using low level features extracted from the images or simple combinations of them [1]. The main features classification includes color, texture and shape image characteristics. CBIR systems extract visual features creating a feature vector and then define the similarities between a query image and images in databases. Next, the system outputs a sequence of images ranked by their decreasing similarity to the search image. Users can get the top-ranked images first, minimizing the time spent on searching useful images. In general, in conventional CBIR systems, it is

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 113–123.
DOI: 10.1007/978-3-319-00029-9_10

often observed that images visually similar to a query image are ranked low in the retrieval results.

In several articles, content-based access to medical images for supporting clinical decision-making has been proposed that would assist the management of clinical data and scenarios for the integration of content-based access methods [2].

A "customized-query" image retrieval approach is described in [3]. A query is classified according to the class labels of the images using the features that best discriminate the classes. Then the most similar images are retrieved within the predicted class using the features customized to distinguish "subclasses" within that class by unsupervised learning. The accuracy of high resolution computed tomography lung images is determined subjectively and they conclude that the suggested approach retrieval doubled the doctors' diagnostic accuracy.

Hersh et al. [4] describe a development and use of a medical image test collection ImageCLEFmed. Such benchmarks are needed by any researcher or developer in order to evaluate the effectiveness of new tools. The content structure of the test collection consists of multiple collections, organized into cases that represent a group of related images and annotations. Each case consists of a group of images and an optional annotation.

The authors in [5] evaluate the global descriptors from MPEG7, GIST and Compact Composite Descriptors for medical image retrieval in the IRMA-2007 collection. First they obtain evaluation results using single descriptors and then combine the descriptors by sorting the sum of each descriptor image rank.

Shyu et al. [6] suggest an approach using a selection of low-level features for medical image content-based retrieval. They use perceptual categories (defined by expert physicians) for disease recognition in images, apply operators to detect the presence / absence of these perceptual categories and develop a retrieval algorithm based on these perceptual categories. They claim that feature extraction based on physicians' perceptual categories (linear and reticular opacities, nodular opacities, high / low-density areas) achieves significantly higher retrieval precision than the traditional approaches for lung images.

Petrakis [**7**] propose a method for approximate searching by image content in medical image databases. Image is represented by attributed relational graphs holding features of objects and relationships between objects. The image objects are divided to expected ones (e.g., heart, lungs) and unexpected (e.g., tumor, hematoma). The images are indexed using method R-trees.

The goal of the presented work is to improve the retrieval effectiveness and accuracy of image search in huge Databases. We give an in-depth comparison with 11 well known image features. The paper recommends two sets of different feature combinations that perform well for medical image categories. For the experiments, IRMA medical image collection with multiple image categories (57) is used and the retrieval performance of the features is analyzed in detail.

The paper is organized as follows. Section 2 describes the basic image retrieval features used in the suggested integration. Section 3 presents the benchmark IRMA 10000 image database. Section 4 provides details of the proposed automatic performance evaluation. Section 5 comprises the experimental results of an image retrieval using the proposed combined features and the particular features in tables and

diagrams. Section 6 presents output results of ten relevant images using Combined Features Sets (Set 1 and 2) and an effectiveness comparison when only one feature is used. Concluding remarks are given in the last section.

2 Features for Image Similarity Search

In our experiments we used an implementation of 11 features available in LIRe: Lucene Image Retrieval (An Extensible Java CBIR Library) [8]. The first five of them are: Brightness Histogram – 256D (Dimension of the feature vector), Tamura Features – 18D, Gabor Features – 60D, Auto Correlogram – 1024D and DCT Coefficients Histogram – 192D. The Brightness Histogram is a well-known image feature [9]. In represents quantized distribution of color information in the image. The Tamura Features describe texture distribution using six local statistical measures. The Gabor Features also represent texture by applying Gabor Wavelet filters over the image. The Auto Correlogram describes both color distribution and spatial correlation of colors. The DCT Coefficients Histogram represents the distribution of the DCT coefficients for different frequencies in the image.

Another three features used in our experiments are defined in the MPEG-7 standard [10, 11] namely: Color Layout – 120D, Edge Histogram – 80D and Scalable Color – 64D. The Color Layout extraction process partitions the image in blocks and selects a representative color in each of them. Discrete Cosine Transform (DCT) is then used to represent the spatial color distribution. The Edge Histogram extraction procedure divides the image into sub-images and classifies the edges found in each sub-image according to their orientation. An Edge Histogram is then built showing the count of each edge type. The Scalable Color feature represents a histogram in HSV (Hue Saturation Value) color space based on the Haar transformation encoding.

The last three tested features belong to a group of so-called Compact Composite Descriptors (CCD)[1]. They combine both color and texture information into an appropriate compact form. These are Color Edge Directivity Descriptor (CEDD) – 144D, Fuzzy Color and Texture Histogram (FCTH) – 192D and a combination of them – Joint Composite Descriptor (JCD) – 168D.

3 Benchmark Image Database

As a benchmark image collection we use the IRMA 10000 database[2] which consists of two parts: A training part of 9000 annotated radiographs divided into 57 categories and a query part of 1000 unclassified radiographs. This database was used for evaluation of categorization of medical images in the automated annotation task ImageCLEF2005[3]. Fig.1 shows some example images from the database and their corresponding class numbers.

---

[1] http://chatzichristofis.info/?page_id=15

[2] http://irma-project.org/

[3] http://ir.shef.ac.uk/imageclef/2005/

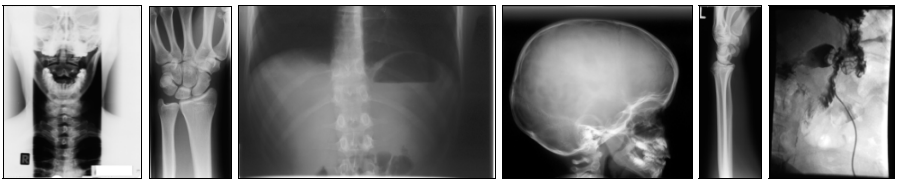


Fig. 1. Example images from IRMA 10000 database and their class number

Since the test part provided is not class labeled we use each image from the training part as a query. Then the goal is to find the most similar images to the query among all remaining 8999 images (Leaving-one-out approach). All images that belong to the same class as the query are considered relevant.

4 Automatic Image Retrieval Performance Evaluation

The proposed process of automatic image retrieval and evaluation is accomplished using the following steps:

1 st. Image Database Indexing. All images from the IRMA database training part are indexed. That is for each *k-th* feature ($k = 1$ to 11) and for each *i-th* image ($i = 1$ to 9000) a corresponding feature vector f_{ki} is calculated and stored in memory.
2 st. Queries Processing. In the image similarity search process for each *k-th* feature and each query image q ($q = 1$ to 9000) a corresponding query feature f_{kq} vector is extracted and the distances $d_{kiq}(f_{kq} - f_{ki})$ to each feature vector f_{ki} is calculated (excluding the case when $q = i$, i.e. the query vector is not compared to itself).
3 st. Sorting Images. As output result the Database images for each *k-th* feature are sorted / ranked in ascending order with respect to the distances d_{kiq} achieved in the previous step. For the combined features approach a linear combination of the distances calculated for each of the features from the set of features considered for combination (Section 5.2) is used to accomplish ranking.
4 st. Image Retrieval Evaluation. Common and well known performance measurements used in the area of information retrieval are the Precision (P) and Recall (R) metrics. These are also used for evaluation of image retrieval effectiveness of CBIR systems [12]. The P and R measures are usually presented in a form of precision-recall curve. Instead of using such curve it is often preferred to use measure parameters. In our experiments the following three parameters are automatically examined for each feature and for the proposed combined features approach (Section 5):

- Precision at N_R number of retrieved image results ($p@N_R$):

First the Precision at N_R results of each n-th query image q_n is calculated:

$$P_{q_n}(N_R) = \frac{Number\ of\ relevant\ images\ from\ the\ first\ N_R\ retrieved}{N_R}. \quad (1)$$

Next, the Precision $p@N_R$ for all queries is evaluated as:

$$p@N_R = \frac{1}{q_N}\sum_{n=1}^{q_N} P_{q_n}(N_R), \quad (2)$$

where q_N is the total number of queries for the Database ($q_N = 9000$). In our tests we calculate the precision of the first 10 retrieved images, i.e. $N_R = 10$;

- Mean Average Precision (MAP):

First the Average Precision $AP(q_n)$ for each n-th query image q_n is calculated:

$$AP(q_n) = \frac{1}{N_{rel}}\sum_{N_R=1}^{N_{rel}} P_{q_n}(N_R).rel(N_R), \quad (3)$$

where N_{rel} is the total number of relevant images in the Database, $rel(N_R)$ takes value of 1 if the current N_R image is relevant to the query and 0 otherwise. After all queries are processed, the value of MAP is given by:

$$MAP = \frac{1}{q_N}\sum_{n=1}^{q_N} AP(q_n) \quad (4)$$

- Error rate (ER):

$$ER = \frac{1}{q_N}\sum_{i=1}^{q_N}\begin{cases} 1 & if\ the\ first\ retrieved\ image\ is\ relevant\ to\ the\ query \\ 0 & otherwise \end{cases} \quad (5)$$

5 Experimental Comparison and Analysis

In Section 5.1 we give the performance results for each of the features. Section 5.2 deals with the selection of appropriate features for combination and performance comparison. The suggested algorithm is implemented using JAVA programming language modifying the open source LIRe for three feature combination approach.

5.1 Feature Performance Comparison

Table 1 presents MAP, Average precision at 10 and Error Rate for each feature and Retrieval time for one query displayed in descending order by precision at 10 (p@10) for the IRMA database. It can be seen that Edge Histogram outperforms the other features by all effectiveness retrieval measures.

Table 1. Mean Average Precision, Precision at 10 [%] and Error rate [%]and Retrieval time for one queryfor each feature of the IRMA Database

Feature	Vector Size	Distance Metric	MAP	p@10	ER	Retr.[s]
Edge Histogram	80	defined by MPEG-7	**44.27**	**72.60**	**19.49**	0.116
Color Layout	120	defined by MPEG-7	**41.17**	**67.84**	**23.54**	0.118
DCT Coefficients	192	Euclidian Distance	**24.06**	**54.83**	**36.34**	1.057
JCD	168	Tanimoto coefficient	**28.37**	**51.81**	**41.97**	0.297
CEDD	144	Tanimoto coefficient	27.38	51.58	42.86	0.167
Tamura Features	18	Euclidian Distance	26.77	46.63	47.90	0.572
Brightness Histogram	256	SAD	22.88	39.56	57.08	0.25
Auto Correlogram	1024	SAD	22.37	35.82	60.81	0.959
FCTH	192	Tanimoto coefficient	23.32	33.39	61.41	0.264
Gabor	60	SAD	18.08	25.67	72.94	0.24
Scalable Color	64	defined by MPEG-7	18.71	20.89	79.12	0.142

5.2 Analysis and Suggesting Combination of Features

Based on the performance results reached by each feature, we consider two sets of features for combination: Set 1 contains the top three best features with maximal Precision at 10 (p@10) from Table 1 (Edge Histogram, Color Layout and DCT Coefficients). Set 2 contains the first, the second and the forth feature from Table 1 (Edge Histogram, Color Layout and JCD). As explained in Section 4 the combination approach utilizes linear combination of the distances calculated for each of the three feature vectors for each set. That is the distance between the query image *q* and the *i-th* image from the database $d_{comb}{}^{1}(q, i)$ for the Combined Features Set 1 is given as:

$$d_{comb}{}^{1}(q,i) = \frac{d(f_q^{EH} - f_i^{EH}) + d(f_q^{CL} - f_i^{CL}) + d(f_q^{JPGC} - f_i^{JPGC})}{3}, \tag{6}$$

and for the second combined features set $d_{comb}{}^{2}(q, i)$ is:

$$d_{comb}{}^{2}(q,i) = \frac{d(f_q^{EH} - f_i^{EH}) + d(f_q^{CL} - f_i^{CL}) + d(f_q^{JCD} - f_i^{JCD})}{3}, \tag{7}$$

where f_q is the query feature vector and f_i is the image feature vector and the abbreviations are: *EH* – Edge Histogram, *CL* – Color Layout, *JPGC* – DCT Coefficients, *JCD* – Joint Composite Descriptor.

On Table 2 are summarized the results from the combined features sets (best values are shown in bold).

Table 2. Mean Average Precision, Precision at 10 [%], Error rate [%] and Retrieval time for one query for Edge Histogram and Combined features sets

Feature	MAP	p@10	ER	Retr.[s]
Edge Histogram – The best single feature	44,27	72,60	19,49	**0,116**
Set 1: Edge Histogram, Color Layout, DCT Coeffs.	**50,68**	**78,30**	14,39	1,2
Set 2: Edge Histogram, Color Layout, JCD	50,53	77,98	**14,04**	0,33
Improvement, %	14,49	7,85	-27,96	

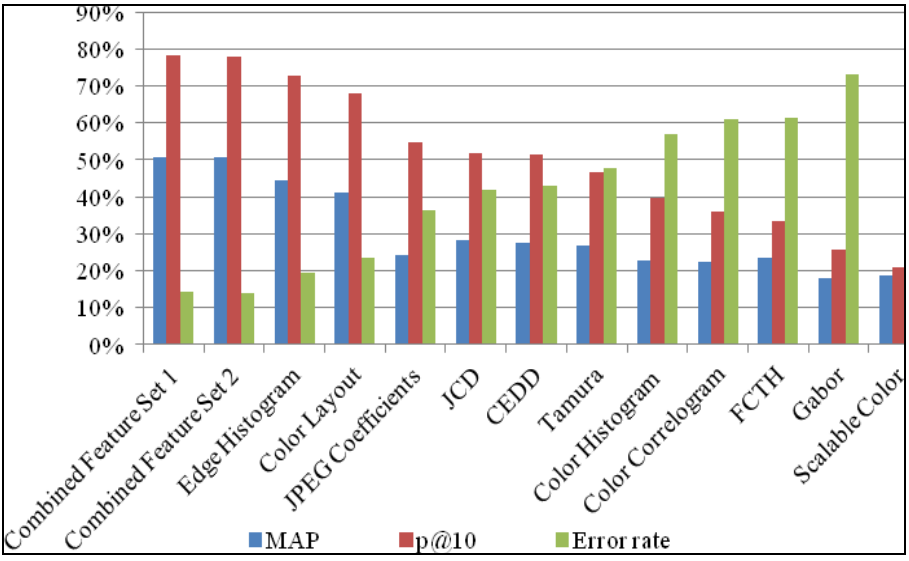


Fig. 2. Comparison of Mean Average Precision, Precision at 10 and Error Rate for each feature and the Combined Features (Set 1 / Set 2) approach for the IRMA Database

6 Retrieved Image Results

The example below (Fig.3) shows that using the features from the suggested Combined Features Set 1 outputs 10 relevant images (from the first ten retrieved) compared to the effectiveness of the Edge Histogram feature where only 5 relevant images are extracted. Fig.4 depicts the results of applying the suggested Combined Features Set 2 where 8 relevant images are extracted, compared to results achieved for the Color Layout feature where only 4 relevant results are found. In both cases our combined features approach achieves 2 times better image retrieval accuracy in comparison to individual features.

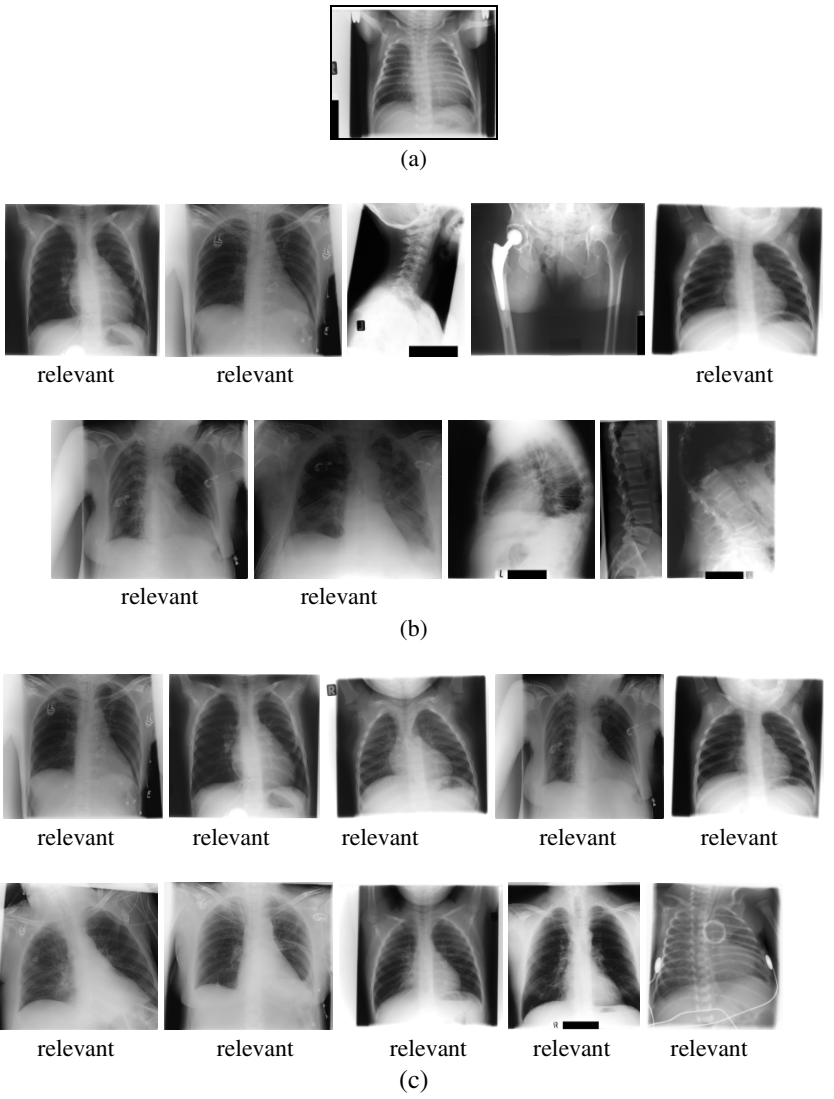


Fig. 3. Top 10 retrieval results for the query (a) from category "x-ray, plain radiography, coronal, chest" in the IRMA Database using the Edge Histogram feature (b) and the Combined Features Set 1 (c)

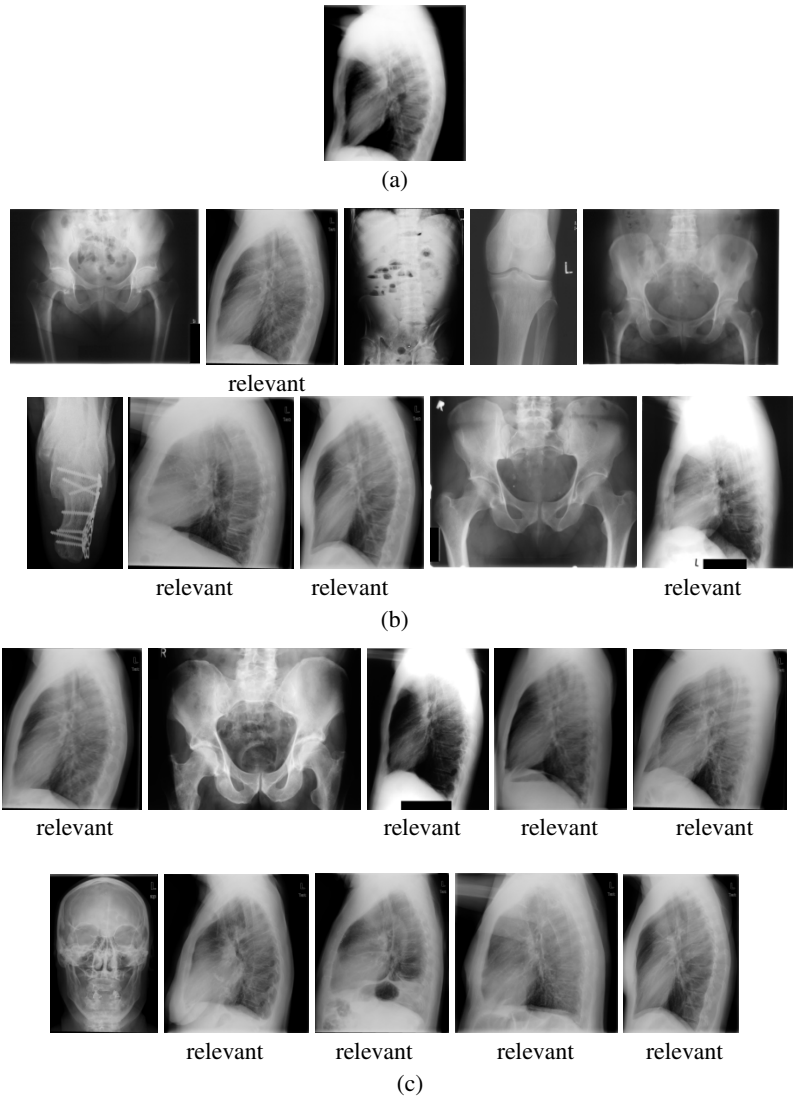


Fig. 4. Top 10 retrieval results for the query (a) from category "x-ray, plain radiography, sagittal, chest" from the IRMA Database using the Color Layout feature (b) and the Combined Features Set 2 (c)

7 Conclusions and Future Work

The paper contributions are as follows:

An evaluation of the individual features for medical radiology image search is done. Then the 3 top features - Color Layout, Edge Histogram and DCT Coefficients - are combined for higher image retrieval ranking than that based on the individual

features. In the experiments we show that our approach can significantly improve the retrieval effectiveness in CBIR systems. The advantage of our method is that the image sorting is done once for the combined features in contrast to the re-ranking methods [5], where sorting is repeated for all individual features and the combined features.

The experiment on IRMA medical images using our approach yields 14.49 % improvement of the retrieval MAP (50.68 % for Combined Features Set 1) whereas the best single feature vector approach (Edge Histogram) reaches retrieval MAP of only 44.27.

For future development we intend to include new features and applying appropriate weights for the particular features used in the combination. For example features with bigger MAP could have higher weights.

Acknowledgments. This paper was supported by the Joint Research Project Bulgaria-Romania (2010-2012): "Electronic Health Records for the Next Generation Medical Decision Support in Romanian and Bulgarian National Healthcare Systems", DNTS 02/19 and by the Project of Technical University of Sofia, Bulgaria, Research and Development Sector (2012): "Classification for Content-based Image Similarity Search for Application in Education and Internet", 121PD0053-07. The IRMA database is courtesy of TM Deserno, Dept. of Medical Informatics, RWTH Aachen, Germany.

References

1. Veltkamp, R., Tanase, M.: Content-Based Image Retrieval Systems: A Survey. UU–CS 2000–34. Utrecht, The Netherlands: Utrecht University: Information and Computing Sciences (2000)
2. Muller, H., Michoux, N., Bandon, D., Geissbuhler, A.: A Review of Content-Based Image Retrieval Systems in Medical Applications – Clinical Benefits and Future Directions. Int. J. Medical Informatics, 1–23 (2004)
3. Dy, J., Brodley, C., Kak, A., Broderick, L., Aisen, A.: Unsupervised Feature Selection Applied to Content-Based Retrieval of Lung Images. IEEE Transactions on Pattern Analysis and Machine Intelligence 25(3) (2003)
4. Hersh, W., Müller, H., Kalpathy-Cramer, J.: The ImageCLEFmed Medical Image Retrieval Task Test Collection. Proceedings of J. Digital Imaging, 648–655 (2009)
5. Coelho, F., Ribeiro, C.: Evaluation of Global Descriptors for Multimedia Retrieval in Medical Applications. In: Database and Expert Systems Applications (DEXA) Workshop, pp. 127–131 (2010)
6. Shyu, C., Pavlopoulou, C., Kak, A., Brodley, C., Broderick, L.: Using Human Perceptual Categories for Content – Based Retrieval from a Medical Image Database. Computer Vision and Image Understanding 88, 119–151 (2002)
7. Petrakis, E., Faloutsos, C.: Similarity searching in medical image databases. IEEE Trans. Knowledge and Data Engineering 9(3), 435–447 (1997)
8. Lux, M., Chatzichristofis, S.: LIRe: Lucene Image Retrieval – An Extensible Java CBIR Library. In: Proceedings of the 16th ACM International Conference on Multimedia, Vancouver, Canada, pp. 1085–1088 (2008)

9. Swain, M.J., Ballard, D.H.: Color Indexing. International Journal of Computer Vision 7(1), 11–32 (1991)
10. Chang, S.F., Sikora, T., Puri, A.: Overview of the MPEG–7 Standard. IEEE Transactions on Circuits and Systems for Video Technology 11(6), 688–695 (2001)
11. Deselaers, T., Keysers, D., Ney, H.: Features for Image Retrieval: An Experimental Comparison. Information Retrieval 11(2), 77–107 (2008)
12. Müller, H., Müller, W., Squire, D.M., Marchand-Maillet, S., Pun, T.: Performance Evaluation in Content-Based Image Retrieval: Overview and Proposals. Pattern Recognition Letters (Special Issue on Image and Video Indexing) 22(5), 593–601 (2001)

Semi-automatic Ultrasound Medical Image Recognition for Diseases Classification in Neurology

Jiří Blahuta[1], Tomáš Soukup[1], Petr Čermák[1], David Novák[1], and Michal Večerek[2]

[1] Slezská univerzita v Opavě, Filozoficko-přírodovědecká fakulta, Bezručovo náměstí 13, 74601 Opava, Czech Republic
[2] VŠB-Technická univerzita v Ostravě, Fakulta kybernetiky a informatiky, tř.17. listopadu, 70200 Ostrava, Czech Republic
jiriblahuta@seznam.cz, tsoukup@centrum.cz, {petr.cermak,david.novak}@fpf.slu.cz, vecerek.michal@email.cz

Abstract. The main aim of this work is semi-automatic ROI positionig in transcranial medical images based on multi-agent systems (MAS) in preprocessing module. Designed approach is based on image processing and is realized by means of artifical intelligence, MAS, which has been experimentally designed in Matlab software environment. Within this processing has been worked with a set of TCS static images in grayscale and binary representation to experimental testing to positioning. This designed application is used for diseases classification in neurology.

Keywords: Agent, MAS, Ultrasound, TCS, image, ROI, DICOM.

1 Introduction

Medical imaging is very important in modern medicine. The main benefit is using of DICOM format for all used modalities; such as ultrasound, CT, etc. We developed a MATLAB-based application for processing of ultrasound TCS images to potential Parkinson´s disease detection. This application is focused on ROI-based processing with artificial intelligence elements, more precisely using of a multi-agent system. This application has been tested for different US devices and artificial intelligence elements help to semi-automatic detection. Whole processing we can divide into 3 parts:

- *loading of an input DICOM image and automatically cut depend on device (see below)*
- *using of MAS to intensity threshold detection on blocks*
- *computing of area inside ROI (defined below)*

Within this paper will be described MAS which is used in this application.

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 125–133.
DOI: 10.1007/978-3-319-00029-9_11

2 TCS Images, Substantia Nigra and Parkinson´s Disease

We will use a set of TCS images in DICOM format throughout the paper. Using of diagnostic ultrasound is very important in neurology. Our developed application is focused on semi-automatic processing of transcranial ultrasound images (TCS), more precisely brain-stem area only. Generally, this application is useful to potential detection of Parkinson´s disease which is generally characterized by damaging substantia nigra (SN) area. SN is an elongated nucleus situated in each cerebral peduncle lateral to the red nucleus with cells containing melanin[1] and produce dopamin to correct function of CNS. From TCS images erudite neurologist can detect parkinsonism, thus changes of SN. The main criterion is the area of SN, risk threshold is 0.20-0.25 cm^2. So, we will measure risk threshold area to detection PD and non-PD cases. SN is well recognizable area in TCS images.

Parkinson's disease (PD) is caused by the death of dopaminergic neurons. It is a degenerative disease of basal ganglias inside the brain. PD has been described by James Parkinson in 19th century. The main symptoms of PD include muscle rigidity, tremors and changes in speech and gait, bradykinesia, sleep disorders and more.[2] More detailed information about SN and PD symptoms and medical US background are available in [4, 5, 9, 14].

The following figure shows the brain-stem area[3] and corresponding TCS image with 50x50 mm area which will be automatically cut.

To further processing we will need 50×50 mm area only which is highlighted on Fig. 1. Application also provides automatically cut the input image depend on device preset – set of images from the same device has the same resolution and we can simply cut 50×50 mm area from axis which is computed from Euclidean distance. The following code shows how to compute distance for window and cut from DICOM information about resolution:

```
d=sqrt((q2-q1)^2+(t2-t1)^2);

imc = imcrop(inp1,[(xres/2)-50 yres/3 d d]);
```

For the input image has been designed different options for each US device which we use with different *imc* vector and also we can use manual mode of 50×50 mm windows defined by `ginput` function.

[1] http://www.medcyclopaedia.com/library/topics/volume_ii/s/substantia_nigra.aspx?s=substantia+nigra&mode=1&syn=&scope=

[2] http://www.mdvu.org/library/disease/pd/

[3] http://cdn.innovateus.net/preset_4/substantia_nigra.jpg

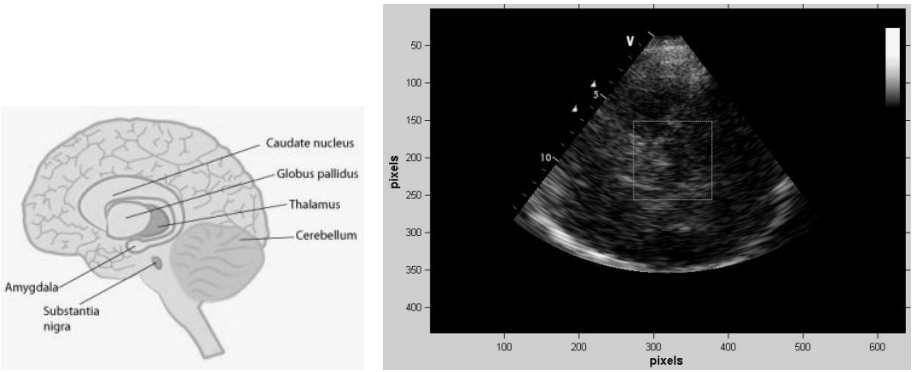


Fig. 1. Position of substantia nigra and TCS image with window

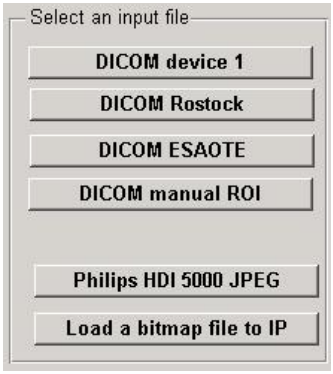


Fig. 2. Buttons for automatic cut window 50x50 mm of input TCS image and manual select

Furthermore, application automatically converts RGB input into grayscale by means of the following equation (1).

$$I = 0{,}299 \times R + 0.587 \times G + 0.114 \times B \tag{1}$$

It is necessary to further processing by MAS and also to measuring of the area. The following flow-chart introduces the main steps of our designed processing.

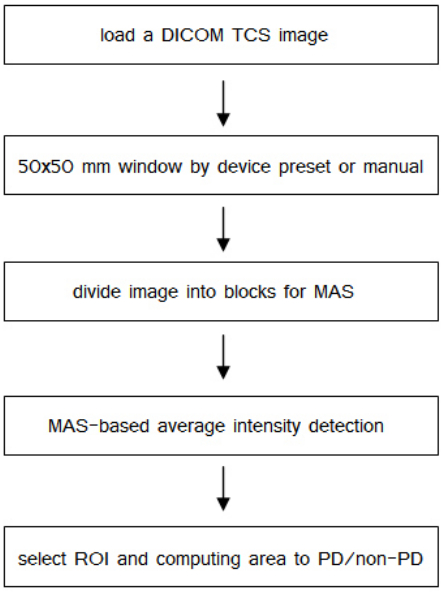


Fig. 3. The flow-chart of the whole processing with MAS approach

3 Intensity Criteria Detection by Multi-agent System

Multi-agent systems (MAS)[4] are strong tool of distributed artificial intelligence for different scopes include image processing. MAS are interdisciplinary science based on cognitive science and related scopes in artificial intelligence. Generally we can describe MAS as a finite set of autonomous agents:

$$MAS = \{a_1, a_2, \ldots, a_n) \tag{2}$$

that can communicate and cooperate and are situated in appropriate environment, in our case agents are situated in 2D image, more precisely image matrix. Each agent is autonomous and can communicate with other agents within MAS and solve the problem. In our research we use software agents only, thus MAS is represented as autonomous source code. More information about MAS and using are available in [10].

After the initial processing including load of the input image and cut, we need the second part which is applicable for MAS. We need to detect local maximal intensity in block. To this processing image is divided into 8×8 blocks and we construct one agent for each block. MAS are well applicable within the image processing tasks to semi-automatic or full automatic processing. In this case we will use the MAS which have 2-D environment as 2-D discrete image described by matrix.

Whole image is represented as discrete matrix D with $m \times n$ elements depending on resolution. We cannot solve that DICOM image contains special tags about modality,

[4] http://multiagent.martinsewell.com/

tissue, patient, device, etc. To this processing we consider image as 2-D matrix D, Each matrix element represents the position (pixel) by 2-D coordinates P(x, y). Also we know that each pixel is represented by 1 value of intensity.

$$D = \begin{bmatrix} P(1,1) & ..... & ..... & P(1,n) \\ ..... & ..... & ..... & ..... \\ ..... & ..... & ..... & ..... \\ P(m,1) & ..... & ..... & P(m,n) \end{bmatrix} \tag{3}$$

Each pixel is represented as $P(x, y)$ up to $P(m, n)$.

We divide the input image, thus matrix D, into 8×8 subregions. Each block represents the average intensity value of pixels which are inside block, thus we get 64 values; each block has 1 value of intensity. So, agent has environment represented by 2D block from image. Inside each region (block) we constructed independently an agent which must detect if inside region average intensity exceeds 25, moreover is computed minimal value of intensity. This value has been MAS is composed by these agents which operate inside each block independently. The following figure shows an original input 50×50 mm and divided into blocks for agent detection.

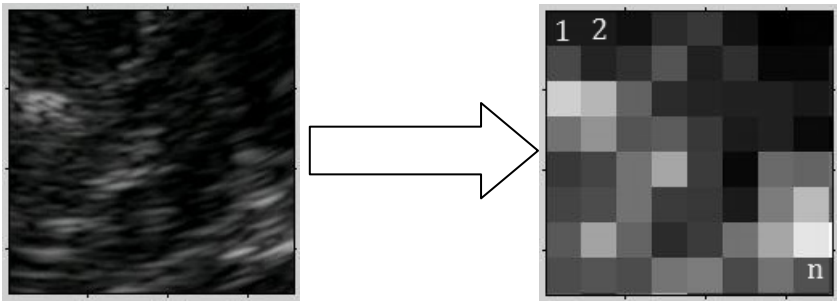


Fig. 4. Blocks 8x8 from input image 50×50 mm with agents *1* up to *n*

From blocks we get 64 values

$$b_1, b_2, \ldots, b_n; n=64$$

where b_i represents value of computed average intensity of i-th block. Furthermore we computed minimal value as variable *minbl*:

$$minbl = \min(b_1, b_2, \ldots, b_n). \tag{4}$$

The algorithm is based on the following steps:

- *create image divided into 8x8 blocks; each block represents average intensity*
- *each agent in block checks intensity value*
- *from all recognized intensity is computed minimal value*
- *if minimal value exceeds 25, application shows the warning dialog*

We get the simple IF-THEN rule:

IF minbl >25 THEN show_dialog ELSE continue.

In other words, we must check that minimal intensity from blocks does not exceed 25. If this minimal intensity exceeds 25, input image may be corrupted by manual set of intensity and lesions may be judged as pathology in spite of that patient is physiological (non-PD). The following figure shows an example of manual intensity level – pathological light lesions can be judged as PD.

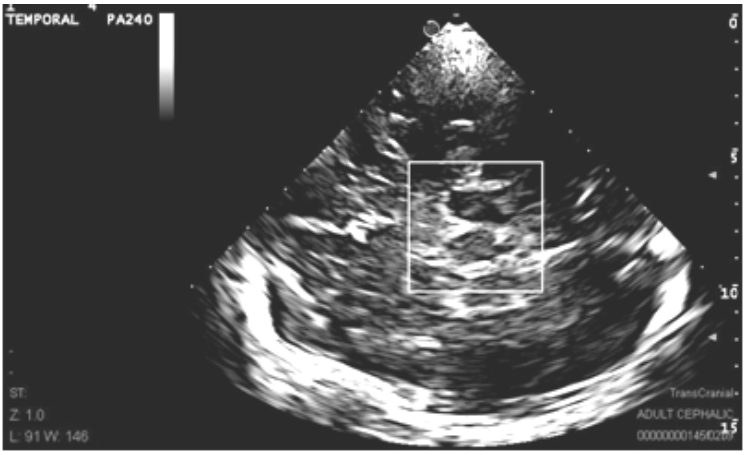


Fig. 5. Potentially corrupted input file by manual high average intensity

4 Implementation in MATLAB

Practical implementation is composed by described parts. MATLAB with Image Processing Toolbox provides powerful tool to image processing, not only for medical imaging. MATLAB has been chosen because provides working with DICOM files as native format and also provide other useful built-in functions for this application, for example ROI-based application.

We described the initial phase of loading an image and MAS which is used to detection the local maxima intensity, for details see below.

To further processing after MAS-based detection we need ROI-based processing. We require the elliptical ROI with area $A = 50$ mm^2 with rake angle of 60°, which is needed to ROI definition. Shape, size and rake angle of ROI were assigned by neurologist. So, we need the ellipse how shows the following figure. Computing of area inside ROI is the main function of the application. Phase of main processing is based on binary thresholding in ROI area and computing of area of defects in this ROI for all intensity levels $T \in \langle 0; 255 \rangle$. We will get a graph with computed area which is computed from block processing 1×1 mm. The area has been computed from binary image $\Rightarrow$ number of blocks in this 1×1 mm grid will get the real area in mm^2. Initial

ROI position is showed on the following figure – elliptical ROI inside SN which is based for area measurement; original and 2x zoom. Inside this ROI will be detected lesions to PD or non-PD cases with graphical representation.

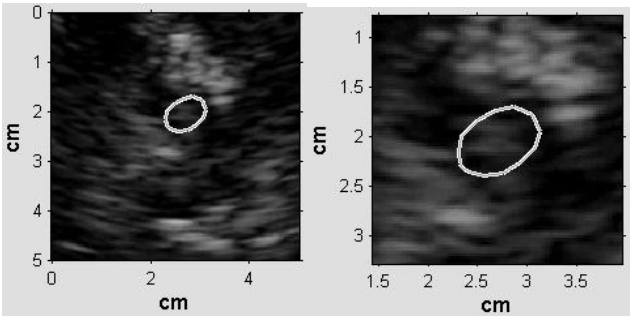


Fig. 6. The used elliptical ROI to measurement of area for all intensity levels 0 up to 255

Application provides to select a different rake angle of ROI. The main goal of this ROI processing is computing of area inside ROI to detection pathology or physiology which is depending on black or light regions inside ROI and we will get the following graph with 90th percentile curve which has been computed from a set of non-PD patients.

