Computer Science & Information Technology

Natarajan Meghanathan
Dhinaharan Nagamalai (Eds)

Computer Science & Information Technology

The Fourth International Conference of Advanced Computer Science &
Information Technology (ACSIT 2016)
Zurich, Switzerland, June 25~26, 2016

AIRCC Publishing Corporation

Volume Editors

Natarajan Meghanathan,
Jackson State University, USA
E-mail: nmeghanathan@jsums.edu

Dhinaharan Nagamalai,
Wireilla Net Solutions, Australia
E-mail: dhinthia@yahoo.com

Preface

The Fourth International Conference of Advanced Computer Science & Information Technology (ACSIT-2016) was held in Zurich, Switzerland, during June 25~26, 2016. The Fourth International Conference on Signal Image Processing and Multimedia (SIPM-2016), The Fourth International Conference of Managing Information Technology (CMIT-2016) and The Eighth International Conference on Computer Networks & Communications (CoNeCo-2016) were collocated with the ACSIT-2016. The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from the West.

The goal of this conference series is to bring together researchers and practitioners from academia and industry to focus on understanding computer science and information technology and to establish new collaborations in these areas. Authors are invited to contribute to the conference by submitting articles that illustrate research results, projects, survey work and industrial experiences describing significant advances in all areas of computer science and information technology.

The ACSIT-2016, SIPM-2016, CMIT-2016, CoNeCo-2016 Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was done electronically. All these efforts undertaken by the Organizing and Technical Committees led to an exciting, rich and a high quality technical conference program, which featured high-impact presentations for all attendees to enjoy, appreciate and expand their expertise in the latest developments in computer network and communications research.

In closing, ACSIT-2016, SIPM-2016, CMIT-2016, CoNeCo-2016 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the ACSIT-2016, SIPM-2016, CMIT-2016, CoNeCo-2016.

We would like to thank the General and Program Chairs, organization staff, the members of the Technical Program Committees and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the conference and wish them every success in their research. It is the humble wish of the conference organizers that the professional dialogue among the researchers, scientists, engineers, students and educators continues beyond the event and that the friendships and collaborations forged will linger and prosper for many years to come.

Natarajan Meghanathan
Dhinaharan Nagamalai

Organization

General Chair

Natarajan Meghanathan   Jackson State University, USA
Dhinaharan Nagamalai   Wireilla Net Solutions, Australia

Program Committee Members

Abdelkrim Khireddine   University of Bejaia, Algeria
Abdul Raouf Khan   King Faisal Unversity, Saudi Arabia
Abhishek P. Narkhede   Sant Gadge Baba Amravati University, India
Alia Ghaddar   Lebanese Universtiy, Lebanon
Ambresh,   Mangalore University, India
Amir Khusru Akhtar   Cambridge Institute of Technology, India
Amr Rekaby Ali   ERSIL Research Lab, Egypt
Andhe Dharani   R.V.College of Engineering, India
Anuj Vadhera   Lovely Professional University, India
Beatriz Otero   Universitat Politecnica of Catalunya, Spain
Benmohammed M   University of Constantine2,Algeria
Chiranjibi Sitaula   Ambition College - Tribhuvan University, Nepal
Damien   Univ. Grenoble Alpes, France
Dires   Fasil Fenta,University of Gondar, Ethiopia.
Ehsan Sardar Torshizi   Urmia University, Iran
El Miloud Ar-Reyouchi   Abdelmalek Essaadi University, Morroco
Elahe Badiee   Payamenoor University, Iran
Esmaeil Kheirkhah   Islamic Azad University, Iran
Fadhil A. Ali   Oklahoma State University,USA
Farhan   University of Indonesia, Indonesia
Fatemeh Alidusti   Islamic Azad University, Iran
Fatih Korkmaz   Cankiri Karatekin University, Turkey
Garje G.V   PVGs's College of Engg. & Tech, India
Gullanar M Hadi   Salahaddin University, Hawler, Iraq
Hamdi M   National Engineering School of Tunis, Tunisia
Hamdi Yalin Yalic   Hacettepe University, Turkey
Hossein Jadidoleslamy   MUT University, Tehran, Iran
Hui Wu   University of New South Wales, Australia
Isa Maleki   Islamic Azad University, Iran
Iti Mathur   Bansali University, India
Jacques Epounde Ngalle   Robert Morris University, USA
Jagdish Bhatta   Tribhuvan University, Nepal
Jayaraj   Bharathidasan University, India
John Tenvile   Sunyani Polytechnic, Ghana
Kamaraju.M   Gudlavalleru Engineering College, India
Keneilwe Zuva   University of Botswana, Botswana
Khaled Merit   University of mascara, Algeria

Khikmat Muminov	Umarov Physical Technical Institute, Tajikistan
Koushik Majumder	West Bengal University of Technology, India
Mansouri Ali	Claude Bernard University Lyon 1, Tunisia
Maryam Soltanalikhalili	Shahed University, Iran
Meyyappan T	Alagappa University, India
Mohamed Ali Mahjoub	University of Sousse,Tunisia
Mohammad Asmat Ullah Khan	Effat University, Saudi Arabia
Mohammad H. Alomari	Applied Science University, Jordan
Mohammed Abufouda	University of Kaiserslautern, Germany
Mounir Gouioues	Multi Agent System, Morocco
Muhammad Imran KHAN	Université de Toulouse, France
Muhammad	Tampere University of Technology, Finland
Murattopaloglu	Trakya University, Turkey
Narayanan N.K	Kannur University Kerala, India
Neda Darvish	Islamic azad university, Iran
Neda Enami	Payame Noor University, Iran
Niloofar Khanghahi	Islamic Azad University, Iran
Nishant Kumar	Silicon Institute of Technology, India
Nishchol Mishra	Rajiv Gandhi Technical University, India
Nisheeth Joshi	Banasthali University, India
NMVuong (IT.TDT)	Free University of Bolzano, Italy
Ola Younes	Philadelphia University, Jordan
Orhan Dagdeviren	Ege University, Turkey
Pradnya Kulkarni	Federation University, Australia
Prakash	A.V.V.M Sri Pushpam College, India
Rajeshwara Rao	Jntuk University College of Engineering, India
Ram Gopal L	University of Massachusetts, USA
Ranjan Kumar	Cambridge Institute of Technology, India
Raveendra Babu	VNR Vignana Jyothi Inst of Engg & Tech, India
Ravilla Dilli	Manipal Institute of Technology, India
Reza Ravanmehr	Islamic Azad University, Iran
Saradhi Varma G.P	S R K R Engineering College, India
Sayyed Majid Mazinani	Imam Reza University, Iran
Shervan Fekri-Ershad	International Shiraz University, Iran
Shish Ahmad	Integral University, India
Shivaputra	Dr Ambedkar Institute of Technology, India
Siddhivinayak Kulkarni	Federation University, Australia
Smitha N Pai	Manipal Institute of Technology, India
Sokyna Qatawneh	Al-Zaytoonah University of Jordan, Jordan
Sultan Alshehri	University of Regina, Canada
Umesh Kumar Lilhore	NIIST Eng College Bhopal MP, India
Ute Riemann	SAP Deutschland AG & Co. KG, Germany
Zeenat Rehena	Aliah university, WSN, India

Technically Sponsored by

Networks & Communications Community (NCC)

Computer Science & Information Technology Community (CSITC)

Information Technology Management Community (ITMC)

Organized By

Academy & Industry Research Collaboration Center (AIRCC)

TABLE OF CONTENTS

APPLICATION FOR LOGICAL EXPRESSION PROCESSING

Marcin Michalak, Michał Dubiel, Jolanta Urbanek

Institute of Informatics, Silesian University of Technology, Gliwice, Poland
Marcin.Michalak@polsl.pl

ABSTRACT

Processing of logical expressions – especially a conversion from conjunctive normal form (CNF) to disjunctive normal form (DNF) – is very common problem in many aspects of information retrieval and processing. There are some existing solutions for the logical symbolic calculations, but none of them offers a functionality of CNF to DNF conversion. A new application for this purpose is presented in this paper.

KEYWORDS

Boolean Formulas, CNF, DNF, Symbolic Expressions

1. INTRODUCTION

In many aspects of machine learning methods it becomes necessary to represent the knowledge extracted from the data in the terms of a logical formulas. Then the proper tool for Boolean expression processing should be used.

This paper is organized as follows: it starts from a brief overview of problems which solutions are found on the basis of some logical expression analysis; then a short discussion about existing platforms of symbolic formulas processing and their limitation is shown. Afterwards, the new application for CNF to DNF conversion is presented and it's functionality is shown on some short example. The present limitations and the computational complexity of the application are also mentioned. The paper ends with some conclusions and perspectives of the further works.

2. FIELDS OF APPLICATIONS

Many algorithms of machine learning are based on processing information in a logical way. Here two of them are presented: induction of reducts and induction of biclusters.

2.1. Induction of Reducts

A notion of reduct [2] is strongly connected with the theory of data representation – rough sets [1] – defined in the 80's of last century. In this theory data is presented in a table called an information system. Information system is defined as an ordered pair of two non-empty and finite sets: set of objects U and set of attributes A. It is generally assumed that all attributes take only discrete values but there are know some approaches that allows for the continuous attributes in the data.

A sample information system is presented in the Table 1. This information system provides information about four objects, for which the following attributes were measured: temperature, weight, height, color of eyes.

Table 1. A sample information system.

	temperature	weight	height	eyes
U1	high	normal	small	blue
U2	high	low	normal	green
U3	high	low	high	brown
U4	high	high	high	blue

One can notice that the knowledge about values of all attributes for all objects introduced a partition of a set of objects – two objects belong to the same element of partition if they have the same value of the same attributes. With the fundamental theorem on equivalence relation we know that every partition of a set forms an equivalence relation. This relation is called an indiscernibility relation as two object are in a relation when there is no ability to discern each other due to the knowledge about their attributes values.

For a presented information system the indiscernibility relation forms a following partition.

$$U \setminus IND_A = \{\{U_1\}, \{U_2\}, \{U_3\}, \{U_4\}\}$$

The partition formed by this relation is being considered as a knowledge in an information system. It means that any join of elements of partition removes some amount of knowledge from our system. For the information system presented in the Table 1 it is obvious that temperature does not provide any new knowledge as all objects have the same value of this attribute. Excluding the temperature from the set of attributes we obtain the same partition of the set of objects:

$$U \setminus IND_{A \setminus \{temperature\}} = \{\{U_1\}, \{U_2\}, \{U_3\}, \{U_4\}\}$$

This leads us to the question, whether all attributes are necessary to deliver the same knowledge? From this point of view a reduct in an information system is a minimal in the sense of inclusion subset of attributes that introduces the same indiscernibility relation as the whole set A. More formal definition states:

A set $B \subseteq A$ is a reduct for the information system if and only if:

i) $U \setminus IND_A = U \setminus IND_B$

ii)   There is no $C \subset B$ that satisfies $U \setminus IND_C = U \setminus IND_B$

Finding reducts is performed with the analysis of the Boolean formula build as follows: for every attribute a Boolean variable is assigned; for every two objects a logical sum of variables corresponding to attribute that differs these two objects is constructed (when two objects are identical no logical sum is built); final formula is an conjunction of alternatives. This formula is called the discernibility function. Finding the discernibility function becomes easier when all alternatives are presented in the discernibility matrix. This matrix has the same number of rows and columns as the number of objects in the information system. Each cell of this matrix contains attributes that differs objects corresponding to the row and the column of the matrix. It becomes from the definition that this matrix is symmetric. A discernibility matrix for the system presented

in the Table 1 has the following form as presented in the Table 2. Logical variables t, w, h, e, correspond to the attributes temperature, weight, height and eyes.

Table 2. Discernibility matrix for the information system.

	U1	U2	U3	U4
U1	$\emptyset$	$\{w, h, e\}$	$\{w, h, e\}$	$\{w, h\}$
U2	$\{w, h, e\}$	$\emptyset$	$\{h, e\}$	$\{w, h, e\}$
U3	$\{w, h, e\}$	$\{h, e\}$	$\emptyset$	$\{w, e\}$
U4	$\{w, h\}$	$\{w, h, e\}$	$\{w, e\}$	$\emptyset$

The discernibility function will take the following initial form (repeating alternatives were omitted):

$$f(t, w, h, e) = (w \lor h \lor e) \land (w \lor h) \land (h \lor e) \land (w \lor e)$$

As we can see the value of the function does not depend on the value of the variable t, corresponding to the temperature attribute. When simplified, according to the absorption rules, and transformed to the disjunctive normal form, the function takes a form:

$$f(t, w, h, e) = (w \land e) \lor (w \land h) \lor (h \land e)$$

which is built from three prime implicants. This means that there are three two-element reducts for the information system. As it was expected temperature is not an element of any of them as it does not provide any new knowledge.

There is a theorem that says that for every reduct in the information system there is a prime implicant of the discernibility function and vice versa. Attributes in a reduct are the attributes corresponding to the variables in the implicant.

From a formal point of view the problem of finding reducts in the information system is a problem of finding prime implicants of a Boolean formula presented as the logical conjunction of alternatives of non-negated variables.

2.2. Induction of Relative Reducts

Let us consider a specific type of an information system with an emphasized attribute called a decision attribute (or just a decision). This kind of information system is called a decision table and is defined as an ordered pair of finite sets of objects and attributes, where one attribute d is denoted as the decision $(U, A \cup \{d\})$. All non-decision attributes are called conditional attributes. Generally, decision tables contain an information about how the value of the decision depends on values of conditional attributes. It is also worth to state a question whether all conditional attributes are necessary to provide the same amount of knowledge? This leads us to the notion of a relative reducts for the decision table.

Let us define the indiscernibility relation build on the set of conditional attributes A. Let us also define a decision class as the subset of objects with the same value of a decision. The positive region of a decision table $(POS_{DT}(A))$ is a union of equivalence classes that are subsets of decision classes. The higher proportion of objects from positive region to all of objects the more knowledge in the decision table (the more objects can be correctly assign to proper classes). Than as a relative reduct a subset of conditional attributes that give the same amount of knowledge as the whole set of conditional attributes can be defined. The formal definition of a relative reduct is as follows:

A set $B \subseteq A$ is a relative reduct for the information system if and only if:

    i) $POS_{DT}(B) = POS_{DT}(A)$

    ii)   There is no $C \subset B$ that satisfies $POS_{DT}(C) = POS_{DT}(B)$

Finding relative reducts can be also done with the analysis of some Boolean formula.This formula is based on an information stored in a matrix called modulo d discernibility matrix. The phrase "modulo d" is due to the fact that we do not want to discern objects from the same decision class. So only information about difference between objects from different decision classes is presented in this matrix. Let there be a following decision table (Table 3).

Table 3 A sample decision table.

	A	B	C	d
1	a	a	a	yes
2	b	a	b	yes
3	b	c	d	no
4	c	c	d	no

A modulo d discernibility matrix for this decision table will take a following form:

Table 4 A modulo d discernibility matrix for a decision table.

	1	2	3	4
1	$\emptyset$	$\emptyset$	$\{a,b,c\}$	$\{a,b,c\}$
2	$\emptyset$	$\emptyset$	$\{b,c\}$	$\{a,b,c\}$
3	$\{a,b,c\}$	$\{b,c\}$	$\emptyset$	$\emptyset$
4	$\{a,b,c\}$	$\{a,b,c\}$	$\emptyset$	$\emptyset$

Due to the absorption rules we obtain a following disjunction form of the formula:

$$f = b \vee c$$

This means that there are two one-attribute relative reducts for this decision table. This also means that the A attribute is completely redundant.

2.3. Induction of Decision Bireducts

A notion of a decision bireduct [3] also refers to the decision table. Let us consider the same decision table as presented in Table 3. It contains the same kind of information as the previous one but derived from a different objects

As it can be easily observed two last objects are identical so they cannot be discerned due to any possible set of attributes. It is worth to be considered whether do there exist some subsets of objects which are completely discernible due to the subset of attribute. A presented information is delivered by a notion of a bireduct, which definition is as follows:

For a decision system represented by an ordered pair $(U, A \cup \{d\})$ of a set of objects U and a set of attributes A a following ordered pair (B, X) of $B \subseteq A$ and $X \subseteq U$ is called an information bireduct if and only if B discerns all pairs of objects in X (i.e. for each $i, j \in X$ there exists at least one $b \in B$ such that objects i and j differs at least on an attribute b and the following properties hold:

1. There is no proper subset $C \subset B$ such that C discerns all pairs in X, where $d(i) \neq d(j)$.

2. There is no proper superset $Y \supset X$ such that B discerns all pairs in Y, where $d(i) \neq d(j)$.

There also exists a construction of a Boolean formula of logical variables corresponding to attributes and objects, which prime implicants correspond to decision bireducts one to each other.

Let $(U, A \cup \{d\})$ be a decision system. Consider the following Boolean formula with variables $\bar{\imath}, i = 1, \dots, |U|$ and $\bar{u}, u \in A$:

$$f = \bigwedge_{i,j:d(i) \neq d(j)} \left(\bar{\imath} \vee \bar{\jmath} \vee \bigvee_{a:a(i) \neq a(j)} \bar{a} \right)$$

An arbitrary pair $(B, X), B \subseteq A, X \subseteq U$, is a decision bireduct, if and only if the Boolean formula $\bigwedge_{a \in B} \bar{a} \bigwedge_{i \notin X} \bar{\imath}$ is the prime implicants for f. An initial form of f for the decision table that is considered is as follows:

$$f = (1 \vee 3 \vee A \vee B \vee C) \wedge (1 \vee 4 \vee A \vee B \vee C) \wedge (2 \vee 3 \vee B \vee C) \wedge (2 \vee 4 \vee A \vee B \vee C)$$

and its DNF form:

$$f = (1 \wedge 2) \vee (3 \wedge 4) \vee (2 \wedge A) \vee (3 \wedge A) \vee B \vee C$$

Two last bireducts are a typical reducts. Two first describes decision classes. A remaining two are decision bireducts.

3. EXISTING SOLUTIONS

3.1 Matlab

Mathworks provides a toolbox for a symbolic operations. It also allows to define a formulas of Boolean variables, to perform some simplifications but it is not possible to process this formulas exactly to the demanded DNF form.

3.2. R Packages

There exist a package in R which make it possible to find prime implicants for a logical statement but it requires a truth table as an input. This truth table defines alternatives of which variables (negated or not) returns true. This cause that the final prime implicants contain also negated variables.

4. APPLICATION DESCRIPTION

Application is developed in Microsoft Visual Studio 2015 and written in C#. It requires a text input, where labels are Boolean variables and logical operators are represented with '*' and '+'. Alternatives are separated with brackets. As the file with a correct content is loaded a text operators '*' and '+' are displayed as a correct logical operators. The application window with a loaded logical expression is presented in the Fig. 1.

Figure 1. The application with a loaded formula.

The result of calculation is presented in the Fig. 2.

Figure 2. The founded solution: a disjunctive normal form
of the expression in a conjunctive normal form.

Result formula can be presented in several ways, including the LaTeX mathematical mode. Let us consider a logical formula as follows:

$$f = (u1 \lor a1 \lor a2) \land (u1 \lor a1 \lor a3) \land (u1 \lor a2 \lor a3) \land (u2 \lor a1 \lor a3) \land$$
$$\land (u2 \lor a2 \lor a3) \land (u1 \lor u2 \lor a2) \land (u1 \lor u3 \lor a2) \land (u1 \lor u2 \lor a3) \land$$
$$\land (u1 \lor u3 \lor a3) \land (u2 \lor u3 \lor a3)$$

Below the following steps of transformation from CNF to DNF are presented. In the first step each two consecutive alternatives are logically multiplied. Expressions being processed are typed with the bold font and the result is underlined. For a better clarity of presentation only a logical sum is presented with a binary operator:

$$f = \underline{(u1 \lor a1 \lor a2a3)}(u1 \lor a2 \lor a3)(u2 \lor a1 \lor a3)(u2 \lor a2 \lor a3)(u1 \lor u2 \lor a2)$$
$$(u1 \lor u3 \lor a2)(u1 \lor u2 \lor a3)(u1 \lor u3 \lor a3)(u2 \lor u3 \lor a3)$$

$$f = (u1 \lor a1 \lor a2a3)(u1 \lor a2 \lor a3)(u2 \lor a1 \lor a3)\underline{(u1a3 \lor u2 \lor a2)}(u1 \lor u3 \lor a2)$$
$$(u1 \lor u2 \lor a3)(u1 \lor u3 \lor a3)(u2 \lor u3 \lor a3)$$

$f = (u1 \lor a1 \lor a2a3)(\boldsymbol{u1} \lor \boldsymbol{a2} \lor \boldsymbol{a3})(u2 \lor a1 \lor a3)(u1a3 \lor u2 \lor a2)(\boldsymbol{u1} \lor \boldsymbol{u3} \lor \boldsymbol{a2})$
$(u1 \lor u2 \lor a3)(\underline{u1u2 \lor u3 \lor a3})$

$f = (u1 \lor a1 \lor a2a3)\underline{(u1 \lor u3a3 \lor a2)}(\boldsymbol{u2} \lor \boldsymbol{a1} \lor \boldsymbol{a3})(u1a3 \lor u2 \lor a2)$
$(\boldsymbol{u1} \lor \boldsymbol{u2} \lor \boldsymbol{a3})(u1u2 \lor u3 \lor a3)$

$f = (\boldsymbol{u1} \lor \boldsymbol{a1} \lor \boldsymbol{a2a3})(\boldsymbol{u1} \lor \boldsymbol{u3a3} \lor \boldsymbol{a2})\underline{(u2 \lor u1a1 \lor a3)}(u1a3 \lor u2 \lor a2)$
$(u1u2 \lor u3 \lor a3)$

$f = \underline{(u1 \lor u3a1a3 \lor a1a2 \lor \sout{u3a2a3} \lor a2a3)}(\boldsymbol{u2} \lor \boldsymbol{u1a1} \lor \boldsymbol{a3})(\boldsymbol{u1a3} \lor \boldsymbol{u2} \lor \boldsymbol{a2})$
$(u1u2 \lor u3 \lor a3)$

$f = (u1 \lor u3a1a3 \lor a1a2 \lor a2a3)\underline{(u2 \lor \sout{u1a1a3} \lor u1a1a2 \lor u1a3 \lor a2a3)}$
$(u1u2 \lor u3 \lor a3)$

$f = (\boldsymbol{u1} \lor \boldsymbol{u3a1a3} \lor \boldsymbol{a1a2} \lor \boldsymbol{a2a3})(u2 \lor u1a1a2 \lor u1a3 \lor a2a3)(\boldsymbol{u1u2} \lor \boldsymbol{u3} \lor \boldsymbol{a3})$

$f = (u1u2 \lor u1u3 \lor u1a3 \lor \sout{u1u2u3a1a3} \lor u3a1a3 \lor \sout{u3a1a3} \lor \sout{u1u2a1a2} \lor$
$\lor u3a1a2 \lor \sout{a1a2a3} \lor \sout{u1u2a2a3} \lor \sout{u3a2a3} \lor a2a3)(u2 \lor u1a1a2 \lor u1a3 \lor a2a3)$

$f = (u1u2 \lor u1u3 \lor u1a3 \lor u3a1a3 \lor u3a1a2 \lor a2a3)(u2 \lor u1a1a2 \lor u1a3 \lor a2a3)$

For the simplification of calculation let the repeating conjunction be bolded:

$f = (u1u2 \lor u1u3 \lor \boldsymbol{u1a3} \lor u3a1a3 \lor u3a1a2 \lor \boldsymbol{a2a3})(u2 \lor u1a1a2 \lor \boldsymbol{u1a3} \lor \boldsymbol{a2a3})$

Now the multiplication becomes more easy:

$f = u1a3 \lor a2a3 \lor (u1u2 \lor u1u3 \lor u3a1a3 \lor u3a1a2)(u2 \lor u1a1a2)$

$f = u1a3 \lor a2a3 \lor u1u2 \lor \sout{u1u2a1a2} \lor \sout{u1u2u3} \lor u1u3a1a2 \lor u2u3a1a3 \lor\lor \sout{u1u3a1a2a3}$
$\lor u2u3a1a2 \lor \sout{u1u3a1a2}$

$f = u1a3 \lor a2a3 \lor u1u2 \lor u1u3a1a2 \lor u2u3a1a3 \lor u2u3a1a2$

In the Fig. 3. the application window with the loaded formula and result of it conversion to DNF are presented. The obtain solution is equivalent to the solution obtained as the result of step-by-step calculations.

