ISSN : 2165-4069(Online)
ISSN : 2165-4050(Print)

IJARAI

International Journal of Advanced Research in Artificial Intelligence

Volume 1 Issue 2

www.ijarai.thesai.org

A Publication of
The Science and Information Organization

IJARAI

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN ARTIFICIAL INTELLIGENCE

SAI

THE SCIENCE AND INFORMATION ORGANIZATION

www.thesai.org | info@thesai.org

Editorial Preface

From the Desk of Managing Editor…

"The question of whether computers can think is like the question of whether submarines can swim." ― Edsger W. Dijkstra, the quote explains the power of Artificial Intelligence in computers with the changing landscape. The renaissance stimulated by the field of Artificial Intelligence is generating multiple formats and channels of creativity and innovation. This journal is a special track on Artificial Intelligence by The Science and Information Organization and aims to be a leading forum for engineers, researchers and practitioners throughout the world.

The journal reports results achieved; proposals for new ways of looking at AI problems and include demonstrations of effectiveness. Papers describing existing technologies or algorithms integrating multiple systems are welcomed. IJARAI also invites papers on real life applications, which should describe the current scenarios, proposed solution, emphasize its novelty, and present an in-depth evaluation of the AI techniques being exploited. IJARAI focusses on quality and relevance in its publications. In addition, IJARAI recognizes the importance of international influences on Artificial Intelligence and seeks international input in all aspects of the journal, including content, authorship of papers, readership, paper reviewers, and Editorial Board membership.

In this issue we have contributions on fuzzy approach to classify learning disability; a novel machine learning approach to de blurring license plate using K-Means clustering method; experiments for navigation behaviours of robotic systems with different scene perception algorithms in real outdoor scenes; a framework for automatic generation of answers to conceptual question; sensor location problems and advocate them as test problems of non-smooth optimization; and also the role of technology and innovation in the framework of the information society; speech recognition of hausa language; automated marble plate classification system; and also analysis and implementation on human fingerprint patterns system

The success of authors and the journal is interdependent. While the Journal is in its initial phase, it is not only the Editor whose work is crucial to producing the journal. The editorial board members , the peer reviewers, scholars around the world who assess submissions, students, and institutions who generously give their expertise in factors small and large— their constant encouragement has helped a lot in the progress of the journal and shall help in future to earn credibility amongst all the reader members. I add a personal thanks to the whole team that has catalysed so much, and I wish everyone who has been connected with the Journal the very best for the future.

Thank you for Sharing Wisdom!

Managing Editor
IJARAI
Volume 1 Issue 2 May 2012
ISSN: 2165-4069(Online)
ISSN: 2165-4050(Print)

Editorial Board

Reviewer Board Members

CONTENTS

A Fuzzy Approach To Classify Learning Disability

Pooja Manghirmalani
Department of Computer Science
University of Mumbai
Mumbai-98, India.

Darshana More
Department of Computer Science
University of Mumbai
Mumbai-98, India.

Kavita Jain
Department of Computer Science
University of Mumbai
Mumbai-98, India.

Abstract -**The endeavor of this work is to support the special education community in their quest to be with the mainstream. The initial segment of the paper gives an exhaustive study of the different mechanisms of diagnosing learning disability. After diagnosis of learning disability the further classification of learning disability that is dyslexia, dysgraphia or dyscalculia are fuzzy. Hence the paper proposes a model based on Fuzzy Expert System which enables the classification of learning disability into its various types. This expert system facilitates in simulating conditions which are otherwise imprecisely defined.**

Keywords- Learning Disability; Dyslexia; Dysgraphia; Dyscalculia; Diagnosis; Classification; Fuzzy Expert System.

I. Introduction

Learning disability refers to a neurobiological disorder which affects a person’s brain and interferes with a person's ability to think and remember [1]. The causes that lead to learning disability (LD) are maturational delay, some unexplained disorder of the nervous system and injuries before birth or in early childhood. Children born prematurely and children who had medical problems soon after birth can also inherit LD[2].LD can be broadly classified into three types. They are difficulties in learning with respect to read (Dyslexia), to write (Dysgraphia) or to do simple mathematical calculations (Dyscalculia) [3].

The term “specific learning disability” means a disorder in one or more of the basic psychological processes involved inunderstanding or in using language, spoken or written, which may manifestitself in an imperfect ability to listen, speak, read, write, spell, or to do mathematical calculations.The term includes such conditions as perceptual handicaps, brain injury, minimal braindysfunction, dyslexia, and developmentalaphasia. The term does not include children who havelearning disabilities which are primarily the result of visual, hearing, or motor handicaps, ormental retardation, or emotional disturbance, or of environmental, cultural, or economicdisadvantage [4].

A child has a specific learning disability if a severe difference between achievement and intellectual ability in one or more of the areas: (1) oral expression; (2) listening comprehension; (3) written expression; (4) basic reading skill; (5) reading comprehension; (6) mathematics calculation; or (7) mathematic reasoning [5].

LD cannot be cured completely by medication. Children suffering from LD are made to go through a remedial study in order to make them cope up with non-LD children of their age. For detecting LD, there does not exist a global method.

This paper proposes a model for diagnosis and classification of LD. Section II of this paper explores in detail different computational methods and models applied in detecting or diagnosing LD. Having elaborately explored different approaches, we have found that there are still possible ways of approaching the given problem. Section III discusses the Fuzzy Expert System designed to classify the problem of LD. Section IV gives the implementation requirement of the system and sections V and VI discusses the results and conclusion respectively.

II. Taxonomy

The different computational methods and models used in detecting LD can be classified into four groups. The grouping is done based on the broader theoretical foundation and computational characteristics of the models and methods applied. The following sub- section deals with such models and their accuracy to diagnose LD. Once a child is diagnosed with LD, the classification of LD can be done. The literature discusses various methodologies likeDigital Signal Processing (DSP), Digital Imageprocessing (DIP), Soft Computing and hybrid form ofthese techniques to diagnose LD.

Reitano [6] used DSP techniques to diagnose LD by comparing pre-recorded and properly pronouncedphonemes with mispronounced phonemes. Fonseca et al[7] conducted electroencephalograms (EEG) to detectabnormalities related to electrical activity of the brain bystudying different brainwaves. He concluded that there isa significant difference in brainwaves of normal andlearning disabled children.

Mico-Tormos et al [8] inferred that eye movements ofeven an infant could indicate LD by analyzing theresponses of the improvement of eye through Oculographic signals. Pavlidis [9] observed that erratic and strikinglylarge number of regressive eye movements pointed todyslexia.

Jain et al [10] proposed a simple perceptron basedartificial neural network (ANN) model for diagnosing LDusing curriculum based test conducted by specialeducators. Bullinaria [11] applied a multi-layer feedforward perceptron to diagnose dyslexia where letterstrings were mapped to phoneme strings in multi- syllabicwords. Wu et al [12] proved that multi-layer perceptron withback propagation gave better results in diagnosing LD. Helater attempted to diagnose LD using support vectormachines (SVM) [13]. Salhi et al [14] used both wavelettransforms and ANN to diagnose LD from pathologicalvoices. Novak et al [15] have calculated a set of featuresfrom signals of horizontal and vertical eye movement using self-organizing map and genetic algorithm (GA). They concluded that the reading speed increased with the probability of the patient being healthy.

Wu et al [16] further combined different feature selection algorithms like brute-force, greedy and GA along with ANN to improve the identification rate of LD. Georgopoulos et al [17] proposed that a hybridization of GA and fuzzy cognitive map was better equipped for accurate diagnosis. Macašet al [18] developed a system for extracting the features of eye movements from time and frequency domain. They concluded that back propagation based classification gave better results than that offered by Bayes' and Kohonen network.

Manghirmalani et al proposed a soft computing technique called Learning Vector Quantization. The model classifies a child as learning disabled or non- learning disabled. Once diagnosed with learning disability, rule based approach is used further to classify them into types of learning disability that is dyslexia, dysgraphia and dyscalculia [19].

Given in Table-1 is the comparative study of different soft computing techniques based on their accuracy.

TABLE I. COMPARATIVE STUDY OF DIFFERENT SOFT COMPUTING MODELSDESIGNED TO DIAGNOSE LD.

Soft Computing Methodology	Accuracy in %
Classification Tree	71.79 %
K Nearest Neighbour Algorithm	76.92 %
Simple Vector Machine Algorithm	80.95 %
Bayes Classifier	82.69 %
Single Layer Perceptron Algorithm	84.00 %
Feature Selection of Multilayer Perceptron Algorithm	86.00 %
Back Propagation Algorithm	86.54 %
Logistic Regression Algorithm	87.18 %
Multilayer Perceptron Algorithm with Genetic Algorithm	88.2 %
Back-Propagation Gradient Descent Learning Algorithm	89 %
Multilayer Perceptron Algorithm	89.74 %
Naive Bayes classifier	94.23%
Decision Tree classifier	96.15%
Learning Vector Quantization Algorithm	91.8 %

III. SOFT COMPUTING APPROACH – FUZZY EXPERT SYSTEM FOR CLASSIFYING LEARNING DISABILITY

The soft computing technique called Fuzzy Expert System provides an alternative way to model human thinking with the concept of linguistic variables [20]. It deals with real world vagueness that isthe observation of parent and teacher. Hence classifying LD using this approach is quite significant.

A. Collection of Exhaustive Parameters

A curriculum-based test was designed with respect to the syllabus of primary-level school going children. This test was conducted in schools for collecting LD datasets for testing. Historic data for LD cases were collected from LD Clinics of Government hospitals where the tests were conducted in real-time medical environments. The system was fed with 11 input units which correspond to 11 different sections of the curriculum-based test.

Table-1 shows the initial 11 inputs corresponding to curriculum-based test. Column 1 represents the name of the parameter, column-2 represents the total marks allocated to a particular section, and column-3 determines the category of LD a section corresponds to.

Dataset consists of 170 cases of LD children. The system was trained using 100 data items and the remaining was used to test the system.

TABLE II. PARAMETERS AND MARKS OF CURRICULUM -BASED TEST.

Input Parameter	Marks	Category of LD
Essay	10	Dysgraphia
Reading	10	Dyslexia
Comprehension	10	Dyslexia,Dysgraphia
Spelling	10	Dysgraphia
Perception	10	Dyslexia
Solve	10	Dysgraphia
Word Problem	10	Dyscalculia,Dyslexia
Mental Sums	10	Dyscalculia
Time	10	Dyscalculia
Calander	05	Dyscalculia
Money	05	Dyscalculia

B. Fuzzy Expert System for LD

The literature survey indicates that most of the work done till now are for diagnosing the problem i.e. diagnosing methods can only detect whether a child is having or not having an LD, which is anbinary output.

The model proposed below assists in the classification oftype of LD that is dyslexia, dysgraphia or dyscalculia. Since fuzzy expert system is the best method for classification in soft computing, it is used in this system which can give an accuracy of approximately 90%.

Fuzzy expert system is able to process incomplete data and provide approximate solutions to problems which other methods find difficult to solve. Also fuzzy expert system uses rules which are the most fundamental part for classification purpose.

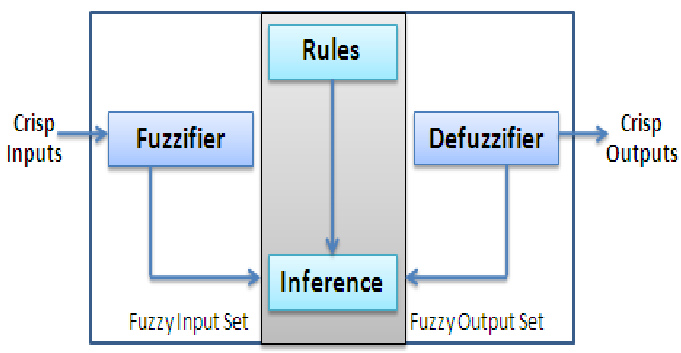
The reasoning in fuzzy expert system is similar to human reasoning. It allows for approximate values and inferences as well as incomplete or ambiguous data (fuzzy data) as opposed to only relying on crisp data.

Terminologies used in fuzzy expert system are not used in other methods [21]. They are: very high, increasing, somewhat decreased, reasonable and very low.

a) A Fuzzy Logic System

A fuzzy logic system (FLS) can be defined as the non-linear mapping of an input data set to a scalar output data. A FLS consists of four main parts: fuzzifier, rules, inference engine, and defuzzifier. FLS is one of the most famous applications of fuzzy logic and fuzzy sets theory. They can be helpful to achieve classification tasks, offline process simulation and diagnosis, online decision support tools and process control[22].

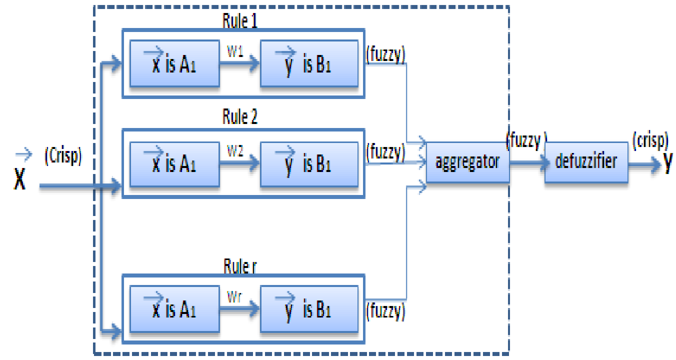
As shown in figure-1, firstly, a crisp set of input data are gathered and converted to a fuzzy set using fuzzy linguistic variables, fuzzy linguistic terms and membership functions. This step is known as fuzzification. Afterwards, an inference is made based on a set of rules. And lastly, the resulting fuzzy output is mapped to a crisp output using the membership functions, in the defuzzification step.



Graph1 Fuzzy Inference System

b) Fuzzy Modeling

As show in the figure-2, given below are the seven steps of fuzzy model applied to the data [23].



Graph2 Fuzzy expert system steps.

1. *Define the linguistic variables and terms.*
2. Construct the membership functions.
3. *Construct the rule base.*
4. *Convert crisp input data to fuzzy values using the membership functions.*
5. *Applying implication method.*
6. *Aggregating all output.*
7. *Convert the output data to non-fuzzy values.*

1) Linguistic variables

- The first step is to take inputs and determine the degree to which they belong to each of the appropriate fuzzy sets via membership functions.
- Since there are 11 inputs out of which 9 are from 0 to 10 and 2 are from 0 to 5. For more classification ofthese inputs are partitioned into 4 linguistic variables Poor, Satisfactory, Good and Excellent.
- Applying Formula for each of these linguistic variables, 4 values in the range of 0 to 1 are returned respectively. Calculating the membership function values all are stored into database.

2) Formula for Membership Function

If variable as “input” is taken for the inputs of the system, then: [23]

a) for linguistic variables "Poor"

memfunc=1.0;	if--input>=0 && input<=4
memfunc=(4.5-input)/0.5	if--input>4 && input<=4.5
memfunc=0.0	if--input>4.5 && input<=10

b) for linguistic variables "Satisfactory"

memfunc=0.0;	if--input>=0 && input<=4
memfunc=(input-4)/0.5	if--input>4 && input<=4.5
memfunc=1.0	if-- input>4.5 && input<=5.5
memfunc=(6-input)/0.5	if-- input>5.5 && input<=6
memfunc=0.0	input>6 && input<=10

c) for linguistic variables "Good"

memfunc=0.0;	if--input>=0 && input<=5.5
memfunc=(input-5.5)/0.5	if--input>5.5 && input<=6
memfunc=1.0	if-- input>6 && input<=7
memfunc=(7.5-input)/0.5	if-- input>7 && input<=7.5
memfunc=0.0	input>=7.5 && input<=10

d) for linguistic variables "Excellent"

memfunc=0.0;	if--input>=0 && input<=7
memfunc=(input-7)/0.5	if--input>7 && input<=7.5
memfunc=1.0	if--input>7.5 && input<=10

In this manner, each input is fuzzified over all the qualifying membership functions required by the rules.

3) Rules of Classification

Case-1: IF score low in fields of essay, reading, comprehension, spelling, perception, solve, word problems, mental sums, time, money;
THEN Dyslexia, Dysgraphia and Dyscalculia

Case-2: IF score low in fields of reading, comprehension, perception, word problem;
THEN Dyslexia

*Case-3:*IF score low in fields of spelling, comprehension, essay;
THEN Dysgraphia

*Case-4:*IF score low in fields of solve, mental sums,word problems, sums related to time, calendar, money;
THEN Dyscalculia

Case-5: IF score low in fields of reading, comprehension, perception word problem, spelling, essay;
THEN Dyslexia and Dysgraphia

Case-6: IF score low in fields of reading, comprehension, perception, solve, mental sums,word problems, sums related to time, calendar and money;
THEN Dyslexia and Dyscalculia

Case-7: IF score low in fields of spelling, comprehension essay, solve, mental sums,word problems, sums related to time, calendar,money;
THEN Dysgraphia and Dyscalculia

Many a times, one type of LD leads to other two types as well; thus most children are detected with all three types of LDs. Unique cases have one type of LD diagnosed or sometimes a combination of two [24].

1) Fuzzy Operators

- Now here the Rules are applied which have two parts[23]:
  - Antecedent (before "Then")
  - Consequent (after "Then").
- Once the inputs have been fuzzified, we know the degree to which each part of the antecedent. The values which are stored into the database are fetched for applying rules. By following the rules, the "Poor" linguistic variable is taken into consideration.
- For the 2nd rule, fields of reading, comprehension, perception, word problem; where the poor scores of these fields are taken and the AND(MIN) operator is applied to the linguistic values of each field.
- For Eg. Suppose Rule 2 has values like {0.2, 0.7, 0.8, 0.1} thus the output will be 0.1. Thus forming a set, for all students, let us consider this set as the 1st array.