In practice we verified a basic rule about decrease of area for *T* for approximate distinction – if decrease is gradual, we can assess that patient is probably diseased (PD features) and if decrease is very fast, patient is probably healthy (small feature-less areas).

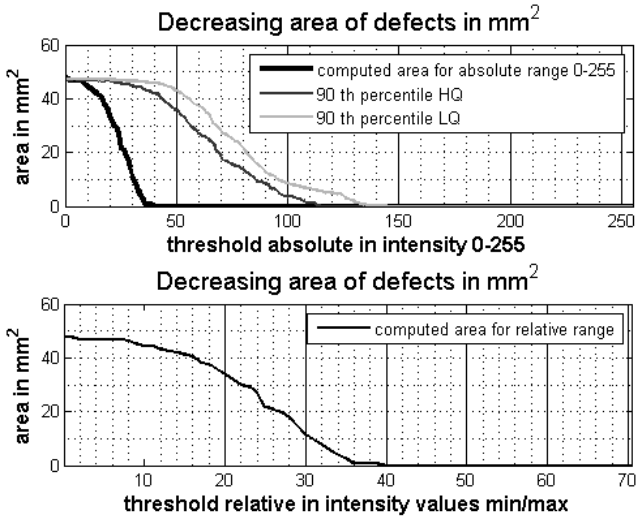


Fig. 7. The result of the measurement with visually comparison with 90th percentile to PD and non-PD cases

More information about reproducibility of the used method is available in [4] and [12].

4.1 MAS-Based Processing

Practical implementation of searching intensity inside subregions is based on block processing in MATLAB, more precisely `blockproc` function with preset 8x8 blocks. Inside each block we detect average intensity by agent. Block processing of image (Fig. 4) is realized by `blockproc` function where we can set the number of block. Theoretically background of used algorithm has been described in previous chapter, practically in MATLAB is the code of agent is constructed:

```
fun = @(block_struct) ...
   mean2(block_struct.data)                                    *
ones(size(block_struct.data));
relg = blockproc(inputim,[d/8 d/8],fun);
```

where *mean2* is 2-D average values of pixels inside block and *d* is 50 mm dimension of the window. The following part of code represents the computing of p_i=*25* criterion as *minbl* variable followed by warning dialog if this value is exceeded (4):

```
minbl = min(relg(:));
if minbl > 25
warndlg (´The minimal intensity exceeded 25, image may be
corrupted.´)
```

This part of the code represented as agent can be changed for different conditions and images. For example, we can change number of blocks or minimum limit for another US images.

5 Conclusions

Designed application based on measurement of lesions from ultrasound TCS images is useful to classification PD and non-PD patients. The used approach is focused on automatic ROI-based measurement with artificial intelligence elements.

The role of MAS to detection average intensity is very useful for medical image processing, in our case of TCS images to PD or non-PD classification. Designed MAS with agents which detect average intensity is useful as autonomous code with changeable size of blocks and intensity level. It is applicable not only for TCS images, but with changes this MAS can be helpful for another images, such from CT, other US images, etc.

Cooperation among agents within MAS may be extension of functionality. Each agent can search average intensity inside block and can use communication acts to send message to other agents within MAS. So, functionality extension is depending on the following adjustable parameters:

- *block size => number of agents for 2D image*
- *add more conditions => groups of autonomous agents which finds different conditions*
- *communication model of agents for interactive sharing of messages*

Designed MAS is an autonomous code to detection intensity conditions which is adjustable for different images, resolution and condition which are adjustable by user. We can create fully automatic MAS system to intensity checking as presets for different cases such as US device model and image format.

Using of MAS is very helpful to practical implementation, because we can control the global intensity of the input image. Furthermore, we correctly computed 90^{th} percentile from healthy patients as reference curve for PD and non-PD classification. This application may be usable for different measuring within US image, generally not only TCS.

References

1. Becker, G.: Degeneration of Substantia Nigra in Chronic Parkinson's Disease Visualized by Transcranial Color-coded Real-time Sonography. Journal of Neuroimaging 45, 182–184 (1995)
2. Blahuta, J., Soukup, T., Čermák, P.: The image Recognition of Brain-stem Ultrasound Images with Using a Neural Network Based on PCA. In: IEEE Intern. Workshop on BARI, pp. 137–142 (May 2011)
3. Blahuta, J., Soukup, T., Čermák, P.: The Image Recognition of Brain-stem Ultrasound Images with Using a Neural Network Based on PCA. Recent Researches in Communications, Electrical and Computer Engineering, EMEH 11, 134–142 (2011)
4. Blahuta, J., Soukup, T., Čermák, P.: The ROI Defect Statistical Analysis of Substantia Nigra to Reproducibility of Designed Experimental Algorithm for Potential PD Diagnosis. In: Mathematical Methods and Techniques in Engineering, pp. 261–266. WSEAS Press, Catania (2011)
5. Gelb, D., O. E. G.-S.: Diagnostic Criteria for Parkinson´s Disease. Archives of Neurology 56(1), 33–39 (1999)
6. Grossberg, S., Carpenter, A.: Neural Networks for Vision and Image Processing (1992)
7. Petrou, M., Wiltshire, S., Image Processing, G.P.: Dealing with Texture (2006)
8. Principe, J.C., Euliano, N.R., Lefebvre, W.C.: Neural and Adaptive Systems: Fundamentals Through Simulations. John Wiley & Sons, Inc. (2000)
9. Schreiber, J., Sojka, E., Ličev, L., Škňouřilová, P., Gaura, J., Školoudík, D.: A New Method for the Detection of Brain Stem in Transcranial Ultrasound Images. Proc. of Biosignals 2, 478–483 (2008)
10. Shoham, Y., Leiton-Brown, K.: Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations. Cambridge University Press (2008)
11. Smagt, P., Kröse, B.: An Introduction to Neural Networks, The University of Amsterdam, 8th edn. (1996)
12. Školoudík, D.: Reproducibility of Sonographic Measurement of the Substantia Nigra. Ultrasound in Medicine & Biology (9), 1347–1352 (2007)
13. Walter, U., Wittstock, M., Benecke, R., Dressler, D.: Substantia Nigra Echogenicity is Normal in Non-extrapyramidal Cerebral Disorders but Increased in Parkinson's Disease. J. Neural Transm. 109, 191–196 (2002)
14. Webster, D.D.: Critical Analysis of the Disability in Parkinson's Disease. Mod. Treat 5, 257–282 (1968)

Classification and Detection of Diabetic Retinopathy

Ahmad Taher Azar[1] and Valentina E. Balas[2]

[1] Faculty of Engineering, Misr University for Science & Technology (MUST),
6th Of October City, Egypt
[2] Aurel Vlaicu University of Arad, Romania
ahmad_T_azar@ieee.org, balas@drbalas.ro

Abstract. Diabetic retinopathy (DR) is the leading cause of blindness in adults around the world today. Early detection (that is, screening) and timely treatment have been shown to prevent visual loss and blindness in patients with retinal complications of diabetes. The basis of the classification of different stages of diabetic retinopathy is the detection and quantification of blood vessels and hemorrhages present in the retinal image. In this paper, the four retinal abnormalities (microaneurysms, haemorrhages, exudates, and cotton wool spots) are located in 100 color retinal images, previously graded by an ophthalmologist. A new automatic algorithm has been developed and applied to 100 retinal images. Accuracy assessment of the classified output revealed the detection rate of the microaneurysms was 87% using the thresholding method, whereas the detection rate for the haemorrhages was 88%. On the other hand, the correct classification rate for microaneurysms and haemorrhages using the minimum distance classifier was 60% and 94% respectively. The thresholding method resulted in a correct detection rate for exudates and cotton wool spots of 93% and 89% respectively. The minimum distance classifier gave a correct rate for exudates and cotton wool spots of 95% and 86% respectively.

Keywords: Diabetic Retinopathy (DR); Blindness; Feature Extraction; Image Processing; Classification; Minimum Distance Classifier (MDC); Microaneurysms (MA), Hemorrhages (HR), Exudates (EX), Cotton Wool Spots (CWS).

1 Introduction

Vision is the most powerful sense that provides a remarkable amount of information about surrounding that enables to interact intelligently with the environment, all without direct physical contact. Nine out of ten people with diabetes eventually develop a complication that affects the eyes known as "diabetic retinopathy" [1] [2]. This complication affects the blood vessels inside the eye and can lead to blindness if untreated [3] [4]. Screening is vital to preventing visual loss from diabetes because retinopathy is often asymptomatic early in the course of the disease [5]-[12]. If the retinopathy is detected in its early stages, blindness can be prevented in more than 50% of the cases [13] [14]. These patients with diabetes mellitus should have their eyes checked by an ophthalmologist at least once a year to see if retinopathy is in progress. These examinations are very time and labour consuming. It is therefore of great interest to develop

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 135–145.
DOI: 10.1007/978-3-319-00029-9_12

an automatic diabetic retinopathy screening system capable of differentiating between people with no retinal abnormalities and those who have, or might have, some kind of abnormalities. This would allow more people to be examined per year and the ophthalmologist to spend more time on those people who are actually in need for their expertise. Medical personnel can also immediately schedule appointments for the patients without further delay for continued diagnosis and follow-up visits at a regular hospital. The present work considers the localization and classification of the retinal abnormalities. There are different kinds of abnormal lesions caused by diabetic retinopathy and the focus is on four retinal abnormalities: Microaneurysms (MA), Hemorrhages (HR), Exudates (EX), Cotton Wool Spots (CWS) in color images from diabetic patients. Figure 1 show the retinal image affected by diabetic retinopathy with Microaneursyms as an abnormal lesion. These abnormalities have been selected as they are used in gradient the retinopathy of diabetics. An algorithm for automatic detection of the retinal dark and bright abnormalities has been developed. It depends on two classification methods; these are thresholding and Minimum Distance Classifier (MDC). It accomplishes several steps for detecting both retinal abnormalities. These are image processing, image smoothing, feature extraction, labeling of binarized images and classifier. The algorithm determines whether the person fundus image is normal (not diabetic person), abnormal (diabetic person), and gives an accurate classification for the two types of abnormalities (dark and bright) automatically and their subclasses.

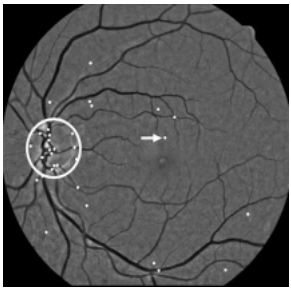


Fig. 1. Microaneursym in retinal image

2 Material and Method

The retinal images were acquired using an ophthalmoscope equipped with a 3-ccd chip RGB-camera from well-known eye hospitals. The image resolution obtained is 640 by 480 in 32 bit RGB. In digital image processing, images are either indexed images or RGB (Red, Green, Blue) images. An RGB image is an M x N x 3 array of colour pixels, where each colour pixel is a triplet corresponding to red, green and blue components of RGB image at specified special location. The range of value of an RGB is determined by its class. An RGB image of class double, has value in the range of [0 1], while class of uint8 is [0 255], similarly for the range [0 65535] is called class uint16. A trained technician performed the image acquisition 100 images are

available (see Fig. 2). Every image is taken with field-of-view of using "Topcon Image Net Camera". The retinal image abnormalities are subdivided basically into two groups black abnormalities (MAs & HEs) and white abnormalities (EXs & CWSs). The methodology adopted for the detection techniques is described in the following subsections.

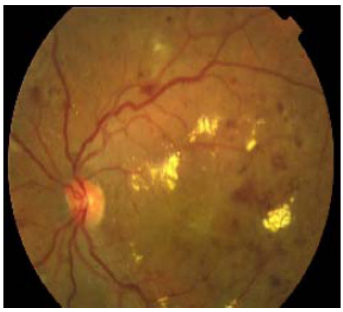


Fig. 2. Retinal fundus image

2.1 Detection of Dark Abnormalities

Dark abnormalities are, Microaneurysms (MAs) which are small dots of equivalent diameter about 1-3 pixel with approximately the same colour as blood vessels (the blood vessels where aneurysm has developed is often visible as the blood vessels are very small compared with aneurysms and/or located beneath the aneurysms, they have well defined edges and usually circular), and Haemorrhages (HEs) having a diameter of about 3-10 pixels, where the diabetic causes structural deformations in the wall of blood vessels that significantly larger depending on the size of the leak [15]-[18]. Detection of dark abnormalities (Mas, HEs) has many steps; they are image preprocessing, feature extraction and classification as shown in Fig. 3.

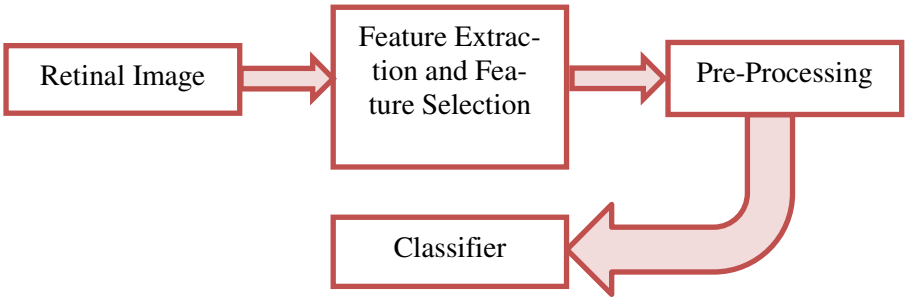


Fig. 3. Classification scheme for retinal images

2.1.1 Image Pre-processing

A colour or a grey level image is the input to the Pre-Processing phase. In this phase some problems are corrected such as the illumination variation of the image, enhancement of the image contrast and detection of image abnormalities. Pre-Processing phase involves sub-phases such as colour space conversion, zero padding of image edges, sharpening filtering and histogram equalization.

The original image (see Fig. 2) is resized to 512×480 pixels and Retinal images are normally captured in full color, using only green band [19] [20]. The green band image is shown in Fig. 4. The raw retinal images are contaminated with noise resulted from the image acquisition and analog-to digital conversion. This noise was reduced using 5×5 average filter. Figure 5 shows the output of the average filter but at the expense of significant image blur as it tends to smear the edges of objects in the image [21].

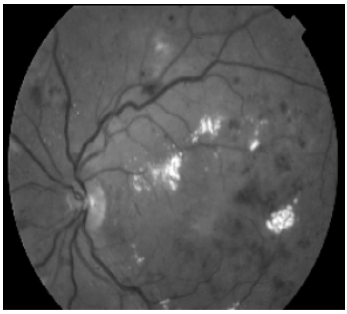


Fig. 4. Green band of fundus image

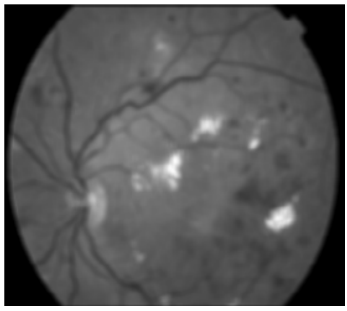


Fig. 5. Green band image smoothed using a 5×5 average filter

2.1.2 Feature Extraction

Some potential abnormalities are so small that they can be hard to see. Therefore the noise from other structures can influence the detection of these abnormalities, whereas the detection of dark abnormalities can be done using two dimensional matched filters (TDMF) [22]. It is capable to detect blood vessels with low contrast, high contrast objects also have a tendency to show up, and this is done by the application of TDMF. The filter employs three primary principles. First, pairs of blood vessel walls may be approximated as linear due to typically small curvature. Second, vessels appear darker than the background as a result of low reflectance in comparison to other retinal surfaces. The profile may be modelled as a Gaussian. Finally, the width of blood vessels is variable [22]. The resulted temple image is thresholded and the output is a binary image containing the blood vessels and other high contrast structures (dark abnormalities, edge, optic disc) as illustrated in Fig. 6.

Fig. 6. Thresholding with TDMF with dark abnormalities

The abnormalities have been appeared on the binary image are all about the same size. Features used to detect dark abnormalities are listed in Table 1.

Table 1. Features for detecting dark abnormalities

Features (pixel)	Microaneurys (ms)	Hemorrhages
Equivalent Diameter	1-3	< 3-10
Major Axis Length	2-3	< 3-15
Minor Axis Length	2-3	< 3-4
Area	4-6	< 6-40

Computing these features indicate the presence of dark abnormalities in the pre-processed image. Since in normal retinal images mostly there are not counted features, while in other diseased retinal images, the number of detected regions in the images with certain features proportional to the state of the diabetic patient. Because of some complications in detecting such as a clear differentiation between MA & HE is not obtained and a tendency for HEs to be classified as MAs, According to these interferences, a threshold level is used to differentiate among the types of dark abnormalities. The automatic detection of Microaneurysms and Hemorrhages are shown in Fig. 7(a) and Fig. 7(b).

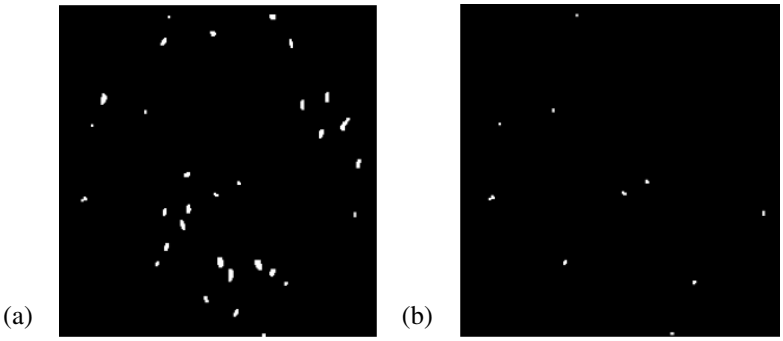


Fig. 7. (a) The automatic detection of Microaneurysms (b) The automatic detection of both Hemorrhages and Microaneurysms

2.1.3 Classification of Dark Abnormalities

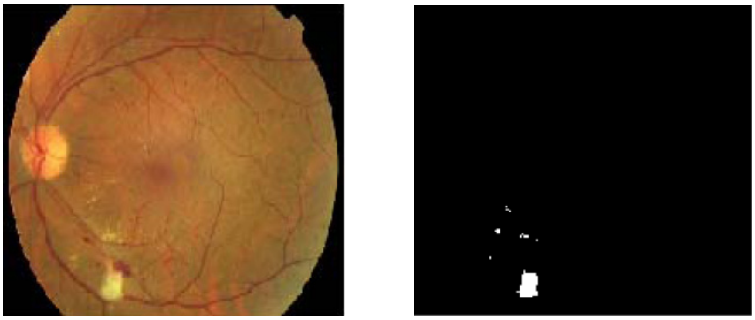
Application of classification techniques for pattern recognition is very effective method to differentiate between various objects in a pattern. Two approaches for classifying dark abnormalities were used: Thresholding and a minimum distance classifier. Thresholding Features used for dark abnormalities are listed in Table 1. They represent the shape of the dark abnormalities to separate them from the dark noise. The noise constitutes a major problem, as the retinal abnormalities are small in size so the noise may be classified as an abnormality and an abnormality may be classified as noise [23] [24]. The thresholding level plays an important role in differentiating the retinal abnormalities from the present noise that leads to a precious decision of retinal images normality. The thresholding levels of dark abnormalities of both MA and HE were chosen so that the normal and abnormal retinal image is diabetic or not. The performance of the algorithm is expressed using sensitivities; i.e. the ratio between the number of correctly classified abnormalities in a class and total number of abnormalities multiplied by 100%. Minimum Distance Classifier (MDC) classifies a pattern based on its distance in a feature space to class prototypes. Although it is one of the earliest methods suggested, the minimum distance classifier is still effective distance classifier and still effective tool in solving the pattern classification problem [25] [26]. Euclidian distance was used as a criterion for measuring distance to class prototype.

2.2 Detection of White Abnormalities

White abnormalities are Exudates (EXs) and Cotton Wool Spots (CWSs). Exudates are similar to Haemorrhages in that the blood vessels leak lipid, bright, with diameter of 1-10 pixels (this matter is not as this build a bright spots with a well-defined edges). Exudates tend to build near each other, sometimes creating circles of exudates [11] [12]. Cotton wool spots (CWSs) are largest of these abnormalities and they are retinal infracts caused by thrombosis and obstruction of the blood vessel. This leads to damage ischemic to the part of the retina supplied by the damaged vessel [11] [12]. The normal red color of the retina (due to the blood vessels supply) appears white, when the blood supply is cut off. Due to the overlap of the blood vessels supply, there are no well edges of cotton spots and the shape is very irregular and the CWSs diameter is about 10- 25 pixels. Detection of bright abnormalities can be done using several steps; they are image pre-processing, image labeling, feature extraction and classifier.

2.2.1 Image Pre-processing

The retinal fundus image is resized into 512×480 (Fig. 8.a) and the retinal image is normally captured in full color, but the green band is only used [9]. The green band retinal image is smoothed with 5×5 average filter and then thresholded into a binary image Fig. 8.b.



(a) (b)

Fig. 8. (a) Original fundus image (b) The automatic detected white abnormalities (Exudates (EXs) & Cotton wool spots (CWSs)) in the thresholded image

2.2.2 Feature Extraction

White abnormalities features are Equivalent diameter, Major axis Length as shown in Table 2.

Table 2. Features Used For Bright Abnormalities

Features (pixel)	Exudates	Cotton wool spots
Equivalent Diameter	1-10	< 10-25
Major Axis Length	1-15	10-15

2.2.3 Classification of Bright Abnormalities

Classification of the two bright abnormalities (i.e.: Exs & CWSs), were performed using the two methods of thresholding and minimum distance classifier mentioned previously

3 Results

3.1 Results of Dark Abnormalities

The application of an algorithm to 100 fundus images and the distribution of the number of detected abnormalities in images with optimum thresholds are shown in Fig. 9 and Fig. 10.

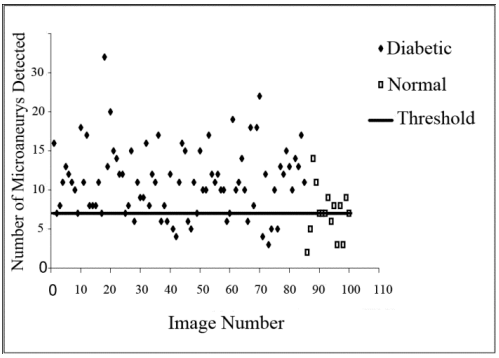


Fig. 9. The number of Microaneurysms in retinal image using the optimum thresholding level

The thresholding method gives a detection rate for the micro aneurysms of 87%, whereas the detection rate for the Haemorrhages was 88%. The minimum distance classifier gives a correct classification rate for micro aneurysms and Haemorrhages of 60% and 94% respectively.

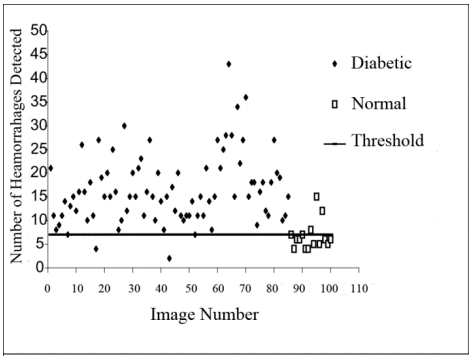


Fig. 10. The number of hemorrhages in retinal image using the optimum thresholding level

3.2 Results of White Abnormalities

The distribution of the number of detected abnormalities in images with optimum thresholds is shown in Fig. 11 and Fig. 12.

The thresholding method results in a correct detection rate for exudates and cotton wool spots were 93% and 89% respectively. The minimum distance classifier gives a correct rate for exudates and cotton wool spots of 95% and 86% respectively.

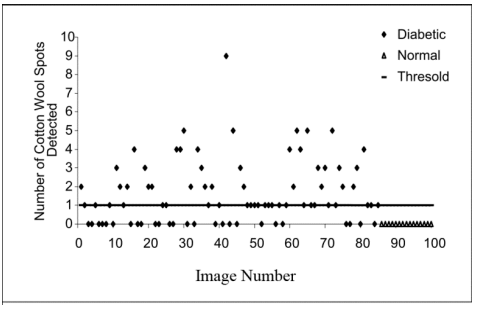


Fig. 11. The number of exudates in retinal image using the optimum thresholding level

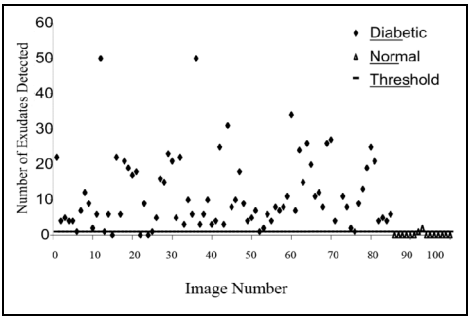


Fig. 12. The number of cotton wool spots in retinal image using the optimum thresholding level

4 Conclusion

The eye diseases mainly contribute to blindness and often can't be remedied because the patients are diagnosed too late with the diseases. The paper presents new automatic approach for detecting retinal abnormalities. The developed algorithm helps in deciding whether the patients with potential sight threatening retinopathy and needs further examination or patients not in need of further referral. The latter would be asked to return in 6-12 months for a new retinal photographing. Classified patients who need further examination will be examined by an ophthalmologist to perform a manual examination in order to provide a detailed diagnosis, and decide on further treatment, e.g. Laser. To facilitate the ophthalmologist decision, the proposed automatic detection algorithm helps in the final decision for the patient infection without confusion.

Acknowledgement. The research of Valentina Emilia Balas was supported by the Bilateral Cooperation Research Project between Romania and Slovakia (2011-2012) entitled: "Hybrid Medical Complex Systems", CogniMediSys, the involved institutions are the Institute of Informatics of the Slovak Academy of Sciences and Petru Maior University.

References

1. Morello, C.: Etiology and Natural History of Diabetic Retinopathy: an Overview. Am. J. Health Syst. Pharm. 64(17 suppl. 12), 3–7 (2007)
2. Gardner, T.W., Antonetti, D.A., Barber, A.J., LaNoue, K.F., Levison, S.W.: Diabetic Retinopathy: More than Meets the Eye. Surv. Ophthalmol. 47(suppl. 2), S253–S262 (2002)
3. Serrarbassa, P.D., Dias, A.F., Vieira, M.F.: New Concepts on Diabetic Retinopathy: Neural Versus Vascular Damage. Arq. Bras. Oftalmol. 71(3), 459–463 (2008)
4. Barber, A.J.: A New View of Diabetic Retinopathy: a Neurodegenerative Disease of the Eye. Prog. Neuropsychopharmacol. Biol. Psychiatry 27(2), 283–290 (2003)
5. Goatman, K., Charnley, A., Webster, L., Nussey, S.: Assessment of Automated Disease Detection in Diabetic Retinopathy Screening Using Two-field Photography. PLoS One 6(12), e27524 (2011)
6. Verma, K., Deep, P., Ramakrishnan, A.G.: Detection and Classification of Diabetic Retinopathy Using Retinal Images. In: Annual IEEE India Conf (INDICON), pp. 1–6 (2011)
7. Jones, S., Edwards, R.T.: Diabetic Retinopathy Screening: a Systematic Review of the Economic Evidence. Diabet. Med. 27(3), 249–256 (2010)
8. Rodgers, M., Hodges, R., Hawkins, J., Hollingworth, W., Duffy, S., McKibbin, M., Mansfield, M., Harbord, R., Sterne, J., Glasziou, P., Whiting, P., Westwood, M.: Colour Vision Testing for Diabetic Retinopathy: a Systematic Review of Diagnostic Accuracy and Economic Evaluation. Health Technol. Assess. 13(60), 1–160 (2009)
9. Farley, T.F., Mandava, N., Prall, F.R., Carsky, C.: Accuracy of Primary Care Clinicians in Screening for Diabetic Retinopathy Using Single-image Retinal Photography. Ann. Fam. Med. 6(5), 428–434 (2008)
10. Bloomgarden, Z.T.: Screening for and Managing Diabetic Retinopathy: Current Approaches. Am. J. Health Syst. Pharm. 64(17 suppl. 12), S8–S14 (2007)
11. Chew, E.Y.: Screening Options for Diabetic Retinopathy. Curr. Opin. Ophthalmol. 17(6), 519–522 (2006)
12. Sinclair, S.H.: Diabetic Retinopathy: the Unmet Needs for Screening and a Review of Potential Solutions. Expert. Rev. Med. Devices 3(3), 301–313 (2006)
13. Xu, J., Hu, G., Huang, T., Huang, H., Chen, B.: Using Multifocal ERG Responses to Discriminate Diabetic Retinopathy. Doc Ophthalmol. 112(3), 201–207 (2006)
14. Jin, X., Guangshu, H., Tianna, H., Houbin, H., Bin, C.: The Multifocal ERG in Early Detection of Diabetic Retinopathy. Conf.Proc. IEEE Eng. Med. Biol. Soc. 7, 7762–7765 (2005)
15. Dupas, B., Walter, T., Erginay, A., Ordonez, R., Deb-Joardar, N., Gain, P., Klein, J.C., Massin, P.: Evaluation of Automated Fundus Photograph Analysis Algorithms for Detecting Microaneurysms, Haemorrhages and Exudates, and of a Computer-assisted Diagnostic System for Grading Diabetic Retinopathy. Diabetes Metab. 36(3), 213–220 (2010)
16. Fleming, A.D., Goatman, K.A., Philip, S., Williams, G.J., Prescott, G.J., Scotland, G.S., McNamee, P., Leese, G.P., Wykes, W.N., Sharp, P.F., Olson, J.A.: The Role of Haemorrhage and Exudate Detection in Automated Grading of Diabetic Retinopathy. Br. J. Ophthalmol. 94(6), 706–711 (2010)
17. Fleming, A.D., Philip, S., Goatman, K.A., Williams, G.J., Olson, J.A., Sharp, P.F.: Automated Detection of Exudates for Diabetic Retinopathy Screening. Phys. Med. Biol. 52(24), 7385–7396 (2007)
18. Fleming, A.D., Philip, S., Goatman, K.A., Olson, J.A., Sharp, P.F.: Automated Microaneurysm Detection Using Local Contrast Normalization and Local Vessel Detection. IEEE Trans. Med. Imaging 25(9), 1223–1232 (2006)

19. Patton, N., Aslam, T.M., MacGillivray, T., Deary, I.J., Dhillon, B., Eikelboom, R.H., Yogesan, K., Constable, I.J.: Retinal Image Analysis: Concepts, Applications and Potential. Prog. Retin. Eye Res. 25(1), 99–127 (2006)
20. Hoover, A., Kouznetsova, V., Goldbaum, M.: Locating Blood Vessels in Retinal Images by Piecewise Threshold Probing of a Matched Filter Response. IEEE Trans. Med. Imaging 19(3), 203–210 (2000)
21. Marrugo, A.G., Millan, M.S.: Retinal Image Analysis: Preprocessing and Feature Extraction. Journal of Physics: Conference Series 274, 012039 (2011), doi:10.1088/1742-6596/274/1/012039
22. Chaudhuri, S., Chatterjee, S., Katz, N., Nelson, M., Goldbaum, M.: Detection of Blood Vessels in Retinal Images Using Two-dimensional Matched Filters. IEEE Trans. Med. Imaging 8(3), 263–269 (1989)
23. Kavitha, G., Ramakrishnan, S.: Abnormality Detection in Retinal Images Using ant Colony Optimization and Artificial Neural Networks - Biomed 2010. Biomed. Sci. Instrum. 46, 331–336 (2010)
24. Sri Madhava, N.R., Kavitha, G., Ramakrishnan, S.: Assessment of Retinal Vasculature Abnormalities Using Slantlet Transform Based Digital Image Processing - Biomed 2011. Biomed. Sci. Instrum. 47, 88–93 (2011)
25. Anitha, J., Selvathi, D., Hemanth, D.J.: Neural Computing Based Abnormality Detection in Retinal Optical Images. In: Proc. of the IEEE International Advance Computing Conf., Patiala, pp. 630–635 (2009), doi:10.1109/IADCC.2009.4809085
26. Anitha, J., Vijila, C.K., Hemanth, D.J., Ahsina, A.: Self Organizing Neural Network Based Pathology Classification in Retinal Images. In: Proc. of the World Congress on Nature and Biologically Inspired Computing, Coimbatore, pp. 1457–1462 (2009), doi:10.1109/NABIC.2009.5393697

Principal Component Analysis Used in Estimation of Human's Immune System, Suffered from Allergic Rhinosinusopathy Complicated with Clamidiosis or without It

Lyudmila Pokidysheva[1] and Irina Ignatova[2]

[1] Siberian Federal University, 79 Svobodny prospect, Krasnoyarsk, 660041 Russia
[2] State Institution Research Institute of Medical Problems of the North, Siberian Department of Russian Academy of Medical Science, 3G Partizana Zheleznyaka, Krasnoyarsk, 660021
pok50gm@gmail.com, ignatovai@mail.ru

Abstract. The immune system is very important in overcoming the influence of harmful factors to human organism. This is one of the three integrative systems of the organism, which provides maintenance of homeostasis together with the nervous and endocrine systems. In this paper, estimation of human's immune system, suffered from allergic rhinosinusopathy, complicated or not complicated by clamidiosis is given. The novelty of method consists in using correlation adaptometry and principal component analysis for this estimation. These methods allowed estimation of changes of the immune system during stress adaptation.

Keywords: Principal component analysis, immune system, allergic rhinosinusopathy, correlation adaptometry.

1 Introduction

The increase of allergic diseases is registered around the world. There is evidence that the prevalence of allergic diseases over the past 50 years has increased by 10 times. The reasons for this rapid increase in prevalence are not entirely clear. Most often they are associated with changes in environmental conditions. Incidence rates allergies are much higher in industrial areas. It is no accident that allergic disease is defined as «diseases of the century». Share allergic rhinosinusopathy (AR) is very high in the structure of allergic diseases (60-70%). It occurs in many countries. Course of the disease depends on the immune system.

The immune system is a unique natural protective mechanism. Due to the coherence of the entire functional system of immunity, the organism is able to confront to a number of factors that have a negative impact. Feature of the immune system is that it contains a regulatory function, capable on the presence of each specific antigen (which is the starting point of the disease) to answer by specific immune response in accordance with its individual characteristics [1, 2, 3].

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 147–156.
DOI: 10.1007/978-3-319-00029-9_13

We have considered the adaptation response of immune system on the presence of Chlamydia infection (CI) in organism of patients with allergic rhinosinusopathy. Mostly young people, who have just entered a period of sexual activity, are ill by Chlamydia. Strong tendency to increase the etiologic significance of intracellular pathogens (Chlamydia, mycoplasma, ureaplasma) and fungi Candida, obviously, can be explained by the evolutionary processes in the microbial world [4-6]. Pathological process may affect different organs and systems, including the eyes, upper, middle and lower respiratory tract, genitourinary system, urinary tract, cardiovascular system, liver and biliary tract, lymph nodes, joints, etc.

The frequency of mucosal lesions CI nasal passages is caused by the rapid increase in the prevalence of acute and chronic diseases of the nose among the child and the adult Russian population [7-10].

Given the diversity of peripheral blood cell parameters which characterize the state of an immune system, it is obvious that statistical mathematical modeling methods need to be applied (L. White, 1988; Rebrov O.Yu, 2002). Method of adaptive adaptometry is one of such approaches. It allows assessing stress changes in functional systems of a body during adaptive stress not judging by values themselves, but by extent to which they are co-related.

Other of these approaches can serve as principal component analysis [11, 12], allowing to replace the initial interrelated features on a certain set of uncorrelated parameters, and describe the object by a smaller number of indicators.

The rest of this work is organized as follows. Section 2 describes the method of estimation of human's immune system during disease, which is a stress for the human organism. In Section 3 the results of our research are discussed. Conclusions are given in Section 4.

2 Description of Methods

In our investigation, a total of 305 people (70 - men and 235 women) from 15 to 79 years old were examined. Those were patients with different pathogenic variants of allergic rhinosinusopathy (AR). The following groups of AR were distinguished: true AR (IAR) - 52 cases (17.05%), pseudo-AR (PAR) - 151 (49.5%) and mixed form of AR (SAR) - 102 cases (33.44%.) Investigations were carried out with the help of Professor, Doctor of Biology Savchenko, A., and Rodina D.V. There are patients with AR and CI being present and without it. State of immune system was assessed by the classic indices of cellular and humoral immunity in patients with different pathogenic variants of AR during periods of exacerbation and remission.

The method of correlation adaptometry allowed estimating the changes of the immune system during stress adaptation not according to the ideas themselves, but according to their correlation.

Previously, we studied the adaptive capacity of blood cells patients with the Allergic Rhinosinusopathy (AR), which is increased by chlamydia and without it. The adaptation processes of blood cells, CD-cells and immunoglobulins in general during the exacerbation and remission of various pathogenetic forms of AR were investigated by correlation adaptometry. The method is based on the effect of increasing

variances and the correlations between indicators in the adaptive response to the impact of harmful factors and decreasing successful adaptation or termination (blocking) of such effects [13-17].

The immune system performance was evaluated by the classical indicators of the cellular and humoral part of immunities in patients with different pathogenic forms of AR in the periods of exacerbation and remission.

To clarify relationships between stress indices for venous blood cells, CD-cells and for immunoglobulins in Eastern Siberia population for groups with different pathogenetic forms of allergic rhinosinusopathy, complicated or not complicated with chlamydia during periods of exacerbation and remission we used method of correlation adaptometry.

Values of average, error of the mean, variance, standard deviation have been computed for each of this group. The hypothesis of normal distribution of the original sample (criterion χ^2) was tested. Correlation coefficients were calculated using the formula:

$$r_{kl} = \frac{cov(x_k, x_l)}{\sigma_{x_k} \cdot \sigma_{x_l}}, \tag{1}$$

where $k,l = 1...8$,

$$cov(x_k, x_l) = \sum_{i=1}^{N} \sum_{j=1}^{N} (x_{k_i} - \overline{x_k})(x_{l_j} - \overline{x_l}) P_{x_k x_l}, \tag{2}$$

$\overline{x_k}$ - Average of x_k,

$\overline{x_l}$ - Average of x_l,

$P_{x_k x_l}$ - Probability of $P(x_{k_i} = \overline{x_k}, x_{l_i} = \overline{x_l})$,

σ_{x_k} - Standard deviation of x_k,

σ_{x_l} - Standard deviation of x_l.

Analysis of cells' parameters pair correlation is carried out for all modified parameters at each stage of the strain (period of exacerbation and remission) and for all the groups.

We identified the number of reliable correlations out of all examined correlation coefficients and also evaluated the intensity of these connections. The extent to which parameters are connected was estimated using a correlation graph's weight, calculated as the sum of its edges' weights (the sum of corresponding pair correlation coefficients):

$$G = \sum_{|r_{ij}| \geq \alpha} |r_{ij}| \tag{3}$$

where r_{ij} - the coefficients of correlation between the i-th and j-th performance, α is determined by the level of confidence r_{ij}. We have taken only reliable correlation coefficients into account, whose values are greater than or equal to α. We took into

account the correlation coefficients greater than 0.5, thus we considered only the strong relations ($\alpha \geq 0.5$).

The principal components define the factor structure of the phenomenon, which is characterized by the correlation coefficients between features and principal components. The internal structure of phenomena does not characterize features themselves, but characterizes the internal factors and properties, which is inherent of this set of objects. They are not always directly measurable. Having obtained these coefficients, we can see what features perform a dominant role under the harmful effects of CI.