Figure 3. A considered formula in CNF and DNF.

A weak point of the application is the high computational complexity. On the Fig. 4 an average time of ten experiments as the function of a input formula length is presented.
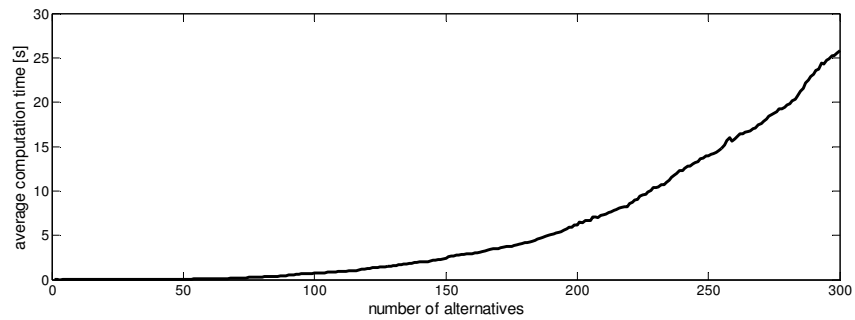

Figure 4. An average time of computation with the increase of the input data length.

As one may observe the computational complexity is polynomial. The degree of the complexity polynomial can be estimated with the Pearson's correlation coefficient between the number of alternatives in the input formula and the appropriate degree of the average computation time root. The results are presented in the Table 1.

Table 1. Correlation between the task complexity and the roots of average time of computation.

Root degree	r
1	0.892489676
2	0.981016972
3	0.998530988
4	0.998626003
5	0.993885497
6	0.987855142
7	0.981611630
8	0.975466952

It can be experimentally estimated that the computational complexity of the algorithm is polynomial with the degree of four.

5. CONCLUSIONS AND FURTHER WORKS

In the paper the new application for logical expression processing is presented. It performs a symbolic Boolean calculations, converting the CNF of the formula to DNF. As results of these computation may be useful in scientific research it provides also a LaTeX format of results. Due to the high computational complexity our further works will focus on reduction of the time of computation.

REFERENCES

[1] Pawlak Z.: Rough Sets. International Journal of Computer and Information Sciences, 11(5):341-356, Warsaw 1982.

[2] Pawlak Z.: Rough Sets. Theoretical Aspects of Reasoning about Data, Springer, 1991

[3] Ślęzak D., Janusz A.: Ensambles of Bireducts: Towards Robust Classification and Simple Representation, Lecture Notes in Computer Science 7105:64-77, 2011

AUTHORS

Marcin Michalak was born in Poland in 1981. He received his M.Sc. Eng. in computer science from the Silesian University of Technology in 2005 and Ph.D. degree in 2009 from the same university. His scientific interests are in machine learning, data mining, rough sets and biclustering. He is an author and coauthor of over 60 scientific papers.

Michał Dubiel is a student at Silesian university of Technology

Jolanta Urbanek received her M. Sc. Eng. in computer science from the Silesian University of Technology. Her scientific interests are in biclustering.

INTENTIONAL BLANK

TURKISH SIGN LANGUAGE RECOGNITION USING HIDDEN MARKOV MODEL

Kakajan Kakayev[1] and Ph.D. Songül Albayrak[2]

[1,2]Department of Computer Engineering,
Yildiz Technical University, Istanbul, Turkey
kkakajan@gmail.com
songul@ce.yildiz.edu.tr

ABSTRACT

In past years, there were a lot of researches made in order to provide more accurate and comfortable interaction between human and machine. Developing a system which recognizes human gestures, is an important study to improve interaction between human and machine.

Sign language is a way of communication for hearing-impaired people which enables them to communicate among themselves and with other people around them. Sign language consists of hand gestures and facial expressions. During the past 20 years, researches were made to facilitate communication of hearing-impaired people with others.

Sign language recognition systems are designed in various countries. This paper presents a sign language recognition system, which uses Kinect camera to obtain skeletal model. Our aim was to recognize expressions, which are used widely in Turkish Sign Language (TSL). For that purpose we have selected 15 words/expressions randomly (repeated 4 times each by 3 different signers) which belong to Turkish Sign Language. We have used 180 records in total. Videos are recorded using Microsoft Kinect Camera and Nui Capture. Joint angles and joint positions have been used as features of gesture and achieved close to 100% recognition rates.

KEYWORDS

Hidden Markov Model, Turkish Sign Language Recognition, Gesture Recognition, Microsoft Kinect, skeleton model

1. INTRODUCTION

In past years, new ways in human-computer interaction have been enabled. The researchers developed various applications on gesture and sign language recognition. Hearing-impaired people use sign language. When hearing-impaired people communicate, they need an interaction point and gestures are used as interaction point. Especially they use upper torso gestures and facial expressions. Handicap of sign language recognition is that signs change in 3D space. Sign language recognition developed using various input devices, such as wearing data gloves, stereo cameras, etc. Madabhushi and Aggarwal [1] developed sign recognition by tracking particular body parts. There are also studies based on depth information acquired from sensors such as Kinect [2, 3]. Ra'eesah Mangera used skeleton model, which is captured with Kinect camera [4]. Each country generally has its own native sign language and some have more than one. It is not clear how many sign languages there are. The 2013 edition of Ethnologue lists 137 sign

languages [5]. There are researches made on Turkish [6, 7], Polish [8], American [9] Sign Language Recognition systems. In this study, we used 3D skeleton information of human skeleton model generated from Microsoft's Kinect sensor using Nui Capture.

In this paper, it is intended to recognize sign language by analysing skeleton model captured with Kinect camera. With the use of Kinect and Nui Capture application, the upper skeleton information of the human participants are recorded and used for training and testing the system. In order to recognize the signs we used K-Means with Hidden Markov Model (HMM). A k-means classifier is used to cluster the data. Every sign or gesture is shown with a series of frames. Features extracted from frames (joint angles, joint distances) and converted into observation sequence by means of k-means method and trained with Hidden Markov Model. Baum-Welch algorithm is used for HMM training [10].

For this work, 18 signs recorded by 3 persons. Each sign repeated 4 times by each signer. The content of recorded signs is shown in Table 1.

Table 1. Dataset used in training and testing phase.

No	Words
1	Let's meet again
2	See you
3	Good bye
4	Good night
5	Who?
6	Hello
7	Where?
8	Sometimes
9	Thanks
10	Yesterday
11	Cook
12	Doctor
13	Pharmacist
14	Baker
15	Driver
16	How?
17	I'm fine
18	Enjoy your meal!

The rest of this paper is ordered as follows: Section 2 describes calculations used for feature extraction from skeleton model and describes algorithm for training and testing procedure used in k-means and HMM. Section 3 provides experimental results obtained from records and Section 4 conclusion.

2. FEATURE EXTRACTION

Kinect Sensor generates depth maps, skeleton model, and RGB images. To interface with the device, NuiCapture and Kinect SDKs are used. NuiCapture is software used to record and analyse Kinect for Windows sensor data easily [11]. NuiCapture can export depth, color, and skeleton data to Matlab, Maya, 3DS Max, and MotionBuilder. The skeleton model extracted with nuiCapture is shown in Figure 1. Kinect camera tracks the 3-D coordinates of these joints. Sign languages consider hand gestures, upper body joints used in feature extraction.

2.1. Joint Distances

For each frame, the 3-D distance between each of the 6 arm joints and the head joint was calculated:

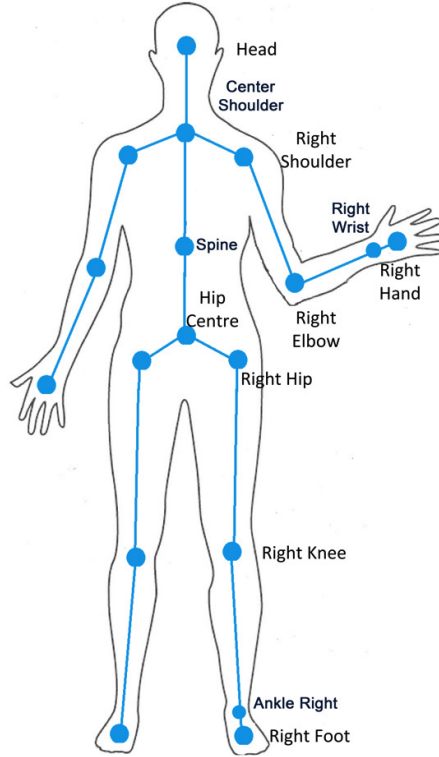$$d(ls, he) = \sqrt{(x_{ls} - x_{he})^2 + (y_{ls} - y_{he})^2 + (z_{ls} - z_{he})^2} \tag{1}$$

Figure 1. Kinect Skeleton Model [4]

Additionally, left hand – spine – right hand distance calculated. 7 dimensional feature vector formed from joint distances:

$$F_{JD} = [d_{ls}, d_{rs}, d_{le}, d_{re}, d_{lh}, d_{rh}, lh - rh]$$

User heights may be various. To reduce for the variation in user height, each distance was divided by the distance between spine and center shoulder [4].

2.2. Joint angles

The distance between joints is affected by the height of the user. Therefore, joint distances are not a scale invariant feature. Joint angles are not dependent on the user height or the distance from the camera. Joint angle also rotation invariant. Seven joint angles were calculated for each frame. Joint angles are shown in Figure 2. To calculate the joint angle, the vector between joints must be calculated. The calculation of shoulder-elbow-hand angle is illustrated in Figure 3. The shoulder – elbow – hand angle equation is given below:

$$\theta = \arccos\left(\frac{\overline{s-e}\cdot\overline{e-h}}{|\overline{s-e}||\overline{e-h}|}\right) \tag{2}$$

$\overline{(s-e)}$ and $\overline{(e-h)}$ is the shoulder – elbow and elbow – hand vector respectively. Numerator of equation 2 is the scalar product of the vectors and the denominator is the product of the magnitudes of the vectors. 7 dimensional feature vector is created from joint angles:

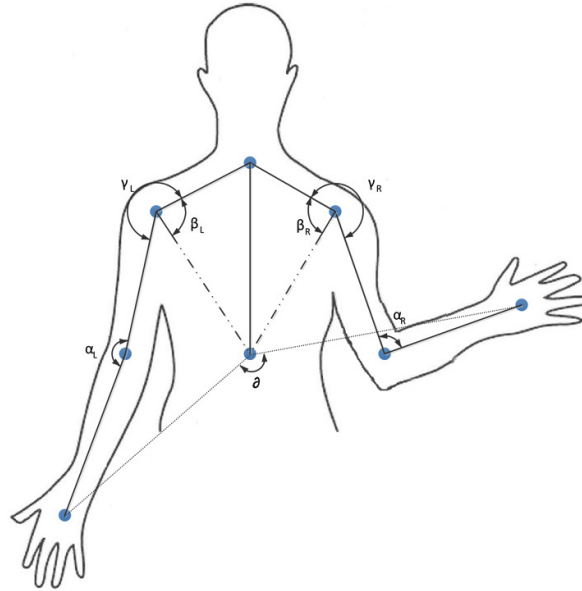$$F_{JA} = [\gamma_L, \gamma_R, \beta_L, \beta_R, \alpha_L, \alpha_R, \partial]$$

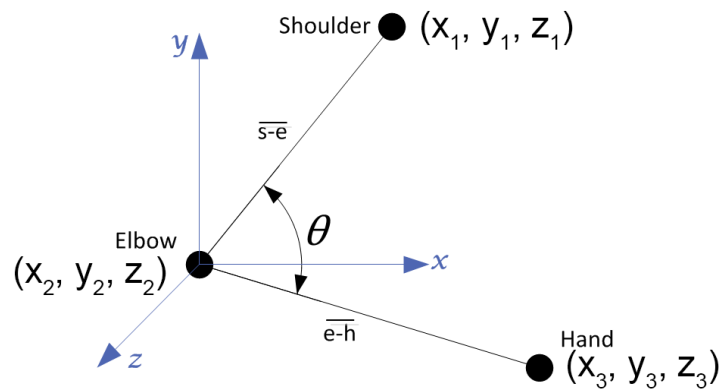
Figure 2. Joint angles calculated from skeleton model

Figure 3. Calculation of shoulder – elbow – hand angle

2.3. Relative Joint Positions

Joint angles are rotation invariant, but a pose with the arms stretched on either side of the spine and arms stretched in front of the spine will have similar feature vectors. Therefore, the relative joint position between the elbow and hand joints and the head joint is calculated for each pose [4]. Figure 4 shows the position of the hand relative to the x-component of the head joint.

$\overline{he - h}$ is head-hand vector and the x-component of the head joint is:

$$he_x = x_1\hat{\imath} + 0\hat{\jmath} + 0\hat{k}$$

We can calculate the position of the hand relative to the head by equation 3

$$\varphi = arccos\left(\frac{\overline{he-h}\cdot\overline{he_x}}{|\overline{he-h}||\overline{he_x}|}\right) \tag{3}$$
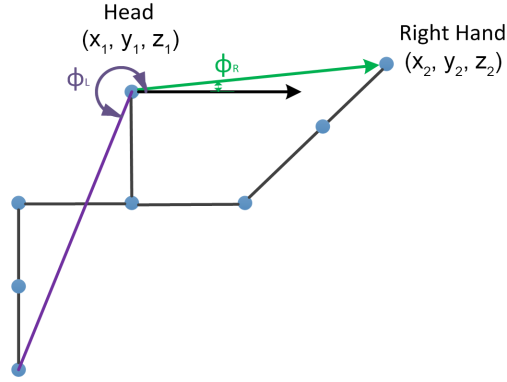

Figure 4. Demonstration of the relative position of the right and the left hands with respect to the head

2.3. Combination of feature vectors

By joining joint angles, joint distances and joint relative position we form 18 dimensional feature vectors for each frame.

$$F_C = [\gamma_L, \gamma_R, \beta_L, \beta_R, \alpha_L, \alpha_R, \sigma_L, \sigma_R, \partial, \varphi_L, \varphi_R, d_{ls}, d_{rs}, d_{le}, d_{re}, d_{lh}, d_{rh}, lh - rh]$$

Description of each feature vector provided in table 2.

Table 2. Description of feature vector elements

γ	Elbow – Shoulder – Neck angle	φ	Relative position of the elbow relative to the head
β	Spine – Shoulder – Neck angle	σ	Relative position of the hand relative to the head
α	Hand – Elbow – Shoulder angle	$lh - rh$	Distance between the left and right hands
∂	Left hand – Spine – Right hand angle	d	Distance between joints and head

2.4. Training and Testing

Feature extraction process is applied to all frames and signs. After feature extraction process, a k-means classifier is trained for each of the signs to obtain cluster centers. In this work, 40 used as a number of cluster (K) for each sign. This value was decided as empirically yielding the best inner-class division. Cluster centers are used for training system. HMM is trained by using Baum-Welch algorithm [10]. To calculate the recognition accuracy rate, the total number of correct recognitions is divided by total number of tests.

3. EXPERIMENTAL RESULTS

The system was tested with 2 different test types as shown in Figure 5. In Test 1, system was trained with all records of 2 signers and tested with all records of the 3rd signer. In Test 2, system was trained with 3 repetitions of each word and tested with the 4th repetition of each word and of each signer.

Figure 5. Test diagram

Dataset was divided into two parts, first part consists of 15 words, second part has 18 words. In the second dataset initial 15 words are the same as in the first dataset along with 3 extra words. Extra words are similar to previous words by means of movement. Test results are shown in Table 3, 4, 5 and 6.

Table 3. 15 words – Test 1

Total signs used for training	15
Total records used for training	120
Total records used for testing	60
Number of correct recognition	59
Number of wrong recognition	1
Recognition rate	59/60*100 = 98%

Table 4. 15 words – Test 2

Total signs used for training	15
Total records used for training	135
Total records used for testing	45
Number of correct recognition	43
Number of wrong recognition	2
Recognition rate	43/45*100 = 95%

Table 5. 18 words – Test 1

Total signs used for training	18
Total records used for training	144
Total records used for testing	72
Number of correct recognition	60
Number of wrong recognition	12
Recognition rate	60/72*100 = 83%

Table 6. 18 words – Test 2

Total signs used for training	18
Total records used for training	162
Total records used for testing	54
Number of correct recognition	50
Number of wrong recognition	4
Recognition rate	50/54*100 = 92%

4. CONCLUSION

In this paper we have presented sign language recognition system based on skeleton model of gestures. We have developed Turkish Sign Language recognition system using the Kinect camera and achieved close to 100% recognition rates. To increase the accuracy of the system, system can be trained by increasing the repetitions and signer number. The system works well for new words and new users.

It is observed that with Kinect's ability of recognition of human gestures adds another aspect to using computer applications. This research sets another example to sign language recognition systems. It will help hearing-impaired people as automatic translation tools.

This research has an advantage of freedom from any external devices for input except Kinect camera. In order to increase the recognition rate, the system needs as many repeated records and signs from different signers as possible. It is observed that when similar movements are used in test the successful recognition rate is reduced because of insufficient training data.

REFERENCES

[1] A. Madabhushi and J. K. Aggarwal, "Using head movement to recognize activity", (2000) "Proceedings of 15th International Conference on Pattern Recognition", vol. 4, pp. 698 – 701.

[2] Yamato, J., Ohya, J. and ISHII, K., (1992). "Recognizing human action in time-sequential images using hidden Markov model", Computer Vision and Pattern Recognition, 379-385

[3] Biswas, K.K. and Basu, S.K., (2011). "Gesture Recognition using Microsoft Kinect", Robotics and Applications

[4] Mangera, R. (2013). "Static gesture recognition using features extracted from skeletal data"

[5]   Wikipedia, Sign Language, https://en.wikipedia.org/wiki/Sign_language

[6]   Haberdar, H., (2005). "Saklı Markov Model Kullanılarak Görüntüden Gerçek Zamanlı Türk İşaret Dili Tanıma Sistemi", Yıldız Technical University, İstanbul

[7]   Memiş, A. and Albayrak, S., (2013). Turkish Sign Language Recognition Using Spatio-temporal Features on Kinect RGB Video Sequences and Depth Maps, Signal Processing and Communications Applications Conference, 1-4

[8]   Oszust, M. and Wysocki, M., (2013). Polish Sign Language Words Recognition with Kinect, Human System Interaction, 219-226

[9]   Z. Zafrulla, H. Brashear, H. Hamilton, T. Starner, and P. Presti, "American sign language recognition with the kinect," in Proceedings of the 13th international conference on multimodal interfaces, ser. ICMI'11, no. September, Sch. of Interactive Computing, Georgia Inst. Of Technology, Atlanta. New York, NY, USA: ACM, 2011, pp. 279–286.

[10]   Wikipedia, Baum-Welch algorithm, http://en.wikipedia.org/wiki/Baum-Welch_algorithm

[11]   Software to Record and analyse Kinect for Windows sensor data easily, http://nuicapture.com/

EFFICIENT APPROACH FOR CONTENT BASED IMAGE RETRIEVAL USING MULTIPLE SVM IN YACBIR

Lakhdar LAIB[1] and Samy Ait-Aoudia[2]

[1,2]National High School of Computer Science ESI, Algiers, ALGERIA
`l_laib@esi.dz`
`s_ait_aoudia@esi.dz`

ABSTRACT

Due to the enormous increase in image database sizes, the need for an image search and indexing tool is crucial. Content-based image retrieval systems (CBIR) have become very popular for browsing, searching and retrieving images in different fields including web based searching, industry inspection, satellite images, medical diagnosis images, etc. The challenge, however, is in designing a system that returns a set of relevant images i.e. if the query image represents a horse then the first images returned from a large image dataset must return horse images as first responses. In this paper, we have combined YACBIR [7], a CBIR that relies on color, texture and points of interest and Multiple Support Vector Machines Ensemble to reduce the existing gap between high-level semantic and low-level descriptors and enhance the performance of retrieval by minimize the empirical classification error and maximize the geometric margin classifiers. The experimental results show that the method proposed reaches high recall and precision.

KEYWORDS

Content Based Image Retrieval, YACBIR, Feature extraction, Multiple Support Vector Machines, Classification

1. INTRODUCTION

With fast and exceptional growth in digital technology, a significant increase in the number of images produced by scientific, educational, medical, industrial, virtual museums, individual photograph collections and other applications raises several practical issues.

Content-based Image Retrieval (CBIR) system [1] [2] [3] [4] is the application of computer vision techniques which is used to extract similar images from an large image database. CBIR system is mainly used in various applications such as art collections, education and training, crime prevention, military field, architectural and engineering design, journalism and advertising, and web searches [5].