2) Applying Implication Method

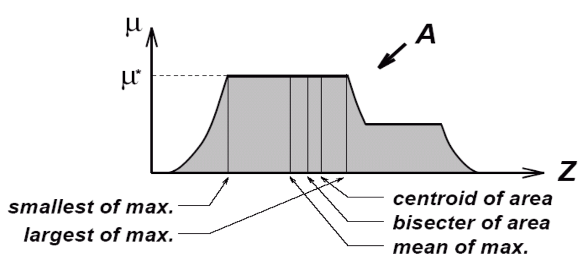
- The input for the implication process is a single number given by the antecedent, and the output is a fuzzy set. Implication is implemented for each rule.
- Two built-in methods are supported, and they are the same functions that are used by the AND method: min (minimum), which truncates the output fuzzyset, and prod (product), which scales the output fuzzy set.
- As the membership values are calculated for antecedents, the membership functions for the consequents are also calculated. Thus a set of values for each input and thus for all the students is computed. Let us consider this set as the 2nd array.
- The implication operator is applied i.e. the Min Operator between the values of the antecedent and the consequent. Thus a fuzzy set is computed for all students, and for each rule fuzzy sets are formed.
- Implication method finds the minimum of the corresponding values from the 1st array and the 2nd array thus forming a new set
- For Eg. The value of 1st array is 0.1 and 2nd array is 0.4, then the output of implication method will be 0.1. Each rule will give this kind of values, to form sets.

3) Aggregrating all Outputs

- Because decisions are based on the testing of all of the rules in a FIS, the rules must be combined in some manner in order to make a decision.
- Aggregation is the process by which the fuzzy sets that represent the outputs of each rule are combined into a single fuzzy set.
- As long as the aggregation method is commutative (which it always should be), then the order in which the rules are executed is unimportant. Three built-in methods are supported[23]:
  - ➢ max (maximum)
  - ➢ probor (probabilistic OR)
  - ➢ sum (simply the sum of each rule's output set)
- The rules are combined using the Sum aggregation method which is formed into a single fuzzy set.
- Since there are 3 LDs, the rules output sets are aggregated to form 3 Aggregation sets, which can be further defuzzified.

4) Defuzzify

- The input for the defuzzification process is a fuzzy set (the aggregate output fuzzy set) and the output is a single number.
- There are five built-in methods supported: centroid, bisector, mean of maximum (the average of the maximum value of the output set), largest of maximum, and smallest of maximum[23].
- Perhaps the most popular defuzzification method is the Centroid calculation, which returns the center of area under the curve.



Graph3 Different methods for Defuzzification

- Centroid Defuzzification formula[27]:

$$z_{COA} = \frac{\int_Z \mu_A(z) z dz}{\int_Z \mu_A(z) dz},$$

where $\mu_A(z)$ is the value from the aggregated set and z is the range of value from 0 to 100.

- Example: If one of the aggregated set contains values

{0.2, 0.5, 0.8, 1}

thenthe defuzzified value is equal to :
=(10+20+30+40)0.2+(50+60+70)0.5+(80+90)0.8+(100)1
(0.2+0.2+0.2+0.2+0.5+0.5+0.5+0.8+0.8+1)
= 72.24

Thus, by using this formula, the threshold values for all 3 LDs, by performing 3 defuzzifications.

c) Classification for LD

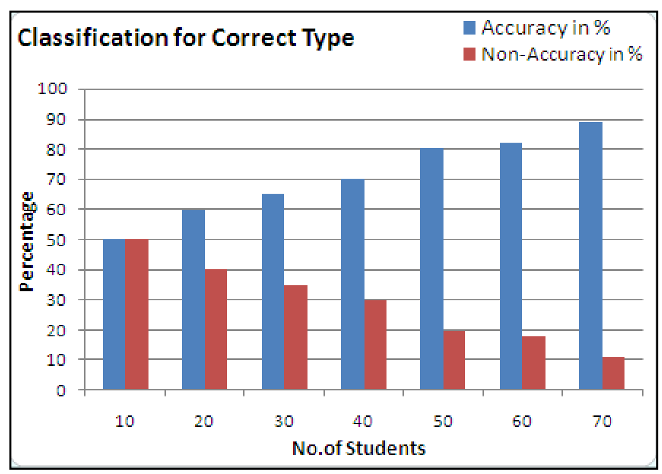
- After some processing, at the aggregation step the rules need to be aggregated.
- While aggregation the classification is done on the basis that the rules having dyslexia as their consequent are aggregated together. (Same is done for dysgraphia and dyscalculia.)
- Thus, the aggregations are defuzzified for all three LDs and thus getting the output which in the form of percentage.
- This value being the required threshold value is used for proper classification.
- Example: If the value getting after inputting new data, is less than or equal to the threshold value, then he/she is suffering from that particular LD.

IV. Implementation

The system is implemented using JAVA. The data collected is stored in Excel sheets. The experiments were conducted on a workstation with an Intel Pentium(R) 4 CPU, 3.06ghz, 2GB of RAM, running on Microsoft Windows XP Home Edition, Version 2002 with Service Pack 3.

V. Results

A. Classification for Correct Type

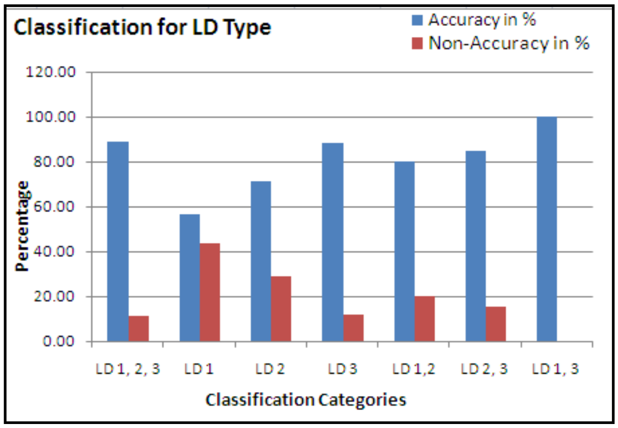


Graph4 Classification for Correct type of LD showing the accuracy and the non- accuracy during the training phase.

Graph-1 gives an overview of the system's accuracy rate. As more data is added to the database, the system's accuracy increases.Initially when the database had 10 records as Training Set, it gave an accuracy of 50%, i.e. the system correctly classified Learning Disabilities 50% accurately.When data in the system training set was increased to 20, the accuracy rate went up to 60% and so on.For 70 data in the Training set, the system is giving an accuracy of 90%.

B. Classification for Specific LD

Graph-2 shows that, after all training and testing of the system, the percentage for all 3 types of LDs is 90%.



Graph5 Graph 2: Classification for specific LD showing the accuracy and the non-accuracy during the training phase.

VI. Conclusion and Future Work

After never-ending deliberations with doctors, special educators, teachers and parents of children with LD, we could understand and monitor the process of classifying LD in India. Soft computing and other techniques are helpful in diagnosing LD with great accuracy, but once diagnosed, it is still difficult to judge the specific LD amongst children. Thus the remedial education given to these cases becomes very general.

We have made an attempt to classify LD using yet another soft computing technique called Fuzzy Expert System. The method is not only simple and easy toreplicate in huge volumes but gives good resultsbased on accepted benchmarks. However, there is scope for further enhancement of system by finding a combination of algorithms so as to build up a model that is satisfactorily more accurate. It is also seen that on increasing the number of data in the training set of the system, the overall accuracy shows a promising growth.

In future we intend to explore the possibility of parameter classification in order to distinguish irrelevant and superfluous variables which might lead to decrease in diagnosis process time and increase in accuracy. This can be beneficial for the special educators, doctors and teachers by providing suggestions that lead to the exclusion of redundant tests and saving of time needed for diagnosing LD. On the whole, the focus of our research is to identify early diagnosis and proper classification of LD and to support special education community in their quest to be with mainstream.

VII. Acknowledgment

Authors acknowledge the contributions of accredited Government hospitals and schools in Mumbai, India. They also express their gratitude to doctors, special educators and parents of LD diagnosed children for their time, support, guidance and patience.

References

[1]. S.A Kirk, Educating Exceptional Children Book, Wadsworth Publishing, ISBN: 0547124139.

[2]. Lisa L. Weyandt; "The physiological bases of cognitive and behavioral disorders"; Blausen Medical Communications, United States.

[3]. Lerner, Janet W, "Learning disabilities: theories, diagnosis, and teaching strategies"; Boston: Houghton Mifflin; ISBN 0395961149.

[4]. Fletcher, Francis,Rourke, Shaywitz & Shaywitz; "Classification of learning disabilities: an evidencebased evaluation"; 1993.

[5]. Kaplan, Robert M.; Saccuzzo, Dennis P. "Psychological Testing: Principles, Applications, and Issues (Seventh ed.)", Belmot (CA): Wadsworth. p. 262 (citing Wechsler (1958) The Measurement and Appraisal of Adult Intelligence), ISBN 978-0- 495-09555-2, 2009.

[6]. Carmen T. Reitano, "System & method for dyslexia detection by analyzing spoken & written words", United States Patent US 6535853 B1 March 18, 2003.

[7]. Lineu C. Fonseca, Glória M.A.S. Tedrus, Marcelo G.Chiodi, Jaciara Näf Cerqueira, Josiane M.F. Tonelotto Arq Neuropsiquiatr, "Quantitative EEG in children with learning disabilities", Analysis of band power, 64(2-B):376-381, 2006.

[8]. Pau Mico-Tormos, David Cuesta-Frau, Daniel Novak,"Early Dyslexia Detection Techniques by means of Oculographic Signals", 2nd European Medical & Biological Engineering Conference, Vienna, Austria. December 2002.

[9]. George Th.Pavlidis, "Eye Movement in Dyslexia: Their Diagnostic Significance", Journal of Learning Disabilities, 2001.

[10]. Kavita Jain, Pooja Manghirmalani, Jyotshna Dongardive, Siby Abraham, "Computational Diagnosis of Learning Disability" , International Journal of Recent Trends in Engineering, Vol 2, No. 3, November 2009.

[11]. John A Bullinaria, "Neural Network Models of Reading Multi-Syllabic Words", International Joint Conference on Neural Networks, 1993.

[12]. Tung-Kuang Wu, Ying-Ru Meng and Shian-Chang Huang, "Application of ANN to the identification of students with LD",IC-AI: 162-168. 9, 2006.

[13]. Tung-Kuang Wu, Ying-Ru Meng and Shian-Chang Huang, "Identifying & Diagnosing Students with LD Using ANN &SVM", IEEE International Joint Conference on Neural Networks,Vancouver, BC, 2006.

[14]. L. Salhi, M. Talbi, A. Cherif, "Voice Disorders Identification usingHybrid Approach: Wavelet Analysis and Multilayer NeuralNetworks",Proceedings of World Academy of Science,Engineering and Technology Volume 35, ISSN 2070-374,November 2008.

[15]. D. Novak, P. Kordk, M. Macas, M.Vyhnalek, R. Brzezny, L.Lhotska "School Children Dyslexia Analysis using SelfOrganizing Maps",IEEE 26th Annual International Conference,San Francisco, CA, 2004.

[16]. Tung-Kuang Wu, Ying-Ru Meng and Shian-Chang Huang,"Effectsof Feature Selection on the Identification of Students with LDUsing ANN", Springer-Verlag Heidelberg, 2006.

[17]. Voula Georgopoulos and Chrysostomos Stylios, "GeneticAlgorithm enhanced fuzzy cognitive maps for medical diagnosis",IEEE International conference on Fuzzy Systems, Hong Kong,2008.

[18]. Martin Macaš, Lenka Lhotská and Daniel Novák, "Bioinspiredmethods for analysis and classification of reading eye movementsof dyslexic children", Department of Cybernetics,Czech TechnicalUniversity in Prague, Czech Republic NiSls Symposium 2005.

[19]. Pooja Manghirmalani, Zenobia Panthaky, Kavita Jain; 'Learning Disability Diagnosis and Classiication-a Soft Computing Approach'; IEEE World Congress on Information and Communication Technologies (WICT); DIO: 10.1109/WICT.2011.6141292, 2011.

[20]. R. Jang , T.Sun, E.Mizutani, Neuro –Fuzzy And Soft Computing

[21]. Novák, V., Perfilieva, I. and Močkoř, J. Mathematical principles of fuzzy logic Dodrecht: Kluwer Academic. ISBN 0-7923-8595-0, 1999.

[22]. M.A. Arjona L.,R. Escarela-Perez,E. Melgoza- 'Vázquezc, C. Hernández.F., Convergence improvement in two-dimensional finite element nonlinear magnetic problems—a fuzzy logic approach'; Finite Elements in Analysis and Design, Science Direct, 2004. DOI: 10.1016/S0168-874X(02)00165-8

[23]. George J. Klir/Bo Yuan, "Fuzzy Sets and Fuzzy Logic: Theory and Applications"

[24]. J. Mendel. Fuzzy logic systems for engineering: a tutorial. Proceedings ofthe IEEE, 83(3):345{377, Mar 1995

AUTHORS PROFILE

Pooja Manghirmalani: She has received her M.Sc. Computer Science degree from University of Mumbai. Currently she is working as an Assistant Professor at the Department of Computer Science, University of Mumbai, Mumbai, India. She has published papers in International Journals and conference proceedings for Learning Disability. Her research interest includes Artificial Intelligence, Machine Learning and Cloud Computing.

Darshana More: She is doing her M.Sc. Computer Science from University of Mumbai, Mumbai, India. She has worked on several government projects and has currently taken up research to support special children of the society. Her areas of interest include Artificial Intelligence, System & Internet Security and Networking.

Kavita Jain: She has received her M.Sc. Computer Science degree from University of Mumbai. Currently she is working as an Assistant Professor at the Department of Computer Science, University of Mumbai, Mumbai, India. She has published papers in Internatinal Journals and Conference Proceedings for Learning Disability. Her research interest includes Bioinformatics, Artificial Intelligence and Machine Learning.

A New Machine Learning Approach to Deblurring License Plate Using K-Means Clustering Method

Sanaz Aliyan[1]
Islamic Azad University, South Tehran Branch, Tehran, Iran

Ali Broumandnia[*2]
Islamic Azad University, South Tehran Branch, Tehran, Iran

***Abstract*-Vehicle license plate recognition (LPR) is one of the important fields in Intelligent Transportation Systems (ITS). LPR systems aim to locate, segment and recognize the license plate from captured car image. Despite the great progress of LPR system in the last decade, there are still many problems to solve to reach a robust LPR system adapted to different environment and condition. The current license plate recognition systems will not effectively work well for blurred plate image. In this paper, to overcome the blurring problem a new machine learning approach to Deblurring License Plate using the K-Means clustering method have proposed. Experimental results demonstrate the effectiveness of the K-Means clustering as a feature selection method for license plate images.**

Keywords- license plate recognition; K-Means clustering; deblurring; machine learning.

I. Introduction

Recently, the vehicle license plate recognition (LPR) technology has developed and applied in our life and has been used extensively in highway and bridge charge, port, airport gate monitoring, and so on. There are many applications for Intelligent Transportation Systems (ITS) such as traffic analysis, parking automation, identification of stolen vehicles, the license plate recognition and, etc. A complete license plate recognition system consists of three major parts: license plate location, license plate segmentation and character recognition [1-3]. The first step locates LP regions and extracts an image containing a plate. The segmentation step separates the symbols or characters from each other in one LP, and the recognition step finally converts the grey-level image block into characters or symbols by pre-defined recognition models. Recognition step refers to image analysis and pattern recognition [4]. Despite the great progress of LPR system in the last decade, there are still many problems to solve to reach a robust LPR system adapted to different environment and condition. Assuming that license plate region is detected in very low resolution (LR), the current license plate recognition systems will not effectively work well [5]. Motion blurring is one of the prime causes of poor image quality in digital imaging. If objects in a scene are moving fast or the camera is moving over the period of exposure time, the objects or the whole scene will look blurry along the direction of relative motion between the object or scene and the camera. Since high resolution (HR) digital cameras are expensive, finding a way to increase the current resolution level is needed. One promising approach is to use signal processing techniques to obtain an HR image (or sequence) from observed multiple low-resolution (LR) images. Recently, such a resolution enhancement approach has been one of the most active research areas, and it is called super resolution (SR) (or HR) image reconstruction. The major advantage of the signal-processing approach is that it may cost less, and the existing LR imaging systems can be still utilized.

In the past, multi-image based approaches have been proposed to solve motion de-blurring, which recovers clear images from motion-blurred images [5]. In multi-image based SR approaches used multiple low-resolution images is assumed same size images while in LPR case, there are different sizes of LR images because a license plate is usually captured when the vehicle is moving. Hence, images captured at different time instances may provide different perspectives because they were captured at different angles in the field of view. This makes the registration task even more difficult. To overcome this problem we work on single image of license plate using a new machine learning approach using K-Means clustering. In general, a machine learning technique using training data set containing low resolution images predicts a high-resolution image corresponding to its low resolution image. Since Noise and redundancy in the feature space increase the likelihood of over-fitting, K-Means clustering is applied to decrease them as a feature selection method.

II. Related Work

The first article that clearly introduced the idea of combing SR and LPR to identify moving vehicles was written by Suresh and Kumar [8]. They proposed a little robust super-resolution algorithm is proposed, in which the HR image is modeled as MRF with a discontinuity adaptive regularization (denoted as “DAMRF”). In [9], an alternative generalized model of DAMRF based on the bilateral filtering which connects the bilateral filtering with the Bayesian MAP is proposed. This DAMRF model has edge-preserving and robust to noise from the bilateral filtering. Furthermore, they applied a method for estimating the regularization parameter automatically to the generalized DAMRF super-resolution reconstruction method. Since the prior was non-convex, they used graduated non-convex (GNC) which is a determining annealing algorithm for performing optimization. However, complex and difficult computational functions of their Maximum a posterior (MAP) based method, prevent it from real-time processing. Later, a fast MAP-based SR algorithm was proposed by Yuan [10]. In this work by using Wiener filter, the computation complexity was decreased to make SR reconstruction, as a de-convolution improvement. But because of removing the priors and registration error, this method became unstable. Later, SR based on non-uniform interpolation of registered LR images was proposed by Gambotto [11].The advantage of this work was the fastness and the disadvantage of that was more sensibility to motion estimates. In [12], a unifying approach of SR license plate localization and for LPR to enhance the robust MAP based SR reconstruction excluding mis-registered images is proposed. The quality of registration before including an image to automatic rejection of mis-registered images is verified before the reconstruction step. A learning-based framework has been proposed in [16] for zooming the digits in a license plate, which have been blurred using an unknown kernel. In this paper, the image as an undirected graphical model over image patches is considered. First, they learn the compatibility functions by nonparametric kernel density estimation, using random samples from the training data. Next, the inference problem is solved by using an extended version of the nonparametric belief propagation algorithm. Finally, the super-resolved and restored images are recognized.