The immune system can be represented as a set of interacting subsystems. The principal components method was used for studying the structure of the relations between indicators of subsystems immunity tension: peripheral blood cells (first subsystem), CD-cells (second subsystem), general immunoglobulin (third sub-system) depending on the presence of CI and without it. When analyzing the groups of adaptive systems to the effects of chlamydial infection, we identified patterns of adaptation and the role of each subsystem (group of cells) with Chlamydia being present and without it.

According to the method of principal components, each internal factor can also be represented as a function of the observed symptoms. Thus, we can move on to new integrated quantitative indicators of generalized properties, which were involved in this factor. In this case, the functional immune system can be represented as a set of interacting subsystems of cells.

The relationship between indicators is greater under conditions of extremely harmful factors. With a strong correlation of the initial signs of all the variability is concentrated in the first l factors. Therefore, setting the proportion of total variance, which should be attributed to the first l components (typically 80-95% of the total variance of symptoms), we can compare the group by the number 1 of the first principal components.

3 Results of Research

Our studies have shown certain regularities of changes in the correlation graph (with a total value of the relationship between the indicators were taken into account the correlation coefficients greater than 0.5), being dependant from the pathogenic forms and stages of AR. The dominant role of lymphocytes and total immunoglobulin compared with other blood cells is evident in restoring homeostasis of patients with AR mixed genesis.

Calculated results of pair correlation between the immune system stress indices have revealed certain regularities in correlation graph changes during periods of exacerbation and remission, irrespective of the pathogenic AR type. The behavior of correlation graph's weight in all surveyed groups is the same: during remission it decreases. In the period of exacerbation the stress increases, enabling recovery of all morphological and functional blood cell parameters, a body utilizes additional resources, what is indicated by increased weight of the correlation graph.

Research data on correlation graph weight G for patients with allergic rhinosinusopathy infected with chlamydia, not infected with chlamydia are given in Table 1.

Table 1. Integrated graph's weight G for patients with allergic rhinosinusopathy infected with chlamydia, not infected with chlamydia and for the entire sample

Presence of chlamydia infection in nasal mucosa in patients with AR	Weight G value	
	exacerbation	remission
G_+ AR infected with chlamydia	11.68	1.84
G_- AR not infected with chlamydia	10.48	1.54

From data in Table 1 it follows that integrated correlation graph's weight G for venous blood cells parameters, CD-cells and total immunoglobulins in different clinical groups behaves similarly, i.e. decreases in remission.

In patients with AR who are infected with chlamydia correlation graph weight G_+ changes in the following way: from 11.68 to 1.84 ($p < 0.5$). In patients with allergic rhinosinusopathy not infected with chlamydia, correlation graph weight G_- decreased from 10.48 down to 1.54 ($p < 0.5$). This suggests that Chlamydia infection is an additional stress for the body and contributes to homeostasis disruption (structural and functional abnormalities of cells). Also it suggests that in this group of patients, an additional suppression of immune cells' functional activity and accumulation of metabolic wastes takes place (increasing body intoxication).

We have also studied a correlation graph weight for venous blood cells parameters, CD-cells and total immunoglobulins in patients with AR who are infected with chlamydia. The results are shown in Table 2.

Table 2. Correlation graph weight for venous blood cells parameters, CD-cells and total immunoglobulins in patients with AR who are infected with chlamydia

Blood parameters	G-weight value	
	exacerbation	remission
Venous blood	1.34	0.93
CD-cells and total immunoglobulins	6.99	0.53

Looking at data presented in Table 2 it can be said that correlation graph weight G for venous blood cell parameters is less than ($p < 0.5$) G-values for CD-cells and total immunoglobulins in patients with AR who are infected with chlamydia. The obtained G-values for venous blood cells, CD-cells and the total immunoglobulins in patients with AR who are not infected with chlamydia are given in Table 3.

Table 3. Correlation graph weight value for venous blood cells, CD - cells and total immunoglobulins in patients with AR who are not infected with chlamydia

Blood parameters	G-weight value	
	exacerbation	remission
G_v venous blood cells	0.93	0.91
G_{CD} CD-cells and total immunoglobulins	9.10	0.64

Based on data in Table 3, one might conjecture that correlation graph weight G_v for venous blood cells' parameters, as well as G_{CD} value for CD-cells and total immunoglobulins is reduced during remission in patients with AR who are not infected with chlamydia. Correlation graph weight G_v decreased from 0.93 down to 0.9 ($p > 0.5$), and G_{CD} value went from 9.1 down to 0.64 ($p < 0.5$). Thus, this indicates that immune cells play a key role in this sample of patients as well.

Our studies have shown that chlamydial infection contributes an additional burden for the organism and contributes to the state of homeostasis (structural and functional abnormalities of cells). There is an additional suppression of the functional activity of immunocompetent cells in these patients.

Organism involves adaptive resources, ensuring the restoration of all the morphological and functional parameters of blood cells, so it shows the increase of the correlation graph weight when AR is complicating by the chlamydial infection. Restoration of normal functioning of all blood cells leads to decrease of the correlation graph weight. Consequently, the removal of the above pathological processes leads to significant improvement in the status and tension release in the body's immune system.

The analysis of the correlation graph dynamics of blood cells, CD – cells and total immunoglobulins in patients with AR, non-infected by Chlamydia, showed a significant decrease in its weight to the indicators of CD-cells and total immunoglobulin ($p < 0.05$).

The correlation graph reflects the degree of accumulation of metabolic toxins within the cells and, consequently, the degree of structural and functional abnormalities of blood cells as well as CD-cells. The organism of the patients with chlamydia is caused by (involves) additional resources, ensuring the restoration of all the morphological and functional parameters of blood cells, the tension increases, and so the value of the weight of the correlation graph increases.

The investigation has shown that the AR chlamydial infection, although it is not the primary etiologic agent of the AR, but implies extra stress on the organism, adding to the disruption of homeostasis (the structural and functional abnormalities of cells).

We calculated the total variance of all components and the contribution of each component in the total variance were, estimated the number l of the first principal component and calculated the variance explained by these components for all groups.

The most significant contribution is the first principal component (its contribution to the overall variability is 30%).

The first principal component correlates with all treatment of the symptoms, but the greatest influence on its magnitude (strong correlation - positive and negative) presupposes the presence of CD-cells (lymphocytes). The biggest factors are peripheral blood cells in the second main component.

In general, the projection of multidimensional data onto the plane of the first two principal components is presented as a data cloud (Fig.1).

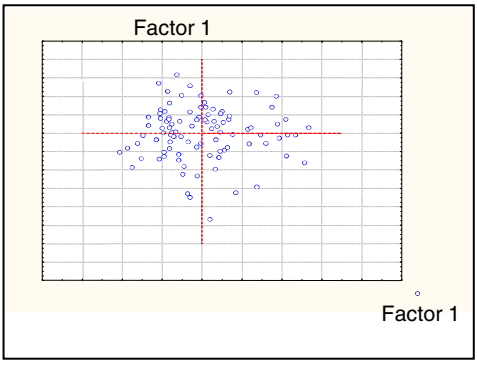


Fig. 1. The projection of multidimensional data (patients with RA) to the plane of the first two principal components (Factor 1: 30.06%, Factor 2: 13.29%)

On Fig.2 below is shown the leading role of CD-cells in the adaptive process of the immune system in patients AR without CI.

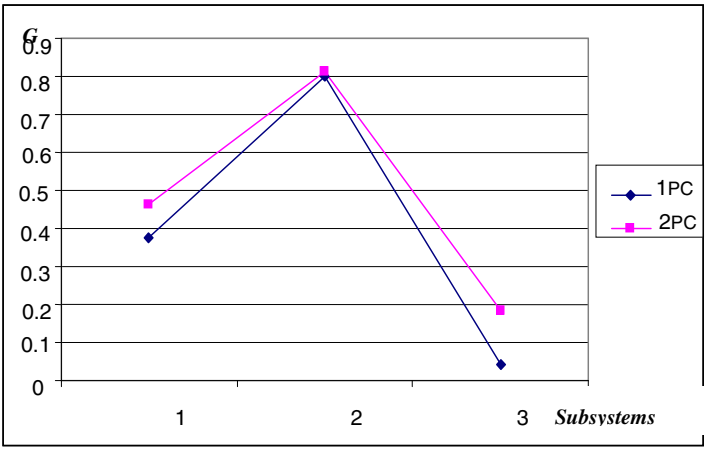


Fig. 2. The integration of individual subsystems of immune system for the first (1 PC) and the second principal component (2 PC) in patients with AR. 1 - a subsystem of peripheral blood cells, 2 - Subsystem of cell lymphocytes, 3 - Subsystem of cell immunoglobulins

The picture changes dramatically when burdening AR by Chlamydia infection. For this group, the proportion variance explained by the first seven principal components is equal to 94.4%.

The share of the first principal component is 31.3% for this period. The correlation of the main component of CD-cells increases.

In the second principal component with the largest correlation coefficients included indicators IgA, IgM and IgG.

While applying the projections of multidimensional data onto the plane of the first two principal components (Figure 3), they are stretched along the second factor.

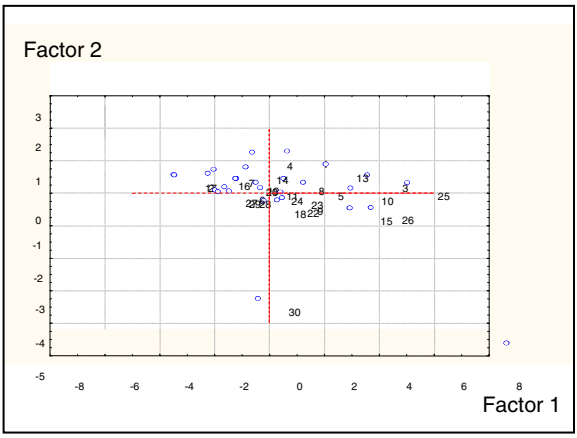


Fig. 3. The projection of multidimensional data (patients with AR, increased by CI) to the plane of the first two principal components (Factor 1: 31.30%, Factor 2: 20.74%)

The dominated role belongs to the subsystem of immunoglobulins (Fig. 4). It should ensure the adaptation process and the resources to overcome the stress state of the organism. The second principal component can be called a factor of immunoglobulins.

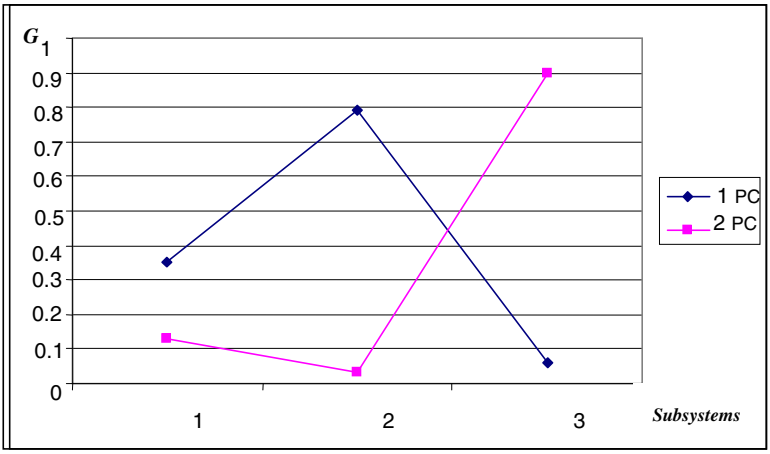


Fig. 4. The integration of individual subsystems of immune system for the first (1 PC) and second (2 PC) of the main component in accordance with AR, increased by CI. 1 – subsystem of blood cells, 2 – subsystem of cell lymphocytes, 3 – subsystem of cell immunoglobulins

4 Conclusion

For analyzing the values of pair correlations between the indicators we have identified certain patterns of voltage changes in immune system cells, as evidenced by the

weight dynamics of the correlation graph. The behavior of the correlation graph weight in all the surveyed groups is the same: during the exacerbation of the disease the organism includes additional resources to the restoration of all the morphological and functional parameters of blood cells. The tension increases, which shows the increase of correlation graph weight. During the period of remission, the weight of the graph decreases.

These results correlate with earlier studies by the correlation adaptometry [13], but it reveals the dominant role of immunoglobulin fractions in the protective adaptation processes to Chlamydia in patients with AR. The method of correlation adaptometry allowed estimating the strength of the immune system in AR and consider at the mechanisms of adaptation at the cellular level deeper.

The analysis of factor structure revealed the hidden patterns of adaptive mechanisms of the immune system of patients with AR with chlamydia and without it.

Principal component analysis allowed describing the internal factors that determine the dominant role of the individual subsystems of immunity at the cellular level, with the AP, is burdened by chlamydia, the periods of compensatory interactions.

References

1. Demchenko, E.V., Ivanchenko, G.F., Prozorovskaya, K.N., et al.: Clinic and Treatment of Chlamydia Laryngitis applying Amiksina. Vestnik Otolaryngologiy (5), 58–60 (2000)
2. Vishnyakova, L.A., Nikitina, M.A., Petrov, S.I., Vydumkina, S.P.: Immune Diagnosis of Acute Pneumococcal and Chlamydial Infections in the Epidemic Rise of Out-Of-Hospital Pneumonia among Children in St. Petersburg in 1998-2004. J. of Microbiology, Epidemiology and Immunobiology (3), 69–73 (2004)
3. Polevschikov, A.V.: Mucosal Immune System: The Molecules, Cells, and the Basic Cooperative Interactions / Ross. Rhinology (1), 22–25 (2004)
4. Falck, G., Engstrang, J., Gad, A.: Scand. Scand. J Infect. Dis. 29(6), 585–589 (1997)
5. Lobzin, Y.A., Liashenko, Y.I., Pozniak, A.: Chlamydia Pneumoniae Infections, St. (2003)
6. Granites, V.M.: Chlamydia. Nizhny Novgorod City (2000)
7. Hook, C.E., Telyatnikova, N., Goodall, J.C., et al.: Effects of Chlamydia Trachomatis Infection on the Expression of Natural Killer (NK) Cell Ligands and Susceptibility to NK Cell Lyses. Clin. Exp. Immunol. 138(1), 54–60 (2004)
8. Gavalova, S.M. (ed.): Chlamydia-Dysbiosis. Integral Correlations. Novosibirsk, RTF LLC, p. 218 (2003)
9. Ignatova, I.A., Smirnov, S.V., Shelkovnikova, E.A., Rodina, D.V.: The New «Mask» of Chlamydial Infection. In: 12th National Congress on Respiratory Diseases: A Compilation of Summaries. Pulmonology, Appendix, vol. S.250, XLIII 4 (2002)
10. Poliakova, T.S., Nechaeva, S.V., Polivoda, A.M.: The Role of Chlamydia and Mycoplasma Infection in the Etiology of Diseases in Otolaryngology. Journal of Otorhinolaryngology (1), 24–27 (2004)
11. Pearson, K.: On Lines and Planes of Closest Fit to Systems of Points in Space. Philosophical Magazine 2, 559–572 (1901)
12. Gorban, A.N., Zinovyev, A.Y.: Principal Graphs and Manifolds, Ch. 2. In: Olivas, E.S., et al. (eds.) Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques, IGI Global, Hershey (2009)

13. Ignatova, I.A., Pokidysheva, L.I., Smirnova, S.V.: Method of the Correlation Adaptometry in Estimating of Immune System's Tension of Patients with Allergic Rhinosinusopathy Complicated Chlamydia without it. Vestnik of Krasnoyarsk State University (4), 256–264 (2004)
14. Sedov, K.R., Gorban, A.N., Petushkova, E.V., Manchuk, V.T., Shalamov, E.N.: Adaptometry as a Method of Clinical Examination of the Population. Vestnik of the Academy of Medical Sciences of the USSR (10), 69–75 (1988)
15. Gorban, A.N., Smirnova, E.V., Tyukina, T.A.: Correlations, Risk and Crisis: From Physiology to Finance. Physica A 389(16), 3193–3217 (2010)
16. Gorban, A.N., Pokidysheva, L.I., Smirnova, E.V., Tyukina, T.A.: Law of the Minimum Paradoxes. Bull. Math. Biology (November 19, 2010), http://arxiv.org/abs/0907, doi:10.1007/s11538-010-9597-1
17. Krasnenko, A.N., Pokidysheva, E.V., Veretnova, K.Y., Tyukina, T.A.: Analysis of Correlations in the Russian Banking System in Adapting to the Economic Crisis of 2007-2008. Journal of Siberian Federal University. Mathematics and Physics 3(4), 521–532 (2010)

Computer-Aided Diagnosis of Laryngopathies in the LabVIEW Environment: Exemplary Implementation

Dominika Gurdak[1], Krzysztof Pancerz[1], Jaroslaw Szkola[1], and Jan Warchol[2]

[1] Institute of Biomedical Informatics, University of Information Technology and Management in Rzeszów, Poland
{dgurdak,kpancerz,jszkola}@wsiz.rzeszow.pl
[2] Department of Biophysics, Medical University of Lublin, Poland
jan.warchol@am.lublin.pl

Abstract. In the paper, we present a computer tool supporting a non-invasive diagnosis of selected larynx diseases. The tool is created on the basis of the LabVIEW environment. LabVIEW enables us to create, in an easy way, a user-friendly graphical interface facilitating both entering input data and visualizing results in order to make the platform ready to use directly in the medical community. Computer-aided diagnosis of laryngopathies, in the presented tool, is based on a family of coefficients reflecting spectrum disturbances around basic tones and their multiples for patients' voice signals.

Keywords: computer-aided diagnosis, laryngopathy, LabVIEW, decision support system.

1 Introduction

Computer support systems for medical diagnosis become increasingly more popular worldwide (cf. [7]). One of their advantages is automation of a diagnosis process. Moreover, such systems are based on objective measurements and observations of selected parameters. There is a big demand for creation of flexible, effective and user-friendly platforms for the intelligent support of diagnostic decisions for the medical community. Such platforms ought not to require the technical knowledge of medical personnel. The relevant element is to design modules of entering and processing medical data as well as modules of data visualization (especially, graphical). These two elements of the platform allow using the system without specialized courses for personnel.

A problem of creating decision support systems for diagnosis of laryngeal diseases was considered earlier in the literature. One of the approaches was based on an automated analysis of vocal cord images (color, shape, geometry, contrast, irregularity and roughness) [17]. Next, an approach based on color images of vocal folds, a voice signal, and questionnaire data was proposed [6]. Our approach is based only on the voice signal analysis. Moreover, we concentrate also on a graphical user interface of the system. Our main goal of research is to deliver to diagnosticians and clinicians an integrated tool supporting a non-invasive diagnosis (based only on voice signal

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 157–167.
DOI: 10.1007/978-3-319-00029-9_14

examination) of patients with selected larynx diseases. The prototype of such a tool has been created using the LabVIEW environment. LabVIEW is a graphical environment enabling programmers to develop sophisticated measurement, test and control systems using intuitive graphical icons and wires that resemble a flowchart [1]. LabVIEW delivers built-in functions for digital signal processing (e.g., signal preprocessing, Fourier transforms, spectrum analysis, etc.) and data visualization. Moreover, LabVIEW enables us to build an application in the modular form due to the so-called SubVIs. SubVI is a subprogram that is embedded in or called from the block diagram window of another program.

Computer-aided diagnosis of laryngopathies, proposed by us in our previous papers, is based on selected parameters of a patient's voice signal (phonation). There exist various approaches to analysis of bio-medical signals (cf. [14]). In general, we can distinguish three groups of methods according to a domain of the signal analysis: analysis in a time domain, analysis in a frequency domain (spectrum analysis), analysis in a time-frequency domain (e.g., wavelet analysis). We have tested several approaches for the computer tool being developed both based on the voice spectrum analysis (e.g. [10], [11]) and based on voice signal analysis in a time domain (e.g. [15], [16]).

In [11], the approach based on extracting features (parameters) reflecting patient's voice spectrum disturbances around a basic tone and its multiples (harmonics) has been presented. That approach has been extended in [10] by adding some new features. In this paper, we show implementation of the proposed approach in the computer tool embedded in the LabVIEW environment.

The rest of the paper is organized as follows. In Section 2, we describe the character of laryngopathy data on the basis of which a diagnosis is made. Section 3 reminds algorithms proposed by us, in our earlier papers, for a decision support of the classification process of laryngopathies and shows their implementation in the LabVIEW environment. Section 4 delivers a description of the created tool, its functionality and graphical user interface. Finally, Section 5 provides conclusions.

2 Laryngopathy Data

Data for testing our tool were extracted from sound samples of the subjects (see an example in Figure 1).

Two groups were taken into consideration [18]. The first group included persons without disturbances of phonation - the control group (CG). They were confirmed by a phoniatrist opinion. The second group included patients of Otolaryngology Clinic of the Medical University of Lublin in Poland. They had clinically confirmed dysphonia as a result of Reinke's edema (RE) or laryngeal polyp (LP). Experiments were carried out by a course of breathing exercises with instruction about the way of articulation. The task of all examined patients was to utter separately different Polish vowels with extended articulation as long as possible, without intonation, and each on separate expiration. Clinical experience shows that harmonics in the voice spectrum of a healthy patient are distributed approximately steadily. However, larynx diseases may

disturb this distribution [18] (see Figure 2 to compare some exemplary spectrums). Therefore, the analysis of a degree of disturbances can support diagnosis of larynx diseases.

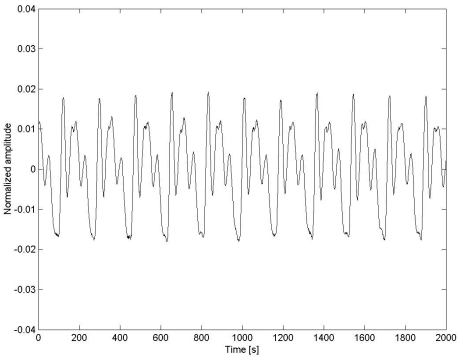
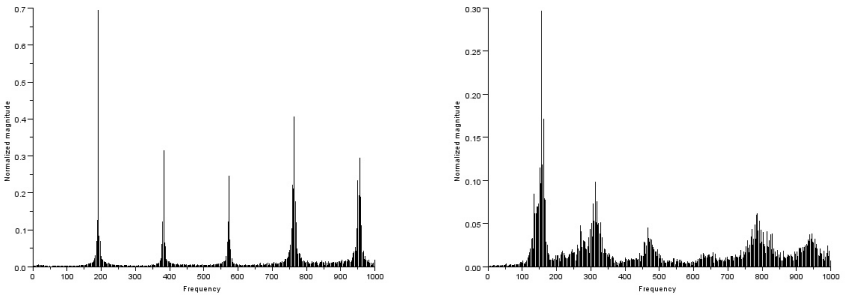


Fig. 1. An exemplary subject's voice signal



a b

Fig. 2. Exemplary spectrums obtained using DTFT (Discrete-Time Fourier Transform) for a patient from the control group (a) and for a patient with a laryngeal polyp (b)

3 Algorithms

Disturbances are expressed by a family of coefficients computed for neighborhoods of a basic tone f_0 and its four multiples (f_1, f_2, f_3, f_4). In a real situation, frequencies f_1, f_2, f_3, etc. are not distributed steadily (cf. [18]). It means that we need to find a real distribution of harmonics. In the presented tool, it is done on the basis of the spectrum of a selected *N*-point time window from the original voice signal. For each original frequency *f*, a maximum magnitude is searched in the range $[f - d_1, f + d_1]$. This maximum value is assumed as a real harmonic. The original basic tone f_0 has been obtained for each patient from histogram created in the Multi-Dimensional Voice

Program (MDVP). It is a software tool for the quantitative acoustic assessment of the voice quality, calculating various parameters on a single vocalization (see [2]). On the basis of f_0, its harmonics (for ideal case) have been calculated.

Each coefficient expresses the distribution of a spectrum around a given frequency *f*. We can distinguish two types of coefficients:

- the regularity coefficient *R* determining a degree of slenderness of this distribution (see Algorithm 1),
- the deviation coefficient *D* determining a relative difference between a real multiple derived from the spectrum and a multiple calculated on the basis of the basic tone f_0 (see Algorithm 2).

Calculations of above coefficients are graphically depicted in Figure 3.

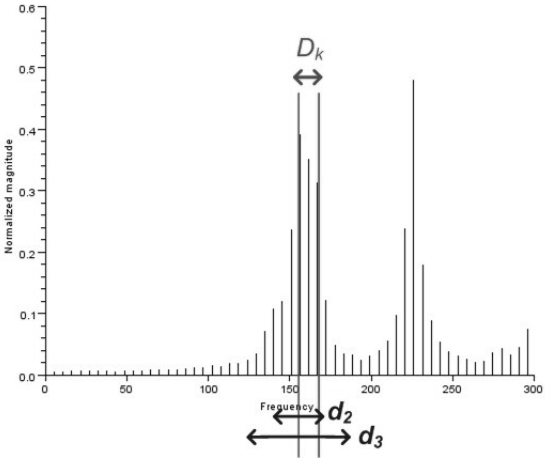


Fig. 3. Calculations of coefficients

Calculating a discrete spectrum *Sp* of a signal *W* is based on the Discrete-Time Fourier Transform (DTFT), see e.g. [14], i.e.:

$$X[k] = \sum_{n=1}^{N} W[n] e^{\frac{-2\pi jkn}{N}} \tag{1}$$

where $k = 0,1,...,N-1$, *N* is the number of points (samples) taken for DTFT. The spectrum *Sp* is a vector of magnitudes, i.e.:

$$Sp\left[\left|X[0]\right|,\left|X[1]\right|,...,\left|X\left[\frac{N}{2}\right]\right|\right] \tag{2}$$

Below, we remind algorithms proposed by us in [10] for calculating both types of coefficients characterizing the spectrum of a patient's voice signal. Both

pseudo-codes and implementations in the LabVIEW environment are given. In the implementations, proper operations on arrays are used (e.g., searching for maximum, getting subarrays, etc.).

Algorithm 1: Algorithm for calculation of regularity coefficients.

Input : Sp - a discrete spectrum of the selected time window in the voice signal of a given patient (a vector of samples), f_0 - a patient's basic tone, d_1 - deviation for searching maximum, d_2, d_3 - deviations for calculating spectrum regularity coefficients ($d_2 < d_3 < d_1$).

Output: R - a family of spectrum regularity coefficients for Sp.

$[f_1, f_2, f_3, f_4] \leftarrow Harmonics(f_0)$;

for *each frequency* f_i $(i = 0,1,2,3,4)$ **do**

  Calculate the index k_i of the strip in Sp corresponding to f_i;

  Find the index m_i of the maximum value in $Sp[k_i - d_1, \ldots, k_i + d_1]$;

  $I_1 \leftarrow Integral(Sp, m_i - d_2, m_i + d_2)$;

  $I_2 \leftarrow Integral(Sp, m_i - d_3, m_i + d_3)$;

  $R[i] \leftarrow \frac{I_1}{I_2}$;

end

Return R;

In Algorithm 1, we use the following functions (procedures):

- $Harmonics(f_0)$ - calculating harmonics (first f_1, second f_2, third f_3, and fourth f_4) of a patient's basic tone f_0, i.e., f_1=2f_0, f_2=3f_0, f_3=4f_0, f_4=5f_0.
- $Integral(Sp, k_1, k_2)$ - calculating a discrete integral I of a fragment (between points k_1 and k_2) of the spectrum Sp, i.e.:

$$I = \sum_{j=k_1}^{k_2} |X[j]| \tag{3}$$

A procedure used in Algorithm 1 for calculating the regularity coefficient $R[i]$ for a given frequency f_i, where i=0,1,2,3,4, has been implemented in LabVIEW as the SubVI (see Figure 4) named *RC*. This SubVI has the following input terminals:

- *FFT* - a discrete spectrum,
- *f*0 - a patient's basic tone,
- *df* - a frequency step in the discrete spectrum *FFT*,
- *d*1 - deviation for searching maximum,
- *d*2, *d*3 - deviations for calculating spectrum regularity coefficients,
- *m* - a frequency multiplier,

and output terminals:

- *R* - a regularity coefficient calculated for mf_0,
- *maxfindex* - an index of a frequency with the maximal magnitude in the discrete spectrum.

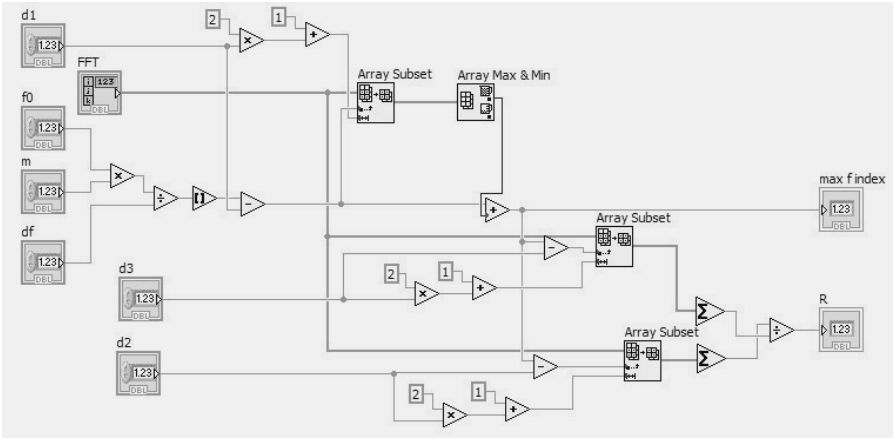


Fig. 4. The RC SubVI

Algorithm 2: Algorithm for calculation of deviation coefficients.

Input : Sp - a discrete spectrum of the selected time window in the voice signal of a given patient (a vector of samples), f_0 - a patient's basic tone, d_1 - deviation for searching maximum.

Output: D - a family of spectrum deviation coefficients for Sp.

$[f_1, f_2, f_3, f_4] \leftarrow Harmonics(f_0)$;

for *each frequency* f_i $(i = 1,2,3,4)$ **do**

  Find the maximum value f_{im} in $Sp[k_i - d_1, \ldots, k_i + d_1]$;

  $D[i] \leftarrow \frac{|f_{im} - f_i|}{f_i}$;

end

Return D;

A procedure used in Algorithm 2 for calculating the deviation coefficient $D[i]$ for a given frequency f_i, where i=1,2,3,4, has been implemented in LabVIEW as the SubVI (see Figure 5) named *DC*. This SubVI has the following input terminals:

- *FFT* - a discrete spectrum,
- f_0 - a patient's basic tone,
- d_f - a frequency step in the discrete spectrum *FFT*,
- d_1 - deviation for searching maximum,
- m - a frequency multiplier,

and output terminals:

- D - a regularity coefficient calculated for mf_0,
- *maxfindex* - an index of a frequency with the maximal magnitude in the discrete spectrum.

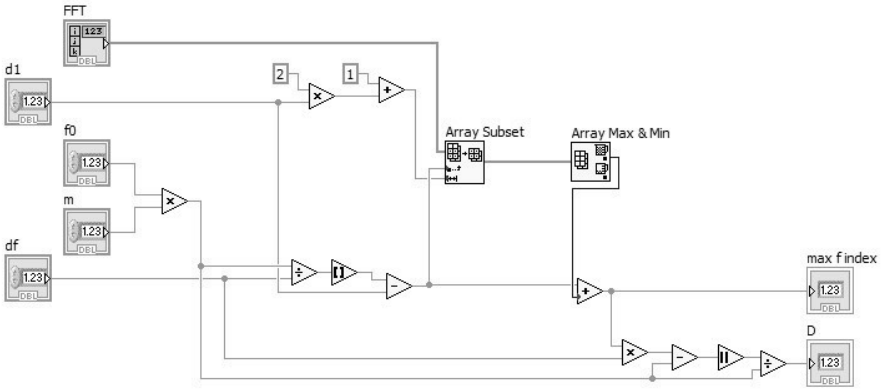


Fig. 5. The DC SubVI

For experiments, we have tested various combinations of input parameters. The best classification results have been obtained for the following parameters: N=8192 - the number of points (samples) taken for DTFT, d_1=12 - deviation for searching maximum, d_2=4, d_3=8 - deviations for calculating spectrum regularity coefficients.

4 A Computer Tool Based on LabVIEW

Our prototype system supporting a non-invasive diagnosis of selected larynx diseases is a tool designed for the LabVIEW platform. The main features of the application are the following:

- User-friendly interface (see Figures 6, 7, and 8). The interface is designed in order to make it possible to use in the medical environment.

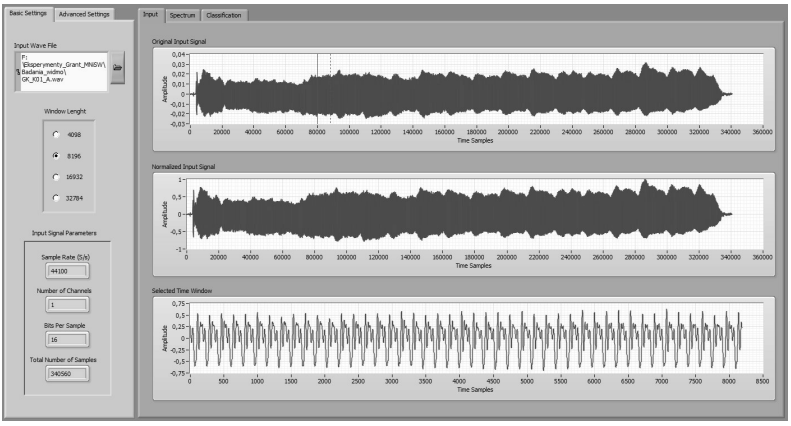


Fig. 6. GUI of the tool - input time signal

- Modularity. The project of the application and its implementation takes into consideration modularity in order to make it possible to extend in the future and enlarge its usage on diagnosis based on other approaches. It is obtained using the architecture based on SubVIs (see Section 3).

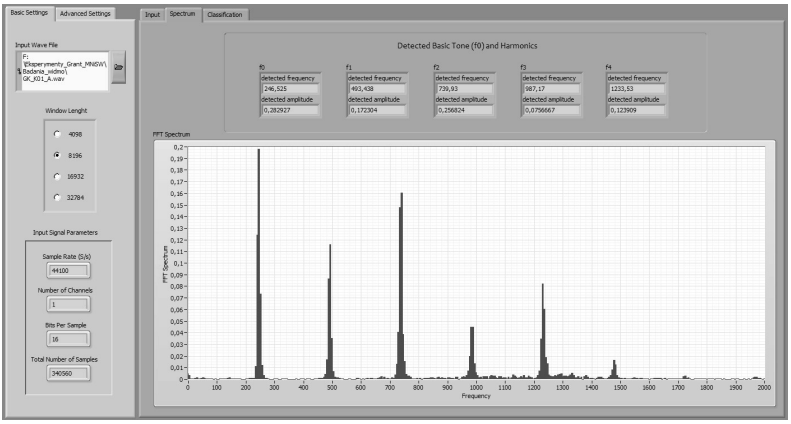


Fig. 7. Figure 7: GUI of the tool – spectrum

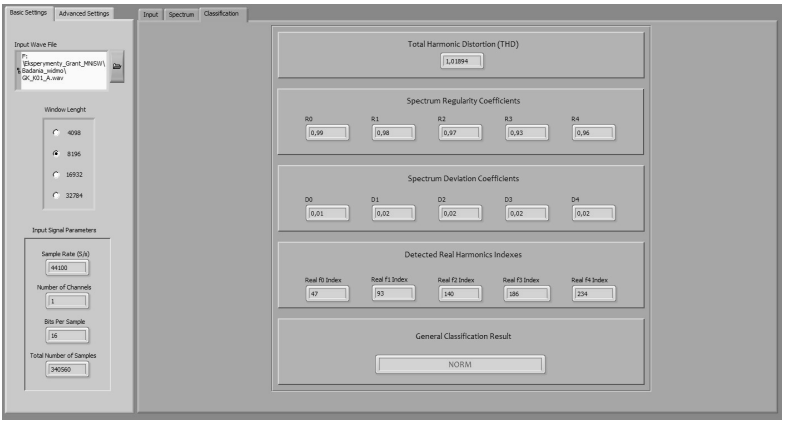


Fig. 8. GUI of the tool - classification results

The input is a patient's voice signal collected in the way described in Section 2 and recorded in the *wave* format. However, in the future, we plan to extend a tool capability to collecting and examining the subjects' voice signals in real time ("on-line"). The user can set some fundamental parameters like:

- a patient's basic tone,
- a starting point at which the voice signal analysis begins (in order to exclude some initial part of the signal which is more deformed),

- a length of the time window for which the Discrete-Time Fourier Transform is calculated,

as well as can modify advanced parameters used in Algorithms 1 and 2, like deviation for searching maximum, deviations for calculating spectrum regularity coefficients.

A general structure of the tool is shown in Figure 9. We can distinguish four main parts of this tool:

- A *signal processing* part. In this part, we make some preprocessing operations and calculate DTFT.
- A *spectrum analysis* part. In this part, we calculate a family of coefficients reflecting spectrum disturbances.
- A *classification* part. In this part, a classification process is made on the basis of obtained coefficients using a proper classification algorithm.
- A *visualization* part. This part supports the idea that visualization plays an important role in professional data mining. Some pictures often represent data better than expressions or numbers. Visualization is very important in dedicated and specialized software tools used in different communities.

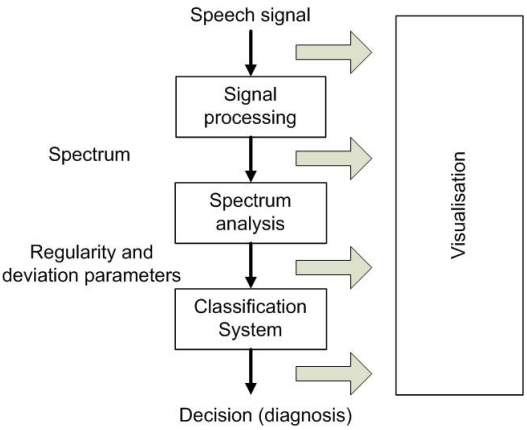


Fig. 9. A general structure of the tool

The user may visualize, among others:

- An original input signal, a signal normalized to the interval [−1.0,1.0] as well as a selected time window of the signal, see Figure 6.
- A spectrum of the signal calculated using the Discrete-Time Fourier Transform (DTFT), see Figure 7.
- Calculated regularity and deviation coefficients as well as a decision (diagnosis), *norm* or *pathology*, proposed by the classification algorithm, see Figure 8.

The approach presented in [10] has been tested using classification algorithms available in the popular data mining and machine learning software tools: WEKA [19], Rough Set Exploration System (RSES) [4], NGTS [9]. Four of them are rule-based

algorithms: exhaustive (RSES) [3], LEM2 (RSES) [8], Genetic (RSES) [20], and NGTS [9]. Two of them are decision-tree based algorithms: J48 (WEKA) - an implementation of C4.5 [13], CART (WEKA) [5].

Exemplary implementation, in the LabVIEW environment, of the classification process based on a decision tree generated using the J4.8 algorithm is shown in Figure 10. For implementation, we have used a formula node, where a decision tree is coded in the C-like language. The J4.8 algorithm delivers a relatively good capability for achieving satisfactory classification accuracy of examined data [10] together with easy implementation. Due to the SubVI architecture, other classification methods can be implemented and added in the proposed tool.