Simply stated, a CBIR system retrieves images corresponding to a query image by examining the image contents (low-level feature). Many image low-level features such as color, texture and shape are widely used [7]. The color aspect can be treated by techniques like averaging and histograms. The texture aspect can be achieved by using four descriptors: contrast, entropy, energy and inverse differential moment [8]. The shape aspect can be extracted by gradient or morphological operators [6] or by considering points of interest. [8].
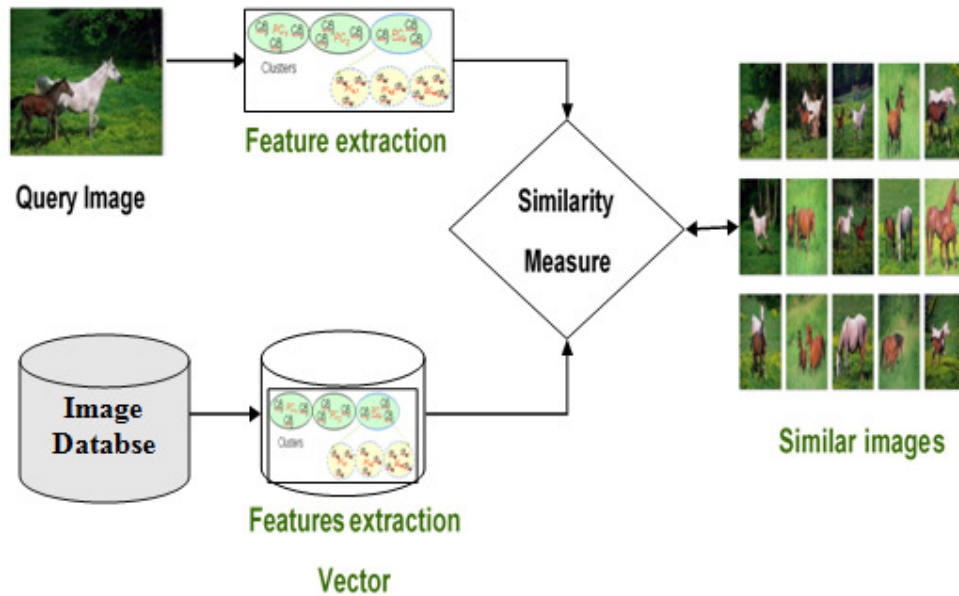
Figure 1: Content Based Image Retrieval Approach.

The main goal in CBIR system is searching the similar images from the database based on their content. To accomplish the retrieval task, CBIR system firstly computes and stores the feature vector for each image in the database. When the query image is introduced in the system, its feature vector will be extracted and compared to those of other images in the database. Then a series of similar images is returned to the user. Different similarity measures are used in the matching process [6].

This paper is organized as follows. We give in section 2 an overview of our proposed system. In section 3, concepts of SVM (Support Vector Machine) are reminded. Experimental results on sample datasets are given in section 4. Section 5 gives conclusions.

2. OVERVIEW OF PROPOSED SYSTEM

2.1. Method summary

Step 1: Feature extraction f (Color, Texture, and Point of interest).

Step 2: Sum of weighted color, texture and points of interest features.

Step 3: Query image feature classification by Multiple   Support Vector Machines (Find the class Label).

Step 4: Compute the distance between the Query image and   images in the dataset (Class Label).

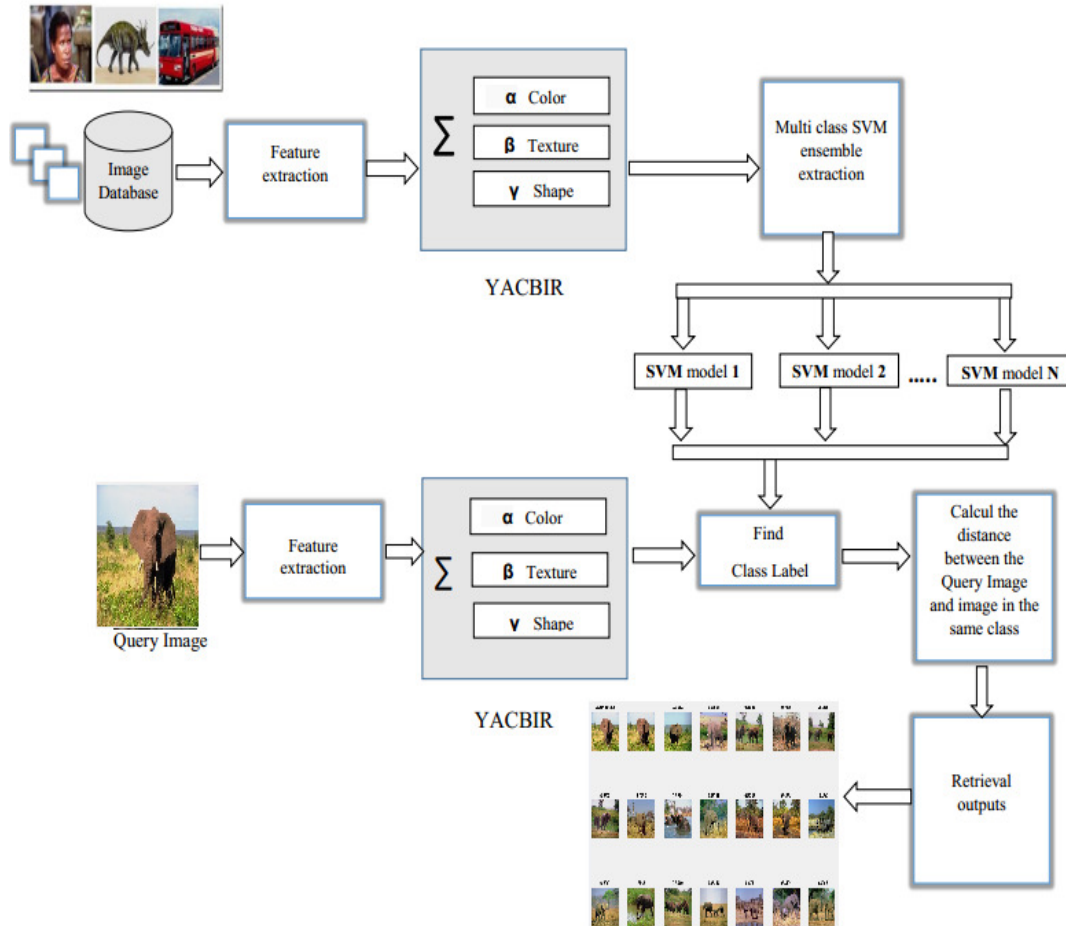**Step 5:** Output Retrieved images (Similar images).

Figure 2: Overview of the process

2.2. Feature extraction in YACBIR

Feature extraction is a common way to improve the performance of the visual features and improve the efficiency of the CBIR systems. The best subset of features is selected by a heuristic search algorithm to improve the precision of CBIR system. Three types of features are used for image indexing [7].

Color - HSV
Texture -contrast, entropy, energy and inverse differential moment
Point of interest -Harris detector

The YACBIR system combines three characteristics of an image to compute a weighted similarity measure. The characteristics of the image are the color, texture and points of interest. Color and texture characteristics used are global while points of interest are local shape characteristics. The similarity measure used in YACBIR is a sum of weighted color, texture and points of interest (shape) similarity measures. This similarity measure is given by:

$$S = \alpha.Sc + \beta.St + \gamma.Ss \quad \text{with} \ (\alpha+\beta+\gamma) = 1$$

2.3. Support Vector Machines (SVM)

In machine learning, support vector machines (SVM) are supervised learning models with learning algorithms. Given a set of data where each data is marked as belonging to one of two categories, a SVM training algorithm constructs a model which classifies new data in one of these two classes. Though SVMs are fundamentally developed for such binary classification case, they are extendable for multi-class problems [9].

The SVM algorithm consists of two phases:

a. Training phase:

Input learning data for a SVM classifier can consist of distance to border vectors, binary images, Zernike moments, and more. Each input data is represented by vector $\mathbf{x}_i$ with label $y_i = \pm 1$, $1 \leq i \leq l$, l is the number of samples.

The decision boundary should classify all points correctly, thus

$$y_i\left(\mathbf{w}^{\mathbf{T}}\mathbf{x}_i + b\right) \geq 1, \ \forall i.$$

The decision boundary can be found by solving the following constrained optimization problem:

$$Minimize \ \frac{1}{2}\|\mathbf{w}\|^2 \quad subjectto \ y_i\left(\mathbf{w}^{\mathbf{T}}\mathbf{x}_i + b\right) \geq 1 \ \forall i$$

The Lagrangian of this optimization problem is:

$$L = \frac{1}{2}\|\mathbf{w}\|^2 - \sum_i \alpha_i\left(y_i\left(\mathbf{w}^{\mathbf{T}}\mathbf{x}_i + b\right) - 1\right) \ \alpha_i \geq 0 \ \forall i \ \mathrm{T}$$

the optimization problem can be rewritten in terms of α_i by setting the derivative of the Lagrangian to zero:

$$Maximize \ W(\alpha) = \sum_i \alpha_i - \frac{1}{2}\sum_{i,j}\alpha_i\alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j$$

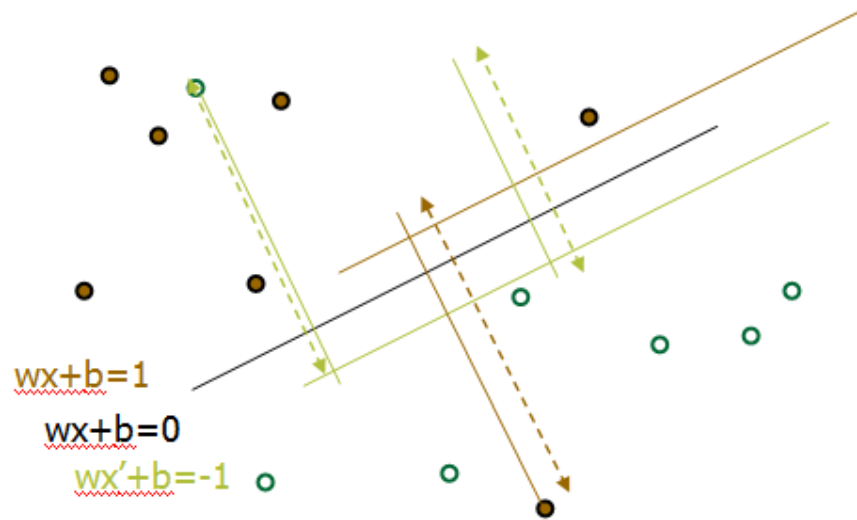$$subjectto \ \alpha_i \geq 0, \sum_i \alpha_i y_i = 0 \ \forall i$$

wx+b=1

wx+b=0

wx'+b=-1

Figure 3: Support Vector Machine (SVM).

This quadratic programming problem is solved when:

$$\alpha_i \left(y_i \left(\mathbf{w}^{\mathbf{T}} \mathbf{x}_i + b \right) - 1 \right) = 0 \quad \forall i \; \mathbf{x}_i \text{ with } \alpha_i > 0$$

b. Testing phase:

The resulting classifier is applied to unlabelled images to decide whether they belong to the positive or the negative category. The label of x is simply obtained by computing:

$$\mathbf{w} = \sum_{j=1}^{s} \alpha_{t_j} y_{t_j} \mathbf{x}_{t_j} \quad \text{with } t_j \left(j = 1, \text{K} , s \right) \text{ the indices of the } s \text{ support vectors. Classify } \mathbf{z}$$

as category 1 if the sum is positive, and category 2 otherwise.

3. EXPERIMENTAL RESULTS

The experiments in this section are performed on a workstation with an Intel Corei3 2.10 GHz CPU and 4 GB of RAM running Win 10.

We evaluate the effectiveness of the proposed method for Corel-1k which is a set of 1000 images that are subdivided into 10 sufficiently distinct semantic classes, including African people, seaside, bus, building, mountain, dinosaur, elephant, horses, flowers and food images some examples are depicted in Figure 4.
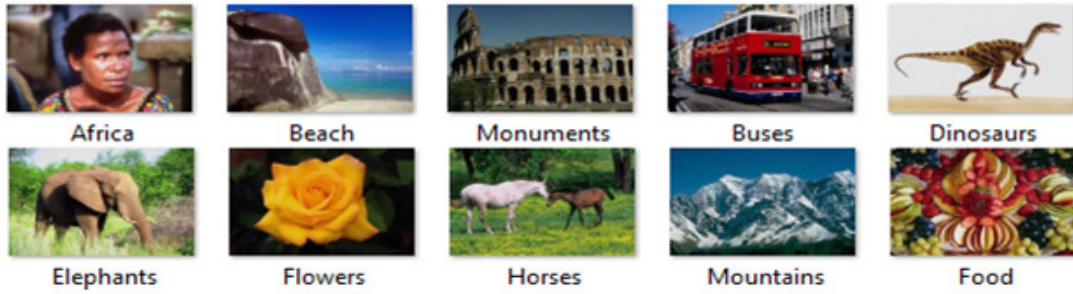
Figure 4: Image database used (Corel-1k).

3. 2 Precision-recall

Precision-versus-recall curve is the basic evaluation measure in information retrieval. They are defined as follow:

$$\text{Precision} = \frac{number\ of\ relevant\ images\ retrieved}{total\ number\ of\ relevant\ images\ retrieved} \ldots\ldots\ldots\ldots [19]$$

$$\text{Recall} = \frac{number\ of\ relevant\ images\ retrieved}{total\ number\ of\ relevant\ images\ in\ the\ database} \ldots\ldots [20]$$

Table 1: Comparison precision proposed method with other methods

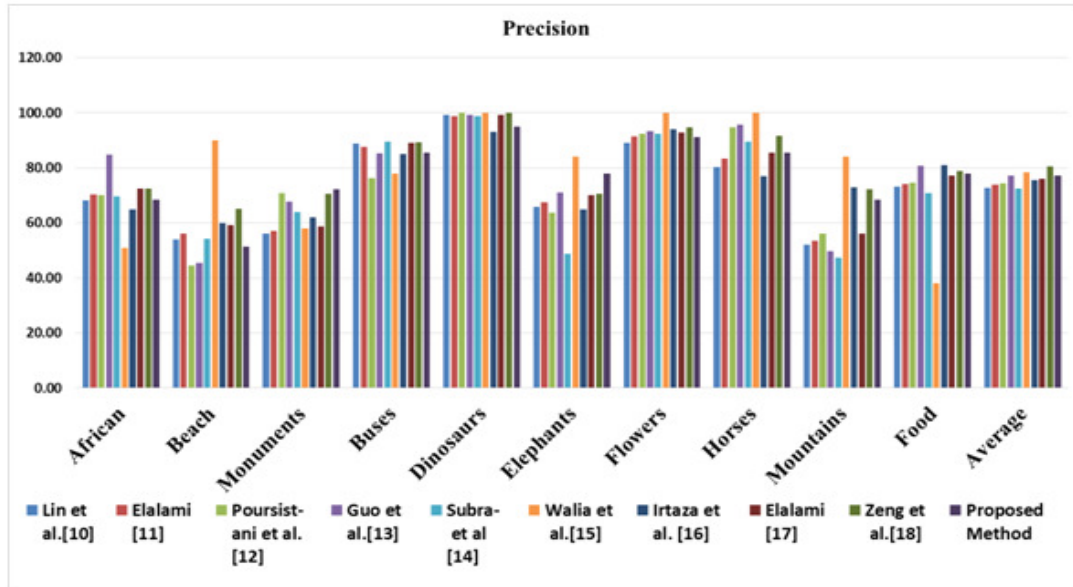| Class	Lin et al.[10]	Elalami [11]	Poursist-ani et al. [12]	Guo et al.[13]	Subra- et al [14]	Walia et al. [15]	Irtaza et al. [16]	Elalami [17]	Zeng et al.[18]	Proposed Method
African	68.30	70.30	70.20	84.70	69.75	51.00	65.00	72.60	72.50	68.40
Beach	54.00	56.10	44.40	45.40	54.25	90.00	60.00	59.30	65.20	51.30
Monuments	56.15	57.10	70.80	67.80	63.95	58.00	62.00	58.70	70.60	72.20
Buses	88.80	87.60	76.30	85.30	89.65	78.00	85.00	89.10	89.20	85.50
Dinosaurs	99.25	98.70	100.00	99.30	98.70	100.00	93.00	99.30	100.00	95.00
Elephants	65.80	67.50	63.80	71.10	48.80	84.00	65.00	70.20	70.50	77.90
Flowers	89.10	91.40	92.40	93.30	92.30	100.00	94.00	92.80	94.80	91.20
Horses	80.25	83.40	94.70	95.80	89.45	100.00	77.00	85.60	91.80	85.50
Mountains	52.15	53.60	56.20	49.80	47.30	84.00	73.00	56.20	72.25	68.40
Food	73.25	74.10	74.50	80.80	70.90	38.00	81.00	77.20	78.80	77.90
Average	72.70	73.90	74.39	77.30	72.51	78.30	75.50	76.10	80.57	77.33

Figure 5. Precision of proposed approach compared to other methods for different semantic classes in the Corel database.

Table 2: Comparison recall proposed method with other methods

| Class	Lin et al.[10]	Elalami [11]	Poursist-ani et al. [12]	Guo et al.[13]	Subra-et al [14]	Walia et al. [15]	Irtaza et al. [16]	Elalami [17]	Zeng et al.[18]	Proposed Method
African	14.10	15.30	14.04	16.94	13.95	10.20	13.00	16.10	14.50	14.46
Beach	19.20	19.80	8.88	9.08	10.85	18.00	12.00	20.30	13.04	10.98
Monuments	17.40	18.20	14.16	13.56	12.79	11.60	12.40	19.10	14.12	14.79
Buses	12.10	11.60	15.26	17.06	17.93	15.60	17.00	12.60	17.84	18.07
Dinosaurs	10.10	9.80	20.00	19.86	19.74	20.00	18.60	10.90	20.00	19.08
Elephants	14.90	15.60	12.76	14.22	9.76	16.80	13.00	16.30	14.10	16.60
Flowers	11.20	11.80	18.48	18.66	18.46	20.00	18.80	12.90	18.96	19.28
Horses	13.40	13.90	18.94	19.16	17.89	20.00	15.40	14.40	18.36	18.37
Mountains	21.30	22.80	11.24	9.96	9.46	16.80	14.60	23.60	14.45	14.34
Food	13.20	13.80	14.90	16.16	14.18	7.60	16.20	14.80	15.76	16.73
Average	14.68	15.20	14.86	15.46	14.50	15.66	15.10	16.10	16.11	16.27
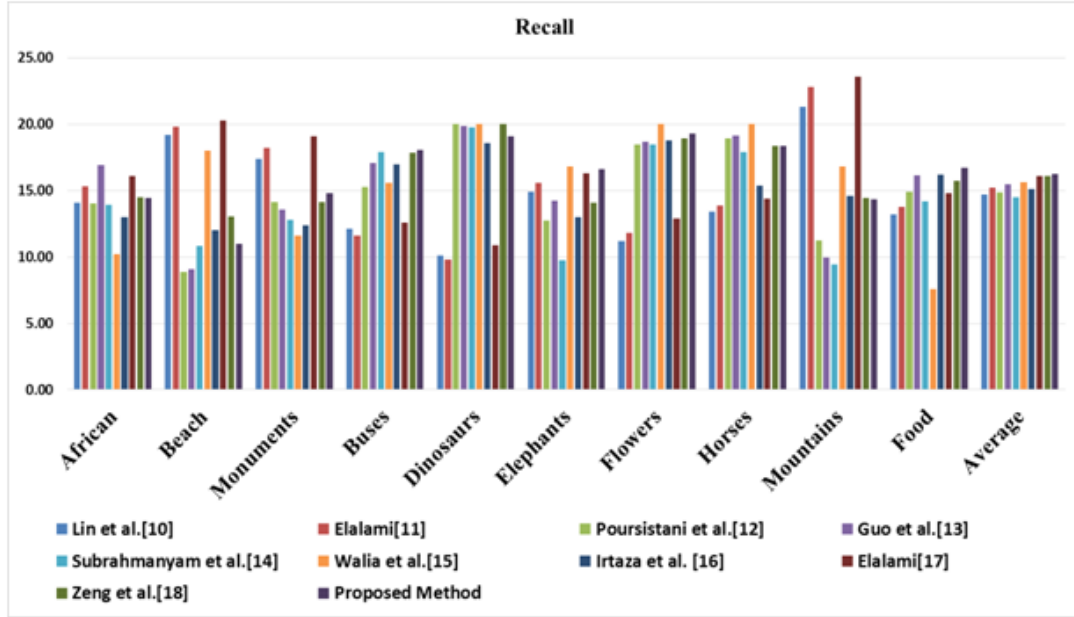
Figure 6. Precision of Proposed Approach Compared to Other Methods for Different Semantic Classes in the Corel Database
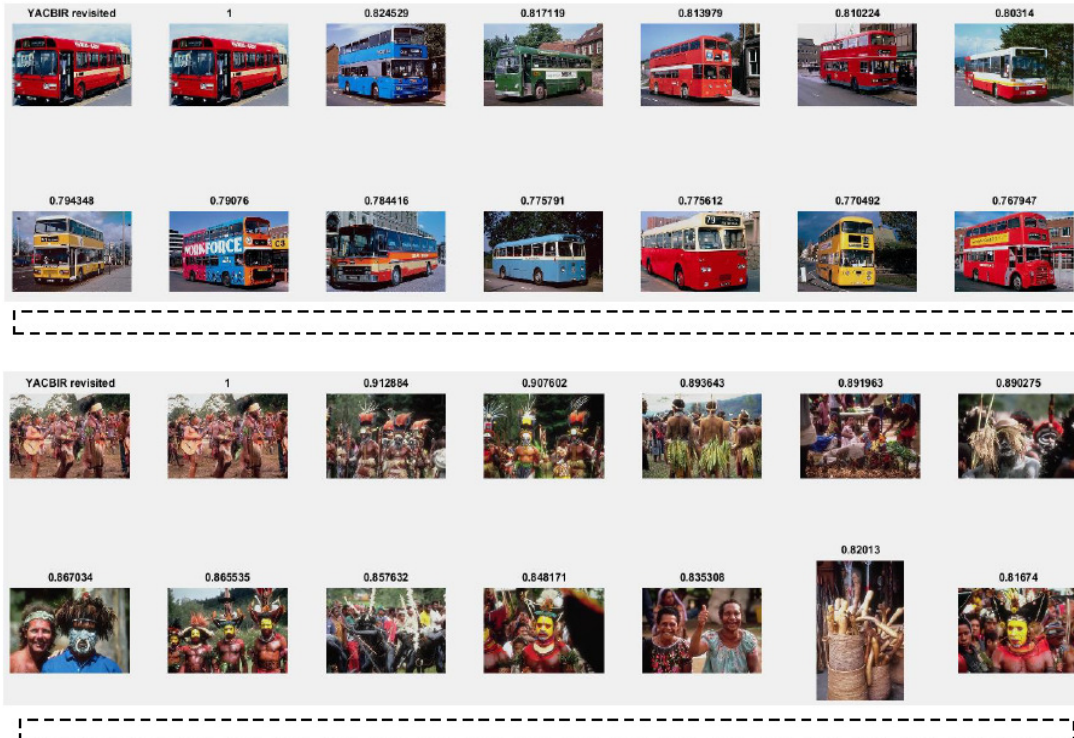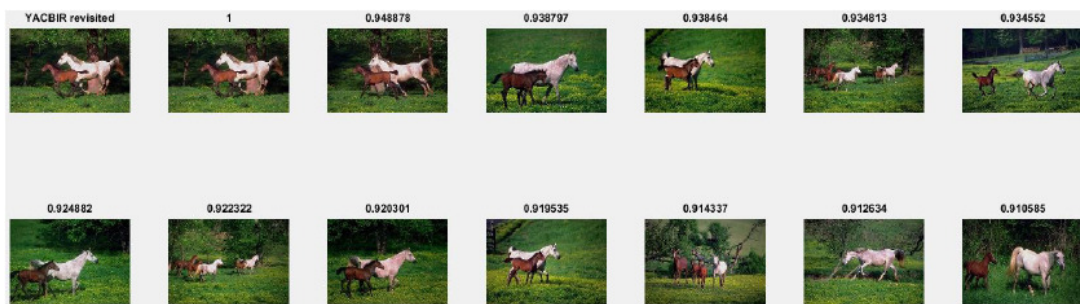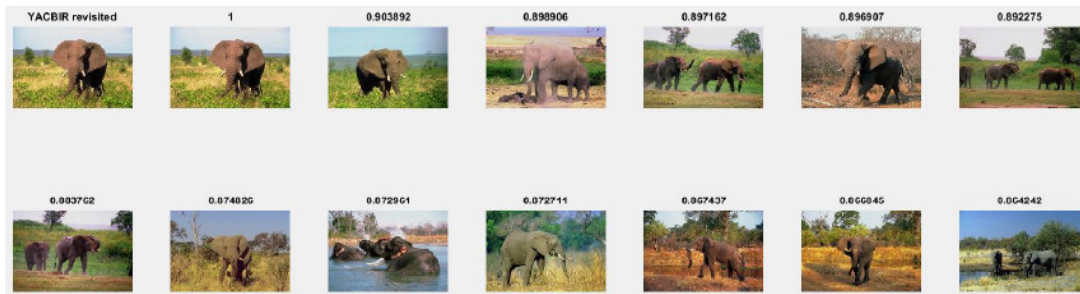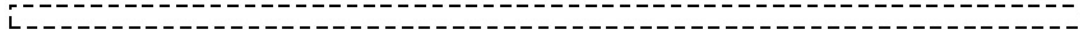
4. CONCLUSION