In multi-image based SR approaches used multiple low-resolution images is assumed same size images while in LPR case, there are different sizes of LR images because a license plate is usually captured when the vehicle is moving. Hence, images captured at different time instances may provide different perspectives because they were captured at different angles in the field of view. This makes the registration task even more difficult. Another drawback of previous LPR de-blurring researches is that the three parts of license plate recognition are not considered, and the plate is cropped manually.

III. THE PROPOSED APPROACH

Learning-based techniques attempt to capture the co-occurrence prior between low-resolution and high-resolution image patches. Recently, some studies represented image patches as a sparse linear combination of elements from an over-complete image patch dictionary [13-15]. In [15] they rely on patches from the input image and learn a compact representation for these patch pairs to capture the co-occurrence prior. Their algorithm tries to infer the high-resolution image patch for each low-resolution image patch from the input. For this local model, they have two dictionaries Dh and Dl, which are trained to have the same sparse representations for each high-resolution and low-resolution image patch pair.

This method is applied for generic and face images. The experimental results have shown that it is very effective for image patches. Sparse representation is not a suitable algorithm for real time system such as license plate recognition that aims to respond fast, since it solves an optimizing problem to generate a high-resolution image that it is time consuming. In this paper to develop a real time deblurring algorithm based on machine learning approaches is used K-Means clustering. In general, machine learning approaches work based on training data set, thus in a de-blurring license plate system, low-resolution plate images and corresponding high-resolution are considered as training data set. The important step to apply machine learning techniques is defining proper features that in this paper 3 * 3 patches are used as features. In fact, a de-blurring trained system reconstructs 3*3 patches of low-resolution image one by one to reach a high-resolution image. The performance of the system is related on the amount of patches as training data set; however, system that uses large data set is not fast because of redundancy in the feature space. Since there are many same and similar patches in data set, it can be grouped them into some cluster and define the median of the cluster as index patch. Proposed approach consists of two phases. First, training tow impacted dictionary simultaneously from data set using K-Means clustering and then de-blurring input low-resolution license plate image.

A. Joint Dictionary Training

Given the sampled training image patch pairs, DS={X,Y} where X = {x1, x2, ..., xn} are the sets of sampled high-resolution image patches and are Y= {y1, y2, ..., yn} the corresponding low-resolution image patches (or features), our goal is to learn dictionaries DH and DL for high-resolution and low-resolution image patches In first place, X is clustered into K clusters using K-means and then cluster Y based on their corresponding low-resolution patches. Finally, by the median operator calculate median for each cluster. Figure 1 shows the framework of joint dictionary training.

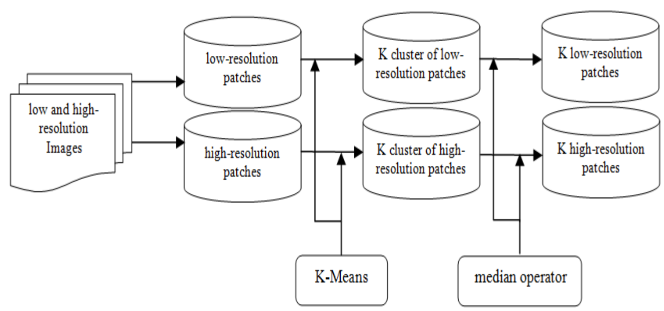


Figure 1. Joint dictionary training framework

B. De-Blurring License Plate Image

In this phase, low-resolution image is divided into 3*3 patches. De-blurring system find the most similar patch in DL for each patch and then use the corresponding patch in DH to construct the high-resolution image. The framework of de-blurring is shown in figure 2.

IV. EXPERIMENTAL RESULTS

In this paper, 151 car images are used as data set as shown in figure 3. A plate location technique is used to extract plate images. This technique work based on edge detection and histogram analysis, as is demonstrated in figure 3 the amount of histograms around the plate is more other parts.

Based on this hypothesize location of plate is detected and extracted. Blurred image is created using a motion filter (Figure 4). These image pairs show our training and testing dataset that is considered 140 and 11 image pairs respectively. The two dictionaries for high-resolution and low-resolution image patches are trained from 272 000 patch pairs sampled from training and 140 testing dataset images and fix the dictionaries size to be 2000. Proposed approach is implemented using MATLAB software.

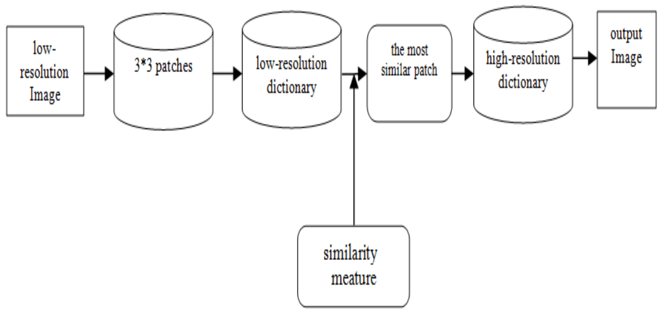


Figure 2. Deblurring License Plate Image Framework

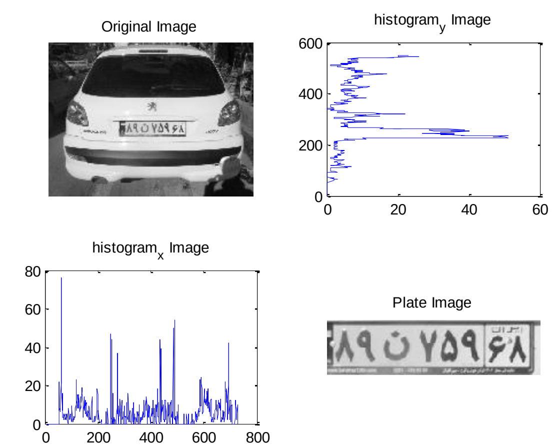


Figure 3. Extracted license plate using plate location technique

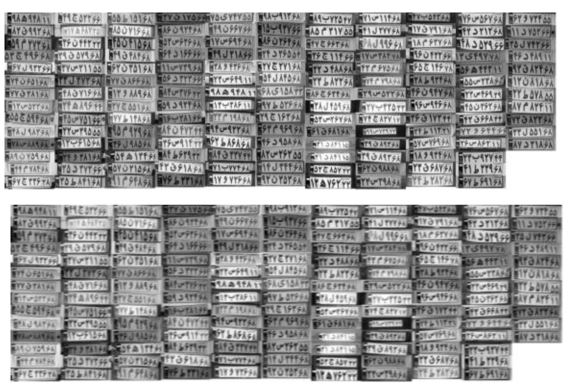
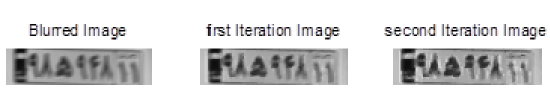


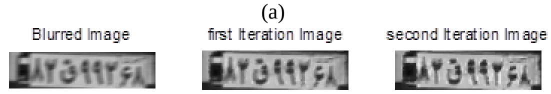
Figure 4. High and low-resolution license plate images

Proposed approach is implemented using MATLAB software.

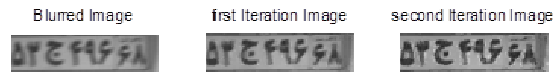
Figure 5 demonstrates the experimental results for two iterations of applying deblurring algorithm.



(a)



(b)



(c)

Figure 5. Deblurring LP images using proposed approach

V. Conclusion

Associations of super resolution techniques to the current license plate recognition systems have been reported. In this paper a novel approach toward single image SR based upon machine learning technique in terms of coupled dictionaries jointly trained from high- and low-resolution image patch pairs is presented. The compatibilities among adjacent patches are enforced both locally and globally. Experimental results demonstrate the effectiveness of the K-Means clustering as a feature selection method for license plate images. However, one of the most important questions for future investigation is to determine the optimal dictionary size for natural image patches in terms of SR tasks. Tighter connections to the theory of compressed sensing may yield conditions on the appropriate patch size, features to utilize and also approaches for training the coupled dictionaries.

References

[1] F. Wang, L. C. Man, B. Wang, Y. Xiao, W. Pan, and X. Lu, “Fuzzy-based algorithm for color recognition of license plates,” Pattern Recognit. Lett., vol. 29, no. 7, pp. 1007–1020, May 2008.

[2] V. Abolghasemi and A. Ahmadyfard, “An edge-based color-aided method for license plate detection,” Image Vis. Comput., vol. 27, no. 8, pp. 1134–1142, Jul. 2009.

[3] H. Caner, H. S. Gecim, and A. Z. Alkar, “Efficient embedded neuralnetwork-based license plate recognition system,” IEEE Trans. Veh. Technol., vol. 57, no. 5, pp. 2675–2683, Sep. 2008.

[4] A. Broumandnia, J. Shanbehzadeh, Fast Zernike wavelet moments for Farsi character recognition, Image and Vision Computing 25 (2007) 717–726.

[5] C. N. E. Anagnostopoulos, I. E. Anagnostopoulos, I. D. Psoroulas, V. Loumos, and E. Kayafas, “License plate recognition from still images and video sequences : A Survey,” IEEE Tran. Intell. Transp. Syatem, vol. 9, No. 3, 2008, pp. 377-391.

[6] B. Zitov´a and J. Flusser, “Image registration methods: a survey,” Image and Vision Computing, vol. 21, no. 11, pp. 977–1000, 2003.

[7] J. Tian and K.-K. Ma, “A survey on super-resolution imaging,” Signal, Image and Video Processing, 2011.

[8] K.V. Suresh, G.M. Kumar, and A.N. Rajagopalan, “Superresolution of License Plates in Real Traffic Videos”, IEEE Trans. Intelligent transportation systems, June 2007, Vol. 8, No. 2.

[9] Z. Weili and L. Xiaobo, “A Generalized DAMRF Image Model for Super-Resolution of License Plates”, in Photonics and Optoelectronic (SOPO), 2010 Symposium, pp. 1-4.

[10] J. Yuan, S.-D. Du, and X. Zhu, “Fast super-resolution for license plate image reconstruction,” in ICPR, IAPR, 2008, pp. 1–4.

[11] J.-P. Gambotto, M. Denis, and R. Livier, “Super-resolution of moving vehicles using an affine motion model,” in AVSS, IEEE, 2009, pp. 535–540.

[12] N. Chu Duc, M. Ardabilian and C. Liming, ” Unifying Approach for Fast License Plate Localization and Super-Resolution”, in Pattern Recognition (ICPR), 2010 20th International Conference, pp. 376-379.

[13] Wang, J., Zhu, S., Gong, Y.: Resolution enhancement based on learning the sparse association of image patches. Pattern Recognit.Lett. 31, 1–10 (2010).

[14] Kim, K.I., Kwon, Y.: Single-image super-resolution using sparse

regression and natural image prior. IEEE Trans. Pattern Anal.Mach. Intell. 32, 1127–1133 (2010).

[15] Yang, J.,Wright, J., Huang, T.,Ma,Y.: Image super-resolution via sparse representation. IEEE Trans. Image Process. 19, 2861–2873 (2010).

[16] Rajaram, S., Gupta, S., Petrovic, N., Huang, S.: Learning-based nonparametric image super-resolution. In: EURASIP Journal on Applied Signal Processing, vol. 2006, Article ID 51306, p. 111doi:10.1155/ASP/2006/5130, (2006).

Experimental Validation for CRFNFP Algorithm

Wang Mingjun, Yi Xinhua, Wang Xuefeng
School of mechanical engineering
Ningbo University of technology
Ningbo, China

Tu Jun
School of mechanical engineering
Shanghai Jiaotong University
Shanghai, China

***Abstract*—In 2010,we proposed CRFNFP[1] algorithm to enhance long-range terrain perception for outdoor robots through the integration of both appearance features and spatial contexts. And our preliminary simulation results indicated the superiority of CRFNFP over other existing approaches in terms of accuracy, robustness and adaptability to dynamic unstructured outdoor environments. In this paper, we further study on the comparison experiments for navigation behaviors of robotic systems with different scene perception algorithms in real outdoor scenes. We implemented 3 robotic systems and repeated the running jobs under various conditions. We also defined 3 criterion to facilitate comparison for all systems: Obstacle Response Distance (ORD), Time to Finish Job (TFJ), and Distance of the Whole Run (DWR). The comparative experiments indicate that, the CRFNFP-based navigating system outperforms traditional local-map-based navigating systems in terms of all criterions. And the results also show that the CRFNFP algorithm does enhance the long-range perception for mobile robots and helps planning more efficient paths for the navigation.**

Keywords- autonomous nagivation; stereo vision; machine learning; conditional random fields; scene analysis.

I. INTRODUCTION

Navigation in an unknown and unstructured outdoor environment is a fundamental and challenging problem for autonomous mobile robots. The navigation task requires identifying safe, traversable paths that allow the robot to progress toward a goal while avoiding obstacles.

Standard approaches to complete the task use ranging sensors such as stereo vision or radar to recover the 3-D shape of the terrain. Various features of the terrain such as slopes or discontinuities are then analyzed to determine traversable regions [2-5]. However, ranging sensors such as stereo visions only supply short-range perception and gives reliable obstacle detection to a range of approximately 5m[6]. Navigating solely on short-range perception can lead to incorrect classification of safe and unsafe terrain in the far field, inefficient path following or even the failure of an experiment due to nearsightedness [7, 8].

To address nearsighted navigational errors, near-to-far-learning-based, long-range perception approaches are developed, which collect both appearances and stereo information from the near field as inputs for training appearance-based models and then applies these models in the far field in order to predict safe terrain and obstacles farther out from the robot where stereo readings are unavailable [9-11]. We restrict our discussion to the online self-supervised learning since the diversity of the terrain and the lighting conditions of outdoor environments make it infeasible to employ a database of obstacle templates or features, or other forms of predefined description collections. The winner of DARPA Grand Challenge[10] combines sensor information from a laser range finder and a pose estimation system to first identify a nearby patch (a set of neighboring pixels) of drivable surface. And then the vision system takes this patch and uses it to construct appearance models to find the drivable surface outward into the far range. Happold and Ollis[9] propose a method for classifying the traversability of terrain by combining unsupervised learning of color models that predict scene geometry with supervised learning of the relationship between geometric features and the traversability. A neural network is trained offline on hand-labeled geometric features computed from the stereo data. An online process learns the association between color and geometry, enabling the robot to assess the traversability of regions for which there is little range information by estimating the geometry from the color of the scene and passing this to the neural network. The system of Bajracharya[11] consists of two learning algorithms: a short-range, geometry-based local terrain classifier that learns from very few proprioceptive examples; and a long-range, image-based classifier that learns from geometry-based classification and continuously generalizes geometry to the appearance.

Appearance-based near-to-far learning methods mentioned above do support the long-range perception which provides the “look-ahead” capability for complementing the traditional short-range stereo- or LIDAR-based sensing. However, appearance-based methods assume that the near-field mapping from the appearance to traversability is the same as the far-field mapping. Such an assumption does not necessarily hold due to the complex terrain geometry and varying lighting conditions in unstructured outdoor environment. Therefore, how to use other strategies or information to compensate for the mapping deviation begins to draw more attention.

Lookingbill and Lieb[12] use a reverse optical flow technique to trace back the current road appearance to how it appeared in previous image frames in order to extract road templates at various distances. The templates can be then matched with distant possible road regions in the imagery. However, trackable features, on which the reverse flow technique is based, are subject to the image saturation and scene elements occurrence patterns. Furthermore, changing illuminant conditions can result in unacceptable rates of misclassification. Noting that the visual size of features scales inversely with the distance from camera, Hadsell and

Sermanet[13] normalize the image by constructing a horizon-leveled input pyramid in which similar obstacles have similar heights, regardless of their distances from the camera. However, the distance estimation for different regions of images introduces extra uncertainties. In addition, this approach does not consider the influence of changing lighting conditions on appearances. Procopio[8] proposes the use of classifier ensembles to learn and store terrain models over time for the application to future terrain. These ensembles are validated and constructed dynamically from a model library that is maintained as the robot navigates terrain toward some goal. The outputs of the models in the resulting ensemble are combined dynamically and in real time. The main contribution of the ensembles approach is to leverage robots' past experience for classification of the current scene. However, since the validation of models is based on the stereo readings from the current scene, this approach is still subject to the mapping deviation.

In 2010, we proposed the model of CRFNFP[1] to incorporate both the spatial contexts and appearance information to enhace the perception robustness and self-adaptability to changing illuminant conditions. And simulation results indicated the superiority of CRFNFP over other existing approaches. In this paper, we further implement the CRFNFP model in a robotic system to study on the navigating behaviors in real scenes.

An outline of this paper is as follows: We first briefly describe the CRFNFP framework in section II. The system implementation of the robot will be detailed in section III and section IV provides the experiment results. We conclude our paper in section V with our further research in this area. And section VI indicates our future research.

II. CRFNFP FRAMEWORK

A. Model Summary

CRFNFP framework is a near-to-far learning strategy to recognize the far-field of the current scene. We first over segment the current scene into superpixels (a superpixel is a set of neighboring pixels) and update the classification database using training samples from stereo readings of near-field of the current scene. Then we incorporate both local appearance of and spatial relationshops (contexts) between regions in the CRFNFP framework to estimate the traversability of regions of the current scene.