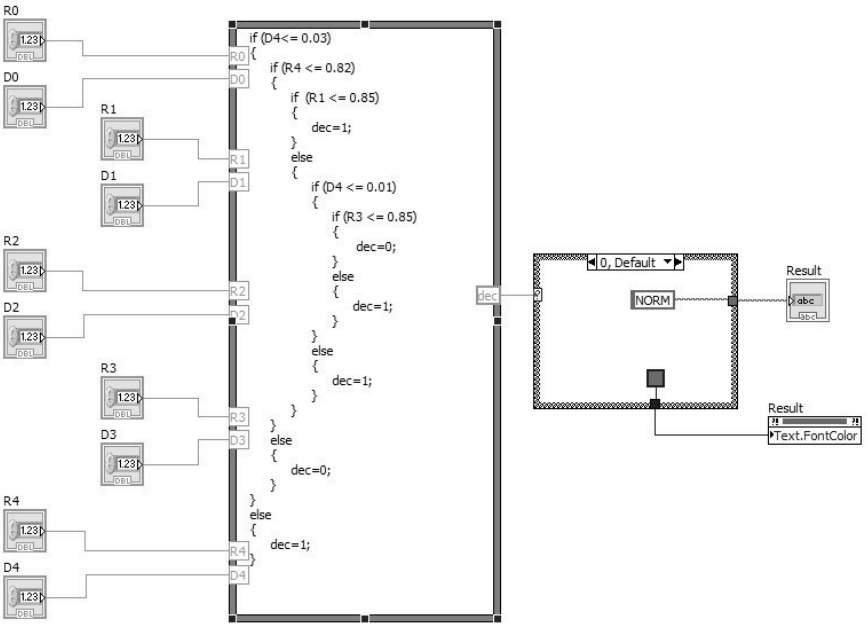


Fig. 10. The classification SubVI

5 Conclusions

In this paper, we have described a computer tool supporting a non-invasive diagnosis of selected larynx diseases, created for the LabVIEW environment. This tool is evolving continuously. So far, algorithms based on the voice spectrum analysis have been implemented. In the future, we plan to add algorithms based on voice signal analysis in a time domain, especially those using recurrent neural networks proposed by us [16]. Hybridization of classification methods of patients with laryngopathies is a main direction of further research. The modular architecture, based on SubVIs for the LabVIEW environment, enables us to extend the tool in an easy way.

Acknowledgement. This research has been supported by the grant No. N N516 423938 from the National Science Centre in Poland.

References

1. LabVIEW, http://www.ni.com/labview/
2. Multi-Dimensional Voice Program, MDVP (2011), http://www.kayelemetrics.com
3. Bazan, J.G., Nguyen, H.S., Nguyen, S.H., Synak, P., Wroblewski, J.: Rough set algorithms in classification problem. In: Polkowski, L., Tsumoto, S., Lin, T.Y. (eds.) Rough Set Methods and Applications, pp. 49–88. Physica-Verlag, Heidelberg (2000)
4. Bazan, J., Szczuka, M.S.: The Rough Set Exploration System. In: Peters, J.F., Skowron, A. (eds.) Transactions on Rough Sets III. LNCS, vol. 3400, pp. 37–56. Springer, Heidelberg (2005)
5. Breiman, L., Friedman, J., Olshen, R., Stone, C.: Classification and Regression Trees. Chapman & Hall, Boca Raton (1993)
6. Gelzinis, A., Verikas, A., Bacauskiene, M.: Automated speech analysis applied to laryngeal disease categorization. Computer Methods and Programs in Biomedicine 91(1), 36–47 (2008)
7. Greenes, R.: Clinical Decision Support: The Road Ahead. Elsevier (2007)
8. Grzymala-Busse, J.: A new version of the rule induction system LERS. Fundamenta Informaticae 31, 27–39 (1997)
9. Hippe, Z.: Machine learning – a promising strategy for business information processing? In: Abramowicz, W. (ed.) Business Information Systems, pp. 603–622. Academy of Economics Editorial Office, Poznan (1997)
10. Pancerz, K., Paja, W., Szkola, J., Warchol, J., Olchowik, G.: A rule-based classification of laryngopathies based on spectrum disturbance analysis - an exemplary study. In: Van Huffel, S., et al. (eds.) Proc. of the BIOSIGNALS 2012, Vilamoura, Algarve, Portugal, pp. 458–461 (2012)
11. Pancerz, K., Szkola, J., Warchol, J., Olchowik, G.: Spectrum disturbance analysis for computer-aided diagnosis of laryngopathies: An exemplary study. In: Proc. of the International Workshop on Biomedical Informatics and Biometric Technologies (BT 2011), Slovak Republic, Zilina (2011)
12. Pawlak, Z.: Rough Sets. Theoretical Aspects of Reasoning about. Data. Kluwer Academic Publishers, Dordrecht (1991)
13. Quinlan, J.: C4.5. Programs for machine learning. Morgan Kaufmann Publishers (1993)
14. Semmlow, J.: Biosignal and Medical Image Processing. CRC Press (2009)
15. Szkola, J., Pancerz, K., Warchol, J.: Computer diagnosis of laryngopathies based on temporal pattern recognition in speech signal. Bio-Algorithms and Med-Systems 6(12), 75–80 (2010)
16. Szkola, J., Pancerz, K., Warchol, J.: Recurrent neural networks in computer-based clinical decision support for laryngopathies: An experimental study. Computational Intelligence and Neuroscience Article ID 289398 (2011)
17. Verikas, A., Gelzinis, A., Bacauskiene, M., Uloza, V.: Towards a computer-aided diagnosis system for vocal cord diseases. Artificial Intelligence in Medicine 36(1), 71–84 (2006)
18. Warchol, J.: Speech examination with correct and pathological phonation using the SVAN 912AE analyser (in Polish). Ph.D. thesis, Medical University of Lublin (2006)
19. Witten, I.H., Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann (2005)
20. Wroblewski, J.: Genetic algorithms in decomposition and classification problem. In: Polkowski, L., Skowron, A. (eds.) Rough Sets in Knowledge Discovery 2, vol. 2, pp. 471–487. Physica-Verlag, Heidelberg (1998)

Analysis and Development of Techniques and Methods in Medical Practice in Cochlear Implant Systems

Svetlin Antonov and Snejana Pleshkova-Bekiarska

Technical University of Sofia, Department of Radio Communications and Video Technologies, Boul. Kl. Ohridsky 8, Sofia 1000, Bulgaria
svantonov@yahoo.com, snegpl@tu-sofia.bg

Abstract. Cochlear implant methods are the area of interest and intensive medical and technical efforts for investigations, analysis, implementation, testing and finally producing more and more precise and effective cochlear implant prosthesis. The goals of this article are first to present the cochlear implant system in historical plan, then to analyze some of the existing signal processing strategies in cochlear implant systems, which leads to the decision of the importance of filter bank design for the precision of the signal processing in the cochlear implant algorithms and practical implementations. It is presented a detailed description and critical comparison of the most useful types of cochlear filter banks and as the results from the analysis some useful conclusion are presented for the cochlear filter banks characteristics, time consuming in calculation, the importance for overall cochlear implants quality of speech sound processing, etc. The defined in conclusion assertions are accepted as the basis of future investigations in area of new cochlear implant model development.

Keywords: Cochlear implant, Signal processing strategies in cochlear implant, Cochlear filter banks, auditory prosthesis, electric stimulation.

1 Introduction

In historical plan the cochlear implant devices are developed as it is shown in Table I. Three phases defining the major events in the development of cochlear implants [1]. The conceptualization phases demonstrated the feasibility of electric stimulation. The research and development phase legitimized the utility and safety of electric stimulation. The commercialization phase saw a wide spread use of electric stimulation in treating sensor neural hearing loss (Table 1).

Table 1. The development of cochlear implants in historical plan

YEAR	EVENT
1800	Volta used his battery to show that electric stimulation can evoke sensations including hearing

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 169–178.
DOI: 10.1007/978-3-319-00029-9_15

Table 1. (*continued*)

1937-1940	Stevens and colleagues identified three mechanisms underlying electric sensations
1973	First International Conference on Electrical Stimulation-San Francisco
1983	First of the Biennial Conferences under the auspices of the Gordon Research Conference
1985	3M/House became the first FDA approved cochlear implant
1988	Nucleus 22 became the first FDA approved multi-channel device

2 Review of Speech Processing Strategies in Cochlear Implants

In general the speech processing strategies are presented in Fig 1. An external unit, also known as the speech processor, consists of a digital signal processing (DSP) unit, a power amplifier, and an RF transmitter.

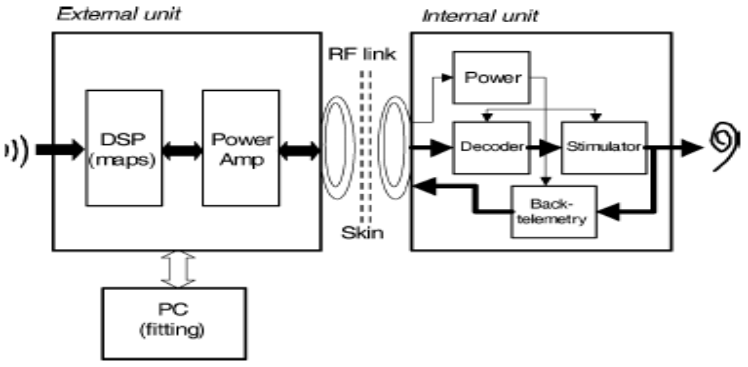


Fig. 1. Architecture and functional block diagram of a modern cochlear implant

Table 2 classifies signal processing used in modern cochlear implants. Most cochlear implants discard the fine structure and encode the coarse features only.

Table 2. Classification for speech processing strategies in cochlear implants

Speech Processing Strategies					
Coarse features				Fine features	
Spectral envelope + Explicit feature extraction	Temporal envelope Implicit feature extraction			Spectral fine structure	Temporal fine structure
F0 & F1 WSP-II	SPEAK n of m, low rate Nucleus	CIS Non-simultaneous pulsatile stimulation Clarion, Med-El, Nucleus	Compressed Analog Simultaneous stimulation	Virtual channels Harmony HiRes 90k Clarion	FAME OPUS FSP Med-El
F0, F1 & F2 WSP II	ACE high-rate Nucleus	PPS & MPS Partially simultaneous Clarion	SAS Bipolar simultaneous Clarion		
Multi-peaks MPEAK					

In the Nucleus MPEAK strategy [5], showed in Fig 2, was added the first formant (*F1*) in the Nucleus WSP III processor and up to six spectral peaks.

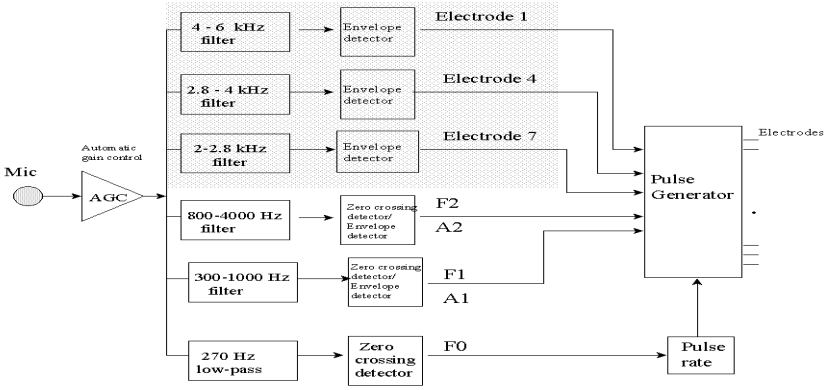


Fig. 2. Block diagram of the signal processing involved in the MPEAK strategy

In Continuous-interleaved-sampling (CIS) strategy the sound is first subject to a number of bandpass filters with the number being as few as 5 in the original CIS implementation and as many as 20 in the Nucleus Freedom device [4]. Block diagram of the signal processing involved in the CIS strategy is shown in Fig.3.

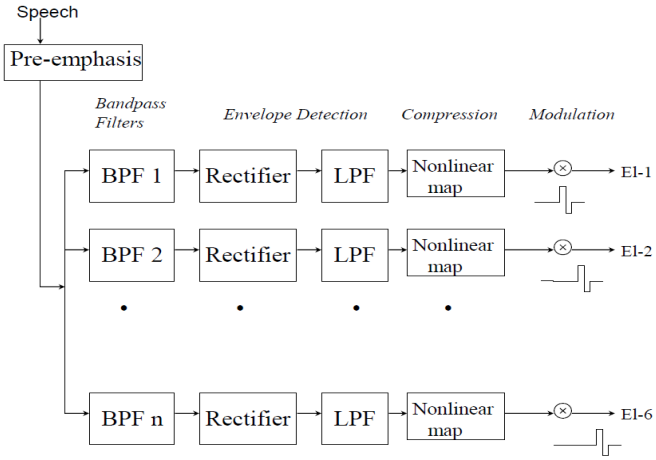


Fig. 3. Block diagram of the signal processing involved in the CIS strategy

At present, signal processing focuses on how to encode spectral and temporal fine structure cues in cochlear implants. To encode the spectral fine structure, more independent electrodes are needed. One way to encode the fine structure is to increase the electric stimulation carrier rate so that the temporal fine structure cue can be represented in the waveform domain, e.g., Med El's FSP processor [2]. A second way is to extract frequency modulation from the temporal fine structure and then use it to frequency modulate the carrier rate [3]. A third way is to use multiple carriers to encode the fine frequency structure. The effectiveness of these new strategies has not been demonstrated in actual cochlear implant users.

3 The Cochlear Filter Banks

3.1 Physical Hydrodynamics Fluid 3D - Model of Cochlear Filter Banks

Many filter banks have been proposed based on different cochlea models [6, 7]. Here is presented the physical model of cochlear filter bank with proposed substantial inter-filter feedback. The created physical model of cochlear filter banks is derived from a geometric model of cochlear. The shape of this model is simplified from the real spiral human cochlear to the stacked parallelepipeds separated by the basilar membrane (Fig. 4).

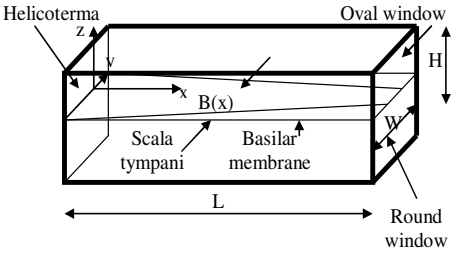


Fig. 4. The proposed shape simplified cochlear geometric model from the real spiral human cochlear to the stacked parallelepipeds separated by the basilar membrane

First it is derived the equation of motion of the basilar membrane as point impedances:

$$-2p(x, y=0, z=0)B(x) = m(x)\frac{d^2 z}{dt^2} + r(x)\frac{dz}{dt} + k(x), \tag{1}$$

where p is the pressure in the scalar vestibule; B - the effective breadth of the basilar membrane; m - the mass; r - the damping; k - the stiffness of the membrane sections.

$P(x,y,z)$ is the solutions of the Laplace equation with the simplifications of the hydrodynamics of the cochlear chamber, presented as incompressible fluid in inviscid flow within rigid walls:

$$p(x, y, z) = \int_{\partial V} g\frac{\partial p}{\partial n} dS\,, \tag{2}$$

where $p(x, y, z)$ is the integral is over the boundary ∂V of the enclosed volume.

Using the Green's functions:

$$g(x, y, z|\xi, \eta, \zeta) = f(\xi - x, y, z|\eta, \zeta) - f(\xi + x, y, z|\eta, \zeta)$$

$$f(\chi, y, z|\eta, \zeta) = \frac{1}{4\pi}\sum_{k=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\frac{(-1)^k}{\left((2kL+\chi)^2 + (mW+(-1)^m\eta - y)^2 + (nH+(-1)^n\zeta - y)^2\right)^{\frac{1}{2}}} \tag{3}$$

is possible to define the effective breadth $B(x)$ as

$$B(x)g(x,0,0|\xi,0,0)\frac{d^2z}{dt^2} = \int_{-\frac{W}{2}}^{\frac{W}{2}} \left(g(x,y,z|\xi,\eta,\zeta).\frac{\partial^2 u_z}{\partial t^2} \right) d\eta \,. \tag{4}$$

In Eq. (1) is used the first approximation. After these operations and space sampling along x at a rate of $\frac{1}{D} = \frac{M}{L}$, for a total of *M* filter sections, the equation of motion, (3) becomes:

$$\left[m^i + 2\rho \frac{D^2 \left(B^i\right)^2}{A} i \right] \frac{d^2 z^i}{dt^2} + r^i \frac{dz^i}{dt} + k^i z^i = -2\rho D B^i i \frac{du_x}{dt} - 2\rho D^2 B^i \sum_{j=0}^{M} \Theta_j^i \frac{B^j}{A} \frac{d^2 z^j}{dt^2}, \tag{5}$$

where $A=W.H$ is the cross-sectional area.

Then sampling in time is applied. This permits to derive the following final equations for the physical filter banks model of cochlear:

$$\begin{aligned} &\tilde{z}^i(n+1) - \tilde{z}^i(n) + a_1^i z_2^i(n) = b^i \left[u_x(n+1) - u_x(n) \right] \\ &z_2^i(n+1) = z_2^i(n) + \left(\frac{T}{2} \right) \left(z_1^i(n+1) + z_1^i(n+1) \right) \\ &\tilde{z}^i(n+1) = \sum_{j=0}^{M} \left(\delta_j^i + c^i d^j \Theta_j^i \right) z_1^i(n+1) \end{aligned} \tag{6}$$

It is seen from the above final equations for the physical filter banks model of cochlear, that filter bank consists of resonators with all-to-all inter-filter feedback with the sampling period T and the following coefficients:

$$\mu^i = m^i + 2\rho \frac{D^2 \left(B^i\right)^2}{A} i + \left(\frac{T}{2} \right) r^i + \left(\frac{T}{2} \right)^2 k^i \qquad a_1^i = \left(\frac{2}{\mu^i} \right) \left[\left(\frac{T}{2} \right) r^i + \left(\frac{T}{2} \right)^2 k^i \right] \tag{7}$$

$$a_2^i = \left(\frac{2}{\mu^i} \right) \left(\frac{T}{2} k^i \right), \; b = -\left(\frac{2\rho}{\mu^i} \right) D B^i i \; (11), \; c^i = \left(\frac{2\rho}{\mu^i} \right) D B^i, \; d^i = \frac{D B^i}{A} \tag{8}$$

The equations derived for the physical filter banks model of cochlear are proved in the time domain using the following physical properties of the average human cochlea:

- $A=0.01$ cm^2;
- B^i tapers from 0.0075 cm at the oval window at a rate of 0.0012(cm/cm);
- 105 filter sections span a frequency range from 187 Hz to 6 KHz;
- the mass, *m(x),* damping, *r(x)* and stiffness, *k(x)* vary exponentially with *x* in such a way as to keep the quality factor constant:

$$Q = \frac{\sqrt{mk}}{r} = const\ ; \tag{9}$$

- input signals are sampled at 13.3Khz;
- the filter bank output is rectified, leakily integrated, compressed and decimated to 1 KHz.

It is seen from the equation of the physical filter banks cochlear model, that the filters are narrowband in the lower frequencies, and become increasingly broadband towards the higher frequencies.

3.2 Physical Gammatone Linear Model of Cochlear Filter Banks

There are also other physical models of cochlear filter banks, using linear models. The gammatone model of cochlear filter banks is one of these linear models and is based on human psychophysical data. The gammatone filters banks are the main part of the computational model of the proposed dual-resonance nonlinear type of filter banks. It is proposed to use banks or cascades of first-order gammatone filters:

$$y[i] = a_0 \cdot x[i] + a_1 \cdot x[i-1] - b_1 \cdot y[i-1] - b_2 \cdot y[i-2], \tag{10}$$

where [*i*] refers to the ith sample of the cochlear signals, *x* and *y* are the input and output cochlear signals to or from the filter, respectively.

The coefficients of first-order gammatone filters are calculated as follows:

$$a_0 = \left| \frac{1 + b_1 \cos\theta - j b_1 \sin\theta + b_2 \cos(2\theta) - j b_2 \sin(2\theta)}{1 + \alpha \cos\theta - j\alpha \sin\theta} \right|, \tag{11}$$

$$a_1 = \alpha \cdot a_0\ ,\ b_1 = 2\alpha\ ,\ b_2 = exp(-2\phi), \tag{12}$$

where *dt* is the sampling period of the digital signal;

$$\theta = 2\pi f_c dt\ ,\ \phi = 2\pi B dt\ ,\ \alpha = -exp(\phi)\cos\theta \ \text{ and } \ j = \sqrt{(-1)}\ . \tag{13}$$

The gammatone filter has an impulse response of the form:

$$h(t) = kt^{(n-1)}\, exp(-2\pi B t)\cos(2\pi f_c t + \varphi)\ , \text{ for } \ t \geq 0 \tag{14}$$

$$h(t) = 0 \ \text{ for } \ t < 0\ , \tag{15}$$

where *n* is the order of the gammatone filter; *B* - the gammatone filter bandwidth; f_c- the center frequency of the gammatone filter; *w* – the phase of the gammatone filter; *k*- the gain of the gammatone filter.

After the described above gammatone filters banks, which are the main part of the computational model of the proposed dual-resonance nonlinear type of filter banks are included banks or cascades of second-order Butterworth lowpass filters:

$$y[i] = C \cdot x[i] + 2 \cdot C \cdot x[i-1] + C \cdot x[i-2] - D \cdot y[i-1] - E \cdot y[i-2], \tag{16}$$

with coefficients:

$$C = \frac{1}{1+\sqrt{2}\cot\theta+\cot^2\theta},\ D = 2C\left(1-\cot^2\theta\right),\ E = C\left(1-\sqrt{2}\cot\theta+\cot^2\theta\right),$$

$$\theta = \pi f_c dt, \tag{17}$$

where f_c is the 3-dB- down cut-off frequency of the low-pass filter; dt - the sampling period of the digital signal.

The model of the proposed dual-resonance nonlinear type of cochlear filter banks can be used as linear or nonlinear mode (Fig. 5). In the linear and non linear parts of dual-resonance nonlinear type of cochlear filter banks the gain g is implemented to the cochlear input *x(i),* respectively as follows to calculate the output *y(i):*

$$\begin{aligned} y(i) &= gx(i) \\ y(i) &= sign(x(i))min(a\,|\,x(i),b\,|\,x(i)^c) \end{aligned}, \tag{18}$$

where *a*, *b* and *c* are the constants or parameters of the chosen piece wise non linear gain function.

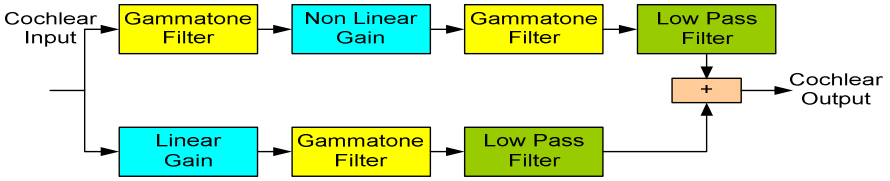


Fig. 5. The model of the dual-resonance nonlinear type of cochlear filter banks used as linear or nonlinear mode

It can be concluded, that the main advantage of physical gammatone model of cochlear filter banks is that it is possible to defined this model as linear or non linear model of dual-resonance nonlinear type filters. The dual-resonance nonlinear model is a model in which there are no connections between the individual filter units. This means, that the distortion products generated in one dual-resonance nonlinear filter unit do not propagate to other units in the dual-resonance nonlinear filter banks.

3.3 The Artificial Models of Cochlear Filter Banks

The artificial models of cochlear filter banks are based not on the physical model of the cochlear, but on the signal processing methods and especially on the time-frequency analysis and their application of a set of filter $w_i(t)$:

$$\omega_0 = \frac{1}{2\pi\|g\|^2}\int_{-\infty}^{\infty}\omega\left|\tilde{g}(\omega)\right|^2 d\omega \tag{19}$$

$$t_0 = \frac{1}{\|g\|^2} \int_{-\infty}^{\infty} t|g(t)|^2 dt \tag{20}$$

$$\sigma_\omega^2 = \frac{1}{2\pi\|g\|^2} \int_{-\infty}^{\infty} (\omega - \omega_0)^2 |\tilde{g}(\omega)|^2 d\omega \tag{21}$$

$$\sigma_t^2 = \frac{1}{\|g\|^2} \int_{-\infty}^{\infty} (t - t_0)^2 |g(t)|^2 dt \,, \tag{22}$$

where $\tilde{g}(\omega)$ is the Fourier transform of $g(t)$ and $\|g\|$ is the norm of

$g(t)$, given by $\|g\|^2 = \int g(t) g^*(t) dt$ (23)

If the following condition, defined from Gabor is satisfied: $\sigma_t \sigma_\omega \geq \frac{1}{2}$,then the fiter is defined as a Gabor filter with the property $\sigma_t \sigma_\omega = \frac{1}{2}$ and in the form:

$$g_G(t) = Ae^{-b(t-t_0)^2} e^{i\omega_0 (t-t_0)^2} \,, \; A, b \in C \tag{24}$$

One other method is to treat the basilar membrane as divided into many “critical bands”, each of which can be modeled by an Equivalent Rectangular Bandwidth (ERB) with the following equation to calculate the bandwidth at a particular frequency:

$$ERB(f) = 24.7\left(\frac{4.37f}{1000} + 1\right) \tag{25}$$

In this model of cochlear filter banks also are uses gammatone filters that are defined by the function:

$$\gamma_{tone}(t) = at^{n-1} e^{-2\pi bt} cos(2\pi f_c t + \phi), \text{ (t>0)}, \tag{26}$$

where *b* and *n* are the parameters used to control the order of the filter, the slope of the magnitude response and are chosen such that the bandwidth is equal to 1.019 times the Equivalent Rectangular Bandwidth (ERB).

3.4 Method for Comparison of Cochlear Filter Banks Models

To compare each of the filter banks model is possible to use the Schroeder’s masking curves, which can be introduce with an equation that accurately describes the response of the ear to various frequencies. To calculate this response, the input frequency is first converted to the bark scale according to the equation:

$$Bark = 13 \cdot atan\left(0.76 \cdot \frac{f\_Hz}{1000}\right) + 3.5 \cdot atan\left(\left(\frac{f\_Hz}{7500}\right)^2\right) \quad (27)$$

The following equation is used then to calculate the response curves at various frequencies:

$$S = 15.81 + 7.5*(dz + 0.474) - 17.5*\sqrt{1 + (dz + 0.474)^2} \quad (28)$$

where *dz* is the difference between the center bark frequency and the bark frequency.

It is known, that the ear does not respond linearly to changes in level and when the level increases, the slope of the curves begins to flatten out. To satisfy this human ear characteristic a modified version of Schroeder's equation exist:

$$S = (15.81 - I) + 7.5*(dz + 0.474) - (17.5 - I)*\sqrt{1 + (dz + 0.474)^2}\ , \quad (29)$$

where *I* is defined as: $I = 5*10^{(Level\ -\ 96)/10}$

The above defined equations for the Schroeder's masking curves are used to compare each of the filter banks model calculating the error for each filter in the banks. Before calculating error it is necessary to apply a threshold to each filter response according to the equation:

$$H = max\{H, Thresh\}, Thresh = 3.64*(f\_kHz)^{0.8} - 6.5e^{-0.6*(f\_kHz - 3.3)^2} + 10^{-3}*(f\_kHz)^4 \quad (30)$$

After that it is much easier to compare the filters, calculating the error by the following equation:

$$Error_i = log_{10}\left[\left(|H_{Mi}| - |H_{Si}|\right)^2\right], \quad (31)$$

where H_M is the magnitude response of the filter; H_S - the magnitude response of the ear according to Schroeder; *i* - the filter number.

The number of subbands can be chosen for example to be 25 in order to coincide with the bark scale. The error is then presented in terms of both subband number and loudness for both filter banks. The comparison of the different cochlear filter banks models is carried out from the condition, that if the error for a particular model is lower, this model outperforms the other model.

4 Conclusion

In this work are presented briefly the history, the presence and future in cochlear implant area. Analysis made of the existing signal processing strategies leads to the conclusions of filter bank importance for the precision of the signal processing in the cochlear implant algorithms and practical implementations. It is presented a detailed description and critical comparison of the most useful types of cochlear filter banks

and as the results from the analysis is presented the following conclusions: the physical models of cochlear as a filter bank describe direct form of the sound waves propagation in inner human ear; these models are closed to the real human cochlear and the work of human hearing system as a very effective sound waves processing system for preparing the suitable nerve impulses sequences transmitting to the human brains; the main disadvantages of the physical models of cochlear as a filter bank are the necessity of solving the differential equations, the difficulty of concrete parameters values definition for basilar membrane characteristics and the need to transform of the results from physical models of cochlear filter bank in an appropriate form of known filter banks for signal processing; the other wide spread models of cochlear as a filter bank like Gamma tone filters, Gabor filters, filter bank based on ERBs, etc. are not suffer from disadvantages of the physical models of cochlear as a filter bank, but are in principle theoretical, artificial and mathematical representation or interpretation of the cochlear processing ability, which gives the uncertainty of their adequateness and the need to objective quantitative and subjective qualitative estimations.

The important comparative inferences for the existing cochlear bank filter models listed above, conduct to the decision for future investigations in new cochlear implant model development and testing area, combining the realism of the physical cochlear filter bank models with the sound signal processing capabilities of the artificial cochlear filter bank models for implementation in cochlear implants prostheses.

Acknowledgments. Research in the subject was recently commenced through a NIS TU-Sofia funded project, No: 121ПД 0063-07.

References

1. Ching, T.Y.C.: Guest Editor: Emerging Trends in Cochlear Implants, 11(58) (November 2005)
2. Hochmair, I., Nopp, P., Jolly, C., Schmidt, M., Schosser, H., Garnham, C., Anderson, I.: MED-EL Cochlear Implants: State of the Art and a Glimpse into the Future. Trends Amplif 10, 201–219 (2006)
3. Nie, K., Stickney, G., Zeng, F.: Encoding Frequency Modulation to Improve Cochlear Implant Performance in Noise. IEEE Trans. Biomed. Eng. 52, 64–73 (2005)
4. Patrick, J.F., Busby, P.A., Gibson, P.J.: The Development of the Nucleus Freedom Cochlear Implant System. Trends Amplif. 10, 175–200 (2006)
5. Skinner, M.W., Holden, L.K., Holden, T.A.: Performance of Postlinguistically Deaf Adults with the Wearable Speech Processor (WSP III) and Mini Speech Processor (MSP) of the Nucleus Multi-Electrode Cochlear Implant. Ear Hear 12, 3–22 (1991)
6. Kates, J.M.: Accurate Tuning Curves in a Cochlear Model. IEEE Trans. Speech and Audio Processing 1(4), 453–462 (1993)
7. Giguere, C., Woodland, P.C.: A Computational Model of the Auditory Periphery for Speech and Hearing Research. J. Acoust. Soc. Am. 95(1), 331–342 (1994)

On Confidentially Connected-Free Graphs

Mihai Talmaciu[1], Elena Nechita[1], and Barna Iantovics[2]

[1] "Vasile Alecsandri" University, Bacau, Romania
[2] "Petru Maior" University, Targu Mures, Romania
{mtalmaciu,enechita}@ub.ro, ibarna@upm.ro

Abstract. Graph theory provides algorithms and tools to handle models for important applications in biology and medicine, such as drug design, diagnosis, or visualization. This paper deals with some theoretical results concerning the relationship between two classes of graphs which may be susceptible of applications in medicine and intelligent systems. The class of *Confidentially Connected*-free graphs is introduced and related to the class of *Asteroidal Triple*-free graphs, as well as to the graphs that have a star-cutset. We give a characterization of *Confidentially Connected* -free graphs using neighborhoods and weakly decomposition.

Keywords: confidentially connected, CC-free graphs, AT-free graphs, asteroidal triple, star-cutset, weakly decomposition, recognition algorithm.

1 Introduction

During the last decades, numerous applications of graph models and techniques have been recorded in biology, environment, public health, and medicine. The study of complex networks (starting from bio-molecular and ending with social networks) makes use of important results in graph theory. This paper deals with some theoretical results concerning the relationship between two classes of graphs which may be susceptible of applications in medicine and intelligent systems. We started from the resemblance between the concepts "asteroidal triple" [15] and "confidential connectivity" [17].

In [6], the authors study the linearity property of the *Asteroidal Triple-free* (AT-free) graphs, property that is also satisfied by the following classes: interval graphs, permutation graphs, and comparable graphs. AT-free graphs generalize the interval graphs and permutation graphs, which are comparable with respect to linearity.

Interval graphs (extensively presented in [12]) appear in models related to real life situations that imply time dependencies or other linear restrictions. Such graphs arise in archeology, molecular biology, genetics, sociology. In [5], some mathematical models of population biology use interval graphs. Recent applications have been found in protein sequencing [14], DNA mapping [19], and macro substitution [10]. Other applications, on various complex optimization problems, are presented in [3, 13, 16].

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 179–187.
DOI: 10.1007/978-3-319-00029-9_16

While Dirac [9] gives a theoretical characterization of triangulated graphs, Fulkerson and Gross [11] give a characterization with algorithms of triangulated graphs. For the recognition of weakly triangulated graphs, Berry, Bordat and Heggernes [2] extended the characterization presented in [15] by Lekkerkerker and Boland. Using this result and weakly decomposition [17] to recognize the triangulated graphs, we studied the relationships between triangulated, interval and Confidentially Connected-free (CC-free) graphs. Other important results in this paper refer to the link between CC-free graphs and AT-free graphs, CC-free graphs and the "star-cutset" concept.

The content of the paper is organized as follows. In Preliminaries, we give the usual terminology in graph theory. In Section 3 we establish the relationship between CC-free graphs and AT-free graphs. Section 4 gives a characterization of the CC-free graphs using neighborhoods. In Section 5 we show that the class of CC-free graphs is equivalent to the class of graphs with star-cutset. In Section 6 we remind the weakly decomposition, give a new characterization of CC-free graphs using this decomposition and a recognition algorithm for CC-free graphs. Ideas for future work conclude the paper.

2 Preliminaries

Throughout this paper, $G=(V,E)$ is a connected, finite and undirected graph [1], without loops and multiple edges, having $V=V(G)$ as the vertex set and $E=E(G)$ as the set of edges. $\overline{G}$ is the complement of G. If $U\subseteq V$, by $G(U)$ we denote the subgraph of G induced by U. By $G-X$ we mean the subgraph $G(V-X)$, whenever $X\subseteq V$, but we simply write $G-v$, when $X=\{v\}$. If $e=xy$ is an edge of a graph G, then x and y are adjacent, while x and e are incident, as are y and e. If $xy\in E$, we also use $x\sim y$, and $x\nsim y$ whenever x,y are not adjacent in G. If $A,B\subset V$ are disjoint and $ab\in E$ for every $a\in A$ and $b\in B$, we say that A,B are *totally adjacent* and we denote by $A\sim B$, while by $A\nsim B$ we mean that no edge of G joins some vertex of A to a vertex from B and, in this case, we say A and B are *non-adjacent*.

The *neighborhood* of the vertex $v\in V$ is the set $N_G(v)=\{u\in V: uv\in E\}$, while $N_G[v]=N_G(v)\cup\{v\}$; we denote $N(v)$ and $N[v]$, when G appears clearly from the context. The *degree* of v in G is $d_G(v)=|N_G(v)|$. The neighborhood of the vertex v in the complement of G will be denoted by $\overline{N}(v)$.

The neighborhood of $S\subset V$ is the set $N(S)=\cup_{v\in S}N(v)-S$ and $N[S]=S\cup N(S)$. A graph is complete if every pair of distinct vertices is adjacent.

By P_n, C_n, K_n we mean a chordless path on $n\geq 3$ vertices, a chordless cycle on $n\geq 3$ vertices, and a complete graph on $n\geq 1$ vertices, respectively.

Let F denote a family of graphs. A graph G is called *F-free* if none of its subgraphs are in F.

The *Zykov sum* of the graphs G_1, G_2 is the graph $G = G_1 + G_2$ having:

$$V(G) = V(G_1) \cup V(G_2),$$

$$E(G) = E(G_1) \cup E(G_2) \cup \{uv : u \in V(G_1), v \in V(G_2)\}.$$

3 The Relationship between *CC*-Free Graphs and *AT*-Free Graphs

In a network, the communication is said to be "confidential" if a message can be passed between any two vertices, without being intercepted by a third vertex. In the language of graph theory, this property stands for confidential connectivity.

At first, we recall the notions of weakly component and weakly decomposition.

Definition 1. ([17],[7],[8]) *A graph $G = (V, E)$ with at least three vertices is called confidentially connected if $\forall (a,b,c) \in V^3$ three distinct vertices, there exists P an a,b-path in G such that $N[c] \cap V(P) \subseteq \{a,b\}$.*

Definition 2. *A graph $G = (V, E)$, where every induced subgraph H of G does not have the confidential connectivity property is called CC-free.*

Definition 3. ([6]) *Three vertices in a graph determine an asteroidal triple if every two of them are joined through a path, avoiding the neighborhood of the third.*

Definition 4. ([6]) *A graph is called AT-free if does not contain any asteroidal triple.*

Remark 1. *In an asteroidal triple, every two vertices are joined by a path whose intersection with the neighborhood of any other third vertex is empty, that is $\forall (a,b,c) \in V^3 : N[c] \cap V(P_{ab}) = \phi$.*

The confidential connectivity property can be translated as follows: every two vertices are joined by a path whose intersection with the neighborhood of any other third vertex is either empty or an extremity or both extremities of the path, that is $\forall (a,b,c) \in V^3 : N[c] \cap V(P_{ab}) \subseteq \{a,b\}$.

Remark 2. *Every graph with the property that any three of its vertices form an asteroidal triple is confidentially connected.*

Remark 3. *A confidentially connected graph with the property that the extremities of every path do not intersect the neighborhood of any other third vertex, called in what follows strongly CC is a graph where every three vertices form an asteroidal triple.*

In what follows we establish the relationship between interval, triangulated and CC-free graphs.

Definition 5. *A graph is called interval graph if its vertices can by put in a one-to-one correspondence with a set of intervals on the real line such that two vertices are adjacent if and only if they correspond to non-disjoint intervals.*

Definition 6. *A graph is chordal if each of its cycles of four or more nodes has a chord, which is an edge joining two nodes that are not adjacent in the cycle.*

Theorem 1. *A graph is an interval graph if and only if it is chordal and strongly CC-free.*

Proof. In [15] Lekkerkerker and Boland demonstrated that a graph is an interval graph if and only if it is chordal and asteroidal triple-free. The conclusion of theorem 1 follows considering remarks 2 and 3.

4 Characterization of *CC*-Free Graphs Using Neighborhoods

The following result is true for CC -free graphs.

Some similar results are stated in [7], but for confidentially connected graphs.

Theorem 2. *Let $G=(V,E)$ be a connected and non-complete graph, with at least three vertices. Then G is CC -free if and only if*

(i) $\exists v\in V$ such that $[\overline{N}(v)]_G$ is non-connected

or

(ii) $\exists v\in V$ such that $N(\overline{N}(v))\neq N(v)$.