In this paper, we proposed an efficient learning approach based on multiple SVMs (Support Vector Machines) and YACBIR system. From the test results, we show that retrieval performance of the proposed algorithm has quite good result in the image retrieval system. In future research, other classification techniques can be tested. Artificial Neural Networks (ANN) transfer functions and multiple SVMs kernel can be used to improve the classifier performance.

REFERENCES

[1]   Query by image and video content: The QBIC system" S. W. Teng and G. Lu, "Codebook Generation in Vector Quantization Used for Image Retrieval," International Symposium on Intelligent Multimedia and Distance Education, 1999.

[2]   Y. Liu, D. Zhang, G. Lu, and W. Ma, "A Survey of Content-based Image Retrieval with High-level Semantics," Pattern Recognition, Vol. 40, No. 1, 2007, pp. 262-282.

[3]   Datta, Ritendra, et al. "Image retrieval: Ideas, influences, and trends of the new age." ACM Computing Surveys (CSUR) 40.2 (2008): 5.

[4]   Lew, Michael S., et al. "Content-based multimedia information retrieval: State of the art and challenges." ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 2.1 (2006): 1-19.

[5]   Ms. K.Arthi, Mr. J. Vijayaraghavan. (2013)"Content Based Image Retrieval Algorithm Using Colour Models", Chennai.

[6]   Meenakshi Madhugunki, Dr. D.S.Bormane. (2011)"Comparsion of Different CBIR Techniques", Pune.

[7]   Ait-Aoudia, Samy, Ramdane Mahiou, and Billel Benzaid. "YACBIR: yet another content based image retrieval system." Information Visualisation (IV), 2010 14th International Conference. IEEE, 2010. Esmat

[8]   Rashedi, Esmat, and Hossein Nezamabadi-pour. "Improving the precision of CBIR systems by feature selection using binary gravitational search algorithm." Artificial Intelligence and Signal Processing (AISP), 2012 16th CSI International Symposium on. IEEE, 2012.

[9]   M. Mr.A.H.Karode and M. W.Pawade, "Texture image classification using support vector machine." International Journel of Computing Technology and Applications, 2012.

[10]   C.H. Lin, R.T. Chen, Y.K. Chan, A smart content-based image retrieval system based on color and texture feature, Image Vis. Comput. 27 (6) (2009) 658–665.

[11]   M.E. ElAlami, A novel image retrieval model based on the most relevant features, Knowl.-Based Syst. 24 (2011) 23–32.

[12]   P. Poursistani, H. Nezamabadi-pour, R.A. Moghadam, M. Saeed, Image indexing and retrieval in JPEG compressed domain based on vector quantization, Math.Comput. Model. 57 (5–6) (2013) 1005–1017.

[13]   J.M. Guo, H. Prasetyo, H.S. Su, Image indexing using the color and bit pattern feature fusion, J. Vis. Commun. Image R 24 (2013) 1360–1379.

[14]   M. Subrahmanyam, Q.M.J. Wu, R.P. Maheshwari, R. Balasubramanian, Modified color motif co-occurrence matrix for image indexing and retrieval, Comput. M. Saeed, Image indexing and retrieval in JPEG compressed domain based on vector quantization, Math.A. Irtaza, M.A. Jaffar, E. Aleisa, T.S. Choi, Embedding neural networks for semantic association in content based image retrieval, Multimed. Tool Appl. 72 (2) (2014) 1911–1931.

[15]   E. Walia, A. Pal, Fusion framework for effective color image retrieval, J. Vis.Commun. Image R 25 (2014) 1335–1348.

[16]   A. Irtaza, M.A. Jaffar, E. Aleisa, T.S. Choi, Embedding neural networks for semantic association in content based image retrieval, Multimed. Tool Appl.72 (2) (2014) 1911–1931.Image retrieval using spatiograms of colors quantized by Gaussian Mixture Models. Neurocomputing Volume 171, 1 January 2016, Pages 673–684 72 (2) (2014) 1911–1931.

[17]   M.E. ElAlami, A new matching strategy for content based image retrieval system, Appl. Soft Comput. 14 (2014) 407–418.

[18]   S. Zeng, R. Huang, H. Wang, Z. Kang Image retrieval using spatiograms of colors quantized by Gaussian Mixture Models. Neurocomputing Volume 171, 1 January 2016, Pages 673–684

[19] K. Mikolajczyk and C. Schmid, Scale and affi ne invariant interest point detectors, Int. J. Comput. Vis. 1 (2004), 63 – 86.

[20] T. P. Minka and R. W. Picard, Interactive learning using a society of models, Pattern Recogn. 30 (1997), 565 – 581.

AUTHORS

Lakhdar LAIB received the B Eng and MSc degrees in computer science from École nationale Supérieure en Informatique (Algeria) in 2010 and 2013, respectively. Since 2014, he has been pursuing PhD studies at École nationale Supérieure en Informatique (Algeria) under the supervision of Professor Samy Ait-Aoudia His primary research interests include statistical models and application to image segmentation, computer vision and pattern recognition.

Samy Ait- Aoudia received a DEA "Diplôme d'Etudes Approfondies" in image processing from Saint-Etienne University, France, in 1990. He holds a Ph.D. degree in computer science from the Ecole des Mines, Saint-Etienne, France, in 1994. He is currently a Professor of computer science at the National High School in Computer Science at Algiers/Algeria, where he is involved in teaching BSc and MSc levels in computer science and software engineering. His areas of research include image processing, CAD/CAM and constraints management in solid modeling.
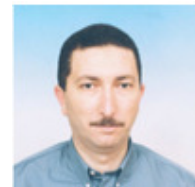
INTENTIONAL BLANK

IMAGE REGISTRATION USING ADVANCED TOPOLOGY PRESERVING RELAXATION LABELING

Chan-IL Kim, Ji-Ae Park, Su Yang, Yoon-Nyun Kim, Hee-Jun Park, and Jong-Ha Lee*

[1]Department of Biomedical Engineering, Keimyung University, Daegu, Korea
segeberg@kmu.ac.kr

ABSTRACT

This paper presents a relaxation labeling technique with newly defined compatibility measures for solving a general non-rigid point matching problem. In the literature, there exists a point matching method using relaxation labeling, however, the compatibility coefficients always take a binary value zero or one depending on whether a point and a neighboring point have corresponding points. Our approach generalizes this relaxation labeling approach. The compatibility coefficients take n-discrete values which measures the correlation between edges. We use log-polar diagram to compute correlations. Through simulations, we show that this topology preserving relaxation method improves the matching performance significantly compared to other state-of-the-art algorithms such as shape context, thin plate spline-robust point matching, robust point matching by preserving local neighborhood structures and coherent point drift.

KEYWORDS

Point pattern matching, graph matching, registration, relaxation labelling

1. INTRODUCTION

The point matching is widely used in computer vision and pattern recognition because point representations are generally easy to extract [1],[2]. Point matching can be categorized as rigid matching and non-rigid matching based on the deformation of objects captured in the images [3]. Compare with the rigid case, non-rigid matching is more complicated. Generally, there are two unknown operations: correspondence and transformation. Most non-rigid point matching approaches use an iterated estimation framework to find appropriate correspondence and transformation [4]. The Iterated Closest Point (ICP) algorithm is one of the most well known heuristic approaches [5]. It utilizes the relationship by assigning the correspondence with binary values zero and one. However, this binary assumption is no longer valid in the case of non-rigid transformation, especially when the deformations are large. The Thin Plate Spline Robust Point Matching (TPS-RPM) algorithm is an Expectation Maximization (EM) algorithm to jointly solve for the feature correspondence as well as the geometric transformation [6]. The cost function that is being minimized is the sum of Euclidean distances between points. In the TPS-RPM, the binary correspondence value of the ICP is relaxed to the continuous value between zero and one. This soft-assign method improves the matching performance because the correspondences are able to improve gradually and continuously without jumping around in the space of binary permutation

matrices [7]. The algorithm is robust compared to the ICP in the non-rigid case, but the joint estimation of correspondences and transformation increases complexity. In addition, the Euclidean distance makes sense only when there are at least rough initial alignments of the shapes. If the initialization is not aligned well, the matching result is poor. Recently, the Shape Context (SC) algorithm has been proposed [8]. It is an object recognizer based on the shape. For each point, the distribution of the distance and orientation are estimated to the neighboring points through a histogram. This distribution is used as the attribute relations for the points. The correspondences can be decided by comparing each point's attributes in one set with the attributes of the other. Because only the attributes are compared, the searching for the correspondences can be conducted more easily compared to the ICP and the TPS-RPM. Generally, the SC performs better in handling complex patterns than the TPS-RPM. Another interesting approach of point matching is a kernel correlation-based point matching [9]. The cost function is proportional to the correlation of two kernel density estimates. The work was extended by using the L2 distance between mixtures of Gaussian representing the point set data in [10]. The Coherent Point Drift (CPD) algorithm is another probabilistic algorithm [11]. They proposed to use the displacement filed between the point sets following the motion coherence theory. They also successfully extend the general non-rigid registration framework, and show that TPS-RPM is its special case. The Robust Point Matching by preserving Local Neighborhood Structures (RPM-LNS) introduces the notion of a neighborhood structure for the general point matching problem [9]. The RPM-LNS uses a relaxation labeling method with binary value coefficient. This approach is based on the assumption that although the absolute distance between two points may change significantly under non-rigid deformation, the neighborhood structure of a point is generally well preserved. The cost function is formulated as an optimization problem to preserve local neighborhood relations.

The relaxation labeling technique is an iterative procedure that reduces local ambiguities and achieves global consistency by exploiting contextual information which is quantitatively represented by "compatibility coefficients" [10], [11]. In this paper, we define optimal compatibility coefficients for a relaxation labeling to solve a non-rigid point matching problem. Our work is inspired by that of [9]. There, point matching problem is formulated to preserve local neighborhood structure during matching process and a relaxation labeling technique is applied to provide optimal solutions. The present work, however, differs from that of [9] in the determination of compatibility coefficients. In [9], the compatibility coefficient values were only considered as binary value "zero" or "one" depending on whether a point and its neighboring point have corresponding points. In this paper, the correlation between points and neighbors is measured by a compatibility function and represents as "n" discrete values. In order to do this, we quantify the amount of smoothness and similarity between points and neighbors using a log-polar diagram. Consider a set of vectors originating from a point to all adjacent points. We compute distance and angle from an origin point to the rest of points by counting the number of bins. This contextual information combined with relaxation labeling process is used to search for an optimized correspondence. After finding the correspondence using several relaxation labeling updates, TPS is used to find the optimal transformation [12]. These two processes are iterated until the optimal correspondence and transformation are found [13]. The proposed relaxation labeling with compatibility function preserves topology, thus we call our method the Topology Preserving Relaxation Labeling (TPRL). The simulation results show that TPRL is stable and robust under all deformations, noises, outliers, rotations, and occlusions.

The reminder of the paper is organized as follows. We review the relaxation labeling problem formulation as an optimization problem in Section II. In Section III, we define new compatibility coefficients and present a relaxation labeling process for point matching problem. We then show the robustness of our approach compared to the other algorithms in Section IV. Finally, Section V discusses the conclusions of the paper.

2. PROBLEM FORMULATION

Let   be a set of control points in a model set and   be a set of control points in the target set. In a matching problem, one-to-one matching is desired, but in general, one-to-one matching is not possible because of outliers. To handle this problem, two point sets are augmented to and by introducing a dummy or a nil point. Then common points can be matched one-to-one and outliers can be matched to a dummy point. Every point has its adjacent points called neighbors. For a given point, one can select adjacent points, is the total number of adjacent points of. Similarly, for a point , adjacent points are and is the total number of adjacent points of .

In this paper, the matching problem is formulated as the graph edge matching problem [5]. Each point is a node of a graph, and a point and its adjacent points make edges of a graph. Then the problem is to maximize the number of matched edges between two graphs. For this purpose, we determine the fuzzy correspondence matrix. Each entry of has continuous value between [0,1] that indicates the weight of the correspondence between and .Then the optimal match is found by maximizing the energy function as follows.

$$\hat{P} = \arg \max_{P} C(\hat{S}, \hat{T}, P) \tag{1}$$

where

$$C(\hat{S}, \hat{T}, P) = \sum_{i=1}^{I} \sum_{x=1}^{X} \sum_{j=1}^{J} \sum_{y=1}^{Y} P_{s_i t_j} P_{a_x^{s_i} a_y^{t_j}} \tag{2}$$

The above equations are subject to the following constraints $\sum_{j=1}^{J+1} P_{s_i t_j} = 1$ for $i = 1, 2, ..., I$ and

$\sum_{i=1}^{I+1} P_{s_i t_j} = 1$ for $j = 1, 2, ..., J$.

Defining adjacent points are very important in maximizing the matching probability and reducing the computational complexity. In this paper, an adjacent point searching boundary   is assumed for any given point. The searching boundary   is defined as a circle with the radius of the median length of all edges. If a point is residing in the circle centered at s_i , it becomes an adjacent point a^{s_i} of s_i . The median length of total edges is an efficient metric for determining the search boundary, if points are selected non-uniformly. If points are placed randomly, however, there are both crowded areas and sparse areas. Logically, a point in a crowded area should have more adjacent points and thus more connected edges. A point in a sparse area should have less adjacent points with less connected edges. If we assign same number of adjacent points to each point, a point in a sparse area may be associated with a distant point in a crowded area as the adjacent point. Furthermore, points in a crowded area may not necessarily form edges, and points in a sparse area may form useless edges. This unnecessary assignment leads to increased computational complexity. The proposed adjacent point definition allows the assignment of more connected edges in a crowded area and fewer connected edges in a sparse area. This flexibility reduces the complexity and increases the reliability of the algorithm.

3. POINT CORRESPONDENCE USING RELAXATION LABELING

Here, we use a relaxation labeling technique to solve an optimization problem of non-rigid point matching. The relaxation labeling is an iterative process to reduce ambiguities in assigning labels to a set of objects by incorporating the contextual information. The contextual information is

usually represented as compatibility coefficients. It is widely known that the relaxation labeling process is greatly affected by the choice of the compatibility coefficients [14],[15]. Thus, we propose compatibility coefficients that preserve a topology during the matching processes.

Initially, each point is assigned with a set of matching probability based on the shape context. After the initial probability assignment, the relaxation labeling process updates the matching probability iteratively. The purpose of the subsequent process is to assign matching probability that maximize $C(\hat{S}, \hat{T}, P)$ under the relaxed condition as $P_{s_i t_j} \in [0,1]$. At the end of the labeling process, it is expected that each point will have one unambiguous matching probability. We follow the relaxation updating rule as follows [16].

1. Compute the compatibility coefficients which impose smoothness and similarity constraints between points.

2. Compute the support function from all compatibility coefficients related to the point.

3. Update the matching probability depending on its support function.

The determination of compatibility coefficients is important because the performance of the relaxation labeling depends on them. In the previous work, the compatibility coefficients are defined as either 0 or 1, a binary value [9]. As a key contribution, we define new compatibility coefficients to relax this binary value into the multiple discrete values. In our approach, the compatibility coefficients quantify the degree of agreement between the hypothesis that s_i matches to t_j and $a_x^{s_i}$ matches to $a_y^{t_j}$. To represent the matching strength between the pair of points and the pair of adjacent points, we use a log-polar diagram. Let the diagram distance and angle originating from s_i to $a_x^{s_i}$ be $d(s_i, a_x^{s_i})$ and $l(s_i, a_x^{s_i})$, respectively. Every point can be an origin of a shape context and the origin varies with points in consideration to calculate the diagram distance and angle. In the diagram, the distance is defined as 0 in the origin and incremented by 1 towards the outer bins. The bins are uniform in log-polar space, which makes the descriptor more sensitive to positions of adjacent points than to those of points far away. For the geometrical angle calculation, the original shape context is not rotation invariant. To initially align, the mass center of a point set is used as a reference point, and we use the direction from a point to the center of mass as the positive x-axis [9]. From this axis, the geometrical angle is incremented by 1 as the counter clockwise direction. An example is shown in Fig. 1. In this example, all adjacent points of s_1 and t_1, which are $a_1^{s_1}$, $a_2^{s_2}$, $a_1^{t_1}$ and $a_2^{t_2}$, have the magnitude of distance of two and the angle of one.
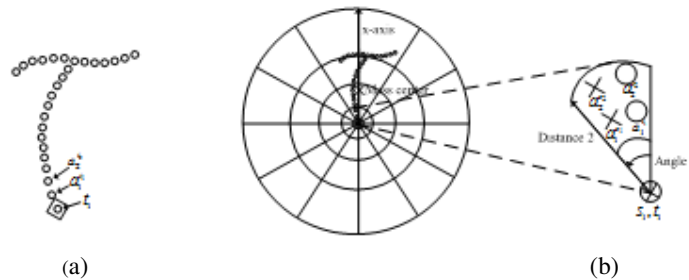

(a)   (b)

Fig. 1. Structure based distance and angle computation. (a) Sample points from the deformed shape "T". (b) Example shape context of a reference sample marked by ◇. The log-polar bins diagram is used to compute structural distance and angle. We use five bins for the distance 0 to 4 from the origin and the angle 0 to 12 from the x-axis.