The problem to be solved by CRFNFP is how to design a specific CRF framework with respect to the self-supervised, near-to-far learning in unstructured outdoor environments. To the best of our knowledge, ours is the first work that introduces and adapts the CRF-based framework to model the spatial contexts and to improve the long-range perception for mobile robots.

B. Model Definition

Let the observed data (local appearance) from an input image be given by $\boldsymbol{X} = \{\boldsymbol{x}_i\}_{i \in \mathbb{S}}$, where $\mathbb{S}$ is the set of sites (one site corresponds to one superpixel in our application) and $\boldsymbol{x}_i$ is the data from the i th site. The corresponding labels at the image sites, which indicate the category of the traversability of a region, are given by $\boldsymbol{L} = \{l_i\}_{i \in \mathbb{S}}$. In this work, we will be only concerned with binary classification, i.e., $l_i \in \{-1,1\}$, -1 for ground and 1 for obstacle.

Our CRFNFP model is based on the Conditonal Random Fields (CRF) model, so we first explain the CRF model.

CRF Definition: Let $\mathbb{G} = (\mathbb{S}, \mathbb{E})$ be a graph such that $\boldsymbol{L}$ is indexed by the vertices of $\mathbb{G}$. Then $(\boldsymbol{L}, \boldsymbol{X})$ is said to be a conditional random field if, when conditioned on $\boldsymbol{X}$, the random variables l_i obey the Markov property with respect to the graph: $\mathbb{P}(l_i \mid \boldsymbol{X}, \boldsymbol{L}_{\mathbb{S}-\{i\}}) = \mathbb{P}(l_i \mid \boldsymbol{X}, \boldsymbol{L}_{\mathbb{N}_i})$, where $\mathbb{S}-\{i\}$ is the set of all nodes in the graph except the node i, N_i is the set of neighbors of the node i in $\mathbb{G}$.

Given the observation $\boldsymbol{X}$, the CRFNFP defines the joint distribution over the labels $\boldsymbol{L}$ as

$$\mathbb{P}(\boldsymbol{L} \mid \boldsymbol{X}) = \frac{1}{\mathbb{Z}} \exp\left\{ \sum_{i \in \mathbb{S}} \mathbb{A}_i(l_i, \boldsymbol{X}) + \sum_{i \in \mathbb{S}} \sum_{j \in \mathbb{N}_i} \mathbb{I}_{ij}(l_i, l_j, \boldsymbol{X}) \right\}$$

(1)

Where Z is a normalizing constant known as partition function, and $\mathrm{A_i}$ and $\mathrm{I_{ij}}$ are the association and interaction potentials respectively.

The association potential is constructed by a Bayes Classifier, which directly maps appearance to traversability. And the interaction potential aims to incorporate spatial relationships and serves as the data-dependent smoothing function.

As a result, the CRFNFP framework not only includes appearance features as its prediction basis, but also incorporates spatial relationships between terrain regions in a principled way. Please refer to the reference[1] for the details of CRFNFP framework.

III. IMPLEMENTATION OF ROBOTIC SYSTEM

A. Summary of Hardware Components

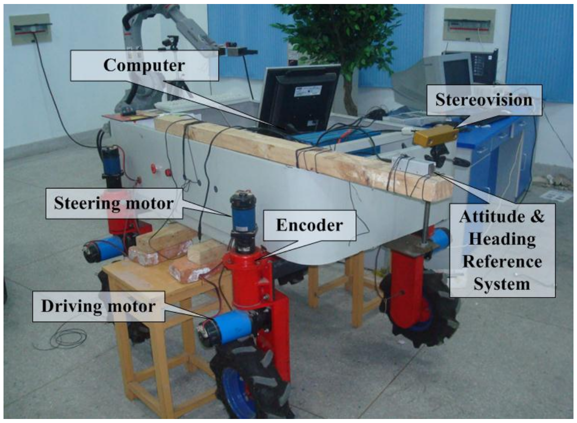


Figure 1. UGV used for navigation experiment.

The UGV is a four-wheeled, 8 DOF mobile robot with each wheel individually driven and steered to obtain the desired maneuverability. And the hardware of the UGV (as shown in Figure 1) mainly consist of vehicle body, an industrial personal computer (IPC), stereovision, AHRS (attitude and heading reference system), GPS (global positioning system). The hardware block diagram in Figure 2 shows the connection relationships among all components.

The GPS offers the global position while the AHRS combined with all the encoder provides the local position of the UGV. The stereo vision continuously takes picture of the current scenes, which is transmitted to the IPC. And the IPC will process all the information and provide the optimal control decision to further drive the UGV.

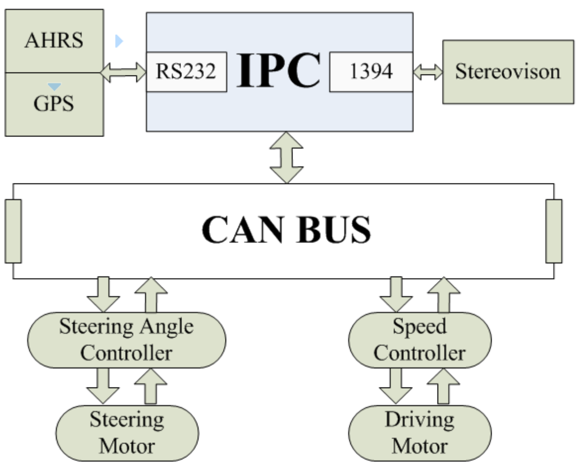


Figure 2. Hardware block diagram of UGV.

B. Flowchart Of Navigation Algorithm

The navigation job can be summarized as follows: given the target point, the robot goes from the start point to the endpoint while intelligently avoiding all the obstacles by taking corresponding actions.

The flowchart of Figure 3 shows that, the robot takes actions based on 3 sources of information:

1) *Near-field local mapping*: This mapping can model the local environment around the robot and provide the guidance for the robot to avoid close-range obstacles such as obstacles within 5 meters.

2) *Far-field path planning*: The inference results of far-field scenes can be used to generate the cost image, which represents the distant-range obstacle distribution (even obstacles up to 100 meters away). So the far-field path planning can lead the robot to avoid the distant obstacles ahead of time. And the correspondinng trajectory can be shorter and smoother while the robot reaches the same target point.

3) *Directional deviation computation:* The directional deviation is defined as the angle between the target point direction and the forward direction of the UGV. The robot needs to approach the target point while minimizing the directional deviation as much as possible.

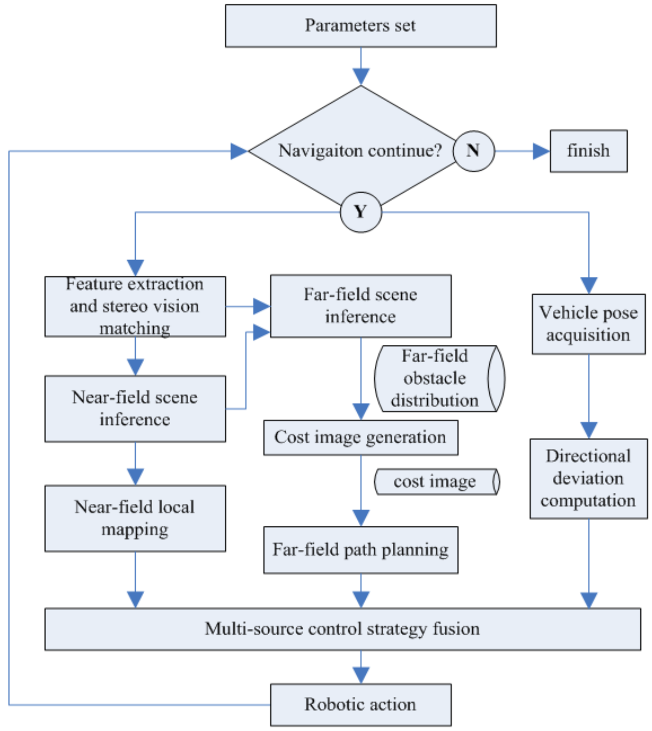


Figure 3. Flowchart of main algorithm of the navigation software.

IV. Experimental Analysis

A. Experimental Design

The experiments were carried out in the playground of Nanjing Agricultural University in december of 2009. The playground was muddy and full of weeds. The corresponding obstacles were manually arranged rectangular banners with the height of 1.2 meters as shown in Figure 4.

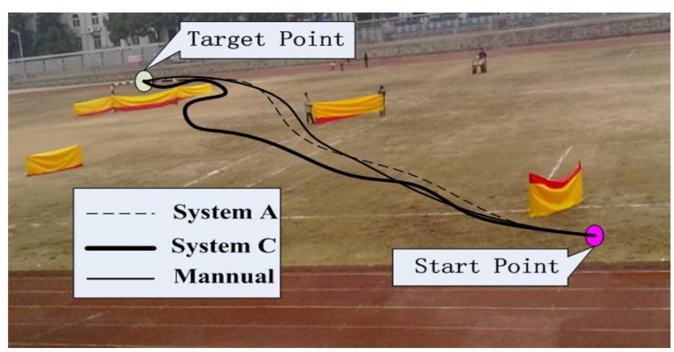


Figure 4. Playground for navigation experiments.

The experiments aim to validate that the CRFNFP model does enhance the long-range perception for mobile robots and helps to plan more efficient paths for the navigation job.

To achieve this goal, we arranged several sets of experiments under conditions of different obstacle color, system configuration and weather. There were 3 colors for obstacles: red(R), yellow(Y), and mixed(M, red banners combined with yellow ones). The weather condition contains sunny(S) and cloudy(C). The 3 system were configured as follows:

1) *System A: CRFNFP*-based far-field scene inference and subsequent far-field path planning, near-field local mapping, directional deviation computation;

2) *System B:* Bayes-classifier-based far-field scene inference and subsequent far-field path planning, near-field local mapping, directional deviation computation;

3) *System C:* Near-field local mapping, directional deviation computation.

The above mentioned Bayes-classifier can be regarded as a simplified CRFNFP model, which doesn't incorporate the spatial contexts for the recognition of scenes. In our implementation, all the algorithms are programmed under Visual C++ 6.0. Our CPU processor in the robot is 2.26 GHz Intel Core Duo P8400. And the running frequencies of A, B, C systems are 2Hz, 7Hz and 15Hz respectively with image resolution 320×240.

We ran system A and system B 5 times respectively for every combination of weather and obstacle color. And we ran system C only 2 times under the sunny and mixed obstacle color condition because system C were run only based on near-field local mapping and directional deviation computation, which means system C is not sensitive to the weather and obstacle color conditions. Furthermore, we mannually drove the robot one time to collect data for comparison. In all runs, the walking speed of the robot is 0.3m/s. We list typical experimental data in TABLE 1.

Figure 5 shows the typical running trajectories of different system configurations and the corresponding obstacles positions. In each subplot, the start point (S) and the circle G represent the position of vehicle and the target point respectively. In our experiments, if the distance between the vehicle and the target point is within 2 meters, we consider the job is already finished. In Figure 5(a), we collected one running trajectory of each system configuration represented by different line styles. And Figure 5(b), Figure 5(c) and Figure 5(d) show two running trajectories of each system configuration respectively. All the quantitative data of the six trajectories are listed in TABLE 1 by bold-type.

B. Evaluation Criterion

To the best of our knowledge, the research of long-range terrain perception for outdoor robots was started just from several years ago. And there is no generally accepted evaluation criterion. In order to better illustrate the significance of our experimental results, we define 3 creterion as follows:

1) *Obstacle Response Distance (ORD for short):* It is defined as the distance between the robot and the obstacle when the robot begins to recognize the obstacle steadily. Take ***Error! Reference source not found.****(a)* as an example, the distance between point C and the longest banner is the ORD of system A while the distance between point A and the longest banner is the ORD of system C.

2) *Time to Finish Job (TFJ for short):* It is defined as the total time for the robot to finish the job.

3) *Distance of the Whole Run (DWR for short):* It is defined as the distance that the robot experienced during the whole run.

Furthermore, we simplified the expressions of experimental conditions in TABLE 1 to facilitate the listing. For example, MIX-A-SUNNY-1 represents the condition of sunny, mixed color banner with the system A and Round 1.

TABLE I. Typical Running Results of Experiments

Experimental Condition	DWR (m)	TFJ (s)	ORD (m)
MIX-A-SUNNY-1	**86.436**	**279**	**47**
MIX-A-SUNNY-2	**84.519**	**273**	**49**
MIX-B-SUNNY-1	84.321	273	48
MIX-B-SUNNY-2	90.556	291	28
MIX-C-SUNNY-1	**94.177**	**304**	**5**
MIX-C-SUNNY-2	**113.460**	**366**	**4**
RED-A-SUNNY-1	87.345	283	46
RED-A-SUNNY-2	86.748	282	47
RED-B-SUNNY-1	88.574	285	40
RED-B-SUNNY-2	84.891	274	45
YELLOW-A-SUNNY-1	88.644	286	39
YELLOW-A-SUNNY-2	85.784	277	47
YELLOW-B-SUNNY-1	**83.525**	**269**	**48**
YELLOW-B-SUNNY-2	**106.880**	**345**	**4**
YELLOW-B-SUNNY-3	93.335	302	16
YELLOW-B-SUNNY-4	86.745	282	45
YELLOW-B-SUNNY-5	84.311	273	48
YELLOW-A-CLOUDY-1	87.675	282	42
YELLOW-A- CLOUDY-2	86.987	280.5	47
YELLOW-B- CLOUDY-1	84.587	274	49
YELLOW-B- CLOUDY-2	88.985	285	40
YELLOW-B- CLOUDY-3	95.366	307.5	15
YELLOW-B- CLOUDY-4	87.961	284	41
YELLOW-B- CLOUDY-5	92.388	298	22
MANUAL	83.500	269	—

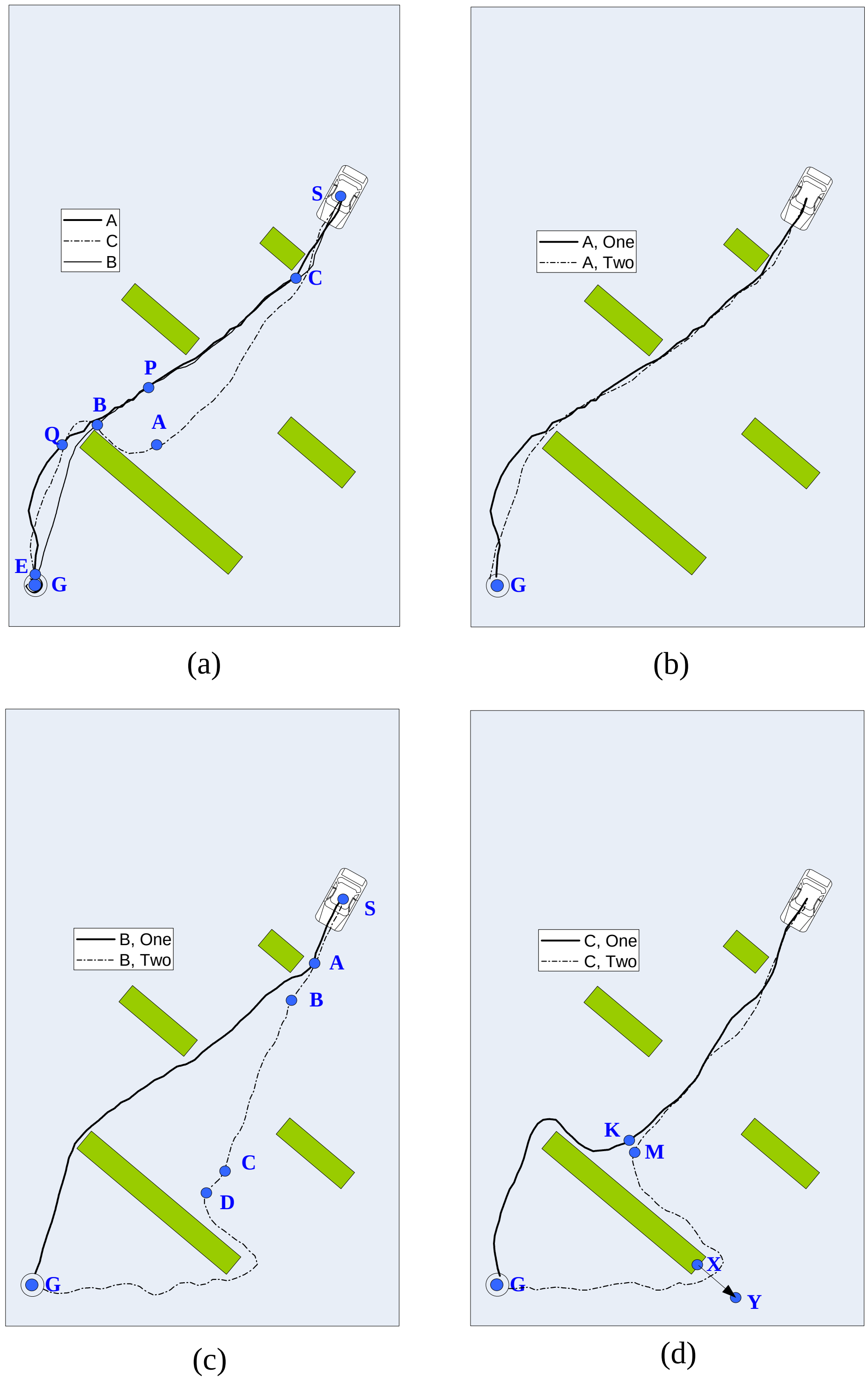


Figure 5. Trajectories of different runs of various system configuration.

C. *Experimental Results*

1) *First, we investigate the navigation behavior of System C:*

The robot based on system C doesn't infer any scenes, and the walking strategy can be summaried as:

a) If there are obstacles within the range of 8 meters around the robot, the robot turns left or right to avoid the obstacles;

b) If there are not obstacles within the range of 8 meters around the robot, the robot needs to caculate the directional deviation and the corresponding the turning angle, which controls the robot to walk towards the target point.