Proof. Let G be CC-free. Then $\exists a,b,c$ such that $\forall P_{ab}: V(P_{ab})\cap N[c]-\{a,b\}\neq\phi$. If $a,b\in\overline{N}(c)$ then $[\overline{N}(c)]_G$ is non-connected.

Let $|\{a,b\}\cap\overline{N}(c)|\in\{0,1\}$. We denote $a'=a$ for $a\in\overline{N}(c)$ or $a'=x\in N(a)\cap\overline{N}(c)$, for $a\in N(c)$. Also, we denote $b'=b$ for $b\in\overline{N}(c)$ or $b'=y\in N(b)\cap\overline{N}(c)$, for $b\in N(c)$. In $[\overline{N}(c)]_G$ there is not path between a' and b', therefore $[\overline{N}(c)]_G$ is not connected, so (i) holds.

Let G be *CC*-free again. If for every $v\in V$ we have that for every x neighbor of v, there exists a non-neighbor y ($y\in\overline{N}(v)$), then there exists a path P_{xy} such that $N[v]\cap V(P_{xy})\subseteq\{x,y\}$. This means that G is not CC -free, so (ii) holds.

Let $[\overline{N}(v)]_G$ not connected, for a given v. In Definition 1, considering a and b in different connected components of $[\overline{N}(v)]_G$ and $c=v$ we obtain that every ab-path (which exists because G is connected) either contains c or has as interior vertex a neighbor of c, therefore G is CC -free. Suppose that $\exists w\in N(v)$ such that $N(w)\cap\overline{N}(v)=\phi$. Taking $a=w$, any b in $\overline{N}(v)$ and $c=v$, we have that $\forall P_{ab}$: $V(P_{ab})\cap N[c]-\{a,b\}\neq\phi$, so G is CC -free.

5 *CC*-Free Graphs and Star-Cutset

Let $G=(V,E)$ be a connected graph. A non-empty set of vertices T is called *star-cutset* (Chvatal [4]) if $G-T$ is not connected and there exists a vertex v in T that is adjacent to any other vertex in T.

Some similar results are stated in [8], but for confidentially connected graphs.

The following result is true for *CC*-free graphs.

Theorem 3. *Graph G is CC-free if and only if G has a star-cutset.*

Proof. Let G *CC*-free. We prove that G has a star-cutset. If G is *CC*-free, then $\exists v\in V(G)$, $\exists x,y\in V(G)-\{v\}$ such that $\forall P_{xy}$ in $G-v$: $V(P_{xy})\cap N_G[v]-\{x,y\}\neq\phi$. Let $T=\{v\}\cup \cup_{P_{xy}\ drum\ de\ la\ x\ la\ y\ din\ G} N_G[v]\cap V(P_{xy})$. From this construction it follows that $T-\{v\}\subset N_G(v)$ and in $G-T$ the vertices x and y are in differrent connected components. So T is a star-cutset.

Let G a graph with a star-cutset. We prove that G is *CC*-free. Let $T\subset V(G)$, T a star-cutset. Then $\exists v\in V(G)$ such that $vx\in E(G)$, $\forall x\in T-\{v\}$ and $G-T$ is not connected. Let C_1,C_2 two connected components of $G-T$. Therefore $\exists v\in T$, $\exists x\in C_1$, $\exists y\in C_2$ such that $\forall P_{xy}: V(P_{xy})\cap N_G[v]-\{x,y\}\neq\phi$, so G is *CC*-free.

6 A New Characterization of *CC*-Free Graphs Using the Weakly Decomposition

The notion of weakly decomposition (a partition of the set of vertices in three classes A, B, C such that A induces a connected graph and C is totally adjacent to B and totally non-adjacent to A) and the study of its properties allow us to obtain several important results such as: characterization of cographs, $K_{1,3}$-free graphs, $\{P_4,C_4\}$-free and paw-free graphs.

Definition 7. ([17], [18]) *A set $A\subset V(G)$ is called a weakly set of the graph G if $N_G(A)\neq V(G)-A$ and $G(A)$ is connected. If A is a weakly set, maximal with respect to set inclusion, then $G(A)$ is called a weakly component. For simplicity, the weakly component $G(A)$ will be denoted with A.*

Definition 8. ([17]) *Let $G=(V,E)$ be a connected and non-complete graph. If A is a weakly set, then the partition $\{A,N(A),V-A\cup N(A)\}$ is called a weakly decomposition of G with respect to A.*

Below we remind a characterization of the weakly decomposition of a graph.

The name of "*weakly component*" is justified by the following result.

Theorem 4. ([17], [18] *Every connected and non-complete graph* $G=(V,E)$ *admits a weakly component* A *such that* $G(V-A)=G(N(A))+G(\overline{N}(A))$.

Theorem 5. ([17], [18]) *Let (G=(V,E) be a connected and non-complete graph and* $A \subset V$. *Then* A *is a weakly component of* G *if and only if* $G(A)$ *is connected and* $N(A) \sim \overline{N}(A)$.

The next result, that follows from Theorem 5, ensures the existence of a weakly decomposition in a connected and non-complete graph.

Corollary 1. *If* $G=(V,E)$ *is a connected and non-complete graph, then* V *admits a weakly decomposition* (A,B,C), *such that* $G(A)$ *is a weakly component and* $G(V-A)=G(B)+G(C)$.

Theorem 5 provides an $O(n+m)$ algorithm for building a weakly decomposition for a non-complete and connected graph.

Algorithm for the weakly decomposition of a graph ([17], [18])
Input: A connected graph with at least two nonadjacent vertices, $G=(V,E)$.
Output: A partition $V=(A,N,R)$ such that $G(A)$ is connected, $N=N(A)$, $A \nsim R=\overline{N}(A)$.

```
begin
  A := any set of vertices such that A ∪ N(A) ≠ V
  N := N(A)
  R := V − A ∪ N(A)
  while (∃n ∈ N, ∃r ∈ R such that nr ∉ E) do
    begin
      A := A ∪ {n}
      N := (N − {n}) ∪ (N(n) ∩ R)
      R := R − (N(n) ∩ R)
  end
end
```

A new characterization of *CC*-free graphs, using weakly decomposition, is given below. Some similar results are stated in [17], but for confidentially connected graphs.

Theorem 6. *A connected and non-complete graph* $G=(V,E)$ *is CC -free if and only if* $\exists v \in V(G)$ *such that* $\overline{N}(v)$ *is not a weakly component.*

Proof. Let G a *CC*-free graphs. Then $\exists v \in V(G)$ such that $[\overline{N}(v)]_G$ is not connected. Therefore, $[\overline{N}(v)]_G$ is not the weakly component. Also, from G *CC*-free it follows that there exists $v, w \in V(G)$ with $vw \in E(G)$ and $N(w) \subseteq N(v)$. Furthermore, $\overline{N}(v) \subseteq \overline{N}(w)$, but this situation is not specific for weakly components because these are maximal with respect to inclusion, so they are not comparable. In both situations, $\exists v \in V(G)$ so that $[\overline{N}(v)]_G$ is not a weakly component.

Suppose that $\exists v \in V(G)$ such that $[\overline{N}(v)]_G$ is not a weakly component. If G would be confidentially connected, then for any v in $V(G)$ the following hold: $N(v) = N(\overline{N}(v))$, $\overline{N}(\overline{N}(v)) = \{v\}$ and $\{v\} \sim N(v)$. So $N(\overline{N}(v)) \sim \overline{N}(\overline{N}(v))$. Because $[\overline{N}(v)]_G$ is connected, it follows that $[\overline{N}(v)]_G$ is a weakly component.

Theorem 6 provides the following recognition algorithm for *CC*-free graphs.

Algorithm Recognition
Input: A connected, non-complete graph $G = (V, E)$.
Output: An answer to the question: "Is G *CC*-free?"
begin
1. Generate L_G, the family of the weakly components of G as follows:
   $L_G \leftarrow \varnothing$
   while $V \neq \varnothing$ *do*
   determine the weakly component A with the weakly decomposition algorithm
   $L_G \leftarrow L_G \cup \{A\}$
   $V \leftarrow V - A$
2. Determine $\overline{N}_G(v)$, $\forall v \in V$
3. *If* $\exists A \in L_G$ such that $A \neq \overline{N}_G(v)$, $\forall v \in V$
   then *Return*: "G is CC-free"
   else
   then *Return*: "G is not *CC*-free"
end
EndRecognition

As Theorem 5 provides an $O(n+m)$ algorithm for building a weakly decomposition for a non-complete and connected graph, it follows that step 1 of the algorithm above is $O(n \cdot (n+m))$. Because steps 2 and 3 perform in smaller time, it follows that the complexity of the recognition algorithm for *CC*-free graphs is $O(n \cdot (n+m))$.

7 Conclusions and Future Work

In this paper we introduced the class of *CC* -free graphs and established its relation with the *AT* -free graphs, as well as with the graphs that have a star-cutset. We gave a characterization of these graphs using neighborhoods and weakly decomposition. Our future work concerns to give some applications of *CC* -free graphs including the medicine. Also we will explore the connection of the *CC*-free graphs with the intelligent systems.

Acknowledgments. The research of Mihai Talmaciu was supported by the project entitled Classes of graphs, complexity of problems and algorithms, "AR-FRBCF", 2012-2013, a Bilateral Cooperation Research Project, involving Romanian Academy, National Academy of Sciences of Belarus, Republican Foundation for Fundamental Research and "Vasile Alecsandri"University of Bacau.

The research of Barna Iantovics was supported by the project Transnational Network for Integrated Management of Postdoctoral Research in Communicating Sciences. Institutional building (postdoctoral school) and fellowships program (CommScie) - POSDRU/89/1.5/S/63663, financed under the Sectorial Operational Programme Human Resources Development 2007-2013.

The research of Elena Nechita was supported by the project entitled Electronic Health Records for the Next Generation Medical Decision Support in Romanian and Bulgarian National Healthcare Systems, abbreviated as NextGenElectroMedSupport, a Bilateral Cooperation Research Project between Romania and Bulgaria.

References

1. Berge, C.: Graphs. Nort-Holland, Amsterdam (1985)
2. Berry, A., Bordat, J.-P., Heggernes, P.: Recognizing weakly triangulated graphs by edge separability, Res. Rep. LIRMM, submittted to SWAT (2000)
3. Carlisle, M.C., Loyd, E.L.: On the k-Coloring of Intervals. In: Dehne, F., Fiala, F., Koczkodaj, W.W. (eds.) ICCI 1991. LNCS, vol. 497, pp. 90–101. Springer, Heidelberg (1991)
4. Chvatal, V.: Star cutsets and perfect graphs. J. Combin. Theory, Ser. B 39, 189–199 (1985)
5. Cohen, J.E.: Food webs and niche space. Monographs in population biology, vol. 11, pp. xv+1 190. Princeton University Press, Princeton (1978) ISBN 978-0-691-08202-8
6. Corneil, D.G., Olariu, S., Stewart, L.: Ateroidal triple-free graphs. SIAM J. Discrete Math. 10(3), 399–430 (1997)
7. Croitoru, C., Olaru, E., Talmaciu, M.: Confidentially connected graphs, The annals of the University "Dunarea de Jos" of Galati. In: Proceedings of the Intern. Conf. "The Risk in Contemporany Economy", Supplement to Tome XVIII, XXII (2000)
8. Croitoru, C., Talmaciu, M.: On Confidentially connected graphs. Bul. Stiint. Univ. Baia Mare, Ser B, Matematica - Informatica XVI(1), 13–16 (2000)
9. Dirac, G.A.: On rigid circuit graphs. Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg 25(1-2), 71–76 (1961), doi:10.1007/BF02992776
10. Fabri, J.: Automatic storage optimization. UMI Press Ann Arbor, MI (1982)
11. Fulkerson, D.R., Gross, O.A.: Incidence matrices and interval graphs. Pacific J. Math. 15(3), 835–855 (1965)

12. Golumbic, M.C.: Algorithmic graph theory and perfect graphs. Academic Press, New York (1980)
13. Hashimoto, A., Stevens, J.: Wire routing by optimizing channel assignment within large apertures. In: Proc., 8th IEEE Design Automation Workshop, pp. 155–169 (1971)
14. Jungck, J.R., Dick, O., Dick, A.G.: Computer assisted sequencing, Interval graphs and molecular evolution. Biosystem 15, 259–273 (1982)
15. Lekkerkerker, C.G., Boland, J.C.: Representation of a finite graph by a set of intervals on the real line. Fund. Math. 51, 45–64 (1962)
16. Ohtsuki, T., Mori, H., Khu, E.S., Kashiwabara, T., Fujisawa, T.: One dimensional logic gate assignment and interval graph. IEEE Trans. Circuits and Systems 26, 675–684 (1979)
17. Talmaciu, M.: Decomposition problems in the graph theory with applications in combinatorial optimization - Ph. D. Thesis, University "Al. I. Cuza" Iasi, Romania (2002)
18. Talmaciu, M., Nechita, E.: Recognition algorithm for diamond-free graphs. Informatica 18(3), 457–462 (2007)
19. Zhang, P., Schon, E.A., Fischer, S.G., Cayanis, E., Weiss, J., Kistler, S., Bourne, P.E.: An algorithm based on graph theory for the assembly of contigs in physical mapping of DNA. Bioinformatics 10(3), 309–317 (1994), doi:10.1093/bioin-formatics/10.3.309

Route Search Algorithm in Timetable Graphs and Extension for Block Agents

Ion Cozac

"Petru Maior" University of Tirgu-Mures
cozac@upm.ro

Abstract. This paper describes an algorithm that determines routes using three graphs: the railway graph, the train timetable graph and the summary timetable graph. The search in the timetable graphs is guided by a subgraph of the railway graph, which is defined by the nodes that form an ellipse around the minimum distance path from departure to arrival. We also present some performance evaluations of our proposed algorithm. Finally we describe an extension of this algorithm that can be used in conjunction with block agents to find routes in large timetable graphs, and some applications for medical domain.

Keywords: timetable graph, maximum allowed distance, bridges in graph, block agents.

1 Introduction

The problem of determining optimum path in timetable graphs has been intensively studied during the last ten years, both from theoretical and practical point of view. This problem is important especially if we want to plan trips using the public transportation system (trains, buses, airplanes etc). All the approaches are based on supplementary data structures which are used to speedup the route search. The speedup is important because the server must be able to solve numerous queries simultaneously.

One approach to speedup the search uses multi-level graphs [9]. Precomputed shortest paths are replaced by single edges whose weights are the cost of the corresponding paths. The paper introduces the concept of multi-level decomposition.

Another approach models the timetable information using some heuristics to speedup the implementation [6]. The authors exhibit important extensions of the time-dependent approach to model the earliest arrival and minimum number of transfer problems.

An overview of known models and efficient algorithms for optimal solving the timetable information problem has been given [5]. A comparison between time-dependent and time-expanded approaches has been made in order to evaluate their relative performance [7].

Some authors considered the planning route problem using timetable information, taking into account the presence of delays [4].

The SHARC approach [1] uses contraction, that is, iteratively removal of non-important nodes, plus edges addition to preserve correct costs between remaining

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 189–201.
DOI: 10.1007/978-3-319-00029-9_17

nodes. It is a fast unidirectional algorithm, which is advantageous for timetable graphs where bidirectional search is not allowed.

New reviews of all known algorithms in this field have been made [2], and the algorithms have been evaluated using large, real data sets that are publicly available.

This paper is structured as follows. Section 2 introduces some basic notions which will be used throughout the paper: railway graph, bridges and blocks in graph, timetable graph. Section 3 describes in detail our contribution to solve the route search problem in timetable graphs. Section 4 presents some experiments based on our approach. Section 5 shows how our approach can be extended to distribute the route search problem in large timetable graphs, using block agents. Section 6 presents some possible applications of our algorithms in medical domain. The last section is reserved for conclusions and discussions about future work.

Our approach is based on three graphs: the railway graph, the original timetable graph and the summary timetable graph. First, we delimit the search space in the railway graph, that is, which are the stations that can be taken into account when searching routes from departure to arrival. Second, we use the summary graph to identify which trains are valid for our planned route. The original timetable graph is used only in some few special cases, if departure and / or arrival station is not in the summary graph.

2 Preliminaries

Railway graphs. A railway graph $G = (V, E)$ is symmetric and weighted: each edge (v,w) has a cost $c(v,w)>0$. We assume the graph G is connected, that is, there is a path from any vertex v to any other vertex w. The cost of a path $P = \langle v_1,\ldots,v_k\rangle$ is the sum of the cost of all its edges:

$$c(P) = \sum_{i=2}^{k} c(v_{i-1}, v_i) \tag{1}$$

$P^* = \langle s,\ldots,t\rangle$ is a shortest path from s to t if there is no other path P' such that $c(P^*) > c(P')$.

Bridges in Graphs. A *bridge* is an edge whose deletion will disconnect the graph [10]. A *block* is a maximal subgraph that doesn't contain bridges, that is, every new edge we can add to this subgraph is a bridge.

Timetable Graphs. Two basic models for timetable information have been used throughout the above mentioned papers. The time-dependent model uses one node for each station, and every connection by train between two neighbor stations has a corresponding arc. The time-expanded model uses one node for each event (departure from / arrival in station), and an arc between two nodes depicts the train route. We used the time-expanded model due to its versatility, especially because transfers are easily managed.

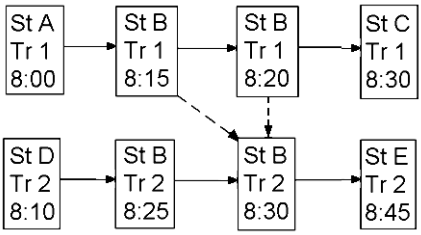


Fig. 1. Partial timetable subgraph, the time-expanded model

Let us examine Figure 1, which depicts the time-expanded model using a partial timetable subgraph. The train T1 starts from station StA at 8:00, arrives in station StB at 8:15, departs at 8:20 and arrives in station StC at 8:30. The train T2 passes through stations StD, StB and StE. If we want to travel from StA to StE, we will take the train T1 from StA to StB, then we will wait for the train T2 to continue the travel.

A *timetable graph* $G_T = (V_T, A_T)$ is directed (not symmetric). Each vertex v_T is a triplet (s,t,h) where s is a railway station, t is a train and h is a time moment. An arc (v_T,w_T) may denote :

- train passing from one station to the next station, if $s(v_T) \neq s(w_T)$ and $t(v_T) = t(w_T)$ (two distinct stations, the same train);
- train halt in a station if $s(v_T) = s(w_T)$ and $t(v_T) = t(w_T)$ (the same station, the same train);
- transfer possibility, if the traveler may gets off in a station to take another train (distinct trains).

The arcs of first two types are marked with continuous links, and the arcs of third type are marked with dashed links. In order to simplify the computations, the field t takes integer values from 0 (for 0:00) to 1439 (for 23:59).

Let us denote by *tim* the cost function which is defined as follows:

$$\text{if } (v_T,w_T) \in A_T \text{ then } tim(v_T,w_T) = (h(w_T) - h(v_T) + 1440) \text{ modulo } 1440 \qquad (2)$$

$tim(v_T,w_T)$ may be the time needed to cover the route from $s(v_T)$ to $s(w_T)$ (different stations), or the waiting time in that station.

There are many solutions to build the timetable graph, if we are talking about how to model the transfers. One solution is to connect each arrival vertex to each departure vertex (Fig 2.a). Another solution is to connect in a ring all the vertices that are related to the same station; the vertices are ordered by time. The last vertex is linked to the first vertex, to allow night transfers [2] [4] [5] [7] [9].

We used a slightly different solution (Fig. 2.b). The arrival vertex is connected to the nearest departure vertex, and the ring contains only departure vertices. This solution allows us to manage easily the transfers, especially if we want to count them, to limit their number, to impose minimum transfer time.

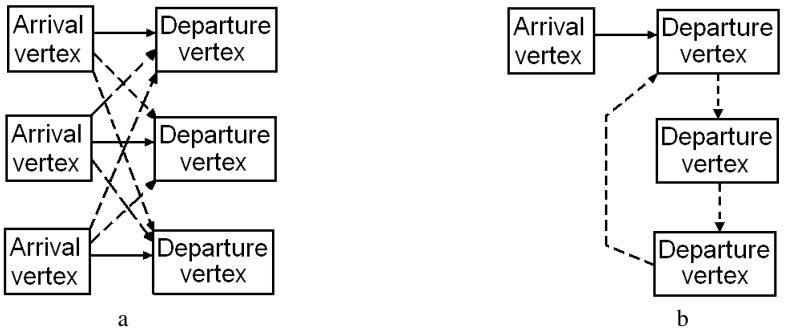


Fig. 2. How to model the transfers: a) from each arrival point to each departure point; b) building a ring with departure points

Suppose k arrival vertices and k departure vertices are related to a particular station. Using the solution depicted in Figure 2.a), the number of auxiliary arcs is k^2. Using the solution depicted in Figure 2.b), the number of auxiliary arcs is k.

Block agents [11] [7]. A *block agent* is a virtual entity that has the following properties:

- autonomy : block agents operate their subproblems without direct intervention of other agents or human;
- social ability : block agents interact with other agents by sending messages to communicate consistent partial states;
- proactivity : block agents perceive their environment and changes in it; they can extend new partial consistent states to more complete consistent states.

3 Route Search Algorithm

3.1 How to Generate the Summary Timetable Graph

Our search algorithm uses the railway graph and two timetable graphs. The first timetable graph depicts only the routes of all trains and nothing else.

If a train starts from station s, or station s has at least three neighbors, then s is set as important. The reason is that the station s may be used to take another train to continue the travel. The second timetable graph depicts a summary of the first timetable, such that every arc links vertices which are related to important stations. This graph includes also rings which are related to important stations.

The built of the summary timetable graph takes about $\boldsymbol{O}(n \log n)$ time, due to some sort and search operations, where $n = |V_T|$.

3.2 Effective Route Search Algorithm

Let dp be the departure station and ar the arrival station. We have to determine as many as possible optimum routes from dp to ar. We should define what an optimum

route is. Let us consider the station *dp* and departure time *h*; we have to determine a route in such a way that we arrive in station *ar* the earliest possible moment. In other words, we solve the earliest arrival problem [2, 4, 5, 6, 7, 9].

The summary timetable graph is used to accelerate the search process, by examining only vertices that are related to important stations. What should we do if *dp*, *ar* or both are not important?

If *dp* is not an important station, we start the search from the departure vertex using the original timetable graph. The start vertex is related to the station *dp* and its time is *h*. The search stops when the first important station is reached. If the next vertex exists, it is of departure type and is related to the same important station. If the search reaches the arrival station, the algorithm returns the path to this vertex. If the search reaches the end of the train route, the algorithm must search for a departure vertex which is related to the end station.

If *dp* is not an important station, the first phase of the algorithm may either return the found path or a departure vertex which is related to an important station.

The destination station *ar* may also be important or unimportant. No matter the type of *ar*, we assign to each train a pointer to the first arrival vertex which is related to *ar*, if such vertex exists in the train route.

Now we have a departure vertex which is related to an important station, so we can use only the summary timetable graph. A variant of Dijkstra's algorithm [3] is used to determine optimum routes in this graph. Every time the search algorithm detects a new train, it checks if this train passes through the arrival station. If this is the case we may insert the arrival vertex in the priority queue and continue the search, or we may simply return the path to it.

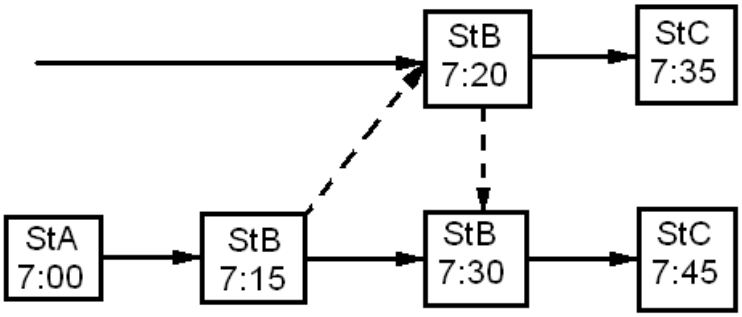


Fig. 3. Optimum versus non optimum path from StA to StC

Let us examine Figure 3 to analyze a particular case. When the algorithm starts, the arrival vertex (time 7:45) is detected on the route of the start train. If this vertex is inserted in the priority queue and the algorithm continues, another arrival vertex will be reached, which gives the best arrival time (7:35). The direct route is not optimum, but it doesn't need any transfer.

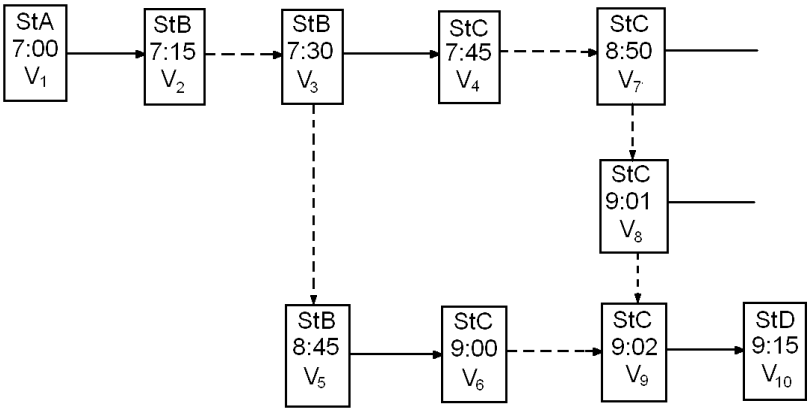


Fig. 4. Determining optimum route from StA to StD with imposed transfer time

Important Remark. Using rings to connect departure vertices may cause troubles when we are searching for routes with imposed minimum transfer time. Let us examine the figure 4, which depicts a critical case in station StC, assuming the waiting time is imposed to be at least 3 minutes. According to the algorithm of Dijkstra, the vertices are extracted from the priority queue in this order: v_1, v_2, v_3, v_4, v_5, v_7, v_6, v_8, v_9, etc. When v_6 is extracted, the successor v_9 is entered in the priority queue with its settings (cost of the route from departure etc). The waiting time is 2 minutes, which is not acceptable. So we have to find another predecessor for v_9, if such predecessor exists. In this case, a better predecessor is v_4, which has been extracted before. If such predecessor doesn't exist, we have to consider the next departure vertex which will replace the vertex v_9 in the current path.

We need to use a supplementary list of vertices, which is assigned to each important station we traverse. This list contains all the arrival vertices which are detected by the search algorithm. When we traverse the ring, this list must also be scanned to link a departure vertex from the best previous arrival vertex, such that the imposed conditions are fulfilled.

This problem does not appear if we are using the method which has been depicted in Fig. 2.a). But that method has another disadvantage: it consumes too much memory and is slower.

3.3 How to Improve the Search Algorithm

This algorithm may have a large search space if stations *dp* and *ar* are far away from each other and there is no direct train to connect them, so the response time may not be acceptable. We know that, using the algorithm of Dijkstra, the search will span a disk around the departure vertex until the destination is reached.

The difficulty of search process is considered using our experimental data. A route is difficult if the search algorithm spans more than quarter the number of arcs.

In order to reduce the search space, we used the following idea (see Fig. 5).

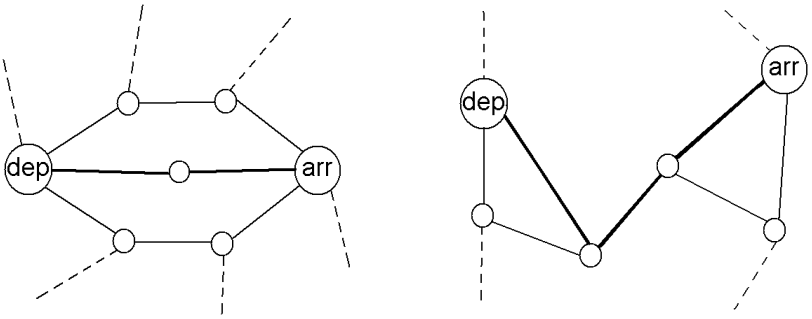


Fig. 5. Determining an ellipse around the minimum cost path

We determine a minimum cost path from *dp* to *ar* in the railway graph. We also determine a subgraph which includes the found path and each vertex z fulfills the following condition:

$$dist(dp,z) + dist(z,ar) \leq f \cdot dist(dp,ar) \quad (3)$$

where $dist(x,y)$ is the cost of the shortest path from x to y (the cost is computed over the railway graph), and $f > 1$ is a convenient chosen parameter (see Experiments). This subgraph will define an ellipse around the minimum cost path and will be used as search space. The value $f \cdot dist(dp,ar)$ is the ***maximum allowed distance*** for any route we are searching.

The search space may have some fragments that will never be used by any route from *dp* to *ar* (see dashed lines in Fig. 5). These fragments are identified using bridges and blocks in the ellipse subgraph. First, we determine the blocks that include the arcs of the shortest path; we may have one block (left) or many blocks (two blocks and one bridge, right). The arcs that don't belong to any of these blocks are dropped.

The usage of this method is motivated by the fact that the railway graph is much lower than the timetable graph (see Experiments). The search space is reduced once, before starting any route search, and every route search will span a disk slice.

4 Experiments

We implemented our algorithms in ANSI C and compiled them using GNU C Compiler version 4.1.0 on an AMD Athlon processor at 1.4 GHz with 2 GB of RAM running Linux Red-Hat 4.1.0-3.

We dealt with the whole Romanian timetable, valid for year 2011, which includes international trains that pass through Romania. The railway graph has 2250 vertices (250 vertices correspond to important stations) and 5150 arcs, the timetable graph has 58300 vertices and 56150 arcs, and the auxiliary graph has 20850 vertices and 33450 arcs. All these data structures need about 1.8 MB to be stored.

The summary timetable graph is generated in about one second. If the routes or the timetables of some particular trains are changed, the original timetable graph is

updated and the summary timetable graph is rebuilt. Whenever the timetable is changed, the railway graph is updated such that it keeps only those links which are traversed by trains.

We queried for the following groups of routes:

- from Bucuresti-Nord to Tirgu-Mures and return;
- from Bucuresti-Nord to Botosani and return;
- from Suceava to Tirgu-Mures and return;
- random departure and arrival.

We chose routes from and to Bucuresti-Nord because it is the most important Romanian station, and has assigned the maximum number of departures and arrivals. We also chose random routes to estimate the general performance of our proposed algorithm.

The parameter *f,* which is used to determine an ellipse around the minimum distance path (to limit the search space), has been set to 1.60. This value has been determined by examining some sinuous routes between Romanian stations. One such route is between Suceava and Tirgu-Mures, which gives three distinct rail paths having distances 365 km, 450 km and 532 km respectively. The ratio between the longest and the shortest path is 1.45. The value of *f* has been increased for safety reasons.

Our railway graph is denser inside Romania than outside. The reason to consider only Romanian stations (the dense part) is that, given a particular pair (*dp*,*ar*) whose shortest distance is *d*, the following condition is fulfilled:

$$(df+e)/(d+e) < f, \text{ for any supplementary distance } e>0 \tag{4}$$

So this value of *f* covers well almost all routes. It is possible, even though very unlikely, that some routes may not be optimal. That is, there may be another route whose distance is greater than the found one, but the travel time is lower. To prevent such cases, an exhaustive method could be used to check routes between any two important stations. This method will imply 62500 queries, and assuming 0.4 seconds on average for one query, this action could take about 7 hours. We may also determine different values for distinct departure stations, to obtain better space reduction. This solution should be improved to eliminate unnecessary queries, for example considering the railway topology.

The first column of Table 1 below shows: departure station, arrival station, minimum and maximum distance for found routes, number of departure vertices. The second column shows: total response time, averaged response time. The total response time is measured from the point of query reception to the point of answer completing. The averaged response time is computed as ratio between total response time and number of departure vertices.

The third column shows the total number of spanned arcs for two cases: with and without space reduction, the unproductive searches being included (see remark below). All these tests used the same triplet: railway graph, original timetable graph, summary timetable graph. The number of spanned arcs depends heavily on the topology of the railway net, which is not uniformly dense.

Table 1. Performance evaluation (samples)

Routes	Response time	Spanned arcs
Bucuresti-Nord -- Tirgu-Mures distances : 449 km / 612 km 109 departure vertices	total : 0.356 seconds per dep vtx : 0.0033 sec	48000 / 185000
Tirgu-Mures -- Bucuresti-Nord distances : 449 km / 612 km 19 departure vertices	total : 0.035 seconds per dep vtx : 0.0018 sec	3200 / 8000
Bucuresti-Nord -- Botosani distance : 477 km 109 departure vertices	total : 0.181 seconds per dep vtx : 0.0017 sec	24000 / 280000
Botosani -- Bucuresti-Nord distance : 477 km 8 departure vertices	total : 0.018 sec per dep vtx : 0.0023 sec	1500 / 2500
Suceava – Tirgu-Mures distances : 365 km / 532 km 61 departure vertices	total : 0.078 sec per dep vtx : 0.0013 sec	16500 / 96000
Tirgu-Mures -- Suceava distances : 365 km / 532 km 19 departure vertices	total : 0.032 sec per dep vtx : 0.0017 sec	2400 / 12000

In order to estimate the general performance of our proposed algorithm, we generated 10000 random queries. The estimation gives the following results: 400 seconds to answer all 10000 queries for 125000 total departure vertices. The averaged response time is 0.04 seconds per query, and 0.0032 seconds per departure vertex.

Generally speaking, the total response time depends on the number of departure vertices, the number of necessary transfers and the size of the ellipse graph which is built using the maximum allowed distance criterion.

It is important to say that the end user is interested about the total response time: he receives several variants of which he can choose. It is also important for us to know the averaged response time (per departure vertex) to evaluate the efficiency of the algorithm.

Remark. In order to find routes from departure to arrival, the application starts a new search from each departure vertex. Some of these vertices are not productive, that is, they will not give any route to destination. But the time consumed with these unproductive nodes is taken into account to measure the response time (total and averaged).

If we don't use any speedup technique (that is, the original timetable graph is extended with supplementary arcs to solve the transfers, so the summary timetable graph is not used), the total response time may be up to 15 times slower. The coefficient is lower if direct trains connect departure and arrival, and higher if three or more trains are needed to cover the route.

The auxiliary graph has been built using rings to connect departure vertices which are assigned to the same important station (see Fig. 2.b). This solution uses low

volume of memory to store this data structure. Another possible technique could be to connect each arrival vertex to each departure vertex (see Fig. 2.a), if the related station is important. Using this technique, the auxiliary graph could have 110000 arcs, and the response time may be three times slower. These estimations are valid for our timetable. If the timetable is higher, the number of arcs and the response time will be higher; the increase factor is worse than linear.

5 Extension for Block Agents

The above solution may be successfully used for medium timetable graphs, for example, if the timetable stores information about a particular country, including its international trains. This solution may not be well suited for continental timetables, due to the following reasons:

- searching for routes traversing two or more countries may take long time to be completed, if three or more trains are needed to cover the route;
- updating the timetable is a time consuming task.

To surpass these difficulties, we propose a solution based on block agents. We didn't find any paper that tackles the subject of finding routes in timetable graphs using block agents. The only paper that tackles a close related subject is, to our knowledge, the work of Salido et al [8], which discusses about how to design a new timetable using constraint satisfactions.

Now let us depict our proposed solution. Every country uses its own graphs (one railway graph and two timetable graphs), which are managed by its own block agent. The timetable graphs include the trains that traverse this country, including the international ones. If both departure and arrival are known, the above solution may be successfully used without any auxiliary tools.

To solve the general problem, we introduce two new block agents. The first agent (let us denote it by CRA - central railway agent) manages the continental railway graph. The second agent (denoted as CTA - central timetable agent) manages a subgraph of the continental timetable graph which is built as following. Each agent who is related to a particular country identifies al the trains that are needed to cover routes between any two distinct frontier stations. The timetable of these trains is sent to CTA, which will build the needed timetable subgraph. The agent CRA receives also the list of stations that are managed by CTA. The central agent CTA will manage the summary timetable graph only, which includes the important stations of the continent.

Now we are able to find routes between any two stations, which are located in different countries. The agent CRA receives a query to identify the ellipse around the minimum distance path and identifies the countries that are covered by this ellipse. It also identifies some important intermediate stations: one set of stations for the departure country and another set of stations for the arrival country (Fig. 6). Three queries are sent to three different agents.

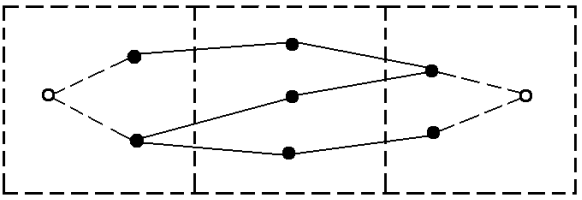


Fig. 6. Determining a route that traverses more than two countries and needs cooperation among agents

The first query is addressed to the agent who has knowledge about the departure station. This agent will determine the first list of routes: if the departure station is not in the list of the agent CTA, it determines trains from departure to the nearest stations that may be managed by CTA (see dashed links). If the departure station is managed by CTA, the first query is not needed.

The second query is addressed to the agent CTA, which will determine the list of routes between departure country and arrival country (see continuous links). This step is always necessary if the departure country is not able to determine the route.

The third query is addressed to the agent who has knowledge about the arrival station. This agent will determine the last list of routes: if the arrival station is not in the map of CTA, it determines trains from the stations that have been set by CRA to the arrival station (see dashed links). If the arrival station is managed by CTA, the third query is not needed.

If the departure and/or arrival station is not in the CTA map, we need to build routes from arrival to departure by mixing the partial routes.

6 Application for Medical Domain

The first and obvious application of the algorithms presented above is for touristic domain, if one is interested to plan its route using the public transportation system. We can also use these algorithms in the medical system, for example to optimize the traffic of the emergency vehicles. A customized application may be designed to watch over this traffic, if the vehicles are endowed with GPS devices.

Using the public transportation system may have a positive impact concerning the public health. If people don't use their own cars, the traffic will be decongested, the pollution will be reduced and the emergency traffic will be improved.

7 Conclusions and Future Work

This paper presented a simple, less memory consuming and fast solution that can be used to solve the problem of searching routes in timetable graphs. We adopted this solution in the hypothesis that short / direct routes are much more frequent than long routes with two or more transfers (difficult routes). Our approach requires the least memory consumption and the shortest time needed to generate the summary timetable graph. Other methods require about several minutes or hours to generate the auxiliary data structures.

We have to search for better solutions to solve the update problem, to eliminate the rebuilt of the summary timetable graph. Removing an existing train or inserting a new train should involve some operations to update the links between different trains in all important related stations. Updating the timetable graph of an existing train may involve a new order of vertices in the related rings. Improved solutions are necessary: if the original timetable graph will have 1000000 vertices, the built of the summary timetable graph will take about 24 seconds - the time complexity is $\boldsymbol{O}(n \log n)$.

In order to reduce the search space, we determine an ellipse around the shortest distance path, which is defined using the maximum allowed distance criterion, based on the parameter *f*. Its value has been determined using an empirical method. We may determine better (possible multiple) values for *f*, to obtain better space reduction. We may also search for other solutions to reduce the search space, maybe with the cost of increasing the preprocessing time and / or the memory consumption for supplementary data structures.