The compatibility coefficients are represented by the smoothness and the similarity between the pair of points and the pair of adjacent points. In the non-rigid deformation, although the absolute distance between two points may change significantly, the adjacent point structure is generally well preserved. This happens more strongly if the distance between point and its adjacent point is closer. Thus the first constraint of the compatibility coefficients, smoothness, is measured by the diagram distance between s_i and $a_x^{s_i}$ as below.

$$\alpha(s_i, a_x^{s_i}) = \left(1 - \frac{d_i(s_i, a_x^{s_i})}{\max(d_i(s_i, a_x^{s_i}))}\right) \tag{4}$$

where $\max(d_i(s_i, a_x^{s_i}))$ is the longest edge in the graph. Two points s_i and $a_x^{s_i}$ are the most salient if $\alpha(s_i, a_x^{s_i})$ is 1 and the least compatible if $\alpha(s_i, a_x^{s_i})$ is 0. The second constraint, similarity, is measured by the combined difference of structural distance and angle of initial edge $(s_i, a_x^{s_i})$ and the final edge $(t_j, a_y^{t_j})$. This constraint imposes that if an edge $(s_i, a_x^{s_i})$ has less distance and angle differences with the edge $(t_j, a_y^{t_j})$, then it has higher probability of matching. The disparity between edges $(s_i, a_x^{s_i})$ and $(t_j, a_y^{t_j})$ is defined as follows.

$$\beta(s_i, a_x^{s_i}; t_j, a_y^{t_j}) = \left(1 - \left|\frac{d_i(s_i, a_x^{s_i}) - d_j(t_j, a_y^{t_j})}{\max\{d_i(s_i, a_x^{s_i}) - d_j(t_j, a_y^{t_j})\}}\right|\right) \tag{5}$$

$$\gamma(s_i, a_x^{s_i}; t_j, a_y^{t_j}) = \left(1 - \left|\frac{l_i(s_i, a_x^{s_i}) - l_j(t_j, a_y^{t_j})}{\max\{l_i(s_i, a_x^{s_i}) - l_j(t_j, a_y^{t_j})\}}\right|\right) \tag{6}$$

If two edges $(s_i, a_x^{s_i})$ and $(t_j, a_y^{t_j})$ have same geometrical distance and angle, then they are the most compatible. We define the total compatibility coefficient by

$$r_{s_i t_j}(a_x^{s_i} a_y^{t_j}) = \alpha(s_i, a_x^{s_i}) \cdot \beta(s_i, a_x^{s_i}; t_j, a_y^{t_j}) \cdot \gamma(s_i, a_x^{s_i}; t_j, a_y^{t_j}) \tag{7}$$

Clearly, $r_{s_i t_j}(a_x^{s_i} a_y^{t_j})$ ranges from 0 to 1. High value of $r_{s_i t_j}(a_x^{s_i} a_y^{t_j})$ corresponds to high matching probability between $(s_i, a_x^{s_i})$ and $(t_j, a_y^{t_j})$, and the low value corresponds to incompatibility. The support function $q_{s_i t_j}$ in the k-th iteration is given by

$$q_{s_i t_j}^k = \sum_{i=1}^{I} \sum_{j=1}^{J} r_{s_i t_j}(a_x^{s_i} a_y^{t_j}) p_{a_x^{s_i} a_y^{t_j}}^k = \sum_{i=1}^{I} \alpha(s_i, a_x^{s_i}) \sum_{j=1}^{J} \beta(s_i, a_x^{s_i}; t_j, a_y^{t_j}) \cdot \gamma(s_i, a_x^{s_i}; t_j, a_y^{t_j}) p_{a_x^{s_i} a_y^{t_j}}^k \tag{8}$$

Note that we use $p_{a_i^m a_j^k}^k$ to weight $r_{s_i t_j}(a_x^{s_i} a_y^{t_j})$ because the support also depends on the likelihood of pair of adjacent points matching probability. Finally, $p_{s_i t_j}^k$ is updated according

$$p_{s_i t_j}^{k+1} = p_{s_i t_j}^k q_{s_i t_j}^k / \sum_{j=1}^{J} p_{s_i a_i^{t_j}}^k q_{s_i a_i^{t_j}}^k \tag{9}$$

The iterative process can be summarized as follows. If a matching probability between s_i and t_j is supported from their adjacent points $a_x^{s_i}$ and $a_y^{t_j}$, then the chance of being matched increases. The probability decreases if they have relatively less supporting data from their adjacent points.

After pre-defined iteration, the estimated matching probability is assigned to every point. The points with maximum matching probability less than 0.95 are labeled as outliers by matching them to dummy points. The outlier rejection is performed throughout the optimization process.

The strategy of iterated point correspondence and transformation estimations is widely used for non-rigid point matching [6]. In this work, we use the thin plate spline (TPS) model which is used for representing flexible coordinate transformations. The matched point pairs from several relaxation updates are used to estimate the parameters of TPS deformation model. The estimated parameters are used to transform the model set to bring to close to the target set. Suppose point (x_i, y_i) is matched to (u_i, v_i) for $i = 1, 2, ..., n$, let $z_i = f(x_i, y_i)$ be the target function value at location (x_i, y_i), we set z_i equal to u_i and v_i in turn to obtain one continuous transformation for each coordinate. The TPS interpolant $f(x, y)$ minimizes the bending energy

$$I(f) = \iint_{\mathbb{R}^2} \left(\frac{\partial^2 f}{\partial x^2} \right)^2 + 2\left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 f}{\partial y^2} \right)^2 dxdy$$

And has the solution of the form

$$f(x, y) = a_1 + a_x x + a_y y + \sum_{i=1}^{n} w_i U(\| (x_i, y_i) - (x, y) \|)$$ where $U(r)$ is the kernel function,

taking the form of $U(r) = r^2 \log r^2$. If error appears in the matching results, we use regularization to trade off between exact interpolation and minimizing the bending energy as follows.

$$H[f] = \sum_{i=1}^{n} (v_i - f(x_i, y_i))^2 + \lambda I_f$$

The regularization parameter λ controls the amount of smoothing. Finding correspondence and transformation process is alternatively iterated. Finally, one finds the desired correspondence matrix and the mapping function

4. VALIDATION, COMPARATIVE ANALYSIS AND PERFORMANCE EVALUATION

In order to access the performance of the Topology Preserving Relaxation Labeling (TPRL) point matching method, we compare our matching results with state-of-the-art algorithms such as the Shape Context (SC), the Thin Plate Spline-Robust Point Matching (TPS-RPM), the Robust Point Matching by preserving Local Neighborhood Structure (RPM-LNS), and Coherent Point Drift (CPD) method. The Iterated Closest Point (ICP), the correlation-based approach, and the L2 distance approach was not considered in this paper since they have much larger registration errors compared to the algorithms mentioned above

4.1. Experiments Based on Synthetic Data

We test algorithms with respect to different degrees of deformation, different amount of noise and outliers, and different rotation. We use two different shape. First model set is totally selected 96 points along the outline of a fish shape, closed contour shape. The second model set uses 108 points points from a Chinese character ("blessing"), an open contour shape. It is a more complex pattern than a fish shape.

In each test, one of distortions is applied to a model set to create a target set. In the first test, the model set is warped through progressively larger degrees of non-rigid warping. The warped data sets are used as the target set. In the second test, we add certain amount of Gaussian noise to the moderately deformed model set and obtain the new target set. The purpose is to test the algorithms' tolerance to noise. In the third test, different amount of outliers were added to the deformed model sets. Finally, we rotate deformed model sets in a counter clockwise direction for the rotation test. Fig. 2. shows examples of synthesized data sets [6]. A total of 100 samples are generated in each distortion level. The matching performance of each algorithm is compared by mean and standard deviation of RMS distance error over 100 trials in each distortion level.
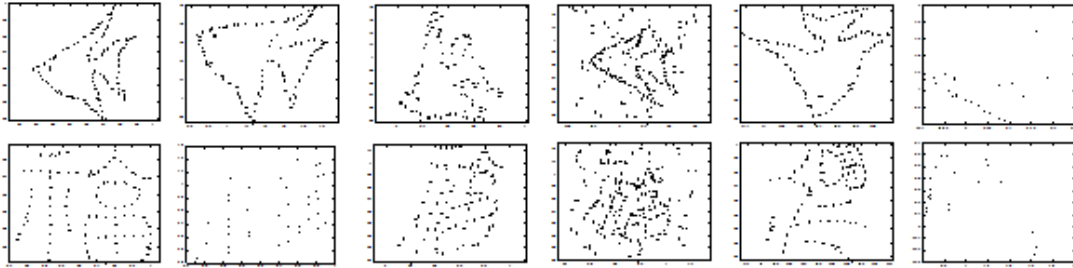

(a) Original   (b) Deformation   (c) Noise   (d) Outlier   (e) Rotation   (f) Occlusion

Fig. 2. Synthesized data sets for statistical tests. Model sets are shown in the first column. Column 2-5 shows the target sets for the deformation, noises, outliers, rotations, and occlusion.

We use root mean square (RMS) error for the registration error metric [19]. It is computed as the registration error between the last warped model set and the target set, and is defined by the following equation:

$$RMS = \sqrt{\frac{1}{M} \sum_{i=1}^{M} (s_i - t_i)^2}$$ (11)

where M is the total number of points. The statistical matching results (error means and standard deviations for each setting) are shown in Fig. 3.
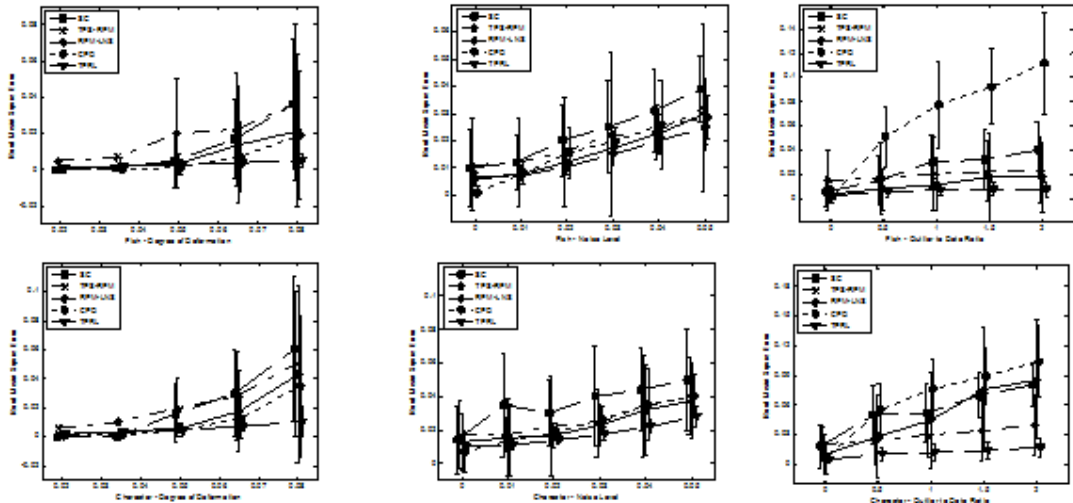
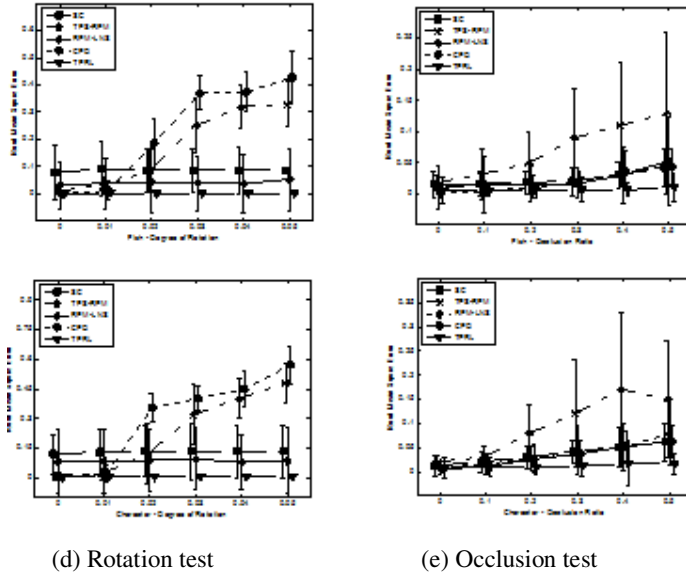(a) Deformation test   (b) Noise test   (c) Outlier test

(d) Rotation test   (e) Occlusion test

Fig. 3. Comparison of the matching performance of TPRL (∇) with shape context ($\square$), TPS-RPM (*), RPM-LNS ($\Diamond$), and CPD (O) for the fish and character shape, respectively. The error bars indicate the standard deviation of the error over 100 trials.

2.1.1. Deformation

In this section, we evaluate algorithm performances under different degree of deformations. Fig. 3a show the results for the four matching strategies on the fish data set and the character data set. According to the results, TPRL method outperforms the other algorithms, SC, TPS-RPM, RPM-LNS, and CPD method. The standard deviation as well as the mean of RMS error is significantly lower than other algorithms. At 0.02 degree of deformation, five algorithms achieve very similar matching performance in both fish and character shape. However, as the degree of deformation increases, we observe that the SC, TPS-RPM, RPM-LNS, and CPD method give poor results while the TPRL method shows the robustness to large deformations. In fact, it seems that the TPRL method is not much affected by deformations. These results show the robustness of the TPRL method under large deformed data sets.

2.1.2. Noise

The presence of noise makes the point's location ambiguous. Therefore, this type of data is more challenging than the deformation data. In this section, the performance is measured with a significant amount of noise. Fig. 3b presents the results for matching strategies for fish and character shapes. The results show that all algorithms are affected by this type of data distortion. However, we notice that the TPRL method compensates the location ambiguity and finds more accurate correspondences than other three. In contrast, the result of the SC method is the most influenced by noise among four algorithms. The descriptor alone in the SC confuses the real data point and noise since there is no mechanism to distinguish data point and noise in the same descriptor bin. The TPRL method reduces this ambiguity by using topology preserving scheme.

2.1.3. Outliers

In addition to the deformation and noise, present outlier further complicates the point matching problem. To evaluate the performance of our method with outliers, we added maximum 405 points outliers in the target set. Fig. 3c displays the RMS error with respect to outlier levels. From

the result, we can notice that SC in both shapes and TPS-RPM in a character shape are easily confused by outliers and start to fail once the outlier level becomes relatively high. TPS-RPM in the fish shape and RPM-LNS in both shapes seem not much affected by outliers but the error is still higher than TPRL. TPRL is very robust regardless of the outlier level.

2.1.4. Rotation

In Fig. 3d, we evaluate the influence of rotation. From the result, we notice that the applied transformation (rotation) does not affect the performance of SC, RPM-LNS and TPRL. All error curves except TPS-RPM are relatively constant at approximately 0.08 mm for SC, 0.03 mm for RPM-LNS and 0.002 mm for TPRL in a fish shape; and 0.12 mm for SC, 0.08 mm for RPM-LNS and 0.005 mm for TPRL in the character shape. Note that until 30 degrees of rotation, the error of TPS-RPM is lower than SC and RPM-LNS. But from 90 degree of rotation, TPS-RPM deteriorates quickly and finally reaches 0.33 ± 0.085 mm for the fish shape and 0.48 ± 0.1 mm for the character shape. TPRL is rotation invariant and consistently outperforms three other algorithms in all degrees of rotations.

2.1.5. Occlusion

Occlusion is also important degradation in real applications. We use six occlusion levels to test the five algorithms. As shown in Fig 3e, the performance of the TPS-RPM algorithm deteriorates quickly compared to the other four algorithms. The TPRL algorithm preserves the structure well under the large occlusion ratio and achieves the best result in both fish and character shape synthetic data.

2.2. Convergence of Relaxation Labeling

According to the consistency theory, the relaxation labeling process increases the overall consistency by reducing local ambiguities [11],[20]. During the experiments, we verified that the labeling update process converge to an unambiguous solution starting from an arbitrary initialization. Fig. 4 below shows the matching probability of four randomly selected points in 105 105 point matching experiments in IV-A. We can observe that while 300 iterations of relaxation labeling updates, the ambiguity of matches decreases and 104 points goes to zero except one point which find their correspondence and goes to one. From the experiments, most elements of matrix do converge to zero to one.

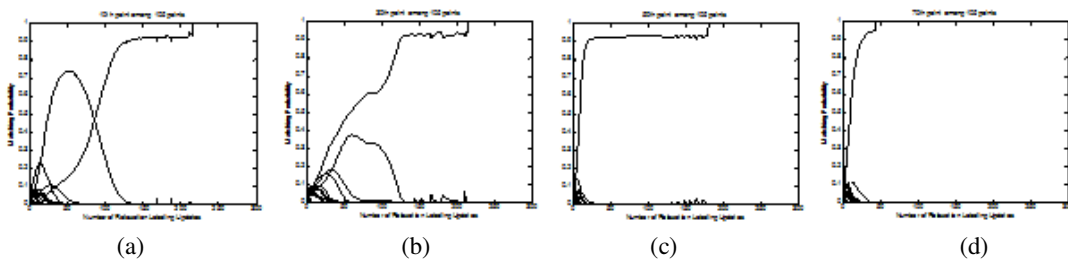| (a) | (b) | (c) | (d) |

Fig. 4. The matching probability of (a) 10th, (b) 30th, (c) 50th, (d) 70th points among 105 points in 105 105 matching experiments.

2.3. Experiments Based on Real World Data Set

We also conducted experiments on real image data sets. For the real world experiments, we obtained ten frames (10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th, 90th, 100th frame) of Carnegie

Mellon University hotel sequence as shown in Fig. 4. The database consists of 100 frames of a moving sequence of a toy hotel. Totally 100 points were manually selected in each of the images. Then we conducted four different experiments, 100 against 100, 100 against 90, 100 against 80, and finally 100 against 70 points for every ten image frames defined above.

Fig. 4. Sequence images of toy hotel. Frames 10, 20, 30, 40, 50 (top). Frames 60, 70, 80, 90, 100 (bottom).

Fig. 5 shows the results for these experiments. The average RMS distance between each pair and its standard deviation is taken for the statistics.
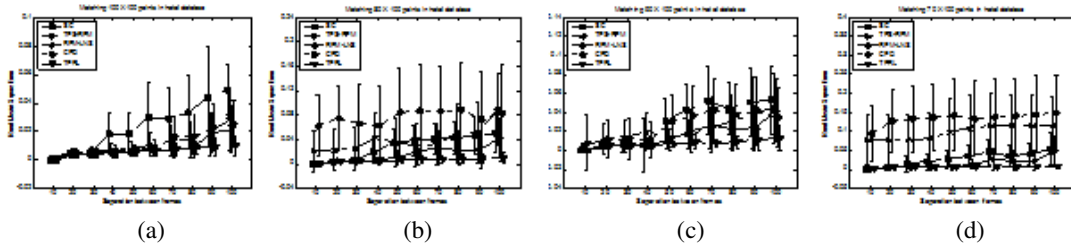

(a)   (b)   (c)   (d)

Fig. 5. Comparison of the matching performance of TPRL (∇) with shape context ($\square$), TPS-RPM (*), RPM-LNS ($\Diamond$), and CPD (O) in the hotels sequence for increasing frame separation and different target set sizes ((a) 100 100, (b) 90 100, (c) 80 100, (d) 70 100). Error bars correspond to the standard deviation of each pair's RMS error.

In these experiments, we notice that as the relative sizes of the patterns become largely different, the TPRL algorithm performs significantly better than the competing algorithms. The outperformance of the TPRL algorithm is also observable when the target image sequence number is increasing.

2.4. Experiments Based on Large Data Set

We also test the TPRL performance under much larger data set. For the model set, 1000 points are extracted from the hey image shown in Fig. 10. To distort the data set, Gaussian random ground function is used. After non-rigid deformation, we have deleted 500 points for one data set and we have added 1000 to 2000 outliers which was generate outliers randomly from a normal zero-mean distribution for other data set. The TPRL performs well under large point sets.

2.5. Processing Times

In order to compare algorithms, it is necessary to analyze the complexity of each algorithm. Assume that both model set and target set have N points. The algorithms are based on the NP-hard problem and have similar computational complexity of $O(N^2)$ for matching in $\mathbb{R}^2$ Among the compared algorithms, the CPD method performs the fastest registration time under large distortion of data set. The TPRL algorithm takes slightly longer time to compute compared

to CPD algorithm. For 105 points of model set and deformed target set matching, the TPRL takes about 2.1 seconds on a desktop PC with Intel Core 2 Duo @ 2.13GHz and 2GB RAM

5. CONCLUSIONS

A novel approach called the TPRL algorithm for the non-rigid point matching problem using a relaxation labelling method with newly defined compatibility coefficients is presented. The matching probability is updated by relaxation process taking into account of the proposed compatibility. A comparative analysis is performed against other algorithms, SC, TPS-RPM, RPM-LNS, and CPD. The TPRL algorithm shows superior performance in both closed (fish) and opens (character) contour examples. We also showed that the TPRL algorithm always obtained the highest matching results compared to other four algorithms under different degree of deformation, noises, outliers, rotation, and occlusion.

ACKNOWLEDGEMENTS

REFERENCES

[1]   L.G. Brown(1992), "A Survey of Image Registration Techniques," ACM Computing Surveys, vol. 24, no. 4, pp. 325-376.

[2]   H. J. Johnson and G. E. Christensen(2002), "Consistent Landmark and Intensity-Based Image Registration," IEEE Trans. Medical Imaging, vol. 21, no. 5, pp. 450-461.

[3]   B. Zitova and J. Flusser(2003), "Image Registration Methods: A Survey," Image and Vision Computing, vol. 21, no 11, pp. 977-1000.

[4]   A. Rangarajan, H. Chui, and F. Bookstein(1999), "The Softassign Procrustes Matching Algorithm," Medical Image Analysis, vol. 4, pp. 1-17.

[5]   P. J. Besl and N.D. Mckay(1992), " A Method for Registration of 3-D Shapes," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 14, no. 2, pp. 239-256.

[6]   H. Chui and A. Rangarajan(2003), "A New Point Matching Algorithm for Non-Rigid Registration," Computer Vision and Image Understanding, vol. 89, no. 23, pp. 114-141.

[7]   A. Rangarajan, S. Gold, E. Mjolsness(1996), "A Novel Optimizing Network Architecture with Applications," Neural Computing, vol. 8, no. 5, pp. 1041-1060.

[8]   S. Belongie, J. Malik, and J. Puzicha(2002), "Shape Matching and Object Recognition Using Shape Contexts," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 4, pp. 509-522.

[9]   Y. Tsin and T. Kanade(2004), "A Correlation-Based Approach to robust Point Set Registration," Proc. 8th European Conf. Computer Vision, pp. 558-569.

[10] B. Jian and B. Vemuri(2005), "A Robust Algorithm for Point Set Registration Using Mixture of Gaussians," Proc. 10th IEEE Int'l Conf. Computer Vision, pp. 1246-1251.

[11] A. Myronenko, X. Song, and M.A. Carreira-Perpinan(2007), "Non-rigid Point Registration: Coherent Point Drift," Advances in Neural Information Processing Systems, vol. 19, pp. 1009-1016.

[12] Y. Zheng and D. Doermann(2006), "Robust Point Matching for Nonrigid Shapes by Preserving Local Neighborhood Structures," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 28, no. 4, pp. 643-649.

[13] A. Rosenfeld, R.A. Hummel, and S.W. Zucker(1976), "Scene Labeling by Relaxation Operations," IEEE Trans. System, Man, and Cybernetics, vol. 6, no. 6, pp. 420-433.

[14] R.A. Hummel and S.W. Zucker(1983), "On the Foundations of Relaxation Labeling Processes," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 5, no. 3, pp. 267-287.

[15] F.L. Bookstein(1989), "Principal warps: Thin-plate Splines and the decomposition of deformations," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 11, no. 6, pp. 567-585.

[16] A.D.J. Cross and E.R. Hancock(1988), "Graph Matching with a Dual-Step EM Algorithm," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 11, pp. 1236-1253.

[17] S. Peleg and A. Rosenfeld(1978), "Determining Compatibility Coefficients for Curve Enhancement Relaxation Processes," IEEE Trans. System, Man, and Cybermetrics, vol. 8, no. 7, pp. 548-555.