As is shown by the dash-dotted line in Figure 5(a), the trajectory of the robot can be divided into 3 segments: SA, AB and BE.

During the segment of SA, there are no obstacles. The robot will generally head the direction of target point G. We found the actual direction of trajectory didn't coincide with that of target point but with small-amplitude oscillations. We considered it's caused by the oscillation of the control input, which, in turn, is resulted by GPS input drift and the large inertia of the UGV.

The robot found the close-range obstacle and decided to turn right sharply to avoid the obstacle during the segment of AB. The obstacle was almost 5 meters before the robot when the robot began to turn right. Therefore, we consider the obstacle response distance of system C in this run to be 5 meters. It's obvious that the total length of SA and AB is larger than that of other two trajectories. And the polyline walking of SA and AB can be thought as the first reason for the inefficiency of system C.

The controlling mode of system C in segment BE is similar to that of SA. The main difference of BE to other trajectories in Figure 5(a) lies in that when at the point B, the robot didn't head the target pointing G and the system C cost additional 30 control cycles to adjust the heading. So the extra time for heading adjustment can be thought as the second reason for the inefficiency of system C.

The third reason for the inefficiency of system C lies in the uncertainty of the turning direction decision when the robot confronted close-range obstacles. If the robot first recognize the banner as a left-anterior obstacle and it will turn right; and if the robot first recognize the banner as a right-anterior obstacle and it will turn left. Figure 5(d) shows the two trajectories of system C under the same experimental condition. The solid and the dash-dotted lines indicated the robot chose different turning directions in two rounds of running. The corresponding data of MIX-C-SUNNY-1 and MIX-C-SUNNY-2 in **Error! Reference source not found.** show that the decision of turning left made the robot cost another 62 seconds to reach the same target point (304 seconds and 366 seconds respectively). An extreme example is that if one end of the banner (point X) extends to the point Y, the robot at the point of M may still choose to turn left with a large probability, which will make the robot take more time to reach the same target point.

In summary, the polyline walking, heading adjustment and the uncertainty of the turning direction decision are the main reasons for the inefficiency of system C. And it's obvious that all these reasons, in turn, are caused by the incompetency of system C to incorporate the global obstacle distribution while making decisions.

2) Second, we invastigate the navigation behavior of System B:

Different from system C, system B makes decision based on the distribution of both close-range and distant-range obstacles, which can help the robot to avoid the obstacle ahead of time. The solid trajectory in Figure 5(c) shows that the obstacle response distance of system B in this round is 48 meters, which is much larger than that of system C.

However, we also found that, the dash-dotted trajectory (YELLOW-B-SUNNY-2 in TABLE 1) in in Figure 5(c) indicated that it's 23 meters longer than the other trajectory (YELLOW-B-SUNNY-1 in TABLE 1). To find the real reason for this inefficiency, we used the data accumulated during the experiment to offline simulate the whole run. The simulation showed that the uncertain recognition of distant-range obstacles during the segment BC caused the oscillation of the turning decision of system B. As a result, the robot randomly turns left or right during the segment BC. And on the other side, the solid trajectory turns right continuously to avoid the longest banner ahead of time, which made the whole trajectory smoother and shorter.

During the segment CD of the dash-dotted trajectory, the far-field scene inference and subsequent far-field path planning cannot provide meaningful guidance for the robot. The reason is that when the obstacle is too close to the robot, the system can't see the whole traversable region in the image plane because of the limitation of field-of-view of the stereo vision. Therefore, the long-range inference is not necessarily suitable for all kinds of scenes. And we plan to leave this problem to further research.

To summarize, the uncertain recognition of distant-range obstacles is the main reason for the inefficiency of system B. And the intrinsic reason for this is that the Bayes-classifier-based far-field scene inference used by system B only incorporates the appearance information while recognizing the scene, which is easily affected by the changing illumination of outdoor environments.

3) Third, we invastigate the navigation behavior of System A:

System A performed better than system B and C under same conditions in terms of DWR, TFJ and ORD as shown in TABLE 1.

The ORD of system C is 4 or 5 meters while that of system A is at least 39 meters, which shows that the CRFNFP-based navigating system does enhance the long-range perception for mobile robots.

The average of DWR and TFJ of system A are about 86.8 meters and 280 seconds respectively, while the average of DWR and TFJ of system C are 103.8 meters and 335 seconds, which indicates that the CRFNFP model does help the robot to plan more efficient paths for the navigation.

V. CONCLUSIONS

We designed comparison experiments to further validate the CRFNFP algorithm, which is proposed by us in 2010, in real challenging scenes. The comparative experiments indicate that, the CRFNFP-based navigating system outperforms traditional local-map-based navigating systems in terms of all criterion defined by us in this paper.

And the results also show that the CRFNFP algorithm does enhance the long-range perception for mobile robots and helps planning more efficient paths for the navigation.

VI. Future Research

Our further research may refer to the incorporation of other kinds of contexts, such as semantic contexts and temporal contexts. And we try to simulate the attention mehcanism of human beings into perception algorithms.

VII. Acknowledgment

This work is supported by the national natural science foundation (Q12F030004) and Ningbo natural science foundation (2011A610198, 2011A610155).

References

[1] Wang, M.; Zhou, J.; Tu, J., et al. Learning Long-range Terrain Perception for Autonomous Mobile Robots[J]. International Journal of Advanced Robotic Systems, 2010, 7 (1), 55-66.

[2] Matthies, L. Stereo vision for planetary rovers: Stochastic modeling to near real-time implementation[J]. International Journal of Computer Vision, 1992, 8 (1), 71-91.

[3] Pagnot, R.; Grandjean, P. Fast cross-country navigation on fair terrains[J]. Proceedings - IEEE International Conference on Robotics and Automation, 1995, 3, 2593-2598.

[4] Singh, S.; Simmons, R.; Smith, T., et al. Recent progress in local and global traversability for planetary rovers[J]. Proceedings - IEEE International Conference on Robotics and Automation, 2000, 2, 1194-1200.

[5] Rieder, A.; Southall, B. Stereo perception on an off-road vehicle[J]. Proc. IEEE Intelligent Vehicle Symposium, 2002, 1, 221-226.

[6] Ollis, M.; Huang, W. H.; Happold, M., et al. Image-based path planning for outdoor mobile robots[A]. IEEE International Conference on Robotics and Automation[C]. Pasadena, CA, 2008. pp 2723-2728.

[7] Jackel, L. D.; Krotkov, E.; Perschbacher, M., et al. The DARPA LAGR program: Goals, challenges, methodology, and phase I results[J]. Journal of Field Robotics, 2006, 23 (11-12), 945-973.

[8] Michael J. Procopio, J. M. G. G. Learning terrain segmentation with classifier ensembles for autonomous robot navigation in unstructured environments[J]. Journal of Field Robotics, 2009, 26 (2), 145-175.

[9] Happold, M.; Ollis, M.; Johnson, N. Enhancing supervised terrain classification with predictive unsupervised learning[A]. Robotics: Science and Systems[C]. Philadelphia, PA. Cambridge, 2006.

[10] Dahlkamp, H.; Kaehler, A.; Stavens, D., et al. Self-supervised monocular road detection in desert terrain[A]. Robotics: Science & Systems[C]. 2006.

[11] Max Bajracharya, A. H. L. H. M. B. T. M. T. Autonomous off-road navigation with end-to-end learning for the LAGR program[J]. Journal of Field Robotics, 2009, 26 (1), 3-25.

[12] Lookingbill, A.; Lieb, D.; Thrun, S. Optical Flow Approaches for Self-supervised Learning in Autonomous Mobile Robot Navigation (Autonomous Navigation in Dynamic Environments)[Z]. Springer Berlin/ Heidelberg: 2007; pp 29-44.

[13] Hadsell, R.; Sermanet, P.; Ben, J., et al. Learning long-range vision for autonomous off-road driving[J]. Journal of Field Robotics, 2009, 26 (2), 120-144.

[14] Lafferty, J. D.; McCallum, A.; Pereira, F. C. N. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data[A]. The Eighteenth International Conference on Machine Learning[C]. 2001.

Authors Profile

Wang Mingjun, born in 1981, majored in machine vision and received phD in 2010. His main research now focus on the machine learning and artificial intelligence.

A Framework For Automatic Generation Of Answers To Conceptual Questions In Frequently Asked Question (FAQ) Based Question Answering System

Saurabh Pal
Dept of CSE and IT,
Bengal Institute Of Technology,

Kolkata, India. Sudipta Bhattacharya
Dept of CSE and IT,
Bengal Institute Of Technology,
Kolkata, India.

Indrani Datta
Kolkata, India

Arindam Chakravorty
Dept of IT,
St Thomas College of Engineering & Technology,
Kolkata, India.

***Abstract*- Question Answering System [QAS] generates answer to various questions imposed by users. The QAS uses documents or knowledge base for extracting the answers to factoid questions and conceptual questions. Use of Frequently Asked Question (FAQ) base gives a satisfying results to QAS, but the limitation with FAQ base system is in the preparation of Question and Answer set as most of the questions are not predetermined.QAS using FAQ base fails if no semantically related questions are found in base, corresponding to the input question. This demands an automatic answering system as backup, which generates answer especially to conceptual questions.The work presented here gives a framework for automatic generation of answers to conceptual question especially "what", "why" and "how" type for frequently asked based, Question Answering System and in terms gradually builds up the FAQ base.**

Keywords- Question Answering System; Frequently Asked Question; Frequently Asked Question base; knowledge Base; semantic net; semantic frame; tagging.

I. INTRODUCTION

In today's society with increasing information requirements the web based Question Answering System [QAS] is needed. The question asked by user is either a factual question or a conceptual question. The factual question demands for name of entity, events that had happened, date and other facts whereas the conceptual question asks for definition of terms, ideas, views, classification, explanation, principles, relationships, reasoning – which explains how and why things. This demand for an automatic QAS which will generate answer to user's input question. The QAS takes the question in natural language as input and by using a question - answer set or document set or knowledge base generates the answer.

The QAS takes the user's question as input in natural language and identifies keywords, Named Entity[2] present in the question by using techniques available in Natural Language Processing like word segmentation, stemming, Part Of Speech tagging, Named Entity Recognition etc. The identified Keywords and Named Entity are then used for searching and selecting relevant documents, passages from the document set. Based on the question type and pattern the expected answer patterns (templates) [6][2] are generated which are subsequently used for searching a matching pattern in the extracted passage or in the relevant document. As generated answer patterns, often may not satisfy the intension or semantic of question, understanding the semantic of question is required. The input question is classified or categorized into what it's asking for – name, place, object, date, event etc called Expected Answer Type [2], and then by using Named Entity Recognition [NER] technique the expected Named Entity is extracted from the text (document). Extracting answer from documents using keyword search, answer pattern matching and named entity recognition founds effective for factoid / factual questions then in comparison to conceptual questions. The difficulty in extracting answer to conceptual question from document is in figuring out, where in the passage the answer is present and up to how much the answer would be. For answering the reasoning or conceptual questions other approaches are followed, such as use of Frequently Asked Question [FAQ] Base [7], Knowledge Base [1]. FAQ Base is a set of predefined question and answer pair made prior to use. Use of FAQ base involves understanding the semantic of input question and matching it against the available questions in the question and answer base. The semantically related matching question's answers are retrieved and ranked. This process is effective than searching answer patterns for conceptual questions [3]. Presently more emphasis is given on the use of Knowledge Base, which includes Ontology – a conceptual map used as a tool for understanding the intension/semantics of input question by knowing the meaning and relation of various concepts or entities present in question as well as in the relevant text document for extracting the answer. To more recently various combinations of FAQ base, document set and knowledge base is used for extracting the answer. If the answer is not found in the FAQ base it is searched in the document set or generated automatically from the knowledge base. These techniques are not found fully satisfying to answer the reasoning (why and how) type questions, as answering these types of question needs to know the intension of question and synthesis of the answer using various concepts.

Full extension use of NLP is often not satisfying to answer the conceptual type questions especially those how and why types. Instead using FAQ base gives a satisfying result, as answers to various forms of questions are available in the FAQ base, including how, why, conditional etc. The problem associated with FAQ base System is in the preparation of Question and answer set, as it not known from previous what type and forms of question will be asked by users. This demands obviously an automatic answering system as backup, which will answer to conceptual question – not found in FAQ base.

A framework for automatic QAS based on tagged document and knowledge base incorporating the semantic net and semantic frame is provided here for answering conceptual [especially reasoning ones] question's answer –as a possible solution for making up a tool which will help in understanding the question's intension and answering it and thus in a process gradually helping in enriching the FAQ base.

The proposed framework is provided with a mechanism that makes clear to what it had understood and what not and thus where it fails. This in terms helps in enriching the knowledge base by assisting the expert, in filling the gaps within the knowledge base so that future queries may be answered properly.

II. Problem Defination

Use of FAQ base in QAS gives satisfying results, as answers corresponding to various forms of factual and conceptual questions are available. The difficulty with FAQ base is in building question and answer set in prior, as it is not known from previously what questions would be asked by users. The FAQ base starts with some standard questions and then eventually grows as users make new queries.

The input question is matched with FAQ base question set and semantically related question's answers are extracted. The system fails if no semantically related questions corresponding to the input question are found in the FAQ base. The unanswered question is then sent back to be answered manually [1] later by an expert. The new question and its answer are then updated in FAQ base [7]. As the unanswered questions could be many, the process will bring burden over the domain expert. To overcome this QAS using FAQ base are often combined with auto answer generation system which generates answer to unanswered question from document set or from knowledge base [5].

With various NLP complexities, it is possible that the auto answer generation system may generate wrong answers or have redundant information in it, or the answer generated may not satisfy the semantics of the question properly. The wrong or improper answer may often lead to confusion in user. This requires for a mechanism where the answers are authenticated for correctness before it can be sent back to the user and saved in the FAQ base.

Generating answer automatically requires extraction of information from document / knowledge base. Different techniques are used in, extracting answer from document set or knowledge base. While searching in a document set, the relevant documents are extracted by using keywords, obtained from the input question. The keyword search particularly identifies and locates the passages and sentences in the document, but it fails to identify which word/phrase or how much text/sentence will make the answer. This generally requires knowing what Named Entity the question is asking for such as name, object, date, time, measurement, event, location etc. The expected Named Entity is searched by NER in the extracted passage and filtering out relevant String and discarding irrelevant data. The Named Entity by this also determines the answer type and probable answer patterns. The generated answer pattern / template [4] are searched in the extracted sentence, passage or paragraph obtained from keyword search to generate the answer. These techniques are very much effective in answering factoid/factual question's answer but fail to answer the conceptual question (definition, reasoning, explanation, difference etc). The answer to conceptual or complex question can be single passage or multi sequenced passages. Due to the natural language complexities it is not identifiable which sentences, passages and how much text contains the information and is thus difficult to answer the conceptual question from document set. The other problem with reasoning based questions with multiple concepts[8] or entities present in it, is in understanding the relation between the concepts, and there after artificially synthesis of answer with different concepts.

The above discussion can be summed up for the following requirements

1) A question's answering mechanism for generating answer to various conceptual questions raised in FAQ based QAS and hence in turn helps in building FAQ base. This in turn demands for :

a) A frame work for extracting conceptual information from document set like definition, explanation, short note, contrast, feature, process, reasoning etc.

b) A frame work for extracting conceptual information from knowledge base such as semantic net, semantic frame and synthesis of the proper answer.

III. Design And Soluition

The present Question Answering System is designed for answering scientific and technical domain questions starting with words "what", "why" and "how". For answering to the question by this system, it is necessary that the question sentence should contain relevant keywords within it.

Technical domain questions generally contain concepts and attributes. The semantics of the question can be well understood by the presence of concepts and their attributes in it. The concepts are the entities about which the question is all about and the attributes define what is needed to be asked/ known about the concepts. For example in the question "why semiconductor have valency 4", the concept "semiconductor"- defines that the question is about semiconductor and the attribute "valency" defines what is needed to be asked or get to be known about the concept "semiconductor". The concepts in the question are the entities represented as keywords, which are of course the Nouns. The attributes are noun like "resistor", "resistance" and verb like "resistivity". Thus noun, verb found in question is assumed to be an attribute.

A technical or scientific question may contain one or more concepts and attributes within it. The system produces answer with the assumption and philosophy that producing the relation among the required concepts and description of relation between attribute with concept, if no relation is found then mere description of the concept will obviously covers lot of the answer part in it.

The relation among the concepts in a domain can be defined by a conceptual map. The concept map can be made by semantic net. The semantic net defines how one concept relates with another concept. It is a directed graph consisting of vertices/nodes representing the concepts and edges representing the relationship between the concepts. There is no standard set of relations in semantic networks, but the following relations are very common:

- INSTANCE: X is an INSTANCE of Y, when X is a specific example of the general concept Y. Example: Johan is an INSTANCE of Human.
- ISA: X ISA Y, when X is a specialization of the more general concept Y. Example: sparrow ISA bird.
- HAS - PART: X HAS - PART Y, when the concept Y is a part of the concept X. (Or this can be any other property) Example: car HAS seat.

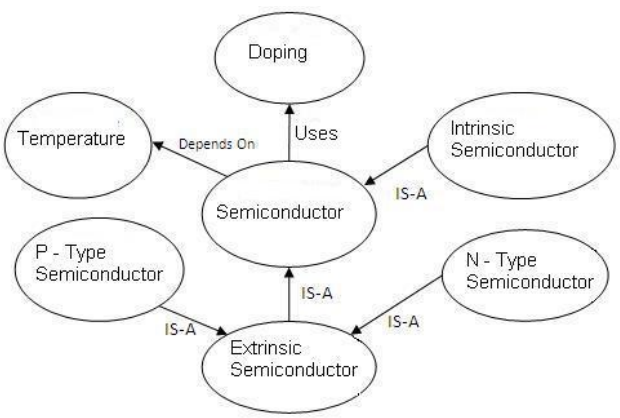


Figure 1. Figure shows a section of semantic net used to represent part of the domain "semiconductor physics".