We also described a theoretical solution to solve the large timetable case, but this solution is not yet implemented. The main difficulty consists in solving some special cases, if some trains have running restrictions. Indeed, it is possible that some intermediate trains or some final trains don't run daily. The queries may be solved sequentially or simultaneously. In the sequential approach, the agent located in the departure country will solve the departure subproblem. CTA will solve the intermediate subproblem using the output of the first train. The last agent will solve the arrival problem using the output of CTA.

Another approach could be to solve simultaneously these three subproblems. We have to obtain multiple solution sets, each set being valid for a certain day. This is because we can not know in advance the arrival day in an intermediate station following a certain route. The arrival day is known only after the search algorithm reaches the destination station.

Using separate agents for different countries will be an advantageous solution, because this way every local agent will almost always manage local queries, that is, departure and arrival are in the same country. Of course, local agents are also able to solve queries for routes between neighbor countries. Anyway, we can assume that local queries are much more frequent than international ones, so this is a good reason to continue this work using block agents.

Acknowledgement. This work was supported by the Bilateral Cooperation Research Project between Romania and Bulgaria, entitled: "Electronic Health Records for the Next Generation Medical Decision Support in Romanian and Bulgarian National Healthcare System", the involved institutions are the Technical University of Sofia and Petru Maior University of Tirgu Mures.

References

1. Bauer, R., Delling, D.: SHARC: Fast and Robust Unidirectional Routing. In: Proceedings of 10th Workshop of Algorithms and Engineering Experiments (ALENEX 2008), pp. 13–26. SIAM (April 2008)

2. Bauer, R., Delling, D., Wagner, D.: Experimental Study of Speed-up Techniques for Timetable Information Systems. Journal Networks 57, 38–52 (2011)
3. Dijkstra, E.W.: A Note on Two Problems Ion Connection with Graphs. Numerische Mathematic 1, 269–271 (1959)
4. Frede, L., Müller-Hannemann, M., Schnee, M.: Efficient On-trip Timetable Information in the Presence of Delays. In: ATMOS 2008, 8th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization and Systems, pp. 1–16 (2008)
5. http://drops.dagstuhl.de/opus/volltexte/2008/1584
6. Müller-Hannemann, M., Schulz, F., Wagner, D., Zaroliagis, C.D.: Timetable Information: Models and Algorithms. In: Geraets, F., Kroon, L.G., Schoebel, A., Wagner, D., Zaroliagis, C.D. (eds.) Railway Optimization 2004. LNCS, vol. 4359, pp. 67–90. Springer, Heidelberg (2007)
7. Pyrga, E., Schulz, F., Wagner, D., Zaroliagis, C.: Toward Realistic Modeling of Timetable Information Through the Time-dependent Approach. In: Proc. of the 3rd Workshop on Algorithmic Methods and Models for Optimization of Railways (ATMOS 2003). ENTCS, vol. 92, pp. 85–103. Elsevier (2004)
8. Pyrga, E., Schulz, F., Wagner, D., Zaroliagis, C.: Efficient Models for Timetable Information in Public Transportation Systems. ACM Journal of Experimental Algorithms 12, Article 2.4 (2007)
9. Salido, M.A., Abril, M., Barber, F., Ingolotti, L., Tormos, P., Lova, A.: Domain-Dependent Distributed Models for Railway Scheduling. Journal Knowledge Based Systems (Elsevier) 20, 186–194 (2007)
10. Schulz, F., Wagner, D., Zaroliagis, C.D.: Using Multi-level Graphs for Timetable Information in Railway Systems. In: Mount, D.M., Stein, C. (eds.) ALENEX 2002. LNCS, vol. 2409, pp. 43–59. Springer, Heidelberg (2002)
11. Tarjan, R.E.: A Note on Finding the Bridges of a Graph. Information Processing Letters 2(6), 160–161 (1974)
12. Wooldridge, M., Jennings, R.: Agent Theories, Architectures, and Languages: A Survey. Intelligent Agents, 1–22 (1995)

From Individual EHR Maintenance to Generalised Findings: Experiments for Application of NLP to Patient-Related Texts

Galia Angelova[1], Dimitar Tcharaktchiev[2], Svetla Boytcheva[1,3], Ivelina Nikolova[1], Hristo Dimitrov[2], and Zhivko Angelov[1]

[1] Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, Sofia, Bulgaria
[2] University Specialised Hospital for Active Treatment of Endocrinology "Acad. I. Penchev" (USHATE), Medical University – Sofia, Bulgaria
[3] American University in Bulgaria, Blagoevgrad, Bulgaria
{galia,iva}@lml.bas.bg,
{dimitardt,svetla.boytcheva}@gmail.com, h_dimitrov@yahoo.com,
angelov@adiss-bg.com

Abstract. Experiments in automatic analysis of free texts in Bulgarian hospital discharge letters are presented. Natural Language Processing (NLP) has been applied to medical texts since decades but high-quality results have been demonstrated only recently. The progress in automatic text analysis opens new directions for secondary use of Electronic Health Records (EHR). It enables also the design and development of software systems which provide better patient access to his/her health records as well as better maintenance of large EHR archives. We report about successful extraction of important patient-related entities from hospital EHR texts and consider several scenarios for application of NLP modules in healthcare software systems.

Keywords: Information Extraction, AI in Health Informatics.

1 Introduction

The performance of automatic text analysis gradually improves during the last decades but these systems are still rarely used outside the research groups where they have been developed [1]. There are several reasons for the slow penetration of the NLP technologies into practical settings: *(i)* developing high-quality NLP is very expensive as it requires much effort for collection of relevant language resources as well as for the design and implementation of software that might process various formats and styles of texts. Moreover, the medical domain is very large and the initial investments for the development of convincing NLP demonstrators are huge; *(ii)* the exploitation of NLP modules requires constant effort for supporting lexicon updates and tuning the systems to new linguistic constructions and text types; *(iii)* it is well

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 203–212.
DOI: 10.1007/978-3-319-00029-9_18

known that the technology works with high accuracy but not 100%; therefore some results might be erroneous and misleading.

On the other hand, the quick adoption of EHR worldwide implies constant growth of electronic narratives discussing patient-related information. Actually the most important findings about the patients are kept as free texts in various documents and languages. These text descriptions are oriented to human readers but the computers might process them for statistical analysis, research purposes, obtaining data to support decision making and health management etc. In this way the so called Information Extraction (IE) becomes the dominating paradigm of NLP application to biomedical texts. The main IE idea is to extract automatically important entities from free texts, with accuracy as high as possible, and to build software systems operating on these entities (skipping the non-analysed text fragments). NLP in general is viewed as a rather complex Artificial Intelligence task so IE is proposed as a technology at the middle between keyword search and deep text analysis; it focuses on surface linguistic phenomena that can be recognised without deep inference. It is expected that IE progress would enable radical improvements in the clinical decision support, biomedical research and the healthcare in general [2].

This chapter is structured as follows. Section 2 briefly reviews related work and presents state-of-the-art figures about IE accuracy for various medical entities. Section 3 discusses the particularities of Bulgarian hospital discharge letters and overviews the IE components that have been developed during the last 3 years. Section 4 sketches potential applications of the IE prototypes and their extensions. Section 5 contains the conclusion.

2 Related Work

We consider quite briefly some state-of-the-art IE results for English clinical texts. The IE performance is measured by *precision* (the percentage of correctly extracted entities among all recognized entities in the test set), *recall* (the percentage of correctly extracted entities among all available entities in the test set) and the harmonic mean *f-score*=2**Precision***Recall*/(*Precision*+*Recall*). Below we refer to these performance indicators in order to present the background results and the context where we develop our IE prototypes for Bulgarian language.

Recent systems for extraction of drug treatment achieve accuracy higher than 90%: f-score: for instance, 91.40% for drug name and 94.91% for dosage [3]; or f-score 89.9% for drug name and 93.6% for dosage [4]. The system MedEx extracts medication events with f-score 93.2% for drug names, 94.5% for dosage, 93.9% for route, and 96% for frequency [5].

The automatic assignment of ICD codes to diagnoses achieves 89.08% accuracy [6]. The three top systems in the coding competition presented in [6] processed the

negation, hypernyms and synonyms in some way and exploited the semantic network of UMLS [7].

Patient status data are also recognized with high precision and recall. This is due to the fact that the IE systems can be carefully trained to identify relatively small sets of predefined words. For instance, the patient smoking status is classified into 5 categories by selecting sentences which contain the relevant information with f-score 92.64% [8].

Research on temporal IE from clinical texts is a relatively recent activity in medical informatics [9]. Temporal IE systems perform reasoning on temporal clinical data for therapeutic assessments; summarize data from temporal clinical databases, and model uncertainty in clinical knowledge and data [10]. Having in mind the complexity of the temporal IE tasks, some leading groups in biomedical NLP develop annotation schemes in order to unify the efforts for creation of training corpora which are tagged with temporal markers [11].

3 IE from Bulgarian Hospital Discharge Letters

Our experiments were performed on 6,200 discharge letters of patients with endocrine and metabolic diseases. The patient records have been anonymised in the USHATE hospital when exporting them to files for research purposes [12].

3.1 Material

Discharge letters in Bulgarian hospitals traditionally contain predefined sections due to centralised regulations which date back to the middle of the last century. Sometimes the sections might be merged, skipped (when they are empty), their headers and the default section sequence might be changed. However, practically the most important sections are always included in the discharge letters: the sections *Diagnoses*, *Anamnesis*, *Patient Status*, *Lab data & clinical tests* and *Debate* are available in 100% of the letters in a training corpus of 1,300 EHR; the section *Past diseases* is available in 88.52% of the letters in the same corpus while the *Family Medical History* and *Allergies & risk factors* are included correspondingly in 52,22% and 43,56% of the EHRs. Using a preliminary prepared list of section headers (about 80 keywords and phrases), most available sections can be automatically recognised with accuracy 99,99%. This enables a precise splitting of the record texts into subsections. We note that structured discharge letters are relatively rare in the healthcare practice worldwide as this requires a long and stable centralized administrative tradition in the preparation of medical documentation.

Another interesting fact concerns the wordforms used in the Bulgarian discharge letters. Table 1 shows some statistics about occurrences of Bulgarian and Latin words and terms. It turns out that 37% of all words in the discharge letters are 'unknown' as

they are Latin. About one half of the Bulgarian words are general lexica while many specific medical terms are not available in the usual (electronic) dictionaries. These figures illustrate the claims that high-quality NLP requires much effort for the development of extensive linguistic resources. Our only electronic resource with Bulgarian medical terms is the International Classification of Diseases (ICD-10) which contains 10970 terms.

Table 1. Distribution of words in the texts of 6,200 discharge letters

Terms and words	Wordforms	Basic words	Basic words / total
Bulgarian terms	601233	12009	63%
Latin with Latin alphabet	18926	560	3%
Latin terms transliterated to Cyrillic alphabet	179589	6465	34%
Total	**799748**	**19034**	100%

The free texts of patient records in our training corpus contain telegraphic phrasal descriptions and incomplete sentences, which excludes the application of deeper syntactic analysis and traditional NLP techniques in general. Therefore, in our IE experiments, we extract some important patient-related entities only and integrate them into more complex scenarios for decision making and structuring patient history.

3.2 Methods

We have developed several extracting modules for various clinical entities in two research projects. Figure 1 illustrates the types of patient-related entities that are extracted at present: diagnoses, drugs, entities from the Anamnesis section, patient status attributes as well as values of clinical test and lab data.

In the project PSIP (Patient Safety through Intelligent Procedures in medication, 2010-2011, PSIP FP7 ICT eHealth: http://www.psip-project.eu) our team extracted *(i)* drugs, *(ii)* diagnoses and *(iii)* values of clinical tests from USHATE hospital EHRs in order to fill in a PSIP-compliant repository and to validate the PSIP rules for Adverse Drug Events (see the bottom interface (d) on Fig 1). Most data items were available in the Hospital Information System (HIS) of USHATE and these items were sent to the PSIP repository directly by the HIS. But USHATE is a specialised hospital which treats endocrine diseases and their complications; thus drugs for accompanying and chronic diseases are often brought in by the patient and taken without records in the HIS. In this way it turned out that a considerable amount of the necessary data is presented only in the free text of the hospital discharge letters. Therefore we have implemented the following IE prototypes which process discharge letters that are split into sections [13]:

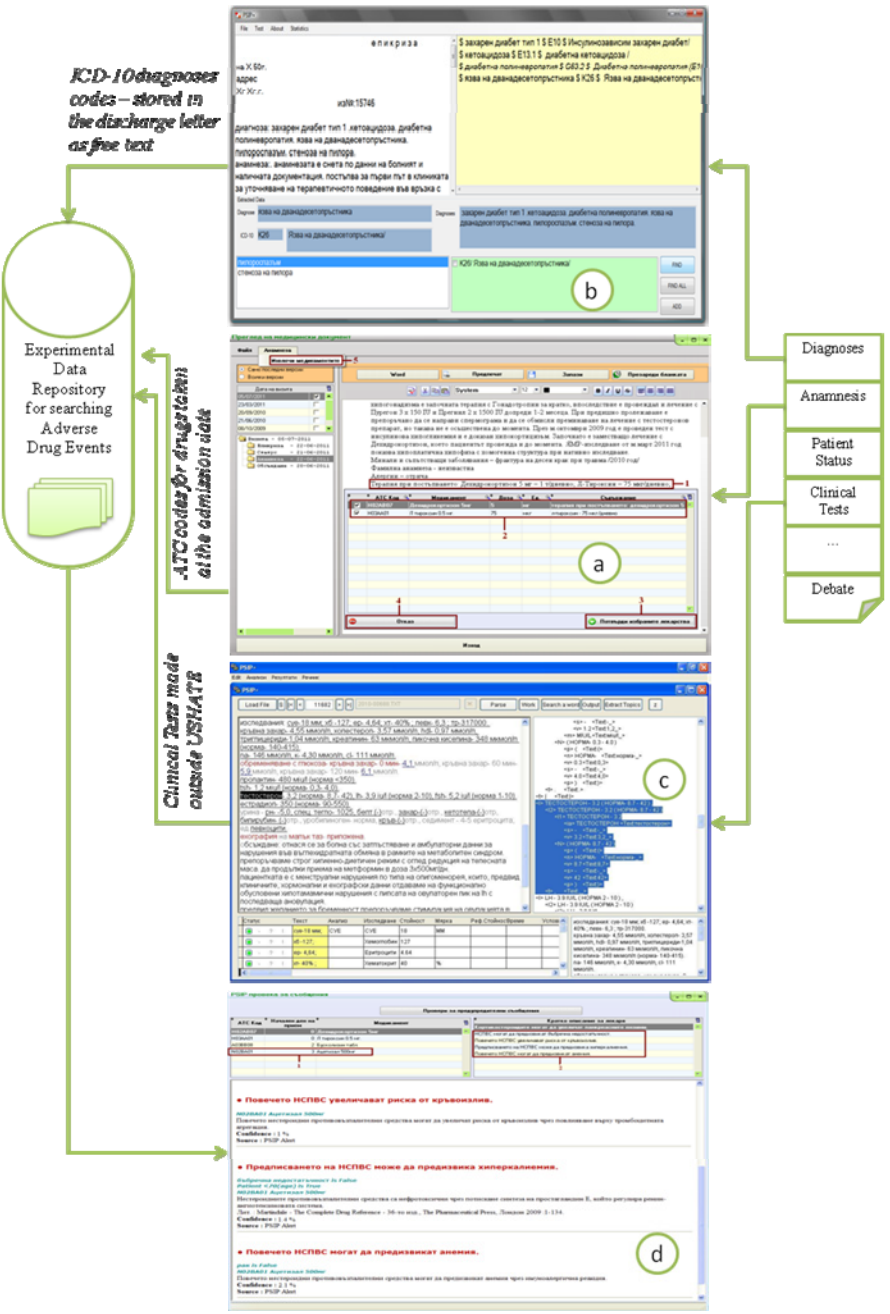


Fig. 1. Extraction of important patient-related entities for supporting clinical decision making and structuring patient histories

(a) Extractor of treatment events, based on ATC[1] codes: the extraction works in several steps, starting by preprocessing which identifies phrases containing drug names, dosages, mode of admission and frequency within their contexts. Some 160,892 occurrences of drug names were automatically recognised in the texts of 6,200 EHRs. The extractor assigns the corresponding ATC code to each medication event. It takes into account negative statements, some elliptical constructions, typical conjunctive phrases, and makes simple inferences concerning temporal constraints. The extraction accuracy (f-score) for drug names is 98.42% and for dose - 93.85% [14]. In general the EHR texts might discuss past, present and future medication events but it is important to allocate the treatment which is relevant for the hospitalisation period. The extractor achieved f-score of 90,17% for the recognition in the *Anamnesis* texts of 355 drugs, taken by the patients during the period of hospitalisation, which are not prescribed via the Hospital Pharmacy [15]. This extractor was online integrated in the HIS during the PSIP validation in USHATE, see interface (a) on Fig. 1.

(b) Extractor for assignment of ICD-10 codes of diagnoses. Bulgarian hospitals are reimbursed by the National Insurance Fund via the 'clinical pathways' scheme. When a patient is hospitalised, medical experts might select from the HIS menu one diagnosis which is sufficient for the association of the desired clinical pathway to the respective patient. Thus most diseases are not formally diagnosed and they are entered in the EHR section *Diagnoses* as free text. We have developed an extractor which matches the phrases, listed in the *Diagnoses* section, to disease names as given in ICD-10 nomenclature [16], see interface (b) on Fig. 1. This extractor processes Bulgarian and Latin terms including Latin terms transliterated to Cyrillic. The major challenge in automatic encoding is that the correspondence between the EHR diagnoses and ICD-10 names/codes is not a one to one correspondence. There is a variety of linguistic expressions which might refer to the same diagnosis and one expression might refer to many diagnoses. For 6,200 EHRs, some 26,826 phrases were found in the *Diagnoses* sections; the extractor assigns correctly ICD-10 codes to 84,5% of them.

(c) Extractor of values of clinical tests and lab data. When a patient is examined in USHATE, the lab data are entered to the EHR via the HIS. However there are a number of tests made outside the hospital and their results need to be extracted from the free text of *Lab data & clinical tests* sections (see interface (c) on Fig. 1.). This extractor works with precision of 98,2%.

The research project EVTIMA[2] (Efficient Search of Conceptual Patterns with Application in Medical Informatics) deals with IE and structuring of patient descriptions in order to build an internal conceptual representation which enables quick search of repeating patterns in diabetes development. Initially we considered the status descriptions as they are narrated in the *Patient Status* sections [17]. Skin status is extracted with f-score 83.33%, neck status – with f-score 91.67%, thyroid gland status – with f-score 92.59%, and limb status – with f-score 89.01%. Age descriptions are extracted

[1] Anatomical Therapeutic Chemical (ATC) Classification System, `http://atc.thedrugsinfo.com/`

[2] Funded by the National Science Fund in Jan. 2009-June 2012, see `http://www.lml.bas.bg/evtima`

from anonymised EHRs with f-score 89.44%. Status extraction needs explicitly-declared domain knowledge which supports the IE algorithms by providing constraints and inference mechanisms. A panel capturing the structured patient status is shown on Fig. 2 interface (a) on the top.

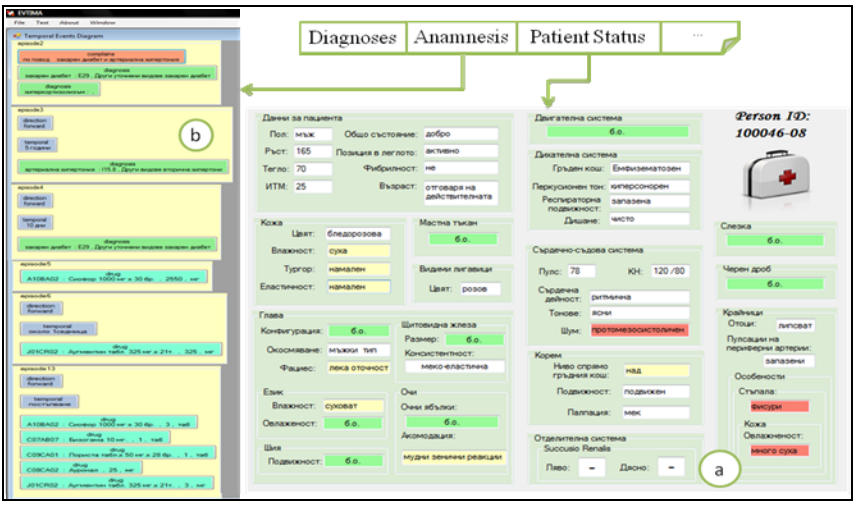


Fig. 2. Extraction of structured patient history

Our recent research is focused on automatic construction of temporal event sequences from the *Anamnesis* sections. We have developed a conceptual model of episodes that happen or occur to the patient; the description of these episodes is summarised in the *Anamnesis* section. When a diabetic patient is hospitallised, the medical experts briefly document the timeline of diabetes progression and the major events in the patient history. Starting with the moment of assigning the principal diagnose and following the complications, they summarise the drugs and procedures that were applied along the years. Episodes have *beginning* and *end* (sometimes without clearly defined time moments) and can be recognised after identification of about 80 types of temporal markers (prepositional and adverbial phrases). About 83.36% of the temporal markers refer to explicitly specified moments of time and can be seen as *absolute* references [18]. The remaining markers express *relative* or *undetermined* references. The markers are identified with precision 87% and recall 68%. The direction of time for the episode events: backwards or forward (with respect to certain moment orienting the episode) is recognised with precision 74.4%. The events happening within the episodes include *(i)* drug admission, *(ii)* diagnosing a disease or complication and *(iii)* changes of the patient condition or status. The extractors for drugs and diagnoses are integrated into the episode structuring. Identifying diabetes symptoms and conditions in free text is uneasy task as no 'canonical forms' exist in any dictionary for the related phrases and paraphrases. Therefore the currently developed extractor achieves lower precision. For instance, blood sugar level and body weight change are identified with f-score 60-96% in the separate processing phases [19]. Fig. 2 (b) on the left displays temporal splitting of patient history into episodes.

4 Potential Applications of the IE Prototypes

The results presented in section 3 were obtained in research projects but they explicate some specific particularities of the contemporary medico-administrative practice in Bulgaria. Having analysed the texts of only 6,200 discharge letters, we see for instance that:

(i) *Diagnoses are not fully encoded in the HIS*. There are 9,321 diagnoses registered in the HIS for the 6,200 EHRs in the test corpus while our extractor found in the texts 22,667 phrases to which ICD-10 codes were assigned;

(ii) *Drugs are not fully encoded in the hospital pharmacy.* According to the HIS entries, there are in average 1,9 drugs per patient, while our drug extractor found in average 5,9 drug names related to the period of hospitalization. The patients in USHATE took some 355 drugs which they brought to the hospital privately (while the hospital pharmacy of USHATE operated with 1182 drugs during the period of our experiments);

(iii) *Many clinical tests are done outside the hospital*, e.g. the hormonal ones, and their results are manually retyped in the corresponding discharge letters; thus the test results can be analysed only by human readers.

These examples show that the NLP extractors, when run over statistically-significant data excerpts, would enable important observations that are impossible at present; for instance the diagnose and drug extractors would enable to quickly construct diagnostically-related homogeneous groups for patients who need to be treated in a similar way. Definition of such groups would facilitate the optimisation of the drug treatment and reimbursement of costs. In this way NLP supports secondary use of the EHR data since it enables quick production of large data resources that might be extracted from millions of patient-related texts.

Regarding the potentially erroneous NLP results, we note that human data processing is not perfect as well. For instance the manual encoding of ICD-10 diagnoses might be erroneous too; [20] reports about 48-49% errors of human coders during the first year of their practice which decrease to 7% errors when the coding experts gain experience. These figures show that the results of our extractor, which assigns ICD-10 codes to diagnoses with accuracy of 84.5%, are within the margins of the usual correctness that is achieved in complex medical tasks as diagnose encoding.

Another direction of potential NLP application is to support the patients when they access their EHR data. The European Commission recommends that the patients should be able to inspect (part of) the content of their EHRs. Having in mind the Bulgarian tradition to use Latin terms both in Latin and Cyrillic alphabets, the first possible application is to ensure automatic translation of the diagnose terms to Bulgarian disease names (we have shown that there are 37% Latin terms in the EHR texts). In this way NLP might provide the 'normalisation' of the text and facilitate the active participation of the patients in their treatment. Last but not least NLP might support also the patient inclusion in the monitoring of his/her treatment by generation of alerts, explanations and recommendations.

5 Conclusion

The technologies for automatic text analysis are developing and deliver gradually improving results which encourage plans for practical NLP application. Despite the fact that much effort is needed to attach complex linguistic phenomena like temporality, negation, conditionality, reference resolution as well as semantic and pragmatic interpretation, we find running projects for biomedical text processing in all countries with advanced information societies. Basic sets of medical terminology are supported in hundreds of natural languages by the World Health Organisation which facilitates the development of electronic dictionaries suitable for NLP. The growing amount of clinical narratives is another stimulating factor that implies the increasing interest in automatic text processing.

Our experience shows that the IE technology is sufficiently mature to be integrated in practical software applications as a tool for delivery of large amounts of extracted entities. The application niches need to be carefully selected and the IE tools should be properly integrated. We are focused on diabetes as a major chronic disease consuming significant budget by the Health Insurance Fund. Our present results show the potential of IE to provide data to healthcare managers and decision makers concerning optimisation of diabetes treatment and reimbursement. Current research was set for Bulgarian Language, but similar methods were successfully used also for other languages in PSIP project.

Acknowledgments. The research tasks leading to these results have received funding from the EC's FP7 ICT under grant agreement n° 216130 PSIP (Patient Safety through Intelligent Procedures in Medication) as well as from the Bulgarian National Science Fund under grant agreement n° DO 02-292 EVTIMA (Efficient Search of Conceptual Patters with Application in Medical Informatics).

References

1. Meystre, S., Savova, G., Kipper-Schuler, K., Hurdle, J.F.: Extracting Information from Textual Documents in the EHR: A Review of Recent Research. In: Geissbuhler, A., Kulikowski, C. (eds.) IMIA Yearbook of Medical Informatics, pp. 138–154 (2008)
2. Demner-Fushman, D., Chapman, W., McDonald, C.: What can NLP do for Clinical Decision Support? J. of Biomedical Informatics 42(5), 760–772 (2009)
3. Patrick, J., Li, M.: A Cascade Approach to Extracting Medication Events. In: Proc. Australian Language Technology Workshop (ALTA), pp. 99–103 (2009)
4. Halgrim, S., Xia, F., Solti, I., Cadag, E., Uzuner, Ö.: Extracting Medication Information from Discharge Summaries. In: Louhi 2010, Proc. of the NAACL HLT 2010 Second Louhi Workshop on Text and Data Mining of Health Documents, pp. 61–67 (2010)
5. Xu, H., Stenner, S., Doan, S., Johnson, K., Waitman, L., Denny, J.: MedEx: a Medication Information Extraction System for Clinical Narratives. J. Am. Med. Informatics Assoc. (17), 19–24 (2010)

6. Pestian, J., Brew, C., Matykiewicz, P., Hovermale, D.J., Johnson, N., Cohen, K.B., et al.: A Shared Task Involving Multi-label Classification of Clinical Free Text. In: ACL 2007 Workshop on Biological, Translational, and Clinical Language Processing (BioNLP 2007), Prague, pp. 36–40 (2007)
7. UMLS, the Unified Medical Language System, http://www.nlm.nih.gov/research/umls
8. Savova, G., Ogren, P., Duffy, P., Buntrock, J., Chute, C.: Mayo Clinic NLP System for Patient Smoking Status Identification. J. Am. Med. Inform. Assoc. 15, 25–28 (2008)
9. Zhou, L., Hripcsak, G.: Temporal Reasoning with Medical Data - a Review with Emphasis on Medical NLP. J. Biom. Informatics 40(2), 183–202 (2007)
10. Adlassnig, K.-P., Combi, C., Das, A., Keravnou, E., Pozzi, G.: Temporal Representation and Reasoning in Medicine: Research Directions and Challenges. AI in Medicine 38(2), 101–113 (2006)
11. Savova, G., Bethard, S., Styler, W., Martin, J., Palmer, M., Masanz, J., Ward, W.: Towards Temporal Relation Discovery from the Clinical Narrative. In: Proc. AMIA Annual Symposium 2009, pp. 568–572 (2009)
12. Nikolova, I., Dimitrov, H., Tcharaktchiev, D.: Ethics and Security in Text Mining of Patient Records in Bulgarian: the EVTIMA Solution. In: ACM Proceedings of CompSysTech (2010)
13. Tcharaktchiev, D., Angelova, G., Boytcheva, S., Angelov, Z., Zacharieva, S.: Completion of Structured Patient Descriptions by Semantic Mining. In: Koutkias, V., Niès, J., Jensen, S., Maglaveras, N., Beuscart, R. (eds.) Studies in Health Technology and Informatics, vol. 166, pp. 260–269. IOS Press (2011)
14. Boytcheva, S.: Shallow Medication Extraction from Hospital Patient Records. In: Koutkias, V., Nies, J., Jensen, S., Maglaveras, N., Beuscart, R. (eds.) Studies in Health Technology and Informatics, vol. 166, pp. 119–128. IOS Press (2011)
15. Boytcheva, S., Tcharaktchiev, D., Angelova, G.: Contextualization in Automatic Extraction of Drugs from Hospital Patient Records. In: Moen, A., et al. (eds.) Proc. of MIE-2011, the 23rd Int. Conf. of EFMI, Studies in Health Technology and Informatics, Norway, vol. 169, pp. 527–531. IOS Press (August 2011)
16. Boytcheva, S.: Automatic Matching of ICD-10 Codes to Diagnoses in Discharge Letters. In: Proc. of Biomedical NLP Workshop, Satellite Event of Int. Conf. RANLP 2011, pp. 19–26 (2011), http://aclweb.org/anthology-new/W/W11/W11-42.pdf
17. Boytcheva, S., Nikolova, I., Paskaleva, E., Angelova, G., Tcharaktchiev, D., Dimitrova, N.: Obtaining Status Descriptions via Automatic Analysis of Hospital Patient Records. Special Issue on Semantic IT of Informatica, Int. J. of Computing and Informatics (Slovenia) 34(4), 269–278 (2010); Fomichov, V. (ed.)
18. Boytcheva, S., Angelova, G., Nikolova, I.: Automatic Analysis of Patient History Episodes in Bulgarian Hospital Discharge Letters. In: Proc. Demonstrations at the EACL 2012, pp. 77–81. ACL, France (2012), http://www.aclweb.org/anthology/E12-2016
19. Nikolova, I.: Unified Extraction of Health Condition Descriptions. In: Proc. of the NAACL HLT 2012 Student Research Workshop, pp. 23–28. ACL, Montreal (2012), http://www.aclweb.org/anthology/N12-2005
20. Ivanov, L., Ganova-Yolovska, M., Konstantinov, B.: Quality of Coding and Reliability of Medical Information for Distribution of Financial Resources into Diagnostically-related Groups. Social Medicine 4, 32–34 (1999) (in Bulgarian)

A Terminology Indexing Based on Heuristics Using Linguistic Resources for Medical Textual Corpus

Ali Benafia[1,3], Ramdane Maamri[2,3], and Zaidi Sahnoun[2,3]

[1] Université El Hadj Lakhdar de Batna 05000 Batna – Algérie
[2] Département d'Informatique Campus Nouvelle ville A. M 25000 Constantine
[3] Laboratoire LIRE,Université de Constantine 25000 Constantine – Algérie
{ali_bnfa,rmaamri,sahnounz}@yahoo.fr

Abstract. The term extraction is an important step in building a resource of indexing and many strong tools are available for many languages. This complex process, which identifies candidate terms may become indexes for annotations, is often subject to the problem of lack of relevance of calculated terms. As a result, extractor terms must be strong to handle the errors and suggest better results, without encumbering the user with too many proposed index. In this respect, we are suggesting a new indexing approach based on a hybrid of terminologies extraction using a filter by removing terms and operates upon corpus of medical texts.

Keywords: Term extraction, medical terminology, Mesh, ADM, syntactic patterns, n-grams.

1 Introduction

The development of the Internet leads to a proliferation of medical information. This medical information is increasingly being written, and available in electronic form. The Internet and websites offer a medical text corpus having a gigantic size. However, this information is often not exploited for the sake of heterogeneous character for this corpus. In medicine, we have medical nomenclatures and thesauri, which are somehow a representation of the most important medical concepts. The famous thesaurus UMLS [17] is a "meta-thesaurus" containing, inter alia, the MeSH [5]. There is also a thesaurus extracted from the ADM [16] which contains about 200,000 concepts (essentially symptoms, syndromes and diseases). This knowledge base contains a dictionary that covers most of the vocabulary medical. In the medical domain, find the most relevant medical information is not an easy task for the user thus requiring the use of a classification specialized and hierarchical, in which one can navigate in order to have the most appropriate response. The indexing result includes a set of concepts detected in the texts. This procedure causes a loss of information, the result of indexing is a simpler model to automatically process that a free language text.

The aim of our work is to use the previous resources to annotate medical texts from the terms (single words, compound words ...).

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 213–224.
DOI: 10.1007/978-3-319-00029-9_19

2 Problematic and General Description of Our Approach

Our problematic is this: on the base of a training corpus, is it possible to learn to system, the way to index medical texts from linguistic resources (Syntex, TreeTagger) and medical resources (Mesh, ADM).To respond to this problematic, we will, at first, segment each text to index into phrases and split each phrase recursively into triplets (syntagm, verb, syntagm) so that both syntagms obtained become nominal (for each triplet). We extract from these syntagms, the relevant terms which become the descriptors for the texts. These multi-terms chosen for the indexation make it possible to designate semantic entities or concepts better than single words, and offer a better representation of the semantic content of a text.

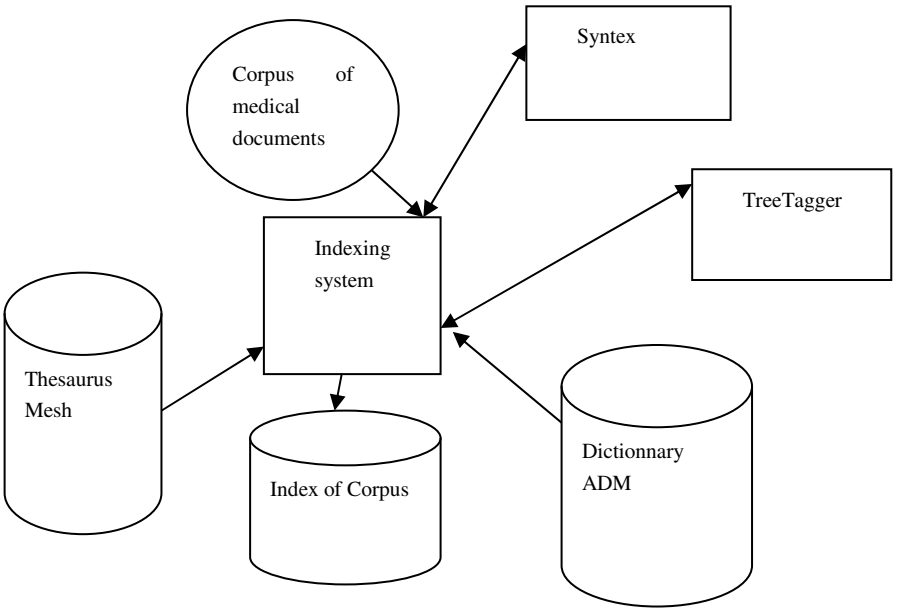


Fig. 1. System architecture

The system presented in Fig.1 integrates the linguistic platforms (Syntex and TreeTagger). The indexer uses two medical resources that are known in the literature (ADM and Mesh).

3 Related Works

Actually, there are two main approaches used for the textual indexing:

- **The semantic indexing** which is based on techniques for contextual disambiguation of words in documents and queries. The idea is that the meaning of a word is completely determined by the other words in the same context. Some authors as

[18] use terminological resources as UMLS1, MeSH2 and SNOMED3 and apply as issue the linguistic approach based on the lexical and syntactic analysis of the discourse structure. In [3] is proposes an indexation using the nucleus of document. This is a set of concepts weighted by their representativeness in the documents and linked together by similarity measures. In [20] is considered a single ontology sufficient to index and to find the informational content of objects. MedicoPort [9] uses relations of UMLS for query expansion. The mass of resources and ontologies to treat make the task very difficult. In addition, these resources and ontologies evolve over time which makes this procedure very tedious

- **The conceptual indexing** which is based on concepts taken from ontologies and taxonomies to index the documents unlike simple word lists. It uses the Terminological and Ontological Resource as hinge of indexing tool.

Khan in [14] proposes an algorithm for attaching the words of text to concepts of the ontology. He establishes the semantic distance between concepts. In [2] is proposed a method in which the terms of a text are attached to concepts of WordNet and is used the notion of semantic similarity between concepts. A similar approach based on CP-Nets (Conditional Preferences Networks) is proposed in [7]. The CP-net formalism is used both for the graphical representation of flexible queries expressing qualitative preferences and secondly for the flexible evaluation of the relevance of documentation. The use of relations between terms is a promising issue for better performance of Information Retrieval.

We opted for the semantic indexing approach because its activity could improve slightly the accuracy of information retrieval tool. On the other side, the conceptual indexing approach is currently difficult to index text at the conceptual level since such indexation generally requires large knowledge bases which make it harder.

4 Pretreatments

4.1 Standardization

Standardization is necessary for a good linguistic and statistical continuation of our treatments. The processing chain corresponding to the normalization is defined as follows:

Fig. 2. Processing chain for the normalization of phrases

The recognition of compound words and their constituents is essential in many applications of medical text analysis , it is necessary to recognize expressions such as 'maux de tête ‘, ‘appareil genital male ‘Syndrome d'immunodéficience acquise’...

For the identification of compound words, it is better to recognize the compound words in texts by consulting the dictionary Delacf [21] often very large in size rather than using an algorithm to find the form of a lemmatized compound word from its form arrowed and operative on the morphological categories and inflectional variables of the simple components [26]. Then we use the dictionary Mesh which is composed of two dictionaries: Delaf for simple words and delacf for compound words. We found in [4] a heuristic which analyses the terms of the text by group "n-grams" and whenever the term is a compound word we mark it, remove it from the analysed sentence, we move to following the term "n-grams" and we update the current sentence and so on until there is no further term to deal with. In this method we privilege the short term (composed of 2, 3, 4 or 5 words) because the majority of medical terms are not long enough in the French language.

4.2 Labeling

To facilitate the extraction of the terms in the sentence to be examined, you need a morpho-syntactic analysis and a syntactic analysis of surface. Considering the tools available on the web for French and the performances of these, we decided to use TreeTagger [13] for morpho-syntactic annotations and Syntex [8] for annotation of the syntactic dependency relations. The processing chain [4] for this type of analysis is:

Legend standardized → **TreeTagger with 2 pass** → Legend labeled → **Syntex** → Legend labeled and structured

Fig. 3. Processing chain for morpho-syntactic labeling

4.3 Formatting of Texts

The purpose of formatting of phrases into noun syntagms is neither used for comprehension purposes nor for translation purposes, but to isolate separately noun syntagms from phrase analyzed. Then, we have proposed a representation model [4] of the phrases in form triplets (nominal syntagm, verb, nominal syntagm). As the filters operate on nominal syntagms, it is then necessary to transform automatically each phrase of text in canonical format expressed in our model. The method adopted for this kind of formatting consists in analyzing each phrase and to locate from its syntactic structure the verb. The position of the verb in the phrase allows splitting this phrase into two parts, the left part and the right part. We proceed by cutting each phrase around the verb, this procedure gives two syntagms. We continue recursively this process while the obtained syntagms are verbal syntagms (contain always a verb).