[18] M. Pelillo and M. Refice(1994), "Learning Compatibility Coefficients for Relaxation Labeling Processes," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 16, no. 9, pp. 933-945.

[19] Q.X. Wu and D. Pairman(1995), "A Relaxation Labeling Technique for Computing Sea Surface Velocities from Sea Surface Temperature," IEEE Trans. Geoscience and Remote Sending, vol. 33, no. 1, pp. 216-220.

[20] A.L. Yulle and N.M. Grzywacz(1989), "A Mathematical Analysis of the Motion Coherence Theory," International Journal of Computer Vision, vol. 3, no. 2, pp. 155-175.

[21] K. Mikolajczyk and C. Schmid(2005), "A Performance Evaluation of Local Descriptors," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 27, no. 10, pp. 1615-1630.

[22] V. Zagrodsky, V. Walimbe, C.R. Castro-Parehja, J.X. Qin, J.M. Song, and R. Shekar(2005), "Registration-assisted Segmentation of Real-time 3-D Echocardiographic Data using Deformable Models," IEEE Trans. Medical Imaging, vol. 24, no. 9, pp. 1089-1099.

[23] S.W. Zucker, E.V. Krishnamurthy, and R.L. Haar(1978), "Relaxation Processes for Scene Labeling: Convergence, Speed, and Stability," IEEE Trans. Systems, Man and Cybernetics, vol. 8, no. 1, pp. 41-48.

ONTOLOGY BASED TELE-HEALTH SMART HOME CARE SYSTEM: ONTOSMART TO MONITOR ELDERLY

L. Nachabe[1,3], M. Girod-Genet[1], B. ElHassan[2] and J. Khawaja[2]

[1]Department of Réseaux et Services Multimédia Mobiles,
Telecom SudParis University, Evry, France
`marc.girod_genet@telecom-sudparis.eu`
[2]Department of Electricity and Electronics, Faculty of Engineering, Branch 1,
Lebanese University, Tripoli, Lebanon.
`bachar_elhassan@ul.edu.lb`
[3]American University of Culture & Education, Faculty of Science,
Beirut, Lebanon.
`Linanachabe@auce.edu.lb`

ABSTRACT

The population ageing is a demographical phenomenon that will intensify in the upcoming decades, leading to an increased number of older persons that live independently. These elderly prefer to stay at home rather than going to special health care association. Thus, new tele-health smart home care systems (TSHCS) are needed in order to provide health services for older persons and to remotely monitor them. These systems help to keep patients safe and to inform their relatives and the medical staff about their status. Although various types of TSHCS already exist, they are environment dependent and scenario specific. Therefore, the aim of this paper is to propose sensors and scenarios independent flexible context aware and distributed TSHCS based on standardized e-Health ontologies and multi-agent architecture.

KEYWORDS

Smart-Home, tele-Health, Multi-Agent, distributed systems, WSN, Ontology, fall, Indoor Localization

1. INTRODUCTION

The percentage of people classified as elderly is increasing in all major regions of the world. The definition of elderly is not standard and may differ from one country to another. While for most developed countries the age of 65 years is adopted as a definition of elderly, this number may decreases to 60 or even 55 in third world countries (Africa for example). The UN has not adopted a standard criterion, but generally uses 60+ years to refer to the older population [1]. According to the World Health Organization (WHO), an increase from 600 million old people in 2000 to 2 billion by 2050 is expected. By 2050, 22% of the world's population will be over 60. In Europe, the percentage of elderly will be over 30%. These numbers are highlighted by the life expectancy at birth parameter. In addition, according to the WHO, approximately, 10% of the population experience some form of disability. Already 21% of people above the age of 50 have severe vision, hearing and/or mobility problems [2].

In general, most elderly people prefer to grow old at their own home, which is known as "aging-in place" [3]. In fact, at this age, the increasing risks of falls and strokes could threaten their live. Thus, to grow at home, vital signals should be remotely monitored to keep the elderly safe. Thanks to the evaluation of wireless senor networks, and the explosion of new embedded systems, especially wearable devices dedicated for biomedical usage, the tele-health home care systems (THCS) are realizable. According to the Market Reports Hub press release, the global tele-health market will reach $6.5 billion in 2020 at an annual growth rate of 24.2% [4]. Medical devices for self- health monitoring are expected to form 80% of wireless devices in 2016 [5].

However, these THCs are facing paramount concerns [6]. On one hand, these concerns are related to the tremendous types of sensor/actuators, communication protocols, scenarios, vendors, and service providers. In other terms, these systems are very heterogeneous, vendors dependent, non-scalable (integration of new sensors/actuators and services need technical support), and hardly upgradable (if new technology appears in the market, the overall system should be re-implemented). On the other hand, these systems should provide context-aware services capable of creating and adjusting the offered services based on the occupant requirements, preferences and situation. Moreover, social implication of tele-health system is still at the heart of the THCs barriers where the privacy and security are mandatory. Finally, these systems should offer reliability, real time monitoring, friendly user interfaces with minimum expert intervention and reasonable resource consumption (in terms of energy, processing and bandwidth). Therefore, "smart" THCS should be introduced (TSHCS). In this context, we define "smart" as the capability of the system to recognize new components (automatic nodes discovery), discover new services, and infer new information from basic collected raw data from sensors.

In this paper, we present a new TSHCS designed using ontologies to describe generated data and services. The usage of ontologies paves the way to an interoperable solution and tackles the problem of heterogeneity within the same system, and between different systems. Moreover, multi-agent architecture is introduced in our proposed TSHCS system in order to distribute the processing among different components. This leads to a flexible and more efficient power consumption solution that the existing TSHCS. All these aforementioned ideas will be clarified in the next sections.

The paper is organized as follow. In Section II, the reasons behind using ontologies and multi-Agent architectures are clarified. Overview of the retained ontologies is also presented in this section. Section III describes the general architecture of our proposed TSHCS system, called OntoSmart system. Section IV details our OntoSmart based applications called OntoSmartHome and dedicated for the patient and his/her relatives. Moreover, tests and analysis are given in this section. Finally, Section V concludes the paper.

2. BACKGROUND

2.1 Reason behind using ontologies

As many THCSs are proposed [7], different terminologies are used to describe the same properties or object. Moreover, various ranges of sensors/actuators (ambient sensors, electrical devices, biomedical sensors, etc.) from wide range of vendors supporting different wireless technologies (Bluetooth LE, LoRa, ZigBee, etc.) are used [7]. This variety caused the problem of heterogeneity management and interoperability in THCSs. In general, these systems are monolithically deployed. The integration of a new device or a new service needs the reconfiguration of the overall system, and some time the deployment of a new system [8]. For that reason, new techniques that deal with the data meaning and enhance the automatic node and service discovery are needed. These techniques are named Semantic techniques where an entity presents an aspect of the real domain described by metadata (data about data). Thus, by using semantic techniques within THCS, the raw data can be annotated by semantic data respecting a

predefined semantic model to unify its description [9]. In that way, the problem of interoperability between different systems can be resolved. Ontologies are the most powerful semantic techniques. Using the formal definition, ontologies are "tools for specifying the semantics of terminology system in a well-defined and unambiguous manner" [10]. W3C recommended the use of OWL language to describe ontologies [11] . In fact, the most powerful point behind using OWL is the capability of reasoning in order to infer more significant data.

Many researches have introduced the usage of ontologies in TSHCSs. Nguyen et al. [12] proposed to formalize smart home context-aware system using ontologies. The ontology was installed on the web server. The ontology describes the habitant, the environment parameters and the activity done by the habitant. But, to effectively resolve the problem of interoperability, advanced description about the devices (especially the sensors/actuators) is needed [13]. For example, the used communication protocols, the source of energy, the memory components, etc. Moreover, a clear mechanism for automatic node/service discovery should be defined. Chen et al. [14] proposed a three layers architecture (Semantic Layer, Intelligent Layer and Application Layer) for TSHCSs. The Semantic layer is responsible for the annotation of raw data based on an ontology that describes: the home entities, the living space, the location, the person and his/her behavior. Moreover, this layer adds relations between different classes. The Intelligent Service Layer offers the capability of interoperability and high level automation. The application Layer is left to the developer. Although Chen et al. proposed a general architecture; they however neither give details about the implementation of this architecture nor evaluate it using a real test-bed. Latfi et al. [9] also introduced ontologies to model TSHCSs in order to assist patient in accordance to his/her behavior. They conceived seven ontologies that mainly describe the patient's history, equipment, software application, and behavior. These proposed ontologies still need to be deployed in real scenarios. Ye et al. [15] focused on modeling TSCHSs using three main classes: Concept, Context and Activity. The novelty of their resulting ontology is that it shows how the use of semantic rule can enhance the inference of new significant knowledge. In summary, all previously mentioned attempts stressed on the importance of using normalized ontologies in order to mask the problem of interoperability in TSCHS, reuse and share generated data, and create a context aware environment. However, none of these attempts give a clear solution for automatic discovery and integration of new components.

Additional researches stressed on the importance of introducing service oriented architecture for smart home. Authors in [16] proposed service oriented ontologies to describe service properties and location in smart homes relying on SUMO ontology. A service oriented rule based framework has been proposed by Leong et al. [17]. However, the service discovery is restricted to the server side (single point of failure) and relies on SOAP (XML structure which lack to the semantic annotation). Solve-D [18] is a framework for disabled users in smart home environment consisting of three main service ontologies: Generic service ontology to describe standard services for appliances, personalized service ontology and service context ontology where complex rules can be defined for advanced decision making. However, Solve-D focused on home appliances rather than fully sensor/actuator description. Meshkora et al. [19] presented ontology based COMANCHE project in order to organize and discover information needed to manage software in home environment. This project enables complex relation modeling between processes and includes rich description of devices (interfaces, location, manufacturer, processing capabilities, etc.). However, it lacks the description of the quality of service, and the data format, accuracy, precision, etc.

It is rigorous that none of these attempts combine the fully description of the sensors/actuators, data, process, and services. While reliability and service consistency are primordial in such systems, the QoS and computational resources as well as remaining battery level are needed for service modeling. Moreover, tele-health systems should integrate the monitoring data in IoT. Thus grounding methods for service accessing will facilitate the service discovery.

Therefore, the contribution of this work is OntoSmart system that relies on both MyOntoSens Ontology [20] (summarized in Figure1) and MyOntoService Ontology (depicted in Figure 3) to model TSCHSs' raw data and information. These ontologies are provided with associated management enablers, within our OntoSmart system, and will enable the handling of all the aforementioned failings of existing TSHCS architectures.

2.2 OntoSmart System Ontologies

TSCHS can be seen as a WSN (Wireless Sensor Network) encompassing different nodes (sensors, actuators, equipment, servers and gateways). Each WSN, in other term the home, can have one or more owner (the patient to be monitored) and contact persons (medical staff or relatives). Each node is used for certain process (temperature, humidity, etc.) and each actuator performs certain action (turn ON/OFF, shuttle Up/Down, etc.). That is why we used MyOntoSens ontology [20] to describe the components of our proposed system since this ontology already takes into account all the aforementioned TSHCS mandatory attributes. MyOntoSens has been designed as a modular ontology and is divided into three parts: MyOntoSensWSN, MyOntoSensNode and MyOntoSensProcess. Figure 1 summarizes the main classes of MyOntoSens ontology, as used in our proposed OntoSmart system, while Figure 2 depicts the main classes and attributes of MyOntoSensProcess part.
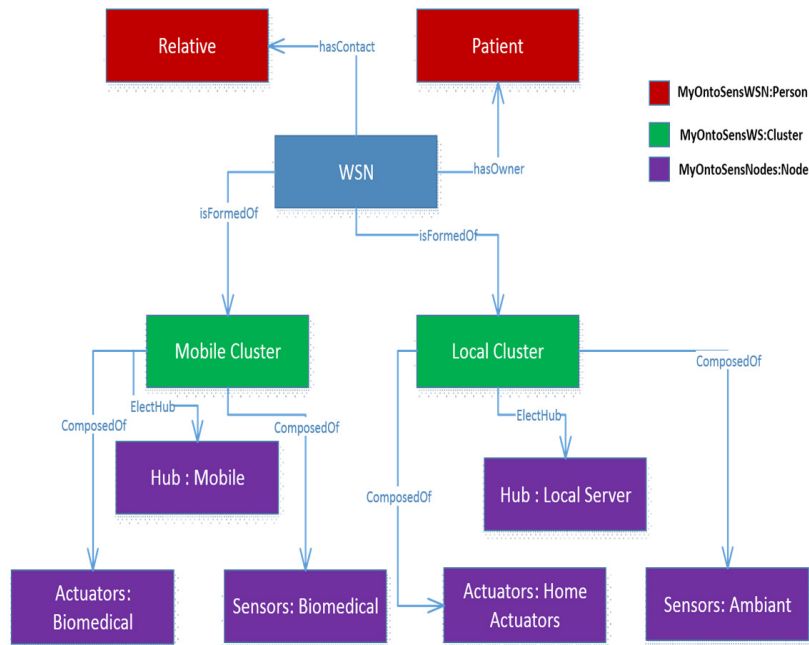

Figure 1-MyOntoSens Ontology as used in OntoSmart System

Moreover, the TSHCS system offers certain services to the patient (monitoring, notification, reminder, etc.). We are modeling these services by introducing original service ontology, called MyOntoService ontology and depicted in Figure 3.

Our new upper service ontology adopts the idea of OWL-S ontology [21]. It is composed of three main sub-ontologies (as depicted in Figure 3):

- MyOntoServiceProfile that describes the service (QoS, QoI, service constraints and service type) and enables their auto-discovery,

- MyOntoServiceProcess that describes how the service is realized based on the Input, Output Precondition, and Effect (IOPE) mechanism,

- And finally MyOntoServiceGrounding that describes how the service can be invoked (or accessed in IoT e.g. HTTP, CoAP, etc.)

To ensure automatic senor/service discovery, the ontologies are described by SWRL [19] rules and the Pellet [22] reasoner is used to infer explicit data (Section 3).

2.3 Reason behind using multi-agent architecture within our TSHCS system

The main aim of our proposed system is modularity, scalability, flexibility and distributed computing. Thus, this architecture should logically be divided in standalone modules offering certain services, capable of interacting with other modules and that can be instantiated when required. This is exactly what a software agent can do. An agent is a piece of software that can run autonomy to perform certain behavior. The composition of many agents is called Multiple Agents System (MAS) [23]. The agents interact by passing predefined messages. The MAS architecture describes how different agents are interconnecting. MAS are widely used in complex system where abstract definition of certain task is required. In our proposed TSCHS system basic functionalities are offered by agents capable of invoking other agents. This is in particular the case for our data collection and information management/sharing/retrieval functionalities. For example, we can cite the data collector agent, query agent, notification agent, writer agent, etc.
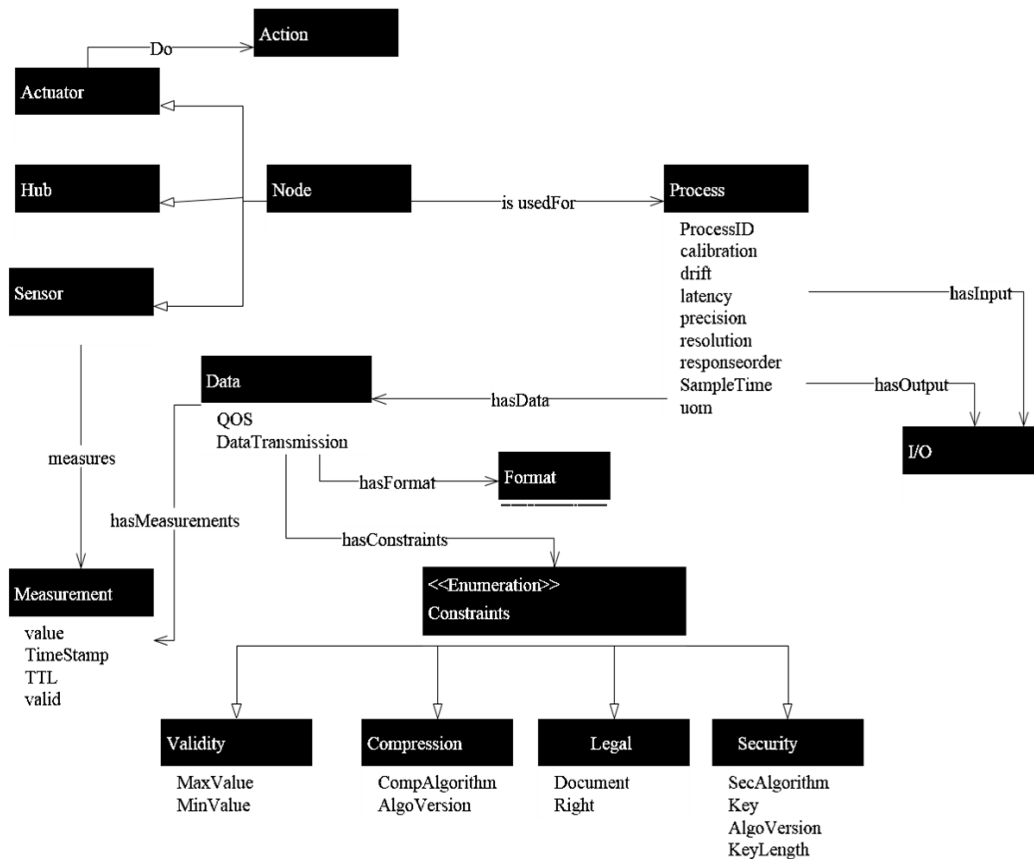

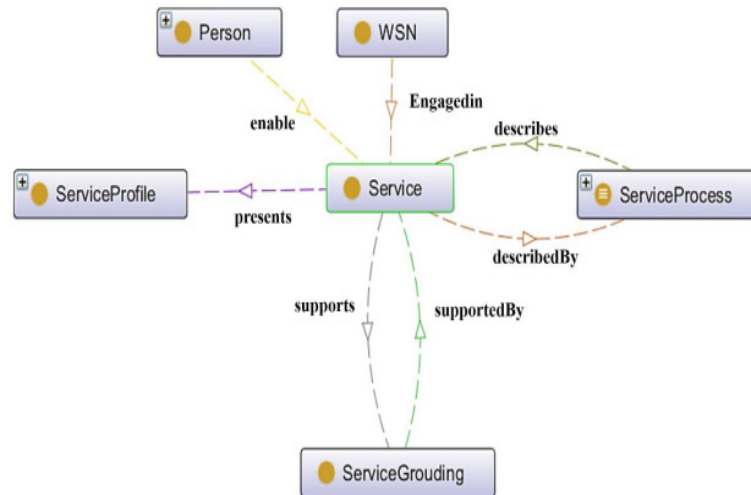Figure 2-MyOntoSensProcess Classes Hierarchy

Figure 3-MyOntoService Ontology

Moreover, using expressive ontologies without defining a general mechanism for data annotation is pointless [24]. That is why we defined three main agents in our overall TSHCS system:

- Scanners: for each communication protocol (Bluetooth LE, WIFI, ZigBee) there is a data scanner that read raw data and call the adequate wrapper for semantic annotation. Moreover, for each grounding protocol (CoAP, HTTP, etc.) there is a service scanner that reads the request from external user and calls the adequate service.

- Writers: To send data from the system to external users, the adequate writer encapsulates the sent data based on the requested protocol.

- Wrappers: To add semantic description of the raw data based on MyOntoSens and MyOntoService ontologies.

Detailed description of how to use these components is depicted in section 3.

3. ONTOSMART GENERAL SYSTEM ARCHITECTURE

Our proposed OntoSmart system comprises a WSN formed of two main clusters: the Wireless Body Area Network (WBAN) encompassing the sensors attached to the body of the patient for vital signals monitoring, and the local cluster formed of a local server and sensors/actuators related to ambient parameters and home appliances. It is worthy to note that the sensors used in our proposed scenario are only for testing purposes. These sensors could be extended to any other nodes required in TSHCSs without affecting the overall system architecture. Figure 4 depicts our proposed OntoSmart system architecture.
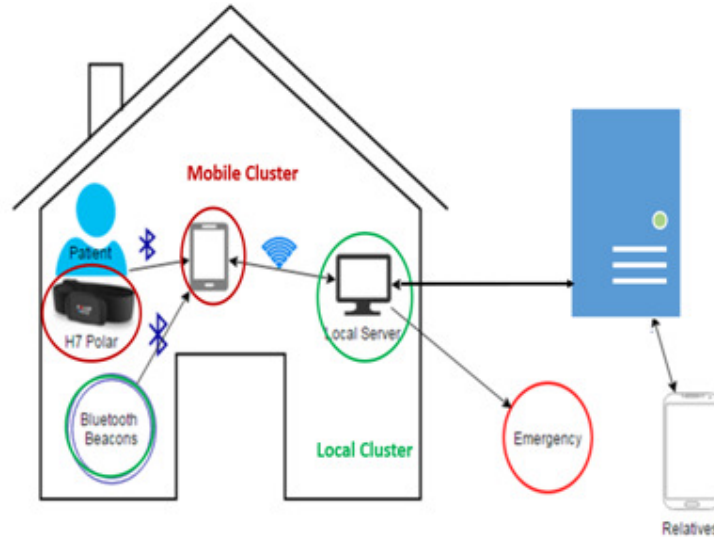
Figure 4- OntoSmart General Architecture

Falls, heart attack and stroke are among the leading causes of hospitalization for the elderly and patients. The chances of surviving a fall, heart attack or stroke are much greater if the senior gets help within an hour [25] For that reasons we mainly focused on monitoring the heart rate of the patient, deducing his/her posture (sitting, moving, falling, etc.), and his/her location (in which room he/she is). Moreover ambient processes are captured such as humidity, light intensity and temperature.

The details of each cluster are given in the next sections.

3.1 WBAN Cluster

The main aim of the WBAN Cluster is to collect patient's vital signals in order to send it to the local server. Here comes the role of the Data Collector Agent (DCA), also called the Cluster Hub. In our scenario, the WBAN Cluster is formed of the heart rate sensor, acceleration sensor and the smart phone of the user which is playing the role of the Cluster Hub. To capture the heart rate, the H7Polar sensor [19] is used. This senor sends the heart rate via its Bluetooth LE interface.

Therefore, a Bluetooth LE Scanner (BLES) agent is installed on the patient's smart phone in order to:

- identify any Bluetooth LE device,

- and recognize the offered service due to the use of UUID defined in standard GATT Profiles [26]

This scanner gives the first step toward an automatic sensor discovery.