Every concept in the semantic net has its own set of attributes, which are described by semantic frame. The semantic frame consists of slots, each slot contains fields called facets which specify attributes, attribute's value and description or comment. The description is a brief note/explanation on why the attribute has that value.

To answer a general conceptual question, explanation of its concept/s is necessary. A conceptual question ("why" & "how" type) is of the form – "Why / How ….P……Q ?", where P and Q are concepts, the question can have only one or more concepts. It is found that one concept has an action on other concept or one concept is the cause and the other concept is its result. To explain this kind of conceptual question it is very important to get the relations between the concepts, like P – relation1 – X – relation2 – Y – relation3 – Q, where X, Y are the other concepts. The relationship states that to explain concepts in the question the other interdependent concepts need to be explained. The relationship among the various concepts in a domain is obtained by a conceptual map. For the question having attribute/s along with the concepts like "why / how….P….a…Q ?", the attribute 'a' defines what need to be asked or get to be known about the concept/s. Here the attribute 'a' may be of concept 'P', 'Q' or of any concept present in between the relationship of 'P' and 'Q'. As it is not known, attribute 'a' belongs to which concept, searching the attribute for a match in the attribute set described in semantic frame for each concept present in the relationship of 'P' and 'Q' for a description regarding the attribute and its value can give the answer to the question. Thus for every concept found in question, relation among the concepts are derived, this would explain in what way the concepts are related and thus interdependency is found. The attributes found in the question are searched against the semantic frame of each concept in the relationship, to know about what is to be known about the concepts.

A Technical Questions may consist of:

- One concept with no attribute
- One concept with one or many attribute.
- More than one concept with no attribute.
- More than one concept with one or many attribute.

For questions having one concept, the concept is searched in semantic net and the description of the concept is obtained from the semantic frame, which provides the answer to the question.

For the questions with having more than one concepts and no attribute, the concepts are located in the semantic net and the relationship between them is obtained. Description of each concept in the relationship is obtained from the corresponding semantic frame and are combined together to form the answer.

For the questions having one concept and one attribute, the concept is searched in the semantic net. The corresponding semantic frame is searched and in its each slot, both the attribute and attribute's value fields are matched against the given attribute. The matching slots description is extracted as the answer to the question.

In questions with having more than one concepts and attributes, the concepts are located in the semantic net and the relation between them is obtained. Each of the question's attributes is searched for a match in the semantic frame of each concept present in the relationship. For match not found in the semantic frame, the description of the semantic frame is obtained and presented as information. The information thus obtained is combined together to make an answer to the question.

In another approach to answer conceptual question of "what" type, the information is extracted from tagged text document. Technical subject/domain document content can be categorized into different sections and parts as introduction, definition, brief, explanation, types, property, features, behavior, function, procedure, components, parts, structure, necessity, application, theorem, principle, formula, example, figure, conclusion etc. Based on these respective tags are made to represent the section like:

1. <introduction>…………</introduction>
2. <def>……………….…</def>
3. <brief>……………….....</brief>
4. <explanation>…………</explanation>
5. <types>……………….....</types>
6. <property>…………..…</property>
7. <feature>………………</feature>
8. <behavior>…………...</behavior>
9. <function>……………</function>
10. <procedure>…………</procedure>
11. <components>………..</components>
12. <parts>……………….</parts>
13. <structure>…………...</structure>
14. <necessity>…………..</necessity>
15. <application>………..</application>
16. <purpose> …………..</purpose>
17. <theorem>…………...</theorem>
18. <principle>………… </principle>
19. <formula>………..….</formula>
20. <example>………..…</example>
21. <figure>…………..…</figure>
22. <conclusion>……..…</conclusion>

Each tag has attribute – "name", "of" with value set as the keyword name (Concept name). The tag should have at least one attribute in it. The attribute "name" defines what is being tagged and "of" define of whom. The tags can be nested. e.g., <def of: semiconductor >A semiconductor is a substance ……………. e.g., <example of: semiconductor>: germanium, silicon, carbon</example></def>

While processing a "what" question, the question is identified for what it's asking and the corresponding tag is recognized. The tagged documents are searched with the identified tag name and tag's attribute value set as the concept's name. The matching tagged information is extracted as answer.

The whole System is developed into three parts – i) Question processing, ii) Answer generation, iii) Answer authentication.

The Question processing process takes the question sentence in natural language text from user, assuming the same or semantically related question is not found in the FAQ base – the question text is processed. The text processing includes spelling check, word trimming, base word conversion, words stemming, and removal of words with least significance, extraction of concepts and attributes.

The Answer generation process takes the concepts and attribute as input from the Question processing process. The process obtains the answer either from the document base or from the knowledge base – consisting of semantic net and semantic frame. Based on the first word of the question sentence its type is determined as to whether "what", "why" or "how" type. The "why" and "how" types of questions are considered as reasoning based question, the information for which are obtained from the knowledge base and are combined together to form the answer. For question of "what" type, the system identifies the tag and the tagged information is extracted from the document set else if no tag is identified the answer is obtained from knowledge base.

The Answer authentication process obtains the unauthenticated answer from Answer generation process and sends it to the expert for authentication. The expert authenticates or corrects the answer with minor alterations. The process saves the question and its authenticated answer into the FAQ base and also sent the answer to the corresponding user, who has raised the question.

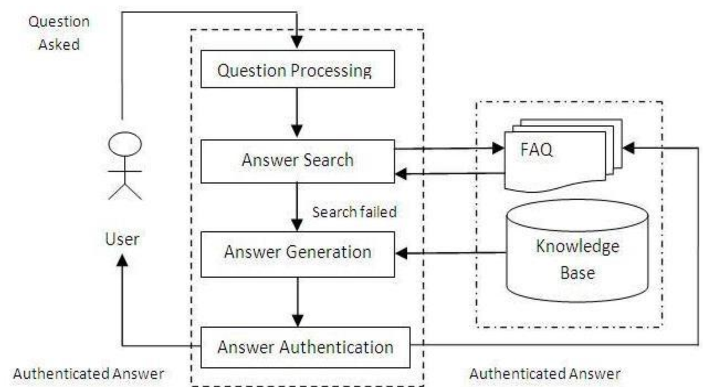


Figure 2. Framework for automatic answer generation system for FAQ based QAS.

IV. ALOGORITHM

Step1 The question is taken as input in the form of string.

Step2 The question is segmented into individual words using white space as delimiter.

Step3 Each segmented word is matched in the domain vocabulary. If the word does not finds a match in the lexicon, its set as wrong spelling.

Step4 If each word spelling is found correct then go to step 8.

Step5 For each wrong spelling founds a corresponding matching spelling with same first character and having 60% same character count is obtained from the domain lexicon.

Step6 words with its wrong spelling and the correct spellings are displayed as possible solution of the word to the user for input.

Step7 Correct spelling words are taken as input from the user as a replacement for the wrong spelling words

Step8 Each segmented word is matched in the domain vocabulary, for a match found it is converted to its base word/root word.

Step9 The words that are found in the domain vocabulary are identified and tagged as "KEYWORD".

Step10 Using a lexicon, part of speech for other words beside keywords in question are identified and tagged [e.g., noun pronoun, verb, adverb, adjective, modal verb etc]

Step11 A normalized question is obtained by eliminating pronoun, auxiliary modal verb, adverb, adjective, article from the question and selecting only keywords, noun and verb.

Step12 If the first word is “what”. Set the question type to “what_type” else go to step 17.

Step13 The normalized question is searched for tag words.

Step14 If the tag word is not present, go to step 18.

Step15 If tag word is present, the tag word is searched in the tagged text file/document against the keywords present in the normalized question.

Step16 If the tag word and the keyword pair is found, the text between the starting and ending tag is displayed,else display “ The tagged information corresponding to the keyword not found ”.

Step17 If the first word is “why”, “how”, set the question type to “why_type”.

Step18 Each word of the normalized question is matched with every nodes of the semantic net.

Step19 If found equal, the name is added to the nodes array else the name is added to the property array.

Step20 Search is performed on the semantic net and semantic frames using the elements of nodes array and property array. Call search (nodes array, property array).

Step21 The duplicate information in the information array is removed.

Step22 The information stored in the information array is displayed.

a) Algorithm for search (nodes array, property array)

Step1 for node array length = 1 and property array length = 0, call search1(nodes array[0]).

Step2 for node array length =1 and property array length >0, for each element in the property array flag = call search2 (nodes array[0] , property array[i]).if flag = false ,

a. Display "Since property array found no match in the properties of nodes array, the description of will probably give the result”

Step3 call search1(nodes array[0]).

Step4 for node array length >1 and property array length >= 0 Call search3(nodes array, property array).

Step5 For node array length = 0 , display “The keywords / words not found in the knowledge base”.

b) Algorithm for search1 (element): searching single node in the semantic net

Step1 The element is searched against every node in the semantic net.

Step6 If the element searched is found, the description of the node from the corresponding semantic frame is retrieved and stored in information array else set flag to false.

c) Algorithm for search2 (element, property): searching single concept with single property in the semantic net

Step1 The element is searched against every node in the semantic net.

Step2 If the element searched is found, the property is matched for each attribute in the corresponding semantic frame.

Step3 If the property gets matched with attribute or its value in the frame, the description of the attribute is retrieved and stored in an information array else the property is searched in base nodes or in the nodes which are one step up in hierarchy to the current node

Step4 If base node found, go to step 2, else set flag to false.

d) Algorithm for search3 (nodes array, property array) : searching multiple nodes with multiple property in the semantic net

Step1 The elements are searched against every node in the semantic net.

Step2 The routes between the nodes are mapped.

Step3 If the route do not exists, set flag to false.

Step4 if flag = true, Relation between the adjacent nodes in the route is stored in an information array.

Step5 For each node in the route/path the corresponding semantic frame is searched for each element of the property array.For each node in the route

{For each element in the property arraycall search2 (node array[i], property array[i])}If search2 (node array[i], property array[i]) did not found a match call search1 (node array[i])}

Step6 If flag = false, display “Route between the nodes doesn’t exist”

V. EVALUATION AND CONCLUSION

An initial evaluation is performed on a technical domain with a set of questions. The question are made strictly using “what”, “why” and “how” as the initial word of the sentence.

After analyzing the answer generated it is found that the answer contains extra information in it. Based on the relevance of information contained in the answer, the answers are categorized into 2 categories as below:

Category “A”: The answer contains exact pieces of information,

Category “B”: The answer contains exact piece of information with addition to that one or more pieces of extra information.

The system is evaluated with 50 questions each on “what”, “why” and “how” types, the following results are obtained.

TABLE I. NNUMBER OF QUESTIONS OF WHAT, WHY AND HOW TYPE IN CATEGORY A AND CATEGORY B

	A	**B**
What	41	9
Why	35	15
How	30	20

The initial evaluation result shows the feasibility of building QAS with this framework, which can automatically produce answers to conceptual questions in a science and technical domain. Thus can be used as tool to answer the Frequently Asked Question and helps in building the FAQ base.

Use of knowledge base (semantic net and semantic frame) and tagged document set is found very helpful. The use of tags has solved the problem of extracting answers from text to an extent as it annotates the text, sentences or passage with meaningful information and thus helps in extracting the required piece of information correctly.

For understanding the meaning of question use of concept present in question is quite effective as it requires much less Natural Language Processing tasks. The use of knowledge/conceptual map is definite requirement as this technique is the very effective in understanding the various concepts present in the question and their relations. Construction of proper conceptual map is very critical and requires domain expertise. Improper or incorrect concept may lead to total failure of answering system.

The program generates message for every instance when

a) The system does not found the required tag for a concept in the document.

b) Searched concepts are not found in the concept map [semantic net].

c) Searched concepts are not related in the concept map [semantic net]

d) Attribute in question don't match against the attributes or its value in the semantic frame of the required concept.

The messages produced by system helps the expert in changing and upgrading the knowledge base [semantic net and semantic frame] and adding tags and other required information to the documents so that the system can suitable produce the right full answer for the future queries.

The tags used are on a single technical subject and needs to be tested more on other domain, so that tag set can be made complete and a standard on tag can be obtained.

REFERENCES

[1] Chun-Chia, W., Jascon, C. H., Che-Yu, Y., Husan – Pu Chang, "A Question Answering System Approach for Collaborative Learning", 10th International Conference on Computer Supported Cooperative Work in Design, IEEE Conferences, May, 2006, pp. 1 – 5.

[2] Boldrini, E. ,Ferrández, S., Izquierdo, R., Tomas, D., Ferrandez O., and Vicedo J.L, "A proposal of Expected Answer Type and Named Entity annotation in a Question Answering context", 2nd Conference on Human System Interaction, IEEE Conferences, May, 2009, pp. 318-322.

[3] Keliang, J., Xiuling, P., Zhinuo, L.,"Question Answering System in Network Education based on FAQ", The 9th International Conference for Young Computer Scientists, IEEE Conferences, Nov, 2008, pp. 2577- 2581.

[4] Parik, J., Narashima Murty, M., "Adapting Question Answering Techniques to the Web", Proceeding of the Language Engineering Conference, IEEE Conferences, Dec, 2002, pp. 163-171.

[5] Qinglin G., "Question Answering System Based on Ontology" , 7th World Congress on Intelligent Control and Automation, IEEE Conferences, June,2008, pp. 3347-3352.

[6] Yongping, D., Ming, H., "Pattern Optimization and the Application in Question Answering", Eighth International Conference on Intelligent Systems Design and Applications, IEEE Conferences, 2008, pp. 35 – 39.

[7] Zheng-Tao, Y., Yan-Xia, Q., Jin-Hui, D., Lu-Han, Cun-Li, M., Xiang-Yan, M., "Research on Chinese FAQ Question Answering System in restricted domain", International Conference on Machine Learning and Cybernetics, IEEE Conferences, Volume 7, Aug, 2007, pp. 3927 – 3932.

[8] Xiaobo, W., Wei,C., Weicun, Z., "Design and Realization of Intelligent Question Answering System Based on Ontology", International Conference on Control, Automation and Systems, IEEE Conference, Oct, 2008, pp. 1313 – 1316.

AUTHORS PROFILE

Saurabh Pal is working as faculty in the Dept of Computer Science and Information Technology at Bengal Institute of Technology, Kolkata, India. He received his B. Sc. from University of Calcutta in 1998, M.Sc. IT from Allahabad Agricultural Institute – Deemed University in 2006, and M.Tech – IT [Courseware Engg] from Jadavpur University in 2010.

Sudipta Bhattacharya is working as Asst Prof in the Dept of Computer Science and Information Technology at Bengal Institute of Technology, Kolkata, India. He received his B.Tech from West Bengal University of Technology in 2006, M.Tech – IT from Bengal Engineering and Science University in 2009.

Indrani Datta received her B.Sc Software System from North Bengal University in 2004, Master of Computer Application from West Bengal University of Technology in 2008. Her primary area of interest is Natural Language Processing, Information Extraction.

Arindam Chakravorty is presently working as Asst Prof in the Dept of Information Technology at St Thomas' College of Engineering and Technology , Kolkata, India. He received his B.Tech(Electrical), M.Tech IT(Courseware Engg) from Jadavpur University.

Sensor Location Problems As Test Problems Of Nonsmooth Optimization And Test Results Of A Few Nonsmooth Optimization Solvers

Fuchun Huang
School of Engineering and Science, Victoria University
Melbourne, Australia

***Abstract*—In this paper we address and advocate the sensor location problems and advocate them as test problems of nonsmooth optimization. These problems have easy-to-understand practical meaning and importance, easy to be even randomly generated, and the solutions can be displayed visually on a 2-dimensional plane. For testing some nonsmooth optimization solvers, we present a very simple sensor location problem of two sensors for four objects with the optimal solutions known by theoretical analysis. We tested several immediately ready-to-use optimization solvers on this problem and found that optimization solvers MATLAB's ga() and VicSolver's UNsolver can solve the problem, while some other optimization solvers like Excel solver, Dr Frank Vanden Berghen's CONDOR, R's optim(), and MATLAB's fminunc() cannot solve the problem.**

Keywords-sensor location problems; mathematical programming; nonsmooth optimization solver; test problems.

I. Introduction

Nonsmooth optimization is an important research field of optimization and has wide applications in real life. Although there are many test problems of nonsmooth optimization [1], some of them are too academic and lack practical backgrounds and importance, while some others are not so flexible in generating random problems of various sizes for testing purposes. Hence it is still good to have more test problems, in particular if the problems have easily understandable practical meanings and importance, and more preferably visual displays. In this paper we address sensor location problems and advocate them as a new group of test problems of nonsmooth optimization solvers. The problems are generally nonsmooth and difficult to solve. We present test results of a simple sensor location problem solved by some nonsmooth optimization solvers, which are: Excel solver developed by FrontLine Solvers [2], CONDOR developed by Frank Vanden Berghen [3], R's optim() function [4], MATLAB's fminunc() function and some other solvers [5], and VicSolver's UNsolver [6]. All these solvers have directly ready-to-use (that is, no need to compile or link by using a compiler) evaluation versions available to anyone, hence the test results reported in this paper can be repeated by anybody.

This paper is organized as following. Section II addresses the sensor location problem from different practical backgrounds, section III explains the abovementioned ready-to-use solvers and their test results and section IV summarizes the main points and results of the paper and points out some future work.

A. Sensor location problems

We use Figure 1 to help us illustrate the sensor location problems. Suppose we want to use three sensors to sense n objects in an area. The locations of the n objects are known, as shown in Figure 1. We want to determine the "best" locations of the three sensors. There could be different criteria for determining the "best" locations. One of them is to minimize the largest squared distance from an object to the nearest sensor, that is,

1) *Sensor Location Problem:* For sensing n objects at locations: (ox(i),oy(i)): i=1,2,…,n, find locations of s sensors: (sx(j),sy(j)): j=1,2,…,s, such that the largest squared distance from an object to the nearest sensor

max(min((ox(i)-sx(j))2+(oy(i)-sy(j))2,j=1,2,…s),i=1,2,…,n)

is minimized.