Text labelled and structured → **Formatage** →Text in form triplets

Fig. 4. Representation of legend in the defined model

5 Filtering of Candidate Terms

5.1 Description of Filters

Our approach for the selection of the relevant terms is described by a processing chain using four filters that are presented in the following:

Text in canonical form → **filtering in cascade** → Candidates terms → **Pruning** → Candidates of relevant terms

Fig. 5. Processing chain for incremental filtering of terms

- 1st filter: *structural dependence between terms*
Each term t in form *n*-gram is written as a series of *n* words $x_1, x_2, x_3, \ldots x_n$ with x_i full or empty word . The measures of dependence terms are defined according to three criteria:

* Criterion of presence: all x_i words must be present except the empty words.
* Criterion of order: all x_i words must appear in the same order x_i without taking into consideration of the empty words.
* Criterion of distance: all x_i words must be within short distance between them.

for each criterion a score measure is evaluated and let *IM-score (T)* , *P-score (T)* and *F-score (T)* be the three scores. The final score which represents the bond strength of dependency between the words m_i of a term *T* is obtained by a linear combination of these three scores:

$$FLD\ (T) = \alpha_1 \times IM\text{-}score\ (T) + \alpha_2 \times P\text{-}score\ (T) + \alpha_3 \times F\text{-}score\ (T) \qquad (1)$$

The coefficients α_1, α_2 and α_3 represent the weightings attributable to criteria defined previously. The heuristic for calculation of the measures of dependence terms is given in [4].

- 2nd filter: *taking into account the syntactic patterns*
This filter is based on the principle of syntactic patterns. We have a base of syntactic patterns which contains a collection of syntactic configurations extracted from an automatic learning of syntactic patterns carried out on a corpus from training. A syntactic pattern is a rule about the order of concatenation of grammatical categories formed for a noun phrase. For our implementation we selected a combination of 128 patterns extracted from a corpus by a manual learning procedure done on a sample of examples of texts. This filter allows to examine each candidate and to check if it is among the existing configurations in the database of patterns built. In the affirmative case, it will be retained otherwise it will be removed.

- 3rd filter: *full words and empty words*
It is possible to build the list of the empty words based on Zipf's law [26].This law has found its application in [23] and considers that the more the word is frequent the more it is shorter. The list of the empty words is built on the basis of the lengths and the

frequencies of the words. We consider empty words, words that are both short and frequent. Starting from this principle, we propose for this filter to combine this law to label firstly the analyzed term and to treat the term in question. The labelling for this filter uses two sets of labels: full or informative word and empty or uninformative word. The criteria used in this filter are "word length" or "word frequency" and "word shortly" or "word scarcity".

- **4th filter:** *semantic potential of terms*

This filter assigns a score for each term regardless of the context to which it relates. The searched score is proportional to the size of the terms, to their distribution and to their type in the corpus, hence:

$$Qinf\,(t_i) = \sum_1^n Q(x_j) \times FLD\,(t_i) \times \alpha_i \quad (2)$$

where: n is the number of words that constitute the term t_i.
$Q\,(x_j)$ - the quantity of information corresponding to the word x_j.
$Q_{inf}\,(t_i)$ – the weighted sum of information quantities for the term t_i.
$FLD\,(t_i)$ - the measure of dependence between the different words of the term t_i (cf. 1st filter).
α_j – the weighting coefficient related to the type of the term t_i.

Coefficients α_j are chosen from the corpus according to the nature of the term and its distribution in the corpus. The value of this coefficient changes according to whether the term is a "single word", a "compound word", a "named entity", a "compound term" consisting of several words or even composed of associated words. The medical named entities, for their part represent a conceptual description which refers to an object whose linguistic representation is often unique; we can describe them as more specific terms, and therefore discriminate. In this case, if a named entity is identified in a noun phrase as a term to be weighted, it will be affected a value sufficient which privilege to be selected as linguistic index term. We can notice from what emerges from this approach of scoring that the longest terms are always preferred. Such a privilege is unquestionably justified by the fact that the longest term remains the most discerning when it is about the choice of descriptors of indexing.

5.2 Conclusion About Filters

The first filter F1 is placed at the beginning in the filtering process. Its role is to assess the dependencies between different combinations of words built from n-gram models and to select the set of terms having a sufficient score. The following two filters (F2 and F3) allow to study the words non-informative (empty words) of linguistic and statistical point of view. The superfluous words will be removed from the terms. The filter F4 is presented at the end of this estate of filtering and is used to calculate the potential of information for each selected term.

6 Pruning of Terms

The pruning procedure applies to a collection lists, generated during the filtering phase, which it receives as input. At output, it sends back a set of classes of terms which constitutes the final index.

We note *L1*, *L2*, *Ln*, a collection of *n* lists containing all the terms selected after filtering. The partitions set of classes of terms which constitutes our index is built progressively and at each iteration we take a term *j* of the class *i* (*i=1,n*) and we look for the closest class in the direction of the semantic term *j* in this set of partitions; if we find it we insert the term *j* in the corresponding class and we update the centroid of conceptual vectors for this class and we repeat the process with the following term in the current list until all lists are treated; in the case where the term *j* does not match any class, we create a new class for this term and we take this class as centroid, the centroid of the term. We note that the construction of this set of partitions of classes is carried out without knowing the number of starting classes in other words it is a clustering of index terms. In each built class we find all terms having the best measure of semantic proximity with the others. We found in [4] all details of applied heuristic.

7 Structure of the Index

To minimize the silence in the query requests, the terms used for indexing medical texts are grouped into classes. Each class is described by the reference term (first level), the context of term (second level), the set of term synonyms (third level), the set of close terms (fourth level) and the term flections (last level). For the context, we are defined 28 semantic classes (in French), we quote:

Diseases, medicaments, micro-organisms, signs and symptoms, examinations, diagnostic and therapeutic actions, microorganisms, anatomical locations, social context, métier and profession…

For example, the class *metier and profession* contains the terms as “internal medicine”, “cardiology service”, “chief doctor”, the class *social context* regroups the terms as “old topic”, “middle-aged adults” and the class *examinations* includes the terms as *angiographic plan*, *c*omputed Radiography *and Digital Radiography.*

The index structure and the information it contains ensures the linguistic variability of terms (lexical variability, semantic, etc.).

8 Experiments and Results

8.1 Corpus and Method Evaluation

Our experimental protocol covers a collection of articles describing texts on diseases written in French. These texts are searched from the Directory CISMeF nd Research engine MedHunt of medical sites located on the url:
http://www.atoute.org/recherche_informations_medicales/bibliographie.htm

For the construction of our corpus and to estimate the smoothing parameters of our indexing approach, we chose some short texts (and long texts). For our study we took 50 medical articles (short+long) for first essay and 100 articles (short+long) for the second essay.

The main problem of the evaluation of the indexing, underlined by [15] is that it doesn't exist an ideal indexation "reference" as a criterion to validate or refute an indexing system, whether human or automatic. Thus, we will use a comparison of the indexing to a "gold standard", a particular indexing taken as reference, developed by an expert indexer. In this case, the texts of the corpus taken as test for our indexing approach are submitted to the expert in order to determine the descriptors .The validation of the indexing is thus carried out by the expert indexer according to his descriptors announced like reference. However, we must remember that the indexing is an open problem: for any document a single set of descriptors constituting an ideal indexing does not exist. On the contrary, several solutions are possible and acceptable. The use of an index reference as evaluation basis is still a plausible method to establish an evaluation of the calculated descriptors compared to those empirically determined. With this principle, we collect in what follows the various measures obtained according to different points of view. The use of an indexing of reference as evaluation bases remains all the same a plausible method and much used to establish and evaluate an indexing system.

For the measure of the quality of our indexing approach, we place ourselves within the framework of an evaluation based on a reference: we have two lists of term descriptor one of these two lists constitute a reference at which the other list (list resulting from our system) must be compared. The distribution of the terms of an index can be represented for a matrix known as confusion matrix.

Table 1. Matrix of confusions for the terms to compare with the terms of reference

Evaluated list	$\in$ reference	$\notin$ reference
Selected	Terms $\in$ to the list of reference and $\in$ to the list to be analyzed	Terms $\notin$ to the list of reference and $\in$ to the list to be analyzed
Not selected	Terms $\in$ to the list of reference and $\notin$ to the list to be analyzed	Terms $\notin$ to the list of reference and $\notin$ to the list to be analyzed

Being P, R and F_m respective measures of precision, recall and for F-measure, we have:

P= (Terms $\in$ to the list of reference and $\in$ to the list to be analyzed)/(Terms $\in$ to the list of reference and $\in$ to the list with analyzer + Terms $\in$ to the list of reference and $\notin$ to the list to be analyzed).

R= (Terms $\in$ to the list of reference and $\in$ to the list to be analyzed)/(Terms $\in$ to the list of reference and $\in$ to the list with analyzer+ Terms $\notin$ to the list of reference and $\in$ to the list to be analyzed).

$F_m = (P.R)/(\alpha.R+P\ (1-\alpha))$ with $0 \leq \alpha \leq 1$; α represents the weight allotted to the precision.

8.2 Measures and Interpretation

The obtained results for Table 2 are not satisfactory because the texts of articles are supposed to be «brute texts» what makes the task of the filters noisy and unfounded since no preliminary analysis (compound words, ambiguities, anaphoras, hooks…) is supposed to be made for this case of evaluation. For the short articles, the results are better compared to long articles because they use a simple medical vocabulary. We also notice that the progressive elimination of candidate's terms by filtering allows to eliminate the noise in the selection of index terms.

Table 2. Calculation of the precisions, recalls and F-measures for a corpus containing 50 articles (first line) and 100 articles (second line) with noisy texts

	Short article				*Long article*		
	P(%)	R(%)	F_m(%)		P(%)	R(%)	F_m(%)
F1+F2+F3+F4	45.21	46.69	45.93		38.01	38.46	38.23
F1+F2+F3+F4	49.96	50.12	50.03		39.87	42.59	37.26

Table 3 contains results on the indexing obtained by our approach by assuming now that the articles have been standardized and labelled manually by the medical expert. The texts submitted to our indexer are treated and containing no errors. The obtained results with manual pretreatment of texts are best compared to the previous case. We obtain for this type of evaluation how to estimate the true character of filters regarding to their efficiency under the ideal conditions .i.e. when the corpus is pre-treated in advance by the medical expert. The success rate for the case of short texts is quite acceptable since it reached the limit of 81% with a corpus containing 100 articles and around 76% for a corpus containing 50 which proves that the contribution of the filtering process is very interesting. The weakness of our approach always remains the long texts indexing since the threshold remains more or less plausible (around 62%) for the first case (n=50) and about71% for the second case (n=100). There is a difference of performance between these two corpus considered that is due to the size of corpus. When this size of corpus is large, our approach becomes efficient. In sum, we can say that the extraction of terms for process filtering improves the quality of results and depends upon manual pre-treatments. This is why the competence of an expert is also important.

Table 3. Calculation of the precisions, recalls and F-measures for a corpus containing 50 articles (first line) and 100 articles (second line) with manual pretreatment of texts

	Short article				*Long article*		
	P(%)	R(%)	F_m(%)		P(%)	R(%)	F_m(%)
F1+F2+F3+F4	76.04	75.65	75.84		62.29	60.56	61.41
F1+F2+F3+F4	81.40	79.31	80.34		70.82	70.29	70.55

Table 4 contains the results obtained from a real test of our approach. We take into account the set of preliminary treatments (standardization, labelling) which precede filters. We note that there is no large difference of results for the sizes of these two corpuses. We reached an acceptable threshold of 75 %(resp. 68%) for the short articles and a threshold of 62% (resp. 59) for the long articles. The results obtained in this case are less reliable compared to previous results (automatic procedure VS manual procedure) this is why we recommend to integrate to our implementation a learning system, operating on the basis of corpus for the phase of pre-treatment.

Table 4. Calculation of the precisions, recalls and F-measures for a corpus containing 50 articles (first line) and 100 articles (second line) with automatic pretreatment of texts

	Short article				*Long article*		
	P(%)	R(%)	F_m(%)		P(%)	R(%)	F_m(%)
F1+F2+F3+F4	68.34	66.05	67.18		59.57	56.94	58.23
F1+F2+F3+F4	73.89	74.63	74.25		63.14	60.82	61.95

The previous results do not hold the step of terms pruning. The processing which constitutes the chain of medical text analysis operates sentence by sentence and extracts the candidate's terms for the indexing. This list of terms can be redundant, including the ones with the others.

The fusion of all the lists resulting from filtering process and their pruning improves the results in terms of friability and precision. The results obtained in Table 5 are almost the same which means that this procedure can correct the abnormalities caused by the filtering process before applying the pruning procedure.

Table 5. Calculation of the precisions, recalls and F-measures for a corpus containing 50 articles (first line) and 100 articles (second line) taking in account the pruning process of terms

	Short article				*Long article*		
	P(%)	R(%)	F_m(%)		P(%)	R(%)	F_m(%)
F1+F2+F3+F4	65.00	64.24	64.61		60.27	60.15	60.20
F1+F2+F3+F4	67.49	66.96	67.22		62.38	64.27	63.31

9 Conclusion and Outlook

We proposed in this work a new paradigm of indexing based on the concept of medical terms. This paradigm supposes the translation of nominal syntagms extracted from the corpus to index in terms of syntagmatic index having an integrated semantic. The used method is hybrid and uses linguistic and medical tools. Most filters presented are inexpensive in calculating time and easy to be implemented; moreover their heuristics allow eliminating the noise gradually in the list of candidate's terms. This is in itself an induction for the improvement of the precision rate. The results of filtering process

remain despite everything relatively good. They depend on the size of the corpus and on the preliminary work carried out on the texts (recognition of the compound words, elimination of ambiguities…) and especially on the nature of the article. From a quantitative point of view, our system answers all needs expressed because the response time of its execution remains acceptable and reasonable. As conclusion and in view of results obtained, we can say that the uses of this indexing are closer to human reasoning. In addition as prospective classification, several aspects must be mentioned:

- Addition of a treatment on the recognition of the medical named entities because these last are considered fixed and thus candidates to be selected like index terms.
- Integration of a treatment on the location of collocations.
- For the long legends, it is highly recommended to define a heuristic on the thematic salience which will allow us to process the rest of the treatments without ambiguities.
- Definition of a system of learning for the construction of the conceptual vectors for our corpus although this problem remains difficult to solve. We note by the way that the procedure of construction carried out for the moment is manual and required a long work.

Acknowledgment. We would like to especially thank Dr. H. Berredjem and Dr. S. Djarallah for their medical skills and their experience.

References

1. Aussenac, N.G., Jacques, M.P.: Designing and Evaluating Patterns for Relation Acquisition from Texts with CAMÉLÉON. Dans: Terminology, J.B.P Amsterdam, Numéro spécial Pattern-Based approaches to Semantic Relations, Vol. 14 N. 1, pp. 45-73 (2008)
2. Baziz, M., Boughanem, M., Aussenac-Gilles, N.: The Use of Ontology for Semantic Representation of Documents (regular paper). In: Semantic Web and Information Retrieval Workshop at SIGIR (SWIR), Sheeld, UK, pp. 38–45 (2004)
3. Baziz, M., Boughanem, M., Aussenac-Gilles, N.: Conceptual Indexing Based on Document Content Representation. In: Crestani, F., Ruthven, I. (eds.) CoLIS 2005. LNCS, vol. 3507, pp. 171–186. Springer, Heidelberg (2005)
4. Benafia, A., Maamri, R., Sahnoun, Z.: An Indexing Approach Based on a Hybrid Model of Terminology-extraction using a Filtering by Elimination Terms. J. of Advances in Information Technology 3(4) (2012)
5. Bethesda, Maryland, MeSH: Medical Subject Headings National Library of Medicine (1986)
6. Biskri, I., Meunier, J., Joyal, S.: L'extraction des Termes Complexes: Une Approche Modulaire Semi-Automatique. In: JADT 2004: 7es Journées Internationales d'Analyse Statistique des Données Textuelles (2004)
7. Boubekeur-Amirouche, F.: Contribution à la Définition de Modèles de Recherche d'Information Flexibles Basés sur les CP-Nets (2008)
8. Bourigault, D., Fabre, C., Frérot, C., Jacques, M.P., Ozdowska, S.: Syntex, Analyseur Syntaxique de Corpus. In: TALN 2005, Dourdan (2005)
9. Can, A.B., Baykal, N.: MedicoPort: A Medical Search Engine for All. Computer Methods and Programs in Biomedicine 86(1), 73–86 (2007)

10. Evans, D.A., Ginther-Webster, K., Hart, M., Lefferts, R.G., Monarch, I.A.: Automatic Indexing Using NLP and First Order Thesauri. In: RIAO 1991 Recherche d'Informations Assistée par Ordinateur, pp. 624–643 (1991)
11. Grimault, F.: Terres d'Innovation Photographie, Indexer et Légender (2009)
12. Jacquemin, C.: Variation Terminologique: Reconnaissance et Acquisition Automatiques de Termes et de Leurs Variantes en Corpus, Mémoire d'habilitation à Diriger des Recherches en Informatique Fondamentale, Université de Nantes (1997)
13. Jacques, M.-P.: Que: la Valse des Etiquettes. Actes de la Conférence Traitement Automatique des Langues Naturelles. In: TALN 2005, Dourdan, France (2005)
14. Khan, L.: Ontology-based Information Selection. Ph.D. Thesis, University of South California (2000)
15. Lancaster, F.W.: Indexing and Abstracting in Theory and Practice. University of Illinois, Champaign (1991)
16. Lenoir, P., Michel, J.-R., Frangeul, C., Chales, G.: Réalisation, Développement et Maintenance de la Base de Données A.D.M. 6 (1), 51–56 (1981)
17. Lindberg, D., Humphreys, B., McCray, A.: The Unified Medical Language System. National Library of Medicine, Bethesda (1993)
18. Liu, H., Hu, Z.-Z., Zhang, J., Wu, C.: BioThesaurus: a Web-based Thesaurus of Protein and Gene Names. Bioinformatics (1), 103–105 (2006)
19. Schwab, D., Lian Tze, L., Lafourcade, M.: TALN 2007, Toulouse, Les Vecteurs Conceptuels un Outil Complémentaire aux Réseaux Lexicaux (2007)
20. Secon, S., Veale, T., Hayes, J.: An Intrinsic Information Content Metric for Semantic Similarity in WordNet. In: Proc. ECAI 2004, 16th European Conf. on Artificial Intelligence, pp. 1089–1090 (2004)
21. Siberztein, M.: Le Dictionnaire Electronique des Mots Composés en Langue Française (1990)
22. Siberztein, M.: Traitement des Expressions Figées avec Syntex. In: Analyse Lexicale et Syntaxique: le Système Intex, Lingvisticae Investigationes XXII, pp. 425–449. John Benjamins Publishing Compagny, Amsterdam (1999)
23. Vergne, J.: Une Méthode Indépendante des Langues Pour Indexer les Documents de l'Lnternet par Extraction de Termes de Structure Contrôlée. In: Actes de la Conférence Internationale sur le Document Électronique (CIDE 8), Beyrouth, Liban (2005)
24. Zipf, G.K.: The Psycho-biology of Language. An Introduction to Dynamic Philology. The M.I.T. Press, Cambridge (1968); Second paperback printing (First Ed. 1935)
25. Yarowsky, D.: Unsupervised Word Sense Disambiguation Rivaling Supervised Methods. In: Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics, pp. 189–196 (1995)
26. Voorhees, E.M.: Using WordNet to Disambiguate Word Senses for Text Retrieval. In: Proceedings of the ACM-SIGIR 1993, pp. 171–180. ACM Press, New York (1993)

Architecture for Medical Image Processing

Rumen Mironov and Roumen Kountchev

Technical University of Sofia, Department of Radio Communications and Video Technologies, Boul. Kl. Ohridsky 8, Sofia 1000, Bulgaria
{rmironov,rkountch}@tu-sofia.bg

Abstract. In this paper, software architecture for medical image processing, analysis and archiving is presented. On the basis of the considered architecture a new task-oriented medical image processing system, which allows imitating of the human visual system, is developed. The basic functions include input/output of halftone images, pre- and post-processing, filtration, compression, enhancement, 2D linear transforms, pseudo-color transforms, analysis and interpolations. Using the system features, various image processing tasks are semantically described in the experimental part. The main advantages of the proposed architecture are the use of adaptive algorithms for processing of medical images, tailored to their specific features.

Keywords: Image Processing, Medical Imaging, Software Architectures, Knowledge Base Systems.

1 Introduction

Medical imaging has been undergoing a revolution in the past decade with the advent of faster, more accurate, and less invasive devices. This has driven the need for corresponding software development which in turn has provided a major impetus for new algorithms in signal and image processing, based on rigorous mathematical foundations that can be integrated into complete therapy delivery systems. Such systems support the more effective delivery of many image-guided procedures such as biopsy, minimally invasive surgery, and radiation therapy [1].

Many different imaging techniques have been developed and are in clinical use. Because they are based on different physical principles [2], these techniques can be more or less suited to a particular organ or pathology. In practice they are complementary, because they offer different point of views for the one and the same medical problems. In medical imaging, these different imaging techniques are called modalities. Anatomical modalities provide insight into the anatomical morphology. They include radiography, ultrasonography or ultrasound (US), computed tomography (CT), magnetic resonance imagery (MRI) There are several derived modalities that sometimes appear under a different name, such as magnetic resonance angiography (MRA, from MRI), digital subtraction angiography (DSA, from X-rays), computed

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 225–234.
DOI: 10.1007/978-3-319-00029-9_20

tomography angiography (CTA, from CT) etc. Functional **modalities**, on the other hand, depict the metabolism of the underlying tissues or organs. They comprise the three nuclear medicine modalities, namely, scintigraphy, single photon emission computed tomography (SPECT) and positron emission tomography (PET) as well as functional magnetic resonance imagery (fMRI). This list is not exhaustive, as new techniques are being added every few years as well [13]. Almost all images are now acquired digitally and integrated in a computerized picture archiving and communication system (PACS) [1, 2, 3].

Medical images typically suffer from one or more of the following imperfections:

- low resolution (in the spatial or/and spectral domains);
- high level of noise;
- low contrast;
- geometric deformations;
- presence of imaging artifacts.

These imperfections can be inherent to the imaging modality (e.g., X-rays offer low contrast for soft tissues, ultrasound produces very noisy images, and metallic implants will cause imaging artifacts in MRI) or the result of a deliberate trade-off during acquisition.

In connection with the fast growing of advanced calculation techniques the software and hardware systems for medical image processing on the base of personal computers are standing topical. They are built modularly and include additional frame grabbers for input and output of digital images, modules for compression, transform, filtration, analysis and visualization [4]. In the related literature some image processing systems are described [5, 6, 7, 8, 9]. All of them include basic software modules for image processing and hardware accelerators for enhancement of the computation operations speed. They are developed on the basis of the following common requirements [10, 11]:

- working on standard computer platforms with different operation systems;
- working with different input/output devices;
- working in interactive mode;
- using of internal system program language;
- using of advanced image processing algorithms;
- large possibilities for upgrades and improvements.

The image processing algorithms in these systems are included as standard libraries and each can be used separately by the user or supervisor. It is a function-oriented approach, which isn't suitable to imitate the human visual system performance [12].

In this paper, new architecture for medical image processing, analysis and archiving is presented. This architecture is task-oriented and shows the complex image processing flows, in the way the natural human visual system is operating.

2 Architecture of the Medical Image Processing System

In Fig.1, the block-diagram of the developed architecture for medical image processing system is shown. This architecture is organized as a set of the following image processing units, based on well-known algorithms [4, 12, 13, 14, 15] and new developed by the authors algorithms [16, 17, 18, 19]:

- *Image Loading Unit* (*ILoadU*). This module comprises functions for loading and converting of compressed and uncompressed basic image file formats (TIFF, BMP, JPEG) and the specially developed image file format, presented in detail in [20]. *ILoadU* also contains functions for image decryption and loading from special image medical database;
- *Image Inputting Unit* (*IInptU*). This module comprises functions for input of medical images and information (text, photos, tables, graphics) from scanners, photo, cameras and other devices;
- *Image Saving Unit* (*ISaveU*). This module comprises functions for saving of compressed and uncompressed basic image file formats (TIFF, BMP, JPEG) and the image file format presented in [11]. *ISaveU* also contains functions for image encryption and saving in a special image database;
- *Image Pre-processing Unit* (*IPrepU*). This module comprises the following functions for image pre-processing – functions for arithmetic and geometric operations, table operations, contrast enhancements, histogram operations, linear and nonlinear noise filtration. They are used for image quality improvement or converting in a form better suited for analysis by human or a machine.
- *Image Presentation Unit* (*IPresU*). This module comprises different image presentation models (positioning, quad-tree, pyramidal, structural) and functions for image converting;
- *Image Compression Unit* (*ICmprU*). This module comprises image compression functions, based on classical algorithms for run-length, Huffman, LZW, arithmetic, scalar and vector quantization, JPEG and a new algorithm for lossless compression;
- *Image Geometrical Transform Unit* (*IGeoTU*). This module comprises functions for translations, scaling, interpolations, rotations, affine and perspective transformations, spatial warping and geometrical resampling.
- *Image Linear Transform Unit* (*ILinTU*). This module comprises functions for 2D linear superpositions and convolutions, correlations, discrete cosine, sine, Hadamard, complex Walsh-Hadamard, Fourier and Karhunen-Loev transforms.
- *Image Pseudo-Color Transformation Unit* (*IPCTU*). This module comprises functions for pseudo-color transforms in the spatial and the frequency area and for image halftoning by adaptive and non-adaptive error diffusion.
- *Image Analysis Unit* (*IAnlyU*). This module comprises functions for image segmentation (contour, brightness, texture) morphological image processing (binary, gray), edge detection and feature extractions.

- *Image Post-Processing Unit* (*IPostU*). This module comprises functions for adaptive and non-adaptive, linear and non-linear filtration of gray level or binary images.
- *Image Visualization Unit* (*IVizuU*). This module comprises functions and drivers for visualization of document images.
- *Image Printing Unit* (*IPrntU*). This module comprises functions and drivers for printing of medical images on documents.

For the proper performance of the developed software system the following additional modules are also needed:

- *System Supervisor* (*SS*). This module comprises the basic functions for system control, diagnostic and interaction between the separated components and the operation system. In the system supervisor analysis and processing of system error messages and receiving of context help information are also achieved.
- *Graphical User Interface* (*GUI*). This module comprises graphical oriented system with menus, dialogs, windows, icons, buttons, fonts, etc., which can be used for dialog input/output of parameters or images, interactive processing and analysis, visualization of documents, graphical presentation of the results from the analyses.
- *Communication Module* (*CM*). This module comprises functions for connection between the System Supervisor and the external modules and the functions of the system interpreter. The usage of the Communication Module facilitates the building of processing tasks – also including new ones, which are not defined in the system. This is a way to use External Algorithms, established by the operator.
- *System Peripheral Module* (*SPM*). In this module the system peripheral drivers are situated.
- *Database Management System* (*DBMS*). In this subsystem the medical information represented by images and texts is archived [21, 22]. As a DBMS the standard MS SQL Server was used.
- *Knowledge Based System* (*KBS*). This subsystem is under development. Knowledge based system comprises artificial intelligent tools working in a narrow domain to provide intelligent decisions with justification. Knowledge is acquired and represented using various knowledge representation techniques rules, frames and scripts. The basic advantages offered by such system are the documentation of knowledge, the intelligent decision support, the self learning, reasoning and explanation.

One of the main research areas in medical decision support is the formulation of biomedical engineering principles based on rigorous mathematical foundations in order to develop general-purpose software methods that can be integrated into complete therapy delivery systems. Such systems support the more effective delivery of many image-guided procedures such as biopsy, minimally invasive surgery, and radiation therapy.

The authors had participated for many years in the development of various systems for image processing, some of which are given in the references [23, 24, 25, 26, 27]. The presented software architecture for medical image processing is synthesized on the basis of this experience and allows overcoming the disadvantages of various types of medical images.

There are different classifications of methods and algorithms for image processing [4, 12, 14]. The methods for converting the input image at the output can be divided into three groups:

- conversion from the input array of integer data into an array of the same type;
- conversion of the input data set in a vector or scalar (separation of the global attributes);
- conversion of a vector or a scalar into an array of integer data;

The most widespread algorithms are from the first group. These are implemented by the following classes of operations:

- Monadic Point Operations – the used input image $U = \{u(x, y)\}$ and the values of each pixel $w(x, y)$ from the output image W depend only on the value of the pixel with the same coordinates in the input image: $F\langle u(x, y)\rangle \Rightarrow \langle \overline{w}(x, y)\rangle$;
- Dyadic Point Operations (used for combining and pseudo coloring of images) – values of each pixel from the output image W depends on function F applied to the input images U and V: $F\langle u(x, y), v(x, y)\rangle \Rightarrow \langle \overline{w}(x, y)\rangle$.

The main image processing operators involve transformations of the type “array into array” that can be described mathematically as follows:

- inversion $\langle Z - u(x, y)\rangle \Rightarrow \langle \overline{w}(x, y)\rangle$;

- sum with a constant C $\langle \overline{w}(x, y)\rangle \Leftarrow \begin{vmatrix} 0 \leftarrow u(x,u) + C < 0; \\ u(x, y) + C \leftarrow 0 < u(x, y) + C \le Z; \\ Z \leftarrow u(x,u) + C > Z; \end{vmatrix}$

- multiplication with C $\langle \overline{w}(x, y)\rangle \Leftarrow \begin{vmatrix} 0 \leftarrow u(x,u).C < 0; \\ u(x, y).C \leftarrow 0 < u(x, y).C \le Z; \\ Z \leftarrow u(x, y).C > Z; \end{vmatrix}$

- log/exp processing $\langle \overline{w}(x, y)\rangle \Leftarrow \begin{vmatrix} 0 \leftarrow u(x,u) = 0; \\ Z.log\{u(x, y)\}/log(Z) \leftarrow 0 < u(x,u) \le Z; \end{vmatrix}$

- square root $\left\langle \overline{w}(x,y) \right\rangle \Leftarrow \left\langle [u(x,y).u(x,y)] / Z \right\rangle;$

- normalizing $\left\langle \overline{w}(x,y) \right\rangle \Leftarrow \left\langle |u(x,y) - Z/2| \right\rangle;$

- highlighting the range $\left\langle \overline{w}(x,y) \right\rangle \Leftarrow \begin{vmatrix} C_3 \leftarrow C_1 \le u(x,u) \le C_2; \\ u(x,u); \end{vmatrix}$

- summation $\left\langle \overline{w}(x,y) \right\rangle \Leftarrow \left\langle [u(x,y) + v(x,y)] / 2 \right\rangle;$

- subtraction $\left\langle \overline{w}(x,y) \right\rangle \Leftarrow \left\langle [u(x,y) - v(x,y)] \right\rangle;$

- multiplication $\left\langle \overline{w}(x,y) \right\rangle \Leftarrow \left\langle [u(x,y).v(x,y)] / Z \right\rangle;$

- maximum $\left\langle \overline{w}(x,y) \right\rangle \Leftarrow MAX \left\langle [u(x,y), v(x,y)] \right\rangle;$

- minimum $\left\langle \overline{w}(x,y) \right\rangle \Leftarrow MIN \left\langle [u(x,y), v(x,y)] \right\rangle;$

Other type of operations “array into array” comprises: histogram transforms, unitary transforms and two-dimensional digital filtrations. These basic image processing operators are used not only to obtain new allocation levels (both brightness distribution and the distribution of red, green and blue components of the image), but also in computer graphics.

The system supports two kinds of data objects: image data objects and image-related, non-image data objects. A system image data object is a multi-dimensional collection of pixels, including *horizontal* and *vertical* position, *brightness* and *color* or *spectral band* values. The system utilizes the following pixel data types: *Boolean*, *Unsigned Integer*, *Signed Integer*, *Real* and *Complex*. The precision and data storage format of pixel data is implementation dependent.

It supports several image related, non-image data objects. These include:

- *Chain*: an identifier of a sequence of operators;
- *Composite identifier*: an identifier of a structure of image arrays, lists and records;
- *Histogram:* a construction of the counts of pixels with some particular amplitude value;
- *Lookup table*: a structure that contains pairs of entries in which the first entry is an input value to be matched and the second is an output value;
- *Matrix*: a two-dimensional array of elements that is used in vector-space algebra operations;
- *Neighborhood array*: a multi-dimensional moving window associated with each pixel of an image;

- *Pixel record:* a sequence of across-band pixel values;
- *Region-of-interest:* a general mechanism for pixel-by-pixel processing selection;
- *Static array*: an identifier of the same dimension as the image to which it is related;
- *Value list*: a collection of pairs of elements in which the first element is a pixel coordinate and the second element is an image measurement.

3 Experimental Image Processing Tasks

The functionality of the human visual system can serve as a reliable guide for breaking up the complex image processing tasks. First, the optical system forms an image of the observed documents. A sensor converts this image into a form that is usable for digital processing with a computer system. The first processing step, denoted as low-level image processing, enhances, restores, or reconstructs the image formed. Further processing extracts features from the images that finally lead to the identification and classification of the objects in the images and can be saved in the documents database. In this way, the circle is closed, starting from documents that are converted into images and processed back to their recognition and description.

The software system is developed on Visual Studio 2010 with the use of .NET Framework and C++ language.

Using the developed system we can define various image processing tasks, which can be semantically described, as follows:

IInptU -> IPrepU -> IPresU -> IPostU -> ISaveU: Inputting an image from a scanner or a digital photo camera, enhancing the image quality by pre-processing, presentation in a matrix form, post-processing and saving in the database. This semantic chain can be used for preliminary record of uncompressed documents in the database.

IInptU -> IPrepU -> IPresU -> ICmprU -> ISaveU: Inputting an image from a scanner or a digital photo camera, enhancing the image quality by pre-processing, presentation in a matrix form, compression and saving in the database. This semantic chain can be used for preliminary record of compressed documents in the database.

ILoadU -> IPrepU -> IPresU -> IPostU -> IVizuU: Loading an image from a file or a database, enhancing the image quality by pre-processing, presentation in a matrix form, post-processing and visualization on the screen. This semantic chain can be used for visualization of saved documents in the database.

ILoadU -> IPrepU -> IPresU -> IPostU -> IPrntU: Loading an image from a file or a database, enhancing the image quality by pre-processing, presentation in a matrix form, post-processing and printing on a laser or ink-jet printer. This semantic chain can be used for printing of saved documents in the database.

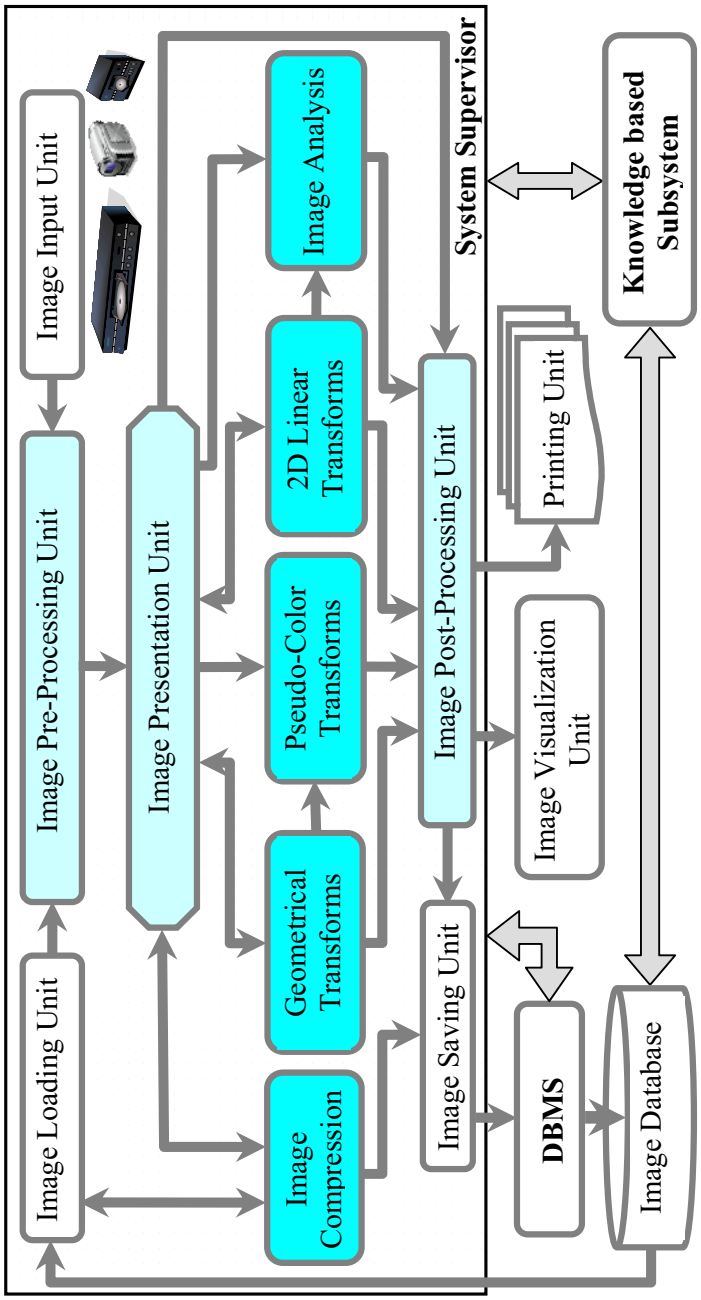


Fig. 1. Architecture of the Medical Image Processing System

4 Conclusions

The presented architecture, control modules, data objects and interfaces build a new task-oriented image processing system, able to imitate the natural human visual system. The so developed architecture has the following advantages:

- module design, based on the image processing tasks;
- possibilities for working in automatic and interactive mode;
- using of an internal system program language;
- using of the advanced image processing algorithms;
- large possibilities for upgrades and improvements.

The presented software system can be used in various areas of the medical image processing, analysis of biomedical images, biotechnologies, ecological monitoring, visual control in the industry, medicine, scientific research, etc.

Acknowledgement. This work was supported by the Joint Research Project Bulgaria-Romania (2010-2012): "Electronic Health Records for the Next Generation Medical Decision Support in Romanian and Bulgarian National Healthcare Systems", DNTS 02/09.