To determine the posture of the patient, the smart phone built-in three acceleration sensor is used. We assume that the smart phone is placed in a vertical position on the hip of the patient. Therefore, the patient is not moving if the acceleration is equal to the gravitational force. We thus have in this case:

- If $|a_y| > 9$ m/s^2 then the patient is standing.

- If $|a_x| > 9$ m/s^2 then the patient is lying.

- If $6 < |a_x| < 8$ m/s^2 and $2 < |a_y| < 5$ and $2 < |a_z| < 5$ then the patient is sitting.

After identifying the node's characteristics, these latter should be sent to the local server for semantic annotation and further treatments. JSON [27] message format was adopted due to its lightness and simplicity. Therefore, a JSON Writer (JS) agent is installed on the smart phone in order to encapsulate data and send it to the local server.

If other types of sensors are used (like ZigBee), the only think that should be done is to install the convenient scanner (like ZigBee Scanner). Likewise, if data from the smartphone is encapsulated using a message format other than the JSON messages, the suitable writer agent should be compelled. In fact, the use of these data agents and writers enable the automatic discovery of any added nodes and enhance the flexibility and scalability of the system where any nodes can be added without the need to rebuild the overall system. What is really required is to ensure that the communication protocol used by the device has the suitable data scanner agent.

3.2 WSN Cluster

The WSN Cluster is formed of the sensors dedicated for indoor localization, the ambient parameters sensors, and the local server (DCA or cluster Hub). In our case, we use the RadBeacon Bluetooth LE sensors to identify the location of the patient. Moreover, the ambient parameters (light intensity, humidity and temperature) are retrieved from the built-in sensors in the patient's smartphone in order to be sent to the remote server using JSON messages. We installed, on the local server, the following agents:

- BLES agent: to identify Bluetooth LE sensors and retrieve the required information in order to invoke the adequate wrapper.

- JSON Scanner (JS) agent: to retrieve data sent from the smartphone and invoke the adequate wrapper.

- Node Wrapper (NW) agent: to add semantic annotation about the sensors/hubs based on MyOntoSensNode ontology.

- Process Wrapper (PW) agent: To add semantic annotation about the process used by the sensors based on MyOntoSensProcess.

In that way, the local server will contain all the semantic description of the home devices. Because more than one patient can use the same WSN sensors (e.g. indoor localization sensors can be used to locate more than one patient in the same home), the object property *"measuredFor"* has been added between the sensors measurements (Measurement class of MyOntoSensProcess ontology) and the patient (Patient Class of MyOntoSensWSN ontology). To this point, all nodes are automatically discovered. Regarding the service discover, we added SWRL [21] rules to infer basic services offered by the system. All processes are considered as service (e.g. heart rate, temperature, etc.). This inference is insured by using Rule 1. Composed services are deduced from additional rules. For example, using the heart rate sensor implies that the maximum heart rate service can be used for notification purposes (Rule 2).

(Rule 1) *Process(?P) -> Service(?P)* .

(Rule 2) Process (HR) -> Service (MaxHR), hasEffect (MaxHR, Notification).

Moreover, the object property *"enable"* between patient individual and service individual is used for allowing the patient to enable/disable any service offered by the system. Only enabled services are retrieved when invoking the query dedicated for service discovery.

3.3 Remote Semantic Storage and Management Server

The main aim of the remote semantic server is to save the fully semantic description of all the TSHCSs. It plays the role of semantic registry. While intelligence treatment for each home is located on its local server, the intelligent cooperation between different TSHCSs, and advanced data analysis for statistical purposes can be conducted on this remote server. In our scenario, the remote server encompasses the following agents:

- SPARQL Agent: to query semantic information saved on the remote/local server in order to retrieve new semantic information.

- CoAP Agent: to publish the data in order to be used by external users for remote monitoring (e.g. patient's relatives and medical staff). The CoAP [28] is a Rest-full protocol used for constraint networks. It uses the UDP protocol as transport layer and the reliability mechanism is driven by the application layer. We used the CoAP Confirmable messages to ensure the required reliability in our proposed TSCHS system. In that way, the Remote server will play the Role of CoAP monitoring server allowing remote users to access the data by using CoAP GET requests.

- Notification Agent: to notify external users about abnormal patient's attitude and dangerous measured values. In our case, we notify the user if the patient is falling or if his/her heart rate exceeds the normal value. It is worthy to note that the security mechanism is not within the scope of this study. That's why we based our test-bed authentication mechanisms on Android GCM due to its implementation simplicity, and by assuming that all our system users have a Gmail account. Let's precise here that for a commercial application, stronger authentication mechanisms will obviously have to be designed and implemented.

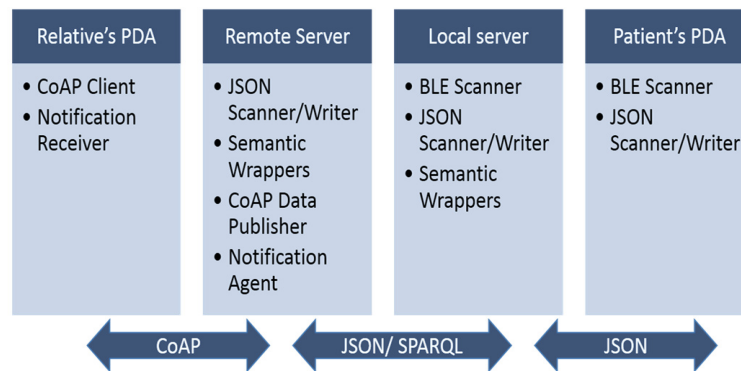Fig. 5 summarizes all the aforementioned agents and building blocks location and interactions.

Figure 5-Implemented Agents and interactions as deployed in OntoSmart System

The remote server will invoke the WSN Wrapper in order to add the new WSN and Patient into its semantic registry. The PatientID (identifier generated by the remote server) is now returned to the smartphone and saved in its shared Preference (local database on the Smartphone) for being used when needed (for patient identification). The patient can now launch its OntoSmartHome application and log into the OntoSmart system using his/her Gmail account.

4. ONTOSMARTHOME APPLICATIONS

Two applications have been developed for OntoSmart system: the first application (detailed in Section A) is dedicated for the patient in order to initialize the system and display the monitoring values; the second application (detailed in Section 4.2) is dedicated for relatives for remote

patient's monitoring. The same monitoring interface (depicted in Figure 6) is retained for the patient and relative applications.

4.1 Patient's OntoSmartHome Application

We have provided our OntoSmart TSHCS system with a patient-dedicated application and corresponding applicative agents (see Figure 5). This application, depicted in Figure 6 and Figure 7, is integrated within the Smartphone of the patient. As already mentioned in the previous sub-section (i.e. 3.3), the patient launches its OntoSmartHome application and logs into the OntoSmart system using his/her Gmail account. Once logged in, the patient can now add relatives (identified by their email). These relatives can thus monitor the patient by sending CoAP Get requests to the remote server and receives notification through a dedicated application (see section 3.5). The initialization phase, where the sensors are discovered, begins now. The Bluetooth LE sensors are recognized and displayed on the screen. For each discovered sensors, the user should precise the purpose of its usage. In particular, for Indoor localization, a setup phase is needed, where the user should walk during 20 seconds in the room in order to collect the minimum and maximum RSSI for each room [29]. When the setup phase is completed, the main page of the application is displayed showing the patient's heart rate and posture, as well as the ambient temperature, humidity and luminosity (see Figure 6). The patient can also add a status (similar to the idea used in social media) to inform the relatives about a special feeling or event, etc. The displayed value will be modified once a new value is measured. Note that a compromise between system accuracy and battery consumption is mandatory in such systems. That's why, after practical testing, we have chosen to acquire 50 samples/sec from the sensors. In addition, these modified values will be sent to the local server and the remote server via the JW in order to be updated on the CoAP server and registered in the remote ontology repository.
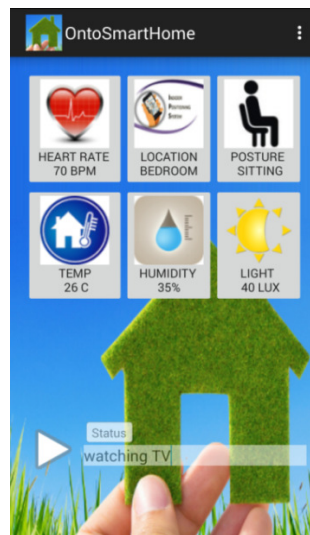

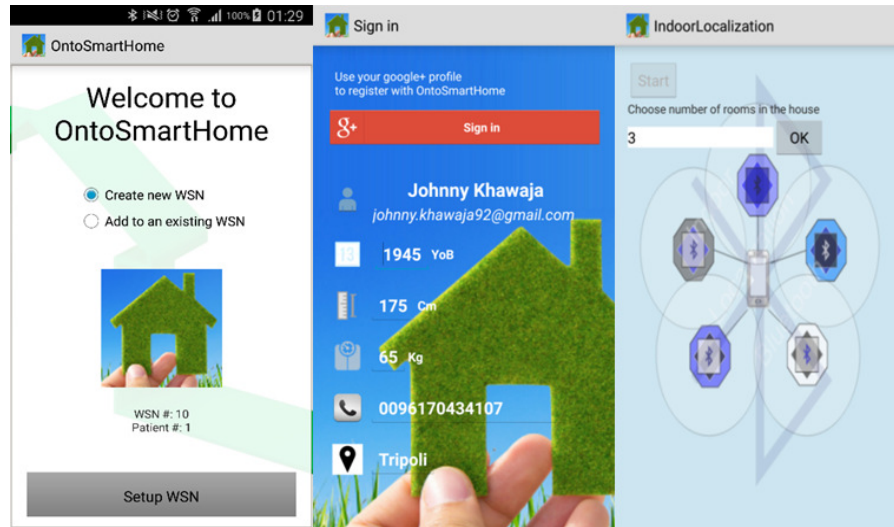Figure 6- Monitoring Interface of OntoSmartHome Application

Figure 7-Patient's OntoSmartHome application

4.2 Relatives OntoSmartHome Application

The first time the application is installed, the relative should successfully log in using his/her Gmail account. Afterward, the relative is asked to enter the patient's email. The patient and relative emails are sent via the JSON Writer (JW) to the remote server where a SPARQL query is invoked to verify the relative permissions. Figure 8 depicts the overall authentication process for the relative's mobile application.

If the relative is able to monitor the patient, CoAP GET requests are sent every minute to the CoAP remote server in order to retrieve the values.

Our OntoSmart system provides the discovery of any used BLE sensor due to the use of BLE Scanner which discovers the node, the node wrapper which adds the node's semantic description to the semantic registry, and the SPARQL agent that executes the node and service discovery queries. Moreover, publishing the monitoring patient's data through a CoAP server permits the relatives to instantly monitor these data by issuing CoAP GET requests. Further analysis and advanced diagnostics can be conducted on the remote server by domain experts in order to help the patient. Moreover, due to the use of semantic registry on the remote server, new domain expert ontologies can be imported to infer severe situations related, for example, to heredity problems or environmental diseases.

4.3 Notifications

Two notification services are implemented and offered within our proposed OntoSmart system: Abnormal Heart Rate and fall detection. The abnormal heart rate will be calculated based on the age of the patient [30].

Regarding the fall detection, the algorithm based on the concept developed by Sposaro *et al.* [31] was adopted. A fall begins with a sudden and significant decrease in the acceleration amplitude "a". The free fall period causes the acceleration's amplitude to drop significantly below 9.8m/s2 threshold during a short time. This represents the period of time when the actual fall is taking place. After this "free-fall-Period", the acceleration experiences a sudden spike as soon as the body hits the floor. The amplitude "a" crosses an upper threshold suggesting a fall. This threshold

is determined to be 3g (around 27m/s2). Consequently, if the acceleration "a" crosses a lower and an upper threshold during certain observation time window, a fall is suspected. This window is often determined to be less than one second. To prevent false alarms, a timer for one minute is enabled allowing the patient to cancel the notification. Otherwise a notification will be sent to relatives mobile phones. In our implementation, the GCM notification server is used as detailed in Figure 9.

4.3 Testing & Results

In fact our testing process focused on two main principles: the requirements (as highlighted in Section I) of THSCs and the mobile application itself. For that purposes, the applications were tested on Samsung GALAXY S4 (Android 4.4.2 KITKAT-API 19 and Android 5.0.1-API 21 Lollipop after upgrade) and SONY XPERIA SL (Android 4.1.2 Jelly Bean- API 16). We first considered one patient per WSN. We installed the application on two different patients' mobile. For each patient, one relative was added. We first insure that the ontology is well classified. Due to the use of incremental reasoner, the services were discovered sequentially. For example, once the H7 Polar is paired, the heart rate service was added on the server. Moreover, the user was able to enable/disable the services by his/her own without the intervention of experts. At certain stage, the Patient 1 decided to disable the indoor localization services. Immediately, on the relative smart phone, the location field was replaced by "Unknown". This proves the ability of our proposed system to offer services based on the occupants needs and requirements.
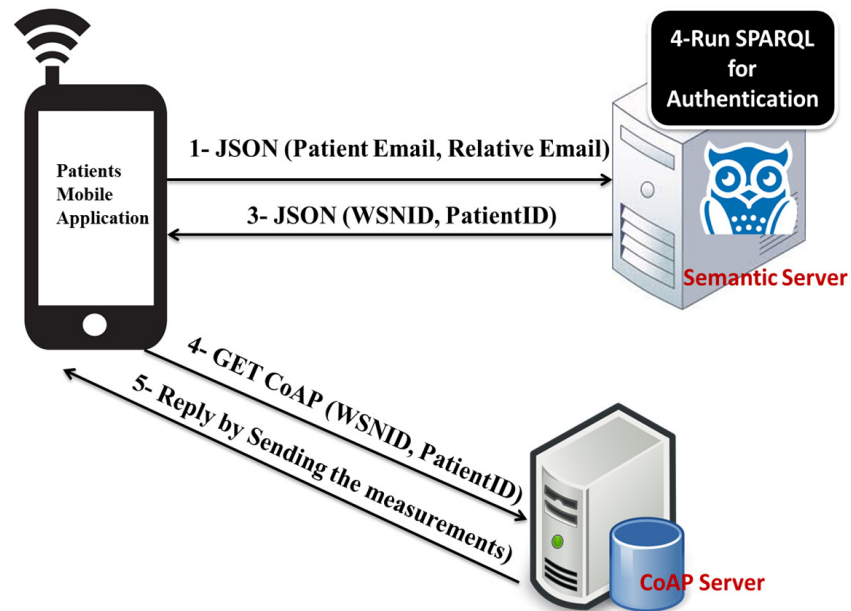

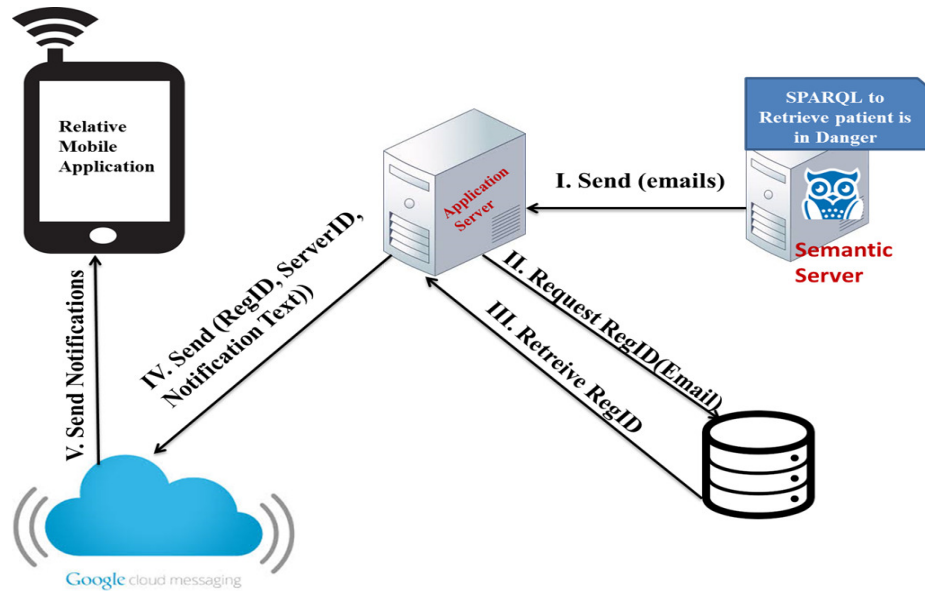Figure 8-Authentication Process for Relative's mobile app.

Figure 9-Notification process using GCM Server.

Then, we considered the scenario where two patients are sharing the same WSN, i.e. sharing the same Rad Beacons for the indoor localization. The distance between rooms was about 20 m and three Rad beacons were carried out and setup in three different zones at the patient's home. For testing purposes only, the remote server and the local server were on the same machine, but working on different port numbers. The ontology was well classified without any ambiguity. This insures the reusability of the proposed system by more than one occupant.

The privacy concerns were addressed by the use of email accounts for application installation and the authentication mechanism when a relative requests to monitor a patient. It is evident that this authentication mechanisms, as well as security concerns, could be enhanced by implementing advanced protocols. However, this enhancement does not affect the overall architecture, but only require the adoption of a new security/authentication agent invoked when needed.

After insuring the well functionality of the system, we focused on measuring the complexity of the system on the server. For both cases, the tests were conducted for about 8 hours. Table 1 depicts the time needed to load and classify the ontology, as well as the CPU and memory usage.

Table 1- System Evaluation on the Server

		Case 1	Case 2
Ontology	Load	24 sec	35 sec
	Classify	90 sec	75 sec
	Model Size	1403	1403
	Inferred Model size	1019	1000
CPU	Max. CPU	76%	60%
	Avg. CPU	35%	27%
	Memory	425 692 K	415 728 K

We can note that the ontology is classified within 2 minutes. The size of the inferred model reflects the importance of SWRL rules used in the proposed system. As few data is required to be

sent for the mobile application to the semantic server, lowest bandwidth will be consumed. The potential of the semantic rules is inferring all needed data to discover the properties of a service. For example, all HR services will be inferred by just sending that the heart process is offered by H7 Polar sensor (see Rule 1 and Rule 2). We are actually working on studying the performance of the server on large scale (more than 10 patients registered in the application) to estimate the probability of crashing.

We passed than to evaluate the performance of the patient's mobile application. The time needed for node/service discovery was maximum 1 second. The patient's application was tested for about 8 hours; the battery consumption was maximum 12 % and minimum 5%, while the average CPU usage was 8%. Table 2 depicts the processing, network and battery usage on the patient smart phone in initiation mode (done only when the application is loaded for the first time) and the normal mode where the measurements are taken periodically.

This evaluation ensures that our proposed application is not dramatically affecting the resources on the smart phone. In fact, these consumptions can be decreased by the use of ambient sensors or biomedical accelerometer sensors instead of the built-in sensors. In that way, the data will be received directly on the server for semantic annotation and treatments. Aiming to determine the falling detection, we tried to test to detect a falling after standing, from bed or from a chair. These fallings were successfully detected. We can note that the posture and the falling status are not 100% accurate because we were using the Smartphone embedded sensors for the measurements (for economy purposes). These measurements can be enhanced by using dedicated accelerometer sensors capable of detecting more precisely the posture of the patient.

It is now the time to evaluate the relatives application. We first tested the authentication process. Only authenticated relatives were able to monitor the patients and receive notifications. We tested the network and battery consumption (depicted in Table 3) of the relatives for about 8 hours.

Table 2- Resources Consumption for the Patient Smart Phone

	Initialization Phase		Main Activity	
	Max	Average	Max	Average
CPU %	5%	2%	11%	8%
Network Usage	16 Kb	12 Kb	40 Kb	36 Kb
Battery Usage	7%	5%	12%	10%

On the relatives' side, the measured values and patient's status and alarms were displayed in real time. The CPU usage on the relative's application attempts as average 0.56 % and as maximum 1 %. These resources consumption are considered insignificant due the use of CoAP messages (1280 bytes). The values were displayed in real time and the notification when abnormal values are detected was received within 1 second.

Table 3- Resources Consumption on the Relative Mobile Application

Average CPU %	0.56%
Average Network Usage	35 Kb
Average Battery Usage	0.40%

5. CONCLUSION

In summary, although our OntoSmart system does not actually provide the best solution for fall detection (only because Smartphone embedded sensors were used instead of dedicated sensors), it offers the flexibility to enhance this detection, whenever needed, by just adding the necessary sensors. Our OntoSmart system is the first step toward a "Plug and Play" system where patients can easily add sensors and actuators. Service developers will only focus on ads-on services and applications without the need to deal with sensor integration and configuration. In addition to the passive intervention devices, active intervention devices for reminder systems or medical assistance can be added to OntoSmart System. More agents are under construction for example, remind patients to take his/her medication or alert him/her if a stove burner is left on. Surely, actuators can be added to the system. All these concerns do not need to re-invent the wheel, but just upgrade the system by adding adequate agents. More tests are being conducted to measure the efficiency of the remote semantic storage and management server when dealing with large amount of TSHCSs. Additional scanners and writers are also under development, especially ZigBee and HL7 scanners/writers. Finally, more security and privacy concerns are under investigations.

They were here when you were a child be there with them when they are going. THSCs will surely help elderly feeling gladder and confortable by staying at home and connected to their family.

REFERENCES

[1]   "WHO | Definition of an older or elderly person."

[2]   "WHO | Ageing and life-course."

[3]   J. Cocco, "Smart Home Technology for the Elderly and the Need for Regulation," pp. 85–108, 2011.

[4]   V. Gruessner, "Global Telehealth Market Expected to Grow Through 2020 - mHealthIntelligence." mHealth Intelligence, 2015.

[5]   M. Memon, S. Wagner, C. Pedersen, F. Beevi, and F. Hansen, "Ambient Assisted Living Healthcare Frameworks, Platforms, Standards, and Quality Attributes," Sensors, vol. 14, no. 3, pp. 4312–4341, Mar. 2014.

[6]   W. K. Edwards and R. E. Grinter, "At Home with Ubiquitous Computing: Seven Challenges," Proc. 3rd Int. Conf. Ubiquitous Comput., pp. 256–272, 2001.

[7]   Q. Ni, A. García Hernando, and I. de la Cruz, The Elderly's Independent Living in Smart Homes: A Characterization of Activities and Sensing Infrastructure Survey to Facilitate Services Development, vol. 15, no. 5. 2015.

[8]   R. Preet, "A Case for Ending Monolithic Apps for Connected Devices," HotOS, pp. 1–7, 2015.

[9]   F. Latfi, B. Lefebvre, and C. Descheneaux, "Ontology-based management of the telehealth smart home, dedicated to elderly in loss of cognitive autonomy," CEUR Workshop Proc., vol. 258, 2007.

[10]   D. Allemang and J. Hendler, Semantic web for the working ontologist: effective modeling in RDFS and OWL. Elsevier, 2011.

[11]   S. U. D. L. M. A. K. S. L. D. S. D. E. Deborah L. McGuinness (Knowledge Systems Laboratory and A. F. van. H. vu. n. Frank van Harmelen (Vrije Universiteit, "OWL Web Ontology Language Overview." .

[12] T. Van Nguyen, W. Lim, H. Nguyen, D. Choi, and C. Lee, "Context Ontology Implementation for Smart Home," 2010.

[13] W. Paper, "Internet of Things : Wireless Sensor Networks Executive summary."

[14] L. Chen, C. Nugent, M. Mulvenna, D. Finlay, and X. Hong, "Semantic smart homes: Towards knowledge rich assisted living environments," Stud. Comput. Intell., vol. 189, pp. 279–296, 2009.

[15] J. Ye, G. Stevenson, and S. Dobson, "A top-level ontology for smart environments," Pervasive Mob. Comput., vol. 7, no. 3, pp. 359–378, 2011.

[16] J. H. Kim, H. Kwon, D. H. Kim, H. Y. Kwak, and S. J. Lee, "Building a service-oriented ontology for wireless sensor networks," Proc. - 7th IEEE/ACIS Int. Conf. Comput. Inf. Sci. IEEE/ACIS ICIS 2008, conjunction with 2nd IEEE/ACIS Int. Work. e-Activity, IEEE/ACIS IWEA 2008, no. March 2016, pp. 649–654, 2008.

[17] C. Y. Leong, A. R. Ramli, and T. Perumal, "A rule-based framework for heterogeneous subsystems management in smart home environment," IEEE Trans. Consum. Electron., vol. 55, no. 3, pp. 1208–1213, 2009.

[18] M. Sohn, S. Jeong, and H. J. Lee, "Self-evolved ontology-based service personalization framework for disabled users in smart home environment," Proc. - 7th Int. Conf. Innov. Mob. Internet Serv. Ubiquitous Comput. IMIS 2013, pp. 238–244, 2013.

[19] E. Meshkova, J. Riihijarvi, P. Mahonen, and C. Kavadias, "Modeling the home environment using ontology with applications in software configuration management," 2008 Int. Conf. Telecommun., pp. 1–6, 2008.

[20] L. Nachabe, M. Girod-Genet, and B. ElHassan, "Unified Data Model for Wireless Sensor Network MyOntoSens Ontology," IEEE Sens. J., 2015.

[21] T. George and B. George, "OWL-S : Semantic Markup for Web Services."

[22] B. Parsia and E. Sirin, "Pellet: An owl dl reasoner," in Third International Semantic Web Conference-Poster, 2004, vol. 18.

[23] J. Ferber, Multi-agent systems: an introduction to distributed artificial intelligence, vol. 1. Addison-Wesley Reading, 1999.

[24] G. Zonta, P. Jr, J. Daltio, and C. B. Medeiros, "A Mechanism for Propagation of Semantic Annotations of Multimedia Content," vol. 5, no. 4, pp. 332–342, 2010.

[25] H. Lentzner, D. Ph, D. Hoyert, D. Ph, K. N. Robinson, and D. Ph, "Trends in Causes of Death," no. 1, pp. 1–9, 2001.

[26] B. Technologies, "Bluetooth ® low energy technology."

[27] M. Lanthaler and C. Gütl, "On using JSON-LD to create evolvable RESTful services," in Proceedings of the Third International Workshop on RESTful Design, 2012, pp. 25–32.

[28] Z. Shelby, K. Hartke, and C. Bormann, "RFC 7252: The Constrained Application Protocol (CoAP)." p. 112, 2014.

[29] X. Zhu and Y. Feng, "RSSI-based Algorithm for Indoor Localization," vol. 2013, no. May, pp. 37–42, 2013.

[30] L. Nachabe, M. Girod-Genet, B. ElHassan, and F. Aro, "Applying Ontology to WBAN for mobile application in the context of sport exercises," in Proceedings of the 9th International Conference on Body Area Networks, 2014, pp. 204–209.

[31] F. Sposaro and G. Tyson, "iFall: An Android Application for Fall Monitoring and Response."

INTENTIONAL BLANK

MEDICAL IMAGING MUTIFRACTAL ANALYSIS IN PREDICTION OF EFFICIENCY OF CANCER THERAPY

Jelena Vasiljevic[1], Natasa Zivic[2], Jelena Pribic[3], Ksenija Kanjer[3], Wojtek Jonakowski[1], Jelena Sopta[4], Dragica Nikolic-Vukosavljevic[3] and Marko Radulovic[3]

[1]Institute "Mihajlo Pupin", Volgina 15, Belgrade, Serbia
[2]University of Siegen, Germany
[3]Department of Experimental Oncology, Institute of Oncology and Radiology of Serbia, 11000 Belgrade, Serbia
[4]Medical Faculty, Institute of Pathology, University of Belgrade, Dr Subotica 1, Belgrade, Serbia

ABSTRACT

Based on pressing need for predictive performance improvement, we explored the value of pre-therapy tumour histology image analysis to predict chemotherapy response. It was shown that multifractal analysis of breast tumour tissue prior to chemotherapy indeed has the capacity to distinguish between histological images of the different chemotherapy responder groups with accuracies of 91.4% for pPR, 82.9% for pCR and 82.1% for PD/SD.

KEYWORDS

breast cancer; prediction; histology; chemotherapy, fractal; multifractal; drug response

1. INTRODUCTION

An ideal predictive marker should reliably predict chemotherapy responses in breast cancer and thus determine the right treatment for each individual patient. Consequently, an effort to improve the accuracy of predictive markers is expected to provide a reduction of relapse rates and prolonged survival.

In parallel to standard molecular approaches, digital pathology emerged as a structure analysis tool to aid in detection, diagnosis(Cross 1997; Vasiljevic, Reljin et al. 2012), risk assessment (Vasilescu, Giza et al. 2012), chemotherapy efficacy assessment (Li, Hu et al. 2014) and therapy prediction for cancer (Laurinavicius, Plancoulaine et al. 2014). It is based on computational analysis of medical images by use of texture or fractal algorithms. Studies indicate an association between multifractal parameters of breast tumour histology and pathologic tumour grade (Braverman and Tambasco 2013), fractal dimension of breast malignant epithelium and survival

(Tambasco, Eliasziw et al. 2010) and breast tumour MRI fractal geometry and response to neoadjuvant chemotherapy (Di Giovanni, Ahearn et al. 2012).

Surprisingly, although fractal analysis is known as powerful morphometry tool for quantitative assessment of complex pathological structures (Huang and Lee 2009), its usefulness in breast cancer therapy prediction has not been investigated. Therefore, in this study we exploit the multifractals to discriminate between breast tumours with different sensitivity to chemotherapy. This task was approached by a neoadjuvant therapy model which has been accepted as an ideal in vivo assessment of therapy response, with a consequent wide use for evaluation of predictive markers (Kanjer, Tatic et al. 2013). The specific objective of this study was to identify the fractal parameter which provides the most accurate prediction of the chemotherapy response.

2. MATERIALS AND METHODS

Patients

Patients received neoadjuvant, anthracycline-based chemotherapy at our institution between 1999 and 2003. Criteria for the selection of patients were as follows: 1) an incisional biopsy of the primary breast cancer confirming invasive carcinoma before commencing the treatment and 2) primary locally advanced breast cancer that was strictly not operable. Patients with bilateral or metastatic disease were not included in the analysis.

Patient treatment and assessment of the response to therapy

Prior to surgery, all patients were treated with standard anthracycline-based chemotherapy (5-fluorouracil 500 mg/m2, doxorubicin 50 mg/m2 and cyclophosphamide 500 mg/m2 intravenously). Breast tumour response was evaluated after chemotherapy completion by pathohistological examination of the resected surgical material including measurement of the residual tumor size, optical microscopy and immunohistohemical analysis according to recommendations of International Expert Panel (Kaufmann, Hortobagyi et al. 2006), as previously described in detail (Kanjer, Tatic et al. 2013).

Multifractal analysis

Sections were cut at 5-μm thickness from the paraffin blocks and stained with haematoxylin / eosin stain as previously described (Kanjer, Tatic et al. 2013). Representative tissue sections were selected for each patient by a pathologist and digital microscopic images acquired at x400 magnification using Olympus BX-51 light microscope and a mounted Olympus digital camera. For each of the three groups, 350 images were again selected by a pathologist (1050 in total). Multifractal analysis of digital medical images was performed by use of free FracLac software and the stratification accuracy of each parameter was subsequently calculated by use of the classification type Single Tree (MATLAB 2010 and DTREG predictive modeling software version 10.3.0). Accuracy refers to the proportion of true results. Other calculated parameters included precision or positive predictive value as the proportion of patients with the disease who are correctly predicted to have the disease. F-Measure is the harmonic mean of precision and recall to give an overall measure of the prediction quality. Recall or negative predictive value is the proportion of patients who do not have the disease who are correctly predicted as not having the disease.

The main multifractal parameter is the Hölder's exponent which depends on local regularity of the observed structure (Evertsz, Mandelbrot et al. 1992).

$$\alpha = \frac{\log \mu \ (box)}{\log \varepsilon}$$

where μ (box) represents the signal measurement within the box and ε is the box size. The distribution of α is known as the multifractal spectrum, f (α), describing the global regularity of observed structure. Multifractal analysis permits the description of structure features from both local and global points of view. For instance, high values of Hölder exponent α denote high local changes, while high multifractal spectrum f (α) values denote frequent events – isolated parts in the whole structure having particular value of α (Reljin, Paskas et al. 2011).

3. RESULTS AND DISCUSSION

3.1 Analysis of microscopic images

The primary tumour of patients with early breast cancer is the main source of information for assessment of the disease recurrence risk and the choice of the most appropriate systemic treatment. Assessment of the predictive potential of multifractal analysis was performed on the patient group which was preoperatively treated with anthracycline-based chemotherapy. Such neoadjuvant model is particularly suitable for characterization of chemotherapy predictors as it allows for exact evaluation of the chemotherapy response with the tumour remaining *in situ* throughout treatment. Multifractal analysis was performed on digital images of tumour tissue sections obtained prior to therapy application. As the study was retrospective, the patient stratification into three response categories was done according to their actual response to therapy: partial pathological response (pPR), pathological complete response (pCR) and progressive/stable disease (PD/SD). Representative images for each of these categories are shown respectively on Figs. 1a, c, e. For multifractal analysis these images were transformed from colour to binary as depicted in Figs. 1b, d, f. At this x400 magnification, the histology of sensitive and resistant tumours was visually similar (Figs. 1a, c, e). Such histological similarity also remained at higher and lower magnifications (not shown).

3.2 Prediction of response to chemotherapy by multifractal analysis

Table 1 indicates that fractal dimension predicts response to chemotherapy with high accuracy. pCR and PD/SD are the extreme groups and thus most important as they give clear guidance regarding the therapy sensitivity or resistance (Cortazar, Zhang et al. 2014), while the clinical use of stratification into partial response pPR group is more limited. Importantly, the prediction accuracies of over 82% achieved for pCR and PD/SD groups are in line with those previously obtained by the standard prediction marker Ki67 and the advanced PET/CT system which even had an advantage of predicting on the basis of the actual response to the first chemotherapy cycle (Sheri and Dowsett 2012; Lee, Bae et al. 2014; Sueta, Yamamoto et al. 2014). Taken together, our results indicate a potentially superior predictive performance of fractal analysis in comparison to existing prediction approaches, with an added advantage of cost effectiveness.

3.3 Importance of multifractal parameters in prediction of chemotherapy response

Multifractal analysis delivers a number of parameters which necessitated a need for evaluation of their relative predictive power. The ranking was based on the information about the role of variables as primary splitters and surrogate splitters. A variable is thus more important if it was selected as a primary splitter early in the Single Tree model.

4. CONCLUSIONS AND FUTURE WORK

Histological examination is mostly used for diagnosis, whereas immunohistochemical and genetic tests are utilized for treatment decisions. We hypothesized that breast tumour histological examination by fractal analysis may be of use for prediction of chemotherapy response, based on its known morphometric discriminating capacity when dealing with irregular tumour tissue structures. Improvements in chemotherapy response prediction are of high clinical relevance due to the major impact of chemotherapy on quality of life and survival. Multifractal analysis of breast tumour tissue prior to chemotherapy was here shown for the first time to differentiate between tissues based on their actual chemotherapy sensitivity. Usefulness of this methodology relies on its high accuracy and also cost-effectiveness deriving from rapid analysis of standard clinical material. Exploration of the predictive usability among obtained multifractal parameters indicated the critical importance of the $f(\alpha)_{max}$, the maximum of multifractal spectrum. It can be speculated that the observed predictive power of the multifractal analysis is based on the detection of unknown structural clues which indicate the response to chemotherapy. Additional investigation is necessary to characterize the value of $f(\alpha)_{max}$ for the prediction of drug resistance.

REFERENCES

[1]   B. BravermanM. Tambasco, Comput Math Methods Med 2013, 262931 (2013).

[2]   P. Cortazar, L. Zhang, M. Untch, K. Mehta, J.P. Costantino, N. Wolmark, H. Bonnefoi, D. Cameron, L. Gianni, P. Valagussa, S.M. Swain, T. Prowell, S. Loibl, D.L. Wickerham, J. Bogaerts, J. Baselga, C. Perou, G. Blumenthal, J. Blohmer, E.P. Mamounas, J. Bergh, V. Semiglazov, R. Justice, H. Eidtmann, S. Paik, M. Piccart, R. Sridhara, P.A. Fasching, L. Slaets, S. Tang, B. Gerber, C.E. Geyer, Jr., R. Pazdur, N. Ditsch, P. Rastogi, W. EiermannG. von Minckwitz, Lancet 384, 164 (2014).

[3]   S.S. Cross, J Pathol 182, 1 (1997).

[4]   P. Di Giovanni, T.S. Ahearn, S.I. Semple, L.M. Lovell, I. Miller, F.J. Gilbert, T.W. Redpath, S.D. HeysR.T. Staff, Breast Cancer Research and Treatment 133, 1199 (2012).

[5]   C.J. Evertsz, B.B. MandelbrotL. Woog, Phys Rev A 45, 5798 (1992).

[6]   A. Goldhirsch, J.H. Glick, R.D. Gelber, A.S. Coates, B. ThurlimannH.J. Senn, Ann Oncol 16, 1569 (2005).

[7]   P.W. HuangC.H. Lee, IEEE Trans Med Imaging 28, 1037 (2009).

[8]   K. Kanjer, S. Tatic, Z. Neskovic-Konstantinovic, Z. Abu RabiD. Nikolic-Vukosavljevic, Pathol Oncol Res 19, 577 (2013).

[9]   M. Kaufmann, G.N. Hortobagyi, A. Goldhirsch, S. Scholl, A. Makris, P. Valagussa, J.U. Blohmer, W. Eiermann, R. Jackesz, W. Jonat, A. Lebeau, S. Loibl, W. Miller, S. Seeber, V. Semiglazov, R. Smith, R. Souchon, V. Stearns, M. UntchG. von Minckwitz, J Clin Oncol 24, 1940 (2006).

[10] A. Laurinavicius, B. Plancoulaine, A. Laurinaviciene, P. Herlin, R. Meskauskas, I. Baltrusaityte, J. Besusparis, D. Dasevi Ius, N. Elie, Y. Iqbal, C. BorI.O. Ellis, Breast Cancer Res 16, R35 (2014).

[11] S.M. Lee, S.K. Bae, T.H. Kim, H.K. Yoon, S.J. Jung, J.S. ParkC.K. Kim, Clinical Nuclear Medicine 39, 882 (2014).

[12] L. Li, W.Y. Hu, L.Z. Liu, Y.C. PangY.Z. Shao, Biomed Mater Eng 24, 163 (2014).

[13] B. Reljin, M. Paskas, I. ReljinK. Konstanty, Diagn Pathol 6 Suppl 1, S21 (2011).

[14] R. Rouzier, P. Pronzato, E. Chereau, J. Carlson, B. HuntW.J. Valentine, Breast Cancer Research and Treatment 139, 621 (2013).

[15] A. SheriM. Dowsett, Ann Oncol 23 Suppl 10, x219 (2012).

[16] Y. Sota, Y. Naoi, R. Tsunashima, N. Kagara, K. Shimazu, N. Maruyama, A. Shimomura, M. Shimoda, K. Kishi, Y. Baba, S.J. KimS. Noguchi, Ann Oncol 25, 100 (2014).

[17] A. Sueta, Y. Yamamoto, M. Hayashi, S. Yamamoto, T. Inao, M. Ibusuki, K. MurakamiH. Iwase, Surgery 155, 927 (2014).

[18] M. Tambasco, M. EliasziwA.M. Magliocco, J Transl Med 8, 140 (2010).

[19] C. Vasilescu, D.E. Giza, P. Petrisor, R. Dobrescu, I. PopescuV. Herlea, Hepatogastroenterology 59, 284 (2012).

[20] J. Vasiljevic, B. Reljin, J. Sopta, V. Mijucic, G. TulicI. Reljin, Biomed Microdevices 14, 541 (2012).
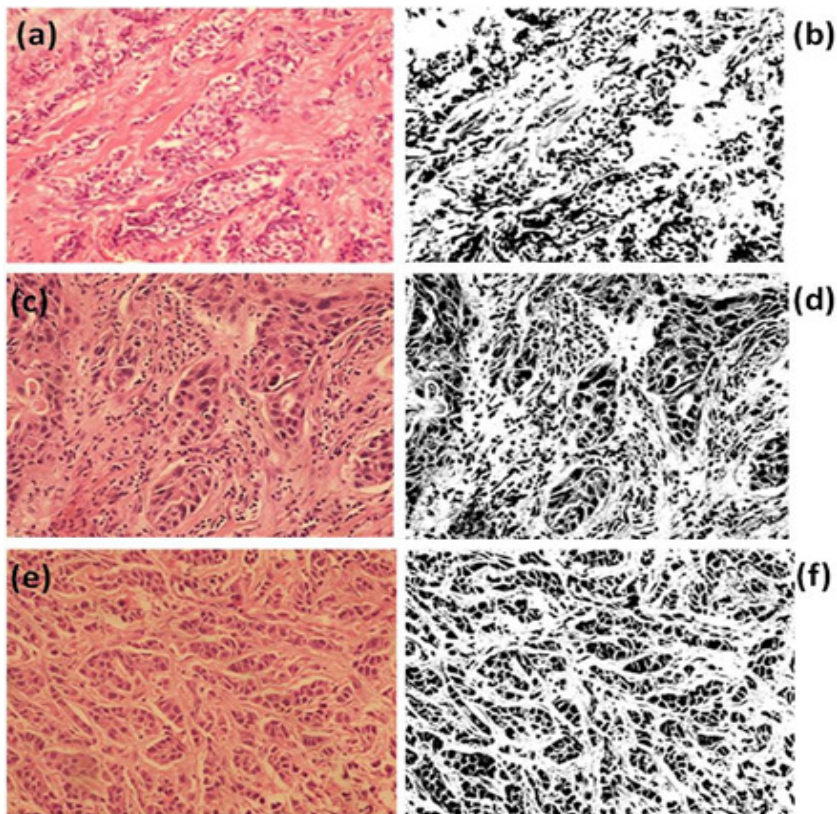
Fig. 1. Primary breast tumour tissue sections before therapy. (a) Partial pathological response in colour and (b) as binary image. (c) Pathological complete response in colour and (d) as binary image. (e) Progressive/stable disease in colour and (f) as binary image. Sections were stained with hematoxylin and eosin, magnification x400.

Table 1 Comparison of groups with different response to chemotherapy by multifractal analysis

	Partial pathological response	Pathological complete response	Progressive/stable disease
Accuracy	91.4 %	82.9 %	82.1 %
Sensitivity	85.1 %	80.3 %	69.1 %
Specificity	94.6 %	84.1 %	88.6 %
Geometric mean*	89.7 %	82.2 %	78.3 %
Precision	88.7 %	71.7 %	75.1 %
F-Measure	0.87	0.76	0.72

*Geometric mean of sensitivity and specificity

AUTHORS

Dr.-Ing. habil. Natasa Zivic

Dr.-Ing. Habil. Natasa Zivic is a lecturer and researcher at the Institute for Data Communication Systems, University of Siegen. She graduated from the Faculty of Electrical Engineering at the Department of Electronics, Telecommunications and Automatics, University of Belgrade in 1999 and defended her Magister Thesis in 2002. She started her PhD Thesis at the Institute for Data Communication Systems, University of Siegen, Germany in 2004 and defended it in 2007. Since 2010 she is employed as a Associated Professor at the University of Siegen where she gives lectures in Basics of Telecommunications, Digital Communication Technologies and Communication Networks. She defended her Postdoctoral thesis in 2012. Her main research areas are robust algorithms for message and image authentication using a combination of channel coding and cryptography. She takes part at international conferences and journals and works as a paper reviewer and editor of one scientific journal. She is an author of 3 books, more than 100 conference and journal papers and 3 (2 German and 1 USA) patents. She took part in more than 10 projects as a project manager or a developer.

Jelena Vasiljevic, PhD of Electrical Engineering, University of Belgrade, Institut Mihajlo Pupin, Docent for Intelligent Systems on Faculty for Computer Science

Jelena Vasiljevic works as Science Assistant in Institute Mihajlo Pupin in Belgrade, Serbia and also as Docent for Intelligent Systems teaching on Faculty for Computer Science. She is teaching Intelligent Systems (Artificial Intelligence), Genetic Algorithms, Neural networks, Fuzzy Logic, Bioinformatics, Knowledge based Systems on PhD, Master and Primary studies.   She is a graduate of Faculty for Electrical Engineering on University in Belgrade, Serbia. PhD research was application of multifractal analysis of digital medical images in cancer diagnostic and for it she has been several times awarded: The First Award on International Medical Fair 2014, The Special Award on International Technical Fair 2014 and many others (also many interviews for TV and newspaper). She is also an author and co-author of many journal and conference peer-reviewed.

AUTHOR INDEX