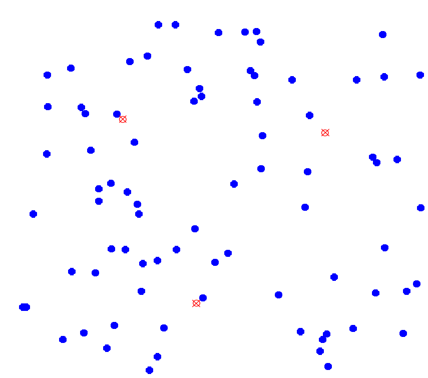


Figure 1. Locations of n objects and 3 sensors.

The same type of location problems could come from different real life backgrounds. The following are just two versions of them among many others, which might be good for teaching purposes.

2) Well Location Problem: For serving n houses at locations: (ox(i),oy(i)): i=1,2,…,n, find locations of s wells:

3) (sx(j),sy(j)): j=1,2,…,s, such that the largest squared distance from a house to the nearest well max(min((ox(i)-sx(j))2+(oy(i)-sy(j))2,j=1,2,…s),i=1,2,…,n) is minimized.

4) *Light Location Problem:* For lighting n target locations: (ox(i),oy(i)): i=1,2,…,n, find locations of s lights: (sx(j),sy(j)): j=1,2,…,s, such that the largest squared distance from a target location to the nearest light max(min((ox(i)-sx(j))2+(oy(i)-sy(j))2,j=1,2,…s),i=1,2,…,n) is minimized.

The sensor location problem may have constraints. For example, we may want to determine the best locations of three sensors on two roads only, as shown below in Figure 2.

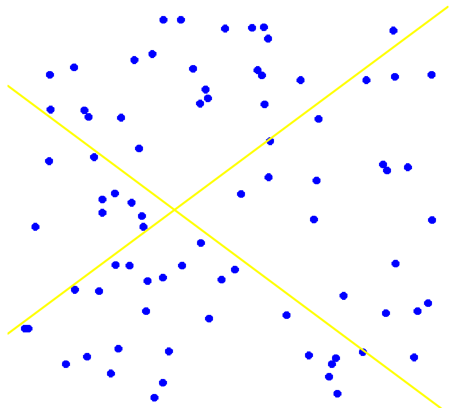


Figure 2. n objects and two roads.

So, in general, the sensor location problem is stated as in the following:

Sensor Location Problem: For sensing n objects at locations: O(i): i=1,2,…,n, find locations of s sensors: S(j): j=1,2,…,s, such that the largest squared distance from an object to the nearest sensor

max(min(distance(O(i),S(j)),j=1,2,…s),i=1,2,…,n)

is minimized.

The largest squared distance from an object to the nearest sensor, as a function of the locations of the sensors, is a continuous but nonsmooth function. The function of a very simple situation yields the nonsmooth surface as shown in Figure 3.

The sensor location problems are like the facility location problems explained in [7] hence in general very difficult to solve. As test problems of nonsmooth optimization, however, they have the merits that they have easily understandable practical backgrounds and importance, the object locations can be randomly generated and there can be a 2-dimensional visual display of the solutions. When the number of objects is small, the optimal solution can be obtained by examining different mappings of objects to different sensors. However, when the number of objects increases, the number of different mappings quickly becomes so huge that checking all different mappings becomes impossible. For example, if there are 100 objects and 5 sensors, then the number of mappings is 5100. Hence smarter algorithms of nonsmooth optimization are necessary for solving medium to large scale sensor location problems.

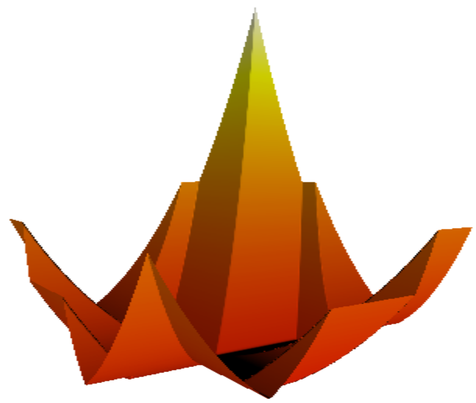


Figure 3. The 3-dimensional surface of a 2-dimensional function of a simple sensor location problem.

II. A SIMPLE SENSOR LOCATION PROBLEM AND TEST RESULTS OF SOME SOLVERS

For testing different solvers of nonsmooth optimization, we have this simple sensor location problem: the four objects are in blue at the corners of a square, that is, (0,0), (0,1), (1,1), (1,0), as shown in Figure 4 below. We want to determine the best locations of two sensors, and apparently we can see there are two optimal solutions: {(0, 0.5), (1, 0.5)} and {(0.5, 0), (0.5, 1)}, as shown in red in Figure 4 below.

Figure 4. One optimal solution of the two sensor location problem

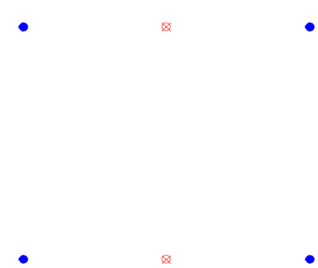


Figure 5. Another optimal solution of the two sensor location problem

There are many published computer programs for nonsmooth optimization, and an incomplete list can be found at http://napsu.karmitsa.fi/nsosoftware/. Here in this paper we focus on only some ready-to-use programs, not those programs in source codes or in a binary library which needs a compiler or linker to compile or link the program to a user's main program. The ready-to-use programs tested in this paper are: Excel solver developed by FrontLine Solvers [2], CONDOR developed by Dr Frank Vanden Berghen [3], R's optim() function [4], MATLAB's fminunc() function and some other solvers [5], and VicSolver's UNsolver developed by Dr Fuchun Huang [6]. In testing these solvers, we use the initial sensor locations {(0.5, 0.5), (0.5, 0.5)}, which is not a local minima as the objective function would decrease if one sensor moves to the left (or up) a little bit and the other moves to the right (or down) a little bit.

A. *Frontline Solvers*

Frontline's Excel solver has three methods or algorithms: GRG nonlinear for solving smooth nonlinear optimization problems; Simplex LP for solving linear problems; and Evolutionary for nonsmooth problems, as shown below in the solvers application interface wizard.

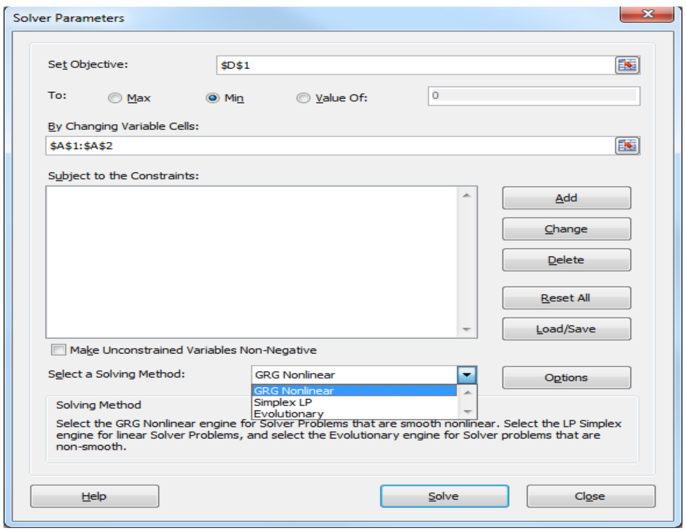


Figure 6. Excel solver's three methods.

For solving the sensor location problem stated at the beginning of this section, we put initial values (x1=0.5, y1=0.5, x2=0.5, y2=0.5) of the locations of the two sensors in cells B1:B4, and the largest squared distance from an object to the nearest sensor is computed by the following formula in cell E1:

E1=MAX(MIN(B1^2+B2^2, B3^2+B4^2), MIN((B1-1)^2+(B2-0)^2, (B3-1)^2+(B4-0)^2), MIN((B1-0)^2+(B2-1)^2, (B3-0)^2+(B4-1)^2), MIN((B1-1)^2+(B2-1)^2, (B3-1)^2+(B4-1)^2)), as shown in Figure 7.

When solve the problem by 'Evolutionary' method as shown in Figure 6, it ends up with the message that 'Solver cannot improve the current solution'.

When solve the problem by 'GRG nonlinear' method, it also ends up with the message that 'Solver cannot improve the current solution'.

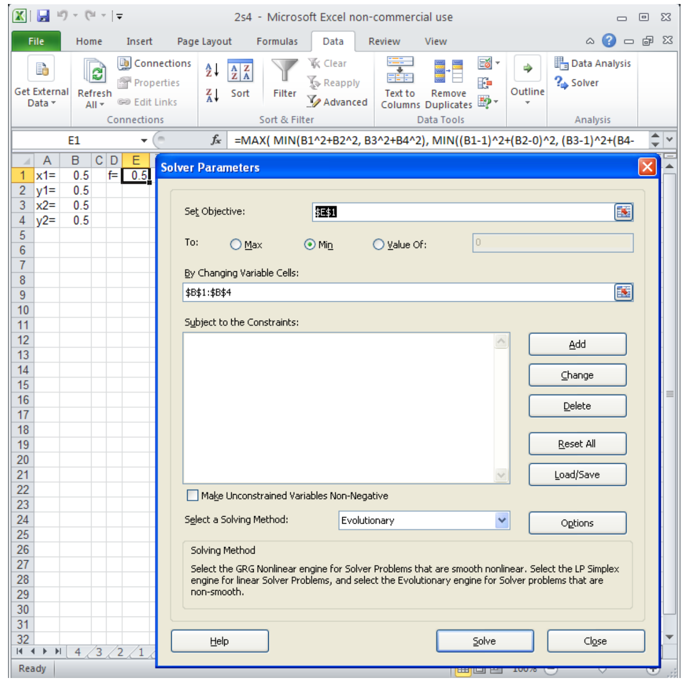


Figure 7. Excel Solver's Evolutionary method to solve the sensor location problem of two sensors.

B. *CONDOR's result*

CONDOR [3] is a constrained, non-linear, derivative-free parallel optimizer for continuous, high computing load, noisy objective functions developed by Dr Frank Vanden Berghen. CONDOR is also available via NEOS Server [8].

The following file is the AMPL [9] code for solving the sensor location problem at the beginning of the section:

```
var x{i in 1..4};

minimize f:
max(
min(x[1]^2+x[2]^2, x[3]^2+x[4]^2),
min((x[1]-1.0)^2+(x[2]-0.0)^2, (x[3]-1.0)^2+(x[4]-
0.0)^2),
min((x[1]-0.0)^2+(x[2]-1.0)^2, (x[3]-0.0)^2+(x[4]-
1.0)^2),
min((x[1]-1.0)^2+(x[2]-1.0)^2, (x[3]-1.0)^2+(x[4]-
1.0)^2)
);

let x[1] := 0.5;
let x[2] := 0.5;
let x[3] := 0.5;
let x[4] := 0.5;

display x;
display f;
```

When the file is submitted to NEOS server to be solved by CONDOR, the following 'optimal' solution is returned:

```
Best Value Objective=2.575139e-01 (nfe=324)
rho=1.000000e-04; fo=2.575139e-01; NF=325
rho=1.000000e-04; fo=2.575139e-01; NF=325
CONDOR 1.06 finished successfully.
325 ( 312) function evaluations
Final obj. funct. Value=0.25751389
_svar [*] :=
1  0.971055
2  0.50182
3  0.0866737
4  0.500002
```

We see the optimal solution and the minimum value

"Final obj. funct. Value=0.25751389"

are not so close to the truly optimal solution and minimum value 0.25.

C. R's optim() function

R [4] is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. The version of R we used is 2.14.1. R's optim() function has five methods for multi-dimensional optimization: "Nelder-Mead", "BFGS", "CG", "L-BFGS-B", "SANN".

The following are R codes of the sensor location problem with optimal value (from theoretical analysis but tested by the code):

```
> # number of sensors:
> ns=2
>
> sxy=rep(0,2*ns);
>
> fr <- function(sxy)
+ {
+ sx=sxy[1:ns];
+ sy=sxy[(ns+1):(ns+ns)];
+
+ dmax=0.0;
+ for(i in 1:no){
+     j=1;    dmin=(sx[j]-ox[i])^2+(sy[j]-oy[i])^2;
+ for(j in 2:ns){d=(sx[j]-ox[i])^2+(sy[j]-oy[i])^2; if(d<dmin)dmin=d;};
+ if(dmax<dmin)dmax=dmin;
+ }
+ return(dmax);
+ };
>
> # optimal
> fr(c(0,1,0.5,0.5))
[1] 0.25
```

The following are R codes of solving the problem by Nelder-Mead method:

```
>   sxy=c(0.5,0.5,0.5,0.5)

>   optres=optim(sxy, fr, NULL, method = "Nelder-Mead", control=list(maxit=999999));

control=list(maxit=999999));

> cat(optres$value, fill=T);

0.5

> cat(optres$par[1:(ns+ns)], fill=T);

0.5 0.5 0.5 0.5
```

We see the solver cannot improve the initial values. The other methods, "BFGS", "CG", "L-BFGS-B", "SANN" all have the same results, that is, none of them can improve the initial values.

D. MATLAB's fminunc() function and other solvers

MATLAB [5] is a numerical computing environment and fourth-generation programming language developed by MathWorks. The version of MATLAB we used is 7.11.1. The following shows the MATLAB m-file of the two-sensor four-object problem stated at the beginning of the section, and running results of the optimization function fminunc() with default option settings:

```
function f=s2o4(x)

f=0.0;

f=max(f,min((x(1)-0.0)^2+(x(2)-0.0)^2,(x(3)-0.0)^2+(x(4)-0.0)^2));

f=max(f,min((x(1)-1.0)^2+(x(2)-0.0)^2,(x(3)-1.0)^2+(x(4)-0.0)^2));

f=max(f,min((x(1)-0.0)^2+(x(2)-1.0)^2,(x(3)-0.0)^2+(x(4)-1.0)^2));

f=max(f,min((x(1)-1.0)^2+(x(2)-1.0)^2,(x(3)-1.0)^2+(x(4)-1.0)^2));

end

>> s2o4([0.5;0.5;0.5;0.5])

ans =

    0.5000

>>

  x = fminunc(@s2o4,[0.5;0.5;0.5;0.5])

Warning: Gradient must be provided for trust-region algorithm;

  using line-search algorithm instead.

> In fminunc at 347

Initial point is a local minimum.

Optimization completed because the size of the gradient at the initial point

is less than the default value of the function tolerance.

<stopping criteria details>

x =

    0.5000

    0.5000

    0.5000

    0.5000
```

We see the solver cannot improve the initial values, and wrongly claims the initial point is a local minimum. Other option settings of the solver yield the same results.

MATLAB has several other solvers and their results are presented and explained below and in Figure 8.

```
>> fminsearch(@s2o4,[0.5 0.5 0.5 0.5])
ans =
    0.5000    0.5000    0.5000    0.5000
Optimization running.
Objective function value: 0.5
Optimization terminated:
 the current x satisfies the termination criteria
using OPTIONS.TolX of 1.000000e-004

 and F(X) satisfies the convergence criteria using
OPTIONS.TolFun of 1.000000e-004

>> simulannealbnd(@s2o4,[0.5 0.5 0.5 0.5])

Optimization terminated: change in best function
value less than options.TolFun.
ans =
    0.5000    0.5000    0.5000    0.5000
>>
```

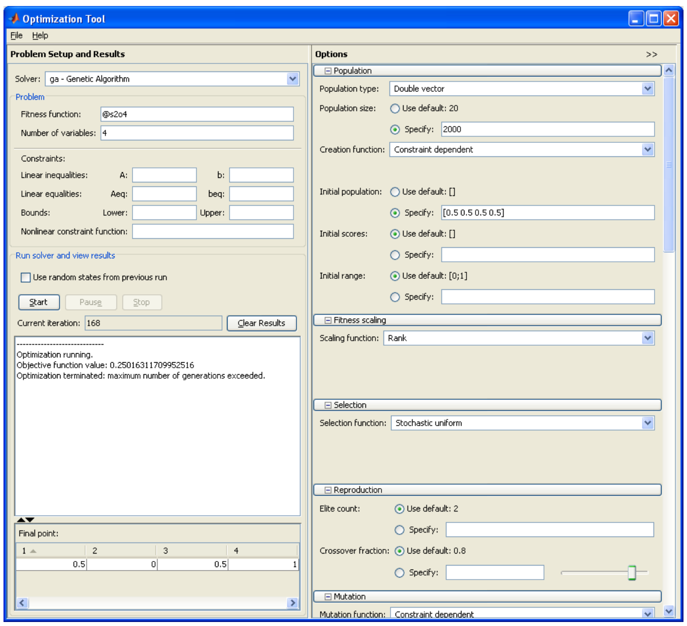


Figure 8. MATLAB's optimization tool wizard.

We see MATLAB's fminsearch() and simulannealbnd() cannot solve the problem, and the solver ga() using genetic algorithm returns the objective function value 0.250163 by using population size 2000 and iteration number 168. As the population size and iteration number increases, the returned objective values become closer and closer to the true minimum, which is 0.25.

E. Unsolver

VicSolver's UNsolver [6] is an unconstrained derivative-free nonsmooth optimization solver developed by Dr Fuchun Huang of Victoria University, Australia. UNsolver uses FEFAR [10] as an application programming user's interface for users to specify an unconstrained smooth or nonsmooth optimization problem. The following is the LEFAR [10] code for the sensor location problem at the beginning of the section.

```
!! optimal 2 sensors for 4 points

function frf(if123,fmin,b)
real                :: fmin
real, dimension(4)  :: b
integer             :: if123

if(if123==1)
      b(1)=0.5
      b(2)=0.5
      b(3)=0.5
      b(4)=0.5
end if

fmin=0.0
if(if123<=2)
fmin=max(fmin,min(b(1)^2+b(2)^2, b(3)^2+b(4)^2));
fmin=max(fmin,min((b(1)-1)^2+(b(2)-0)^2, (b(3)-
1)^2+(b(4)-0)^2));
fmin=max(fmin,min((b(1)-0)^2+(b(2)-1)^2, (b(3)-
0)^2+(b(4)-1)^2));
fmin=max(fmin,min((b(1)-1)^2+(b(2)-1)^2, (b(3)-
1)^2+(b(4)-1)^2));
end if

if(if123==3)
   print,fmin
   print,b(1)
   print,b(2)
   print,b(3)
   print,b(4)
end if

end function
```

The following is the running result of UNsolver:

```
C:\ISMtalk>UNsolver.exe
 Input the file name of the function:
2s4a.far
  fmin=  0.50000000000000000      at
0.50000000000000000       0.50000000000000000
0.50000000000000000       0.50000000000000000

. . . . . .

 fmin=  0.25000000000000000      at
0.0000000000000000       1.0000000000000000
0.50000000000000000       0.50000000000000000
 fmin=  0.25000000000000000      at
0.0000000000000000       1.0000000000000000
0.50000000000000000       0.50000000000000000

. . . . . .

 fmin=  0.25000000000000000      at
0.0000000000000000       1.0000000000000000
0.50000000000000000       0.50000000000000000
 fmin=  0.25000000000000000      at
0.0000000000000000       1.0000000000000000
0.50000000000000000       0.50000000000000000
```

We see UNsolver finds one of the two optimal solutions. An evaluation version and some other test results of UNsolver are available from http://sites.google.com/site/VicSolver.

III. Summary Remarks And Future Scope

We present the sensor location problems and advocate them as test problems of nonsmooth optimization.

These problems have easy-to-understand practical meaning and importance, easy to be even randomly generated, and the solutions can be displayed visually on a 2-dimensional plane. For testing some ready-to-use nonsmooth optimization solvers, we present a very simple sensor location problem of two sensors for four objects with the optimal solutions known by theoretical analysis. We tested several optimization solvers on this problem and found that optimization solvers MATLAB's ga() and VicSolver's UNsolver can solve the problem, while some other optimization solvers like Excel solver, Dr Frank Vanden Berghen's CONDOR, R's optim(), and MATLAB's fminunc() cannot solve the problem. In the near future some medium to large scale "standard" sensor location problems will be generated and put online for researchers testing nonsmooth optimization solvers.

ACKNOWLEDGMENT

The author thanks Professor Yosihiko Ogata, Professor Satoshi Ito, Professor and Director Tomoyuki Higuchi and The Institute of Statistical Mathematics, Tokyo, Japan, for supporting his research visit to the institute from the 5th of January to the 28th of February in 2012.

REFERENCES

[1] L. Luksan and J. Vlcek, "Test Problems for Nonsmooth Unconstrained and Linearly Constrained Optimization," Technical Report, No. 798, Institute of Computer Science, Academy of Sciences of the Czech Republic, 2000.

[2] Frontline Solvers, http://www.solver.com/sdkplatform2.htm, retrived on March 29, 2012.

[3] F. V. Berghen, "CONDOR, a parallel, direct, constrained optimizer for high-computing-load, black box objective functions", Proceeedings of the third *MIT* conference on Computational Fluid and Solid Mechanics, Elsevier, June 14-17, 2005.

[4] R Development Core Team, "R: A language and environment for statistical computing", R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, URL http://www.R-project.org/, 2011.

[5] Wikipedia, "MATLAB", http://en.wikipedia.org/wiki/MATLAB, retrived on March 29, 2012.

[6] F. Huang, "UNsolver: a new solver of unconstrained nonsmooth optimization problems", http://sites.google.com/site/VicSolver.

[7] F. V. Berghen, J. Cardinal and S. Langerman, "Min-max-min Geometric Facility Location Problems", 22nd European Workshop on Computational Geometry, Delphi, March 27-29, 2006.

[8] J. Czyzyk, M.P. Mesnier, and J.J. More, "The NEOS Server", IEEE Computational Science and Enginnering, Vol. 5, pages 68-75, 1998.

[9] R. Fourer, D.M. Gay, and B.W. Kernighan, AMPL: A Modeling Language for Mathematical Programming. Scientific Press, San Francisco, CA, 1993.

[10] F. Huang, "A New Application Programming Interface And A Fortran-like Modeling Language For Evaluating Functions and Specifying Optimization Problems At Runtime", International Journal of Advanced Computer Science and Applications, vol. 3, issue 4, pp. 1-5, April 2012.

AUTHORS PROFILE

Dr Fuchun Huang is a Senior Lecturer in the School of Engineering and Science at Victoria University, Melbourne, Australia. He was awarded a PhD degree by The Graduate University of Advanced Studies, Tokyo, Japan, and has published papers on computational statistics, in particular Monte Carlo methods, pseudo-likelihood and generalized pseudo-likelihood methods, and developed solver software for solving smooth and nonsmooth optimization problem. He is a member of The Japan Statistical Society.

The Role Of Technology and Innovation In The Framework Of The Information Society

Peter Sasvari
Institute of Business Sciences, Faculty of Economics
University of Miskolc
Miskolc-Egyetemvaros, Hungary

Abstract— **The literature on the information society indicates that it is a still-developing field of research. It can be explained by the lack of consensus on basic definitions and research methods. There are also different judgments on the importance and the significance of the information society. Some social scientists write about a change of era, others emphasize parallelism with the past. There are some authors who expect that the information society will solve the problems of social inequalities, poverty and unemployment, while others blame it on the widening social gap between the information haves and have-nots. Various models of the information society have been developed so far and they are so different from country to country that it would be rather unwise to look for a single, all-encompassing definition. In our time a number of profound socio-economic changes are underway. Almost every field of our life is affected by the different phenomena of globalization, beside the growing role of the individual; another important characteristic of this process is the development of an organizing principle based on the free creation, distribution, access and use of knowledge and information. The 1990s and the 21st century is undoubtedly characterized by the world of the information society (as a form of the post-industrial society), which represents a different quality compared to the previous ones. The application of these theories and schools on ICT is problematic in many respects. First, as we stated above, there is not a single, widely used paradigm which has synthesized the various schools and theories dealing with technology and society. Second, these fragmented approaches do not have a fully-fledged mode of application to the relationship of ICT and (information) society. Third, SCOT, ANT, the evolutionary- or the systems approach to the history of technology – when dealing with information society – does not take into account the results of approaches (such as information science or information systems literature or social informatics, information management and knowledge management, communication and media studies) studying the very essence of the information age: information, communication and knowledge. The list of unnoticed or partially incorporated sciences, which focuses on the role of ICT in human information processing and other cognitive activities, is much longer.**

Keywords-Information society; Social Construction of Technology; Actor-Network-Theory.

I. INTRODUCTION

Many theories can be found in the literature on the information society. The theories of the knowledge or information economy, postindustrial society, postmodern society, information society, network society, information capitalism, network capitalism etc. show that it is an important sociological issue to understand what role is played by technology and information in the society we live in. Both aspects - the form of society and the role of technology and information - belong to the central question of the theory of the information society [28] [29].

II. THE DEFINITION OF INFORMATION

In everyday use, the term "information" has meant a kind of guidance for a long time: when someone goes to the railway station to be informed on the content of the timetable or to the information desk to find out where a product or a counter can be found in the department store. Such information exchange works only if the right piece of information, the one that fits and makes sense for both parties of the communication is available.

Information as a term became more and more popular in the last 30-40 years; it has started to have an increasingly important role in everyday language while its strict meaning mentioned above has gradually faded away. At the same time, there has been a growing uncertainty about the true meaning of the term 'information'. All this doubtfulness is mainly caused by the so-called 'information-centered' world we are living in and by the widespread expansion of information and communication technology as almost everyone living in developed western societies can experience the phenomenon called the Information Revolution. All this suggests that information has become an essential part of our society and plays a centre role in our lives.

In information studies, rather complex definitions can be found on the nature of information.

Informatics or computer science is a discipline that deals with the storage, processing and distribution of information as well as planning computer networks and determining their operation principles.

Determining the exact subject of computer science is rather difficult because it is extremely hard to define what information is. According to the German physicist and philosopher, Weizsacker information should be regarded as the third universal elementary quantity beside matter and energy in science and technology [34].

According to the ninth volume of the Pallas Nagy Lexikon (Great Pallas Encyclopedia) information is a term with Latin origins meaning report, enlightment, inform, let somebody know; informant, instructor, messenger [17].

In the Dictionary of Foreign Words and Expressions the following meanings can be read:

1. enlightment, announcement, communication;
2. message, data, news, bulletin.

The fourth edition of American Heritage Dictionary of the English Language distinguishes seven meanings of the term 'information':

1. Knowledge derived from study, experience, or instruction,
2. Knowledge of specific events or situations that has been gathered or received by communication; intelligence or news,
3. A collection of facts or data: statistical information,
4. The act of informing or the condition of being informed; communication of knowledge,
5. Computer Science Processed, stored, or transmitted data,
6. A numerical measure of the uncertainty of an experimental outcome,
7. Law a formal accusation of a crime made by a public officer rather than by grand jury indictment.

The theory of communication states that information is the objective content of the communication between objects conversely affecting each other which are manifested in the change of the condition between these objects.

According to the science of telecommunication information is a series of signals structured in time and space, which is made up of a signal set having a specific statistical structure. The sender transmits data on the condition of an object or on the course of an event and the receiver perceives and interprets these signals. Everything can be regarded as information that is encoded and transmitted through a definite channel.

From the perspective of social science, information is the communication of useful knowledge that is created and transmitted in the intellectual communication system of the society. It is characteristic to the society as a whole, belonging to one of the global issues of world together with energy and environment protection.

According to the economic approach, information is partly a form of service, partly a product but, not as in the case of exchange of goods, both parties can keep their information. The content of material, energy and living labour is gradually decreasing in manufactured goods, while the amount of product information input is increasing at the same rate.

In summary, information is an expression related to enlightment, data, report, learning, communication and news. In certain cases, it can be identified with these items (knowledge, data, enlightment, news); in other cases it is the object of these listed items (conveyance of knowledge, learning, communication).

Despite the fact that it may still sound uncertain, the group of the terms 'data', 'knowledge' and 'communication' can be highlighted for giving an interpretation of information. According to the literature, the transformation of data into information needs knowledge. There are many definitions trying to find a link between information and communication, which also can have an importance when looking for a definition of the information society. Communication is a process of transferring information from one entity to another through a specific medium. If we link these two different approaches together, the picture we are given is a very complex one, where the four terms 'data', 'information', 'knowledge' and 'communication' must be interpreted in one compound definition. The same connection was made by Michael Buckland in his book on information systems.

TABLE I. FOUR ASPECTS OF INFORMATION [3]

	Intangible	**Tangible**
Entity	Information as knowledge	Information as thing
Knowledge	Information as process	Information as process

Information as knowledge is subjective in every case, it is linked to a given individual and it gains its exact meaning in a specific environment. It is intangible as an entity but it can be communicated, made to be known to others. Information as a thing exists similarly to knowledge, however, it is tangible. In this regard, data can be regarded as a kind of recorded knowledge because it is necessary to know the context of its creation (or the record structure), without having this context, the data cannot be interpreted.

III. INFORMATION AND SOCIETY

By definition, society refers to

1. human relations and relationships taken as a whole,
2. any community of human beings is able to perpetuate itself, more or less linked to a specific region or country, sharing a distinctive culture and institutions.

Whether a human community is regarded as a society depends on the extent to which its members are able to interact with each other, thus the capacity and extension of interaction is essential.

The most recent trends show that the definition of society has become less important in trying to understand the world surrounding us because if we examine only individual societies, we may not notice social (multilateral and global) phenomena between and over societies.

If we accept that the key feature of social existence is the development of relations, then the information society may bring a significant change in this very context: a lot more individuals have the opportunity to get in contact with other people in a simpler way and at a lower cost.

A question comes up here immediately: is it possible to call every human society an information society? Information is the essential condition of the functioning of every society, including their subsystems as well. It played an important role in every social formation in the agricultural and industrial societies of previous ages. Information flow is needed in every society but none of the previous societies were labeled "information society" by contemporary analysts and historians. The reason for this is that the communication, reception, processing, storage, interpretation and flow of information never determined earlier societies to such a high extent as today's. The activities relating to information have become more valuable in present day societies and that is what distinguishes them sharply from the societies of the past. This fundamental difference is convincingly described by theoretician of various interests, views and attitudes and orientation in the following five fields:

1. technology,
2. occupation structure,
3. the operation of economy,
4. spatial structure,
5. culture.

Frank Webster's book published in 1995 synthesizes the 1960s and 1970s information society theories in order to analyze the concept and its characteristics within the context of social science [32]. These theories designate the potential directions of what might be a comprehensive research project, which can clarify the concept and exploit these theories as starting points for further exploration. Webster's typology is the following [33]:

A. *Technology*

From the technological perspective we live in an information society since information and telecommunication technologies play a constantly expanding role in all fields of social existence, which has shaken the foundations of social structures and processes and resulted in profound changes in politics, economy, culture, and everyday life. [10]

Most of the attempts made to define information society approach the idea from a technological point of view hence the central question of such explorations sounds like: What kind of new information and communication technology was constructed in recent decades that determined the infrastructure of information society?

B. *Occupation structure and economy*

Studies of occupational structure and economy show that we live in an information society because, when we have passed through the agricultural and industrial stages, the information sector and information oriented jobs dominate the economy. The main questions raised by this approach are: How have the proportions of employed workers changed in the industrial and service sectors in recent decades? How have their performance and the knowledge they use changed qualitatively? Have the so-called informational occupations begun to dominate production?

The question is similar to that which we posed by the technological approach: What is the point at which we can claim that the logic of capitalism, that is, its structure of production has qualitatively changed? Is the often cited "new economy" indeed so different from the old one? Where is the turning point? Is it possible to identify the point at which the former was replaced by the latter?

C. *Spatial structure*

As the spatial theorists see it we live in an information society because through the use of information technologies and globalization physical space tends to lose its determining function. People are participating in networks that determine such social processes as production, division of labor, discussing politics for example.

The main theoretical questions are the following: Does the world follow the logic of networks? Does global society exist? Can it come to life? What is the inherent logic of global networks? Who belongs to them, and why do they wish to do so? What kind of social and economic capital is needed to gain access to a network and how can membership then be maintained? What are the innate social relations of the network, and what part do the new information and communication technologies play in those relations?

D. *Culture*

The cultural perspective also states that we live in an information society because our life is infiltrated by the globalised, extensively digitalized media culture that has become the primary means of providing sense and meaning for us and predominantly determines our lifestyle.

Theories attempting to explain the cultural aspects of information society describe such a global cultural context that may be adopted universally as a referential framework for the media. This approach also suggests that the media enjoy a unique status in the age of information and that they are the most prominent determining factors of social relations.

However, the question remains: whether life exists beyond media culture or not? Does the illusory game of signs have any connection to reality? The catchphrase of the information age is "virtual reality" which reality very often turns out to be more fundamental than the world that created it.

IV. THE INFORMATION SOCIETY

Many theories can be found in the literature on the information society. The theories of the knowledge or information economy, postindustrial society, postmodern society, information society, network society, information capitalism, network capitalism etc. show that it is an important sociological issue to understand what role is played by technology and information in the society we live in. Both aspects - the form of society and the role of technology and information - belong to the central question of the theory of the information society.

One of the first social scientist to develop the concept of the information society was the economist Fritz Machlup [15]. In his breakthrough study, "The production and distribution of knowledge in the United States", he introduced the concept of the knowledge industry by distinguishing five sectors of the knowledge sector:

- education,

- research and development,
- mass media,
- information technologies,
- information services.

Peter Drucker has argued that there is a transition from an economy based on material goods to one based on knowledge [7] [8] [9].

Marc Porat distinguishes [19] [20]

- a primary sector (information goods and services that are directly used in the production, distribution or processing of information) and
- a secondary sector (information services produced for internal consumption by government and non-information firms) of the information economy.

Porat uses the total value added by the primary and secondary information sector to the GNP as an indicator for the information economy. The OECD has employed Porat's definition for calculating the share of the information economy in the total economy. Based on such indicators the information society has been defined as a society where more than half of the GNP is produced and more than half of the employees are active in the information economy.

For Daniel Bell the number of employees producing services and information is an indicator for the informational character of a society [1] [2]. A post-industrial society is based on services. What counts is not raw muscle power, or energy, but information. A post industrial society is one in which the majority of those employed are not involved in the production of tangible goods.

1. economic sector,
2. resource,
3. strategic resource,
4. technology,
5. knowledge base,
6. methodology,
7. time perspective,
8. planning,
9. guiding principle.

Alain Touraine already spoke in 1971 of the post-industrial society [30]. "The passage to postindustrial society takes place when investment results in the production of symbolic goods that modify values, needs, representations, far more than in the production of material goods or even of 'services'. Industrial society had transformed the means of production: post-industrial society changes the ends of production, that is, culture. The decisive point here is that in postindustrial society all of the economic system is the object of intervention of society upon itself. That is why we can call it the programmed society, because this phrase captures its capacity to create models of management, production, organization, distribution, and consumption, so that such a society appears, at all its functional levels, as the product of an action exercised by the society itself, and not as the outcome of natural laws or cultural specificities".

In the programmed society also the area of cultural reproduction including aspects such as information, consumption, health, research, education would be industrialized. That modern society is increasing its capacity to act upon itself means for Touraine that society is reinvesting ever larger parts of production and so produces and transforms itself. This idea is an early formulation of the notion of capitalism as self-referential economy.

Similarly to Bell Peter Otto and Philipp Sonntag assert that an information society is a society where the majority of employees work in information jobs, i.e. they have to deal more with information, signals, symbols, and images than with energy and matter.

Radovan Richta argues that society has been transformed into a scientific civilization based on services, education, and creative activities. This transformation would be the result of a scientific-technological transformation based on technological progress and the increasing importance of computer technology. Science and technology would become immediate forces of production [21].

Nico Stehr says that in the knowledge society