References

1. Angenent, S., Pichon, E., Tannenbaum, A.: Mathematical Methods in Medical Image Processing. Bulletin of the American Mathematical Society 43, 365–396 (2006)
2. Hendee, W., Ritenour, R.: Medical imaging physics, 4th edn. Wiley-Liss (2002)
3. Thonnat, M.: Knowledge–based Techniques for Image Processing and for Image Understanding. Journal de Physique, IV France EDP Science, Les Ulis 12, 189–236 (2002)
4. Pratt, W.K.: Digital Image Processing, 4th edn. John Wiley & Sons (2007)
5. Pandit, H., Shah, D.M.: Application of Digital Image Processing and Analysis in Healthcare Based on Medical Palmistry. In: International Conference on Intelligent Systems and Data Processing (ICISD), pp. 56–59 (2011)
6. Vaida, M.F., Todica, V.: Medical Image Processing using a Service Oriented Architecture. In: 1st International Conference on Advancements of Medicine and Health Care through Technology, MediTech2007, Cluj-Napoca, Romania, September 27-29 (2007)
7. Lin, Y.-D., Tsao, H.-V., Chong, F.-C.: An Image Processing Architecture to Enchance Image Contrast. Biomed. Eng. Appl. Basis Commun. 14, 215 (2002)
8. Lead Technologies Inc. LEADTOOLS Medical Image Processing SDK. WWW Site: `http://www.leadtools.com/sdk/image-processing/medical.html`
9. Yoo Terry, S.: Insight into Images: Principles and Practice for Segmentation, Registration, and Image Analysis. A K Peters/CRC Press (2004)
10. Maher, A., Sid-Ahmed: Image Processing: Theory, Algorithms, and Architectures. McGaw-Hill, Inc. (1995)
11. Chang, S.-K.: Principles of Pictorial Information Systems Design. Prentice-Hall (1989)
12. Gonzalez, R.C., Woods, R.E.: Digital Image Processing, 3rd edn. Pearson Prentice Hall (2008)

13. Ware, C.: Information Visualization – Perception for Design, 2nd edn. Morgan Kaufmann (2004)
14. Jähne, B.: Practical Handbook on Image Processing for Scientific and Technical Applications, 2nd edn. CRC Press (2004)
15. Myler, H.R., Weeks, A.R.: Computer Imaging Recipes in C. Prentice Hall, Englewood Cliffs (1993)
16. Mironov, R.: Algorithms for Local Adaptive Image Processing. In: XXXVII International Scientific Conference on Information, Communication and Energy Systems and Technologies. ICEST 2002, Niš, Yugoslavia, October 1-4, pp. 193–196 (2002)
17. Mironov, R., Kountchev, R.: Method for Watermarking of Medical Images Based on Fast Complex Hadamard Transform. In: Intern. Conf. on Communications, Electromagnetics and Medical Applications (CEMA 2010), Athens, Greece, October 7-9, pp. 58–61 (2010)
18. Kountchev, R., Mironov, R., Kountcheva, R.: Efficient Compression of Medical Images Based on Adaptive Histogram Modification. In: XLVI Intern. Scientific Conf. on Information, Communication and Energy Systems and Technologies (ICEST 2011), Serbia, Niš, June 29- July 1, vol. 1, pp. 13–16 (2011)
19. Mironov, R.: Adaptive Interpolation and Halftoning for Medical Images. In: International Workshop on Next Generation Intelligent Medical Decision Support Systems, MedDecSup 2011, Proc., Targu Mureş, Romania, September 18-19 (2011) (in press)
20. Kountchev, R., Todorov, V.: File Format organization for effective Still Image Transfer with IDP. In: 37th Intern. Scientific Conf. on Information, Communication and Energy Systems and Technologies, Proc., Nis, Yugoslavia, vol. 1, pp. 287–290 (2002)
21. Mironov, R.: Data Structures for Image Presentation. Proc. of National Conference with foreign participation. In: Development of Telecommunication Networks and Systems, TELECOM 1999, Varna, Bulgaria, pp. 547–553 (1999) (in Bulgarian)
22. Mironov, A., Mironov, R.: Using Databases for Presentation of Multimedia Information. In: Proc. of National Conference with foreign participation, Development of Telecommunication Networks and Systems, TELECOM 1999, Varna, Bulgaria, pp. 642–649 (1999) (in Bulgarian)
23. Kountchev, R., Mironov, R.: Program System for Metallographic Images Processing and Analysis. In: XXX Science Session: Communication, Electronic and Computer Systems 1995, Sofia (May 1995) (in Bulgarian)
24. Kountchev, R., Mironov, R.: System for Image Processing, Analyses and Recognition. Journal for Automatics and Informatics 4, 41–44 (1996) (in Bulgarian)
25. Mironov, R., Sirakov, N., Muge, F.: An Architecture of Virtual Multimedia Library. In: Proc. of V Ibero – American Symposium on Pattern Recognition, SIARP 2000, Lisbon, September 11-13, pp. 103–111 (2000)
26. Mironov, R., Kountchev, R.: System Architecture for Distance Learning of Deaf People. In: XLIV Intern. Scientific Conf. on Information, Communication and Energy Systems and Technologies (ICEST 2009), Bulgaria, pp. 191–194 (2009)
27. Mironov, R., Kountchev, R.: Architecture of Image Processing System for Documents. In: XLIV Intern. Scientific Conf. on Information, Communication and Energy Systems and Technologies (ICEST 2009), Bulgaria, pp. 195–198 (2009)

A New Generation of Biomedical Equipment: FPGA

Marius M. Balas

Aurel Vlaicu University, 2 Elena Dragoi, 310330 Arad, Romania
marius@drbalas.ro

Abstract. This chapter aims to broaden the bridge that covers the gap between the biomedical science community and engineers, by encouraging the developers and the users of biomedical equipment to apply at a large scale and to promote the Field-Programmable Gate Array technology. One provides a brief recall of this technology and of its key advantages: high electrical performances (great complexity, high speed, low energy consumption, etc.), extremely short time-to-market, high reliability even in field conditions, flexibility, portability, standardization, etc. A tutorial description of the FPGA development methodology is also provided. Some of the most successful FPGA applications in this field are mentioned: medical imaging, minimally-invasive surgery platforms, radio-frequency identification, etc.

Keywords: Biomedical equipment, healthcare, FPGA.

1 Introduction

A shocking title is drawing our attention when reading the second quarter 2009 issue of IEEE Circuits and Systems Magazine: "A Wake-Up Call for the Engineering and Biomedical Science Communities" [1]. The authors are trying to sensitize the readers on the gap that is drawn between the new achievements of electronics and information technology on one hand and the biomedical science community on the other hand. In order to fill this gap the authors envisaged a coherent action: the establishment of the *Life Science Systems and Applications Technical Committee* (LISSA) within the IEEE Circuits and Systems Society, supported by the *USA National Institutes of Health* (NIH). LISSA organizes several workshops each year. After each workshop a white paper is published by the IEEE Circuits and Systems Magazine in order to present the major challenges in various chosen theme areas.

Inspired and alerted by this message, we observed a dominant idea that is now leading towards a comprehensive modernization and improvement of the biomedical equipment: the large scale applying of *the Field-Programmable Gate Array* (FPGA) technology. The unification of the major part of the equipment used by a whole scientific-technical domain by means of FPGA is already experienced with very positive results by the military and the aerospace industries [2, 3], etc.

Since the very FPGA concept is fundamentally different of the classical digital systems, replacing the programmed operation with a programmed architecture and the bus oriented architectures with a parallel one, we can affirm that a new generation of equipment is born.

R. Kountchev & B. Iantovics (Eds.): *Adv. in Intell. Anal. of Med. Data & Decis.*, SCI 473, pp. 235–246.
DOI: 10.1007/978-3-319-00029-9_21

2 The FPGA Technology and Its Place in the Electronic Industry

FPGA is an electronic technology that is essentially implementing very large scale custom digital circuits by means of the software controlled reconfiguration of a large integrated array, composed of identical Configurable Logic Blocks (CLB). A CLB cell may contain different types of Look-Up-Tables (LUT), flip-flops and elementary logic gates. The FPGA routing system ensures the reprogrammable CLBs configuration that defines each application, by a wide network of horizontal and vertical channels that may be interconnected in any possible way by transistorized interconnecting matrices. In other words, FPGA is a field reprogrammable Very Large Scale Integrated Circuit (VLSI), merging some leading hardware and software technologies. A FPGA generic structure is shown in Fig. 1 [4].

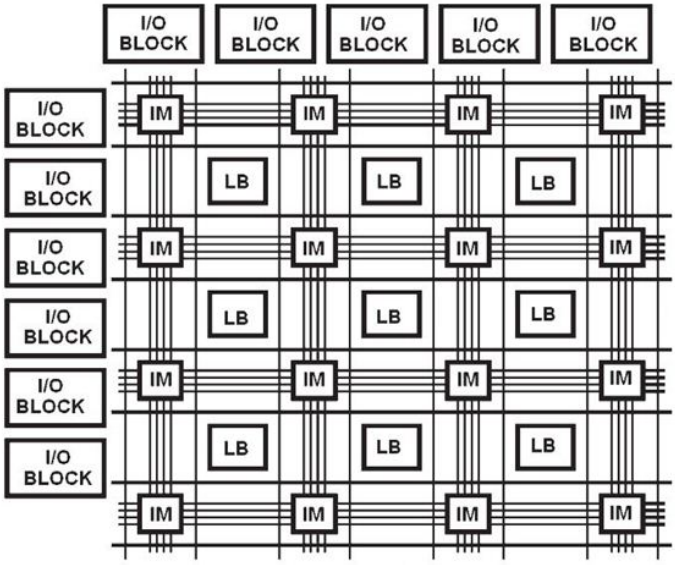


Fig. 1. FPGA generic structure, with logic blocks LB, wired channels, interconnecting matrices IM and input/output interface I/O

In the late 1980s the US Naval Surface Warfare Department initiated a research project that succeeded to build a computer out of 600,000 reprogrammable gates - the direct ancestor of the FPGA hardware. The FPGA software's origin is alike: although the first modern Hardware Description Language (HDL), Verilog, was introduced by Gateway Design Automation in 1985, the decisive leap occurred in 1987, when the US Department of Defense funded a research aiming to lead towards a unified language for description and simulation of the digital circuits, especially ASICs (Application-Specific Integrated Circuits). This approach generated VHDL (Very High Speed Integrated Circuit Hardware Description Language). Using VHDL all the military equipment, no matter its producers, could be treated in the same optimal manner (conception, design, testing, implementation, maintenance, upgrading, etc.). Verilog

and VHDL were used basically to document and simulate circuit-designs initially described in another forms, such as schematic files or even analytic functions. After that, a lot of synthesis tools were developed in order to compile the HDL source files into manufacturable gate/transistor-level netlist descriptions. A key contribution to the FPGA development belongs to IEEE that marked the concept's evolution by conceiving the initial version of VHDL (the 1076-1987 standard), followed by many other subsequent related standards.

The company that imposed in 1985 the first commercial FPGA - XC2064 is Xilinx (co-founders Ross Freeman and Bernard Vonderschmitt). XC2064 gathered for the first time programmable gates and programmable interconnects between gates. Xilinx is the tenure leader of the FPGA market, strongly involved into the aerospace industry [2, 3]. A turning point in the public perception of FPGAs happened in 1997, when Adrian Thompson merged the genetic algorithm technology and FPGA to create the first sound recognition device. Other popular FPGA producers are: Altera, Actel, Lattice Semiconductor, Silicon Blue Technologies, Achronix and QuickLogic.

It is useful for us to relate FPGA with the other leading electronic and IT technologies, in order to point its pros and contras if used in biomedical equipment.

Gate array is a fundamental concept for digital integrated circuits. Since the TTL (*Transistor Transistor Logic*) times, the integrated circuits were realized starting from gate arrays – NAND gates for instance, simply by executing the wired connections that were configuring the desired electric schematics. Wirings were initially mask-programmed at factory, mainly by photolithography or other ROM (*Read Only Memory*) technologies. Given the initial gate array, the design of a new integrated circuit comprised essentially only the design of the photolithographic mask of the wiring system. Replacing this rigid technology with the flexible software controlled configuration of the today's FPGA followed the next steps, always in connection with the developments of the ROM memories:

The *Programmable Array Logic* (PAL), sharing with PROM memories (*Programmable ROM*) the mask-programmable feature.

The *Field Programmable Array Logic* (FPAL) issued providing PALs with appropriate programmers. FPALs are one time programmable: users can write their custom program into the device only once;

The *Complex Programmable Logic* device (CPLD) were higher capacity PALs;

In respect of the parallelism with the memory technologies, FPGAs are *field reprogrammable* devices, similar to EEPROMS (*Erasable PROM*): *users can repeatedly write and erase the stored data referring to the configuration*. However FPGAs outreached by far the simple field reprogrammable devices; their reconfiguration circuits became extremely complex and performing.

Besides the PAL-CPLD connection, FPGAs is organically linked with ASICs. The ASICs are VLSIs customized for a particular use. Customization occurs by the appropriate design of the metal interconnect mask, there are no reprogrammable features. This ROM like technology was developed in terms of sheer complexity (and hence functionality) increasing the number of integrated gates to over 100 millions.

ASICS are meant to be customized for large scale applications or standard products, cell phones for instance. Their design and development are laborious and expensive, but their electrical performances are rewarding: fast operation, low energy consumption,

extremely high reliability, etc. The high end of the ASIC technology belongs to the full-custom ASIC design, which defines all the photolithographic layers of the device. The full-custom design minimizes the active area of the circuit and improves its functional parameters by minimal architectures and the ability to integrate analog components and pre-designed components, such as microprocessor cores.

However the amount of applications that can use exactly the same circuit in huge number is limited, so usual ASICs are realized by *gate array design*. This manufacturing method uses unconnected wafers containing all the diffused layers, produced by full-custom design, in extremely large quantities. These wafers are stocked and wait to be customized only when needed, by the interconnection layout. Thus the non-recurring engineering costs (research, development, design and test) are lower, as photolithographic masks are required only for the metal layers, and production cycles are much shorter. On the other hand this technique is absolutely rigid, the smallest design or implementation error compromises the whole batch of circuits.

This is why the first motivation of the PFGA developers was to create *rapid prototyping* ASICs, in order to avoid the wastage and to further reduce the time to market and the nonrecurring costs. FPGAs can implement any logical function that an ASIC could perform, so when developing an ASIC, the first mandatory stage is to realize its FPGA version. The FPGA prototypes can be easily reconfigured and tuned, until obtaining the desired performances and successfully completing all the testing specifications. Only the final versions of the FPGA netlists (all the connections of the circuits) will be automatically implemented into ASICs. As a matter of fact this technological itinerary holds only for large scale applications, very often the medium or small scale applications remain as FPGAs.

A key issue when comparing FPGA to other electronic technologies is the FPGA versus μC-DSP relationship. The FPGA that is stemming from the digital integrated circuits achieved today a strong position into *the embedded systems* field. By contrast with a general-purpose computer, such as a PC, which is meant to be flexible and to meet a wide range of end-user needs, an embedded system is a computer system designed to do one or a few dedicated and/or real-time specific functions, very often in difficult environment and operation conditions. Each time when we need powerful algorithms, complex data and signal processing or intelligent decisions at the level of physical applications, embedded systems become inevitable. The embedded systems were developed by the *microprocessor* μP – *microcontroller* μC – *digital signal processor* DSP connection. A legitimate question for the biomedical equipment developers is *why to replace the μP-μC oriented generation with a new FPGA oriented one*.

Comparing the two solutions one can observe that computer/μC binary architectures are build around a central spine bone: the data/address/control bus (see Fig. 2). All program instructions use the bus one by one, in order to accomplish the four steps that form the instruction cycle: fetch, decode, execute and writeback of the resulting data. Except the decoding, all the other steps demand complex bus traffic and maneuvers, which creates a fundamental limitation of the computer's speed.

The inconvenient of the bus centered architecture is that since only one instruction is executed at a time, the entire CPU must wait for that instruction to complete before proceeding to the next instruction. Caches memories, pipelines and different parallel

architectures such as multi-core systems are addressing with more or less success this inconvenient. For the embedded systems frame a suitable solution represent the parallel bus that carries data words in parallel, on multiple wires. If we compare Fig. 1 and Fig. 2 we may appreciate FPGAs as fully distributed and parallel, totally opposed to the bus centered architectures. Comparing to the *programmed operation* of a μC, the FPGA presents a *programmed architecture*. The FPGAs are replacing the virtual computing of the logic functions that follow the synchronized succession of the instructions and all their steps, with *a real wired circuit*, able to operate continuously, with no complicated synchronization constraints. This is fundamentally accelerating the FPGA performance. The same algorithm may be performed hundred times faster by a FPGA circuit compared to usual bus oriented architecture devices.

However the FPGA shift is paid by rather laborious programming, which implies a specific vision. A relevant synthesis of the digital computer vs. FPGA dichotomy was provided by Ed Klingman in [6]: "When gates were precious entities and tools 100% proprietary, it made ultimate sense to arrange these limited gates into universally used objects, such as CPU registers, ALUs, instruction decoders (*Arithmetic Logic Unit*), and address decoders. You would then provide a set of instructions that linked and relinked these elements, so that, for example, two CPU register outputs could be connected (via buses) to an ALU input, then the ALU output connected to a destination register, and then the ALU input connected (via buses) to a specific memory address, and the ALU output connected to a different register, and so on and so on

When gates are no longer precious but are commodities, the fixed elements approach no longer makes as much sense. The monstrous development in languages, tools, I/O devices, standards, and so on will keep CPU development and implementation alive for decades, if not centuries, but the economics are now and trending more so in the FPGA direction."

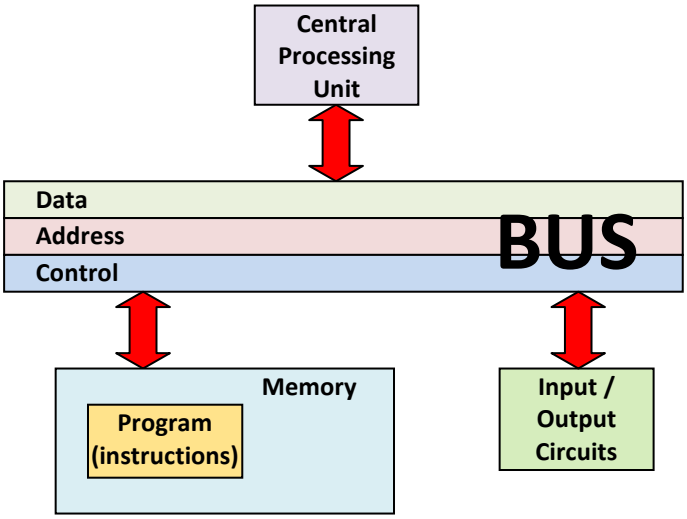


Fig. 2. The generic computer/microcontroller architecture

3 FPGA Design Tools – A Brief Tutorial

The FPGA's intrinsic complexity, pointed by Ed Klingman in 2010, is addressed in 2012 by design tools that become more and more friendly. The last edition of Xilinx ISE Design Suite for instance, is justifying its name by gathering a bunch of previously stand alone software tools, covering all the operations demanded by the design and the implementation of the FPGA application. The main stand-alone tools embedded into ISE are: ISim (behavioral simulation), PlanAhead (mapping and post-layout design analysis), iMPACT (creates PROM files) and ChipScope (post analysis of the design).

The procedural itinerary shown in Fig. 3 is fully automatic, a simple application can be achieved basically just by direct commands (Generate Post-Place&Route for instance), although the experienced designers have full access to the software resources (such as: Manually Place & Route). The main stages of the FPGA application development following the previous itinerary are illustrated in figures [7, 8].

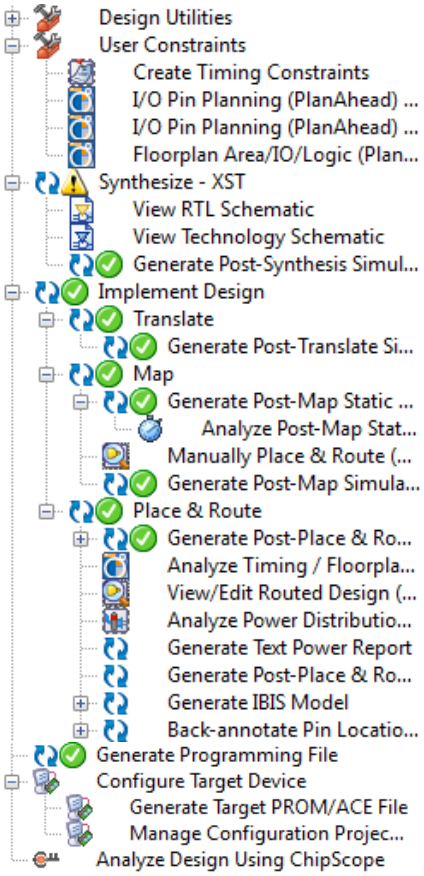
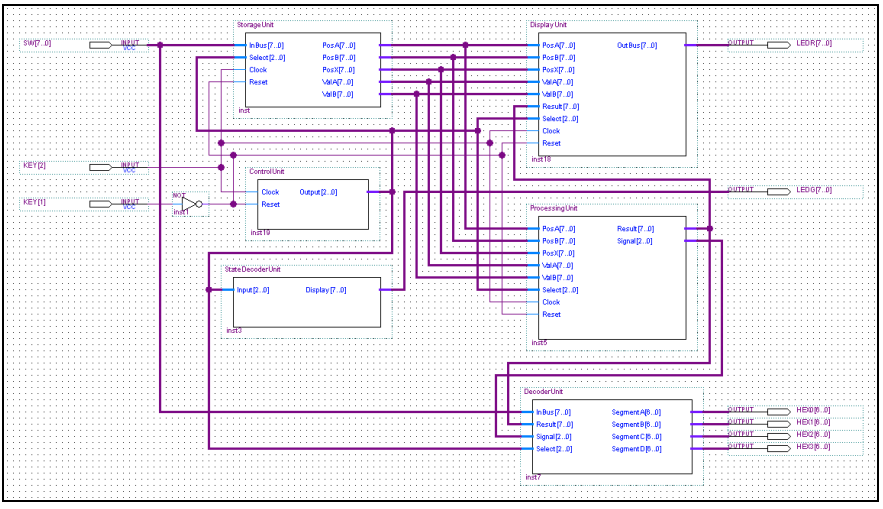


Fig. 3. The Xilinx ISE Design Suite procedural itinerary

```
15  LIBRARY ieee;
16  USE ieee.std_logic_1164.ALL;
17  USE ieee.numeric_std.ALL;
18  LIBRARY UNISIM;
19  USE UNISIM.Vcomponents.ALL;
20  ENTITY RSlatch_RSlatch_sch_tb IS
21  END RSlatch_RSlatch_sch_tb;
22  ARCHITECTURE behavioral OF RSlatch_RSlatch_sch_tb IS
23
24     COMPONENT RSlatch
25     PORT( XLXN_1   :  OUT   STD_LOGIC;
26           XLXN_2   :  OUT   STD_LOGIC;
27           XLXN_3   :  IN STD_LOGIC;
28           XLXN_4   :  IN STD_LOGIC);
29     END COMPONENT;
30
31     SIGNAL XLXN_1   :  STD_LOGIC;
32     SIGNAL XLXN_2   :  STD_LOGIC;
33     SIGNAL XLXN_3   :  STD_LOGIC;
34     SIGNAL XLXN_4   :  STD_LOGIC;
35
36  BEGIN
37
38     UUT: RSlatch PORT MAP(
39        XLXN_1 => XLXN_1,
40        XLXN_2 => XLXN_2,
41        XLXN_3 => XLXN_3,
42        XLXN_4 => XLXN_4
43     );
```

a) HDL code (VHDL in this case)



b) Schematic

Fig. 4. The synthesis of the HDL code/schematic source of the application

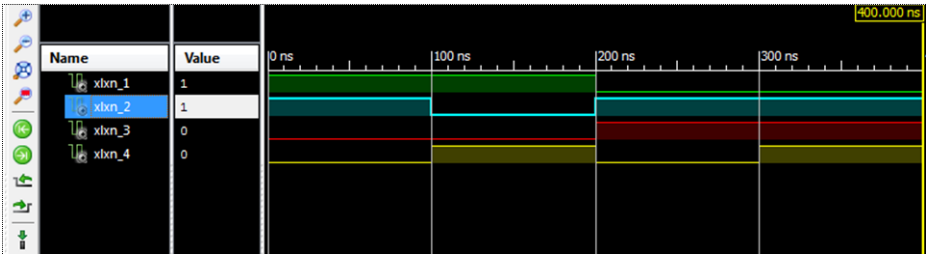


Fig. 5. The behavioral simulation (ISim)

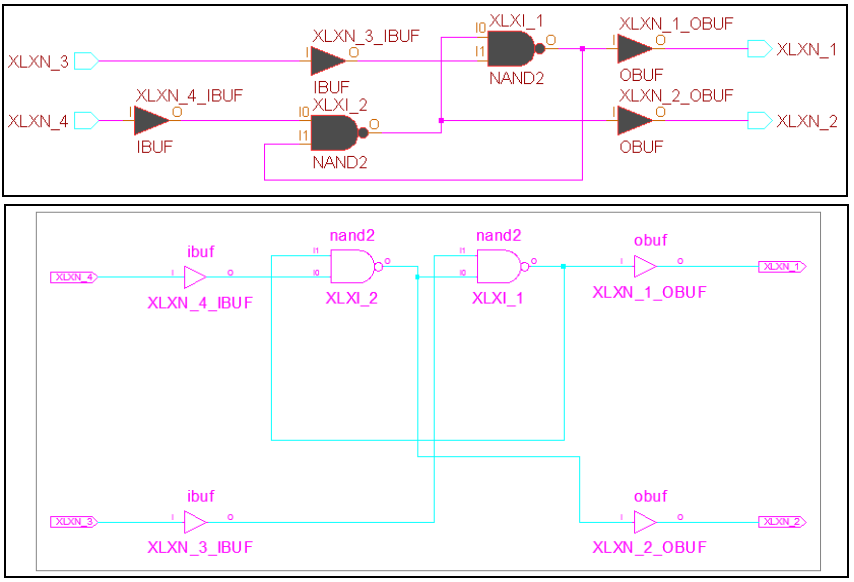


Fig. 6. The RTL schematic (Register Transfer Level) and the reviewed schematic

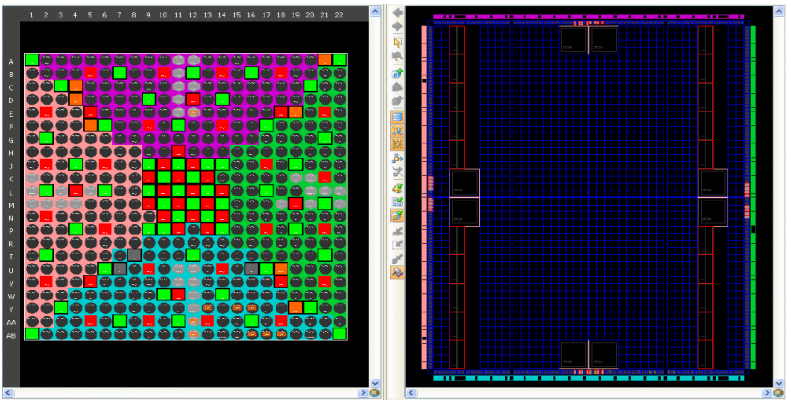


Fig. 7. The Mapping of the design (assigning I/O pins) and the Place&Route

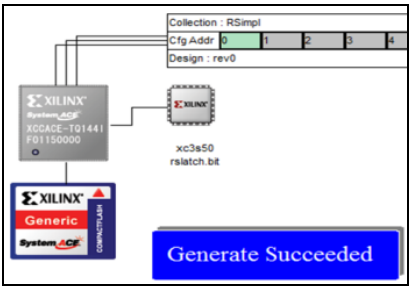
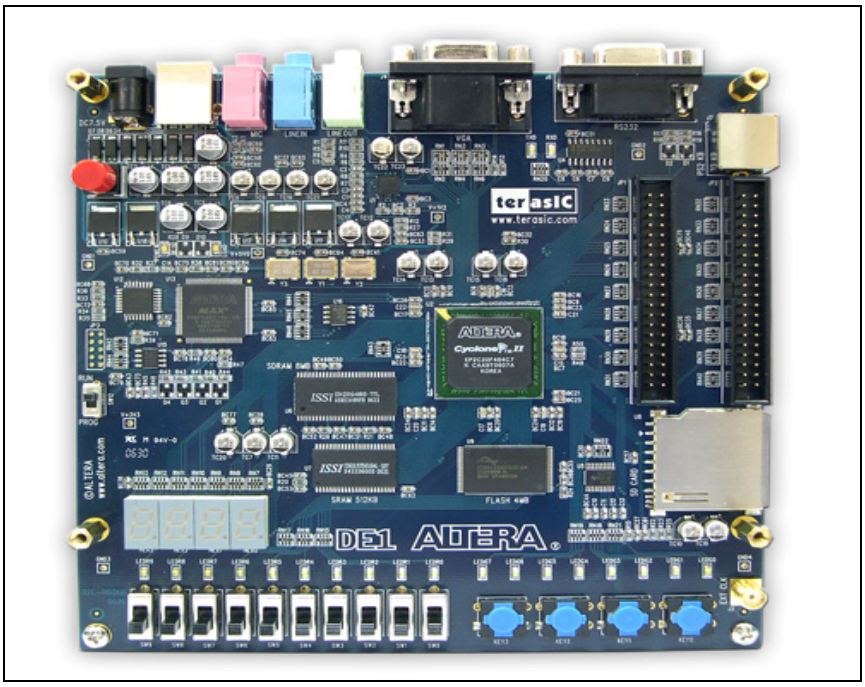
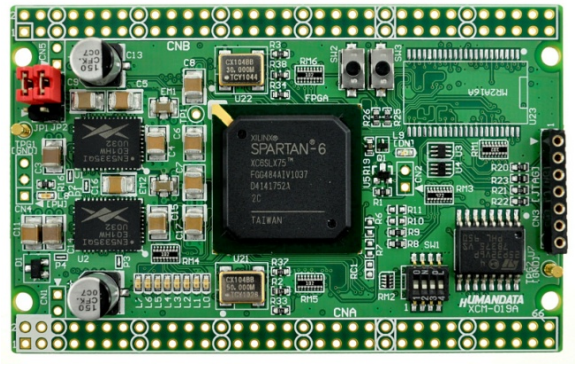


Fig. 8. Generating the netlist and installing the bitstream file into the FPGA's ROM



a) Altera DE1



b) Xilinx Spartan 6

Fig. 9. Popular FPGA Boards

It is to mention that thanks to the HDL high standardization, the conversions HDL code ↔ schematic or VHDL ↔ Verilog are very fast and reliable, as known from the literature and as we personally observed [8, 9].

Another proved approach is the automatic conversion from the well known C code to HDL and consequently to the low-level FPGA programming [7]. This conversion creates conditions for a massive transfer of knowledge (algorithms, procedures, etc.) from the C written applications camp towards FPGA, although critical and optimized FPGA applications need the direct intervention of experienced designers.

4 Existing Healthcare and Biomedical FPGA Applications

The leading producers (Xilinx, Altera, Digilent, etc.) have reached a maturity level that enables FPGAs to deal with a great amount of applications, stemming from virtually any possible domain, including the Artificial Intelligence. The biomedical field was taken seriously into consideration by all producers. In 2007 the Altera team promoted FPGAs as mathematical parallel coprocessors in different models of medical equipment, boosting the speed hundred times, lowering up to 90% the power consumption comparing to conventional coprocessors, with better utilization of the resources in many types of algorithms [10].

Image processing is one of the best suited applications for FPGAs, due to the parallel architectures, the logic density and the generous memory capacity. Besides, driving any type of high resolution display devices, the FPGAs are capable to boost the performance of all the fast imaging algorithms: image construction, enhancement and restoration, image analysis and pattern recognition, image and data compression, color space conversion, etc. Virtually any significant medical imaging technique can be successfully coped by FPGAs: X-rays, magnetic resonance imaging, computed axial tomography (CT scanners), portable ultrasound echographic devices, 3D-imaging, surgical microscopes, optical measuring and analyze instruments, telemedicine systems, etc. [11, 12], 4D imaging-based therapies [13], etc.

Besides imaging, a wide range of clinical applications are mentioned in the FPGA oriented literature: electro surgery equipment [12], heart assisting devices [13], robotic surgery, portable patient monitoring systems, drug delivery systems, automatic external defibrillators (AED) [14], bio-sensors using SmartFusion programmable analog devices [15], etc.

Due to their flexibility and fast upgradeability and to their extremely short time-to-market, FPGAs are now called to support the development of the most advanced medical concepts, as for instance the *minimally-invasive surgery platforms* (MIS), which means significantly less trauma and faster recovery for patients. A MIS platform combines the above mentioned techniques (robotics, endoscopy, 3D visualizations, etc.) in order to minimize the surgical actions and to maximize their precision and sharpness. Such a platform, called *da Vinci Robotic Surgical System*, is now produced by Intuitive Surgical. The platform is composed by an ergonomically designed console where the surgeon sits while operating, a patient-side cart where the patient lays during surgery, four interactive robotic arms, a high-definition endoscopic 3D vision system, and a set of surgical instruments, adapted to the seven degrees of freedom robotic wrists. So far this platform has enabled methodologies for cardiac and general surgery, urology, pediatric, gynecology, colorectal, and head and neck surgeries [12].

The portable ultra-low-power FPGAs enable the implementation of a variety of devices that minimize the power consumption down to 2 μW when the system is not in active use, an important issue for systems such as automated external defibrillators, which may be left unattended for weeks or months between tests, or for the portable insulin pumps [15].

The Radio-Frequency Identification (RFID) is a technology that uses radio waves to transfer data from a RFID tag which is attached to an object, through a reader, with the purpose to identify and to track the object. The RFID tag includes a small *radio frequency* (RF) transmitter and receiver. An RFID reader transmits an encoded radio

signal to interrogate the tag. The tag receives the message and responds with its identification information. The range and the directivity of an RFID system can be largely controlled by the designers.

The Healthcare industry is beginning to use RFID to track patients, personnel, instruments, drugs, etc. All the possible RFID are used: the active technology tracks high-value, or frequency moved items, while passive technology tracks smaller, lower cost items that only need room-level identification. One of the most successful applications is the tracking of the specimens in the pathology laboratory, where the error rate reduced from 9% under 0.05% [16], which was performed in 2007.

In 2008, Clear Count Medical Solutions introduced the Smart Sponge System, an RFID-based system for use in the operating room. The system automatically provides a device-reconciled count by directly matching the unique identifier on each tagged item both entering into and then out of the surgical case [17].

In Nov. 2009 the SocioPatterns platform was used to collect face-to-face proximity between persons with a fine spatial and temporal resolution. They used wearable active RFID devices to detect face-to-face contacts among individuals with a spatial resolution of about 1.5 meters, and a time resolution of 20 seconds. The study was performed in a general pediatric ward of the Bambino Gesù Hospital in Rome, Italy, during a one-week period, and included 119 participants, with 51 health care workers, 37 patients, and 31 caregivers [18].

The mobile RFID units, that are producing, emitting, receiving and decoding the RFID signals need specific radio communication, energy intensive and portable technology, which may be perfectly ensured by the 7 Series FPGAs of Xilinx for instance (up to 96 transceivers operating at 13.1 Gbps, and 16 transceivers operating at 28.05 Gbps raise, I/O bandwidth to 2.8 Tbps, etc.) [19].

5 Conclusions

The FPGA industry reaches now the potential that enables it to deeply revolutionize all our technologies. The healthcare and biomedical equipment industry is facing a strategic challenge: the replacement of conventional electronic equipment characterized by bus centered architectures with FPGA/ASIC based systems, bringing substantial advantages: low costs, flexibility, interoperability, reliability, speed, etc.

The positive FPGA experience issued from the military and the aerospace domains is now beginning to spread into the biomedical and healthcare field, where the personnel should be aware and prepared for this substantial and long term advance.

Acknowledgement. The research of Marius Mircea Balas was supported by the Bilateral Cooperation Research Project between Bulgaria and Romania (2011-2012) entitled: “Electronic Health Records for the Next Generation Medical Decision Support in Romanian and Bulgarian National Healthcare Systems”, NextGenElectroMedSupport. The chapter is a development of a previous contribution entitled “A New Generation of Biomedical Equipment Based on FPGA. Arguments and Facts”, presented at the MedDecSup 2011 Seminar.

References

1. Chen, J., Wong, S., Chang, J., Chung, P.C., Li, H., Koc, U.V., Prior, W., Newcomb, R.: A Wake-Up Call for the Engineering and Biomedical Science Communities. IEEE Circuits and Systems Magazine, 69–77 (second quarter, 2009)
2. *** Xilinx and LynuxWorks Demonstrate Avionics Application Solution at the Military and Aerospace Forum and Avionics USA Conf. LynuxWorks (2009), http://www.lynuxworks.com/corporate/press/2009/xilinx.php
3. Gerngross, J.: Intellectual Property Offers New Choices for Avionics Design Engineers as MIL-STD-1553 and FPGA Technologies Converge. Military and Aerospace Electronics 14(6) (2003), http://www.militaryaerospace.com/index/display/article-display/178482/articles/military-aerospace-electronics/
4. Rodriguez-Andina, J.J., Moure, M.J., Valdes, M.D.: Features, Design Tools, and Application Domains of FPGAs. IEEE Trans. on Industrial Electronics 54(4), 1810–1823 (2007)
5. Hennesy, J.L., Patterson, D.: Computer Architecture: a quantitative approach. Morgan Kaufmann (2006)
6. Klingman, E.: FPGA Programming Step by Step. Electronic Engineering Times Times, http://www.eetimes.com/design/embedded/4006429/FPGA-programming-step-by-step
7. Xilinx: ISE In-Depth Tutorial. 18. Ian. (2012)
8. Balas, M.M., Sajgo, B.A., Belean, P.: On the FPGA Implementation of the Fuzzy-Interpolative Systems. In: 5th Intern. Symp. on Computational Intelligence and Intelligent Informatics, ISCIII 2011, Floriana, Malta, September 15-17, pp. 139–142 (2011)
9. Bălaş, M.M., Socaci, M., Olaru, O.: A FPGA Floating Point Interpolator. In: Balas, V.E., Fodor, J., Várkonyi-Kóczy, A.R., Dombi, J., Jain, L.C. (eds.) Soft Computing Applications. AISC, vol. 195, pp. 331–336. Springer, Heidelberg (2012)
10. Strickland, M.: Medical Applications Look Towards FPGA-based High-performance Computing. Hearst Electronic Products 18 (December 2007)
11. Altera Corp. white papers: Medical Imaging Implementations using FPGAs (2010)
12. Khan, K.: FPGAs Help Drive Innovation in Complex Medical Systems. Medical Electronics Design (2012)
13. Microsemi: Size, Reliability and Security (January 2011), http://www.acaltechnology.com/
14. Actel: Incredible Shrinking Medical Devices (October 2008)
15. Microsemi: Intelligent Mixed Signal FPGAs in Portable Medical Devices. Application Brief AC 242 (December 2010)
16. Bacheldor, B.: At Mayo Clinic, RFID Slashes Error Rate. RFID Journal, http://www.rfidjournal.com/article/view/4422
17. ClearCount. Smart Sponge System, http://clearcount.com/products/smartsponge-system
18. Isella, L., Romano, M., Barrat, A., Cattuto, C., Colizza, V., et al.: Close Encounters in a Pediatric Ward: Measuring Face-to-Face Proximity and Mixing Patterns with Wearable Sensors. PLoS ONE (Public Library of Science), 6 (2) (2011), http://www.plosone.org/article/~ifo:doi/10.1371/journal.pone.0017144
19. Xilinx: 7 Series FPGAs, http://www.xilinx.com/publications/prod_mktg/7-Series-Product-Brief.pdf

Author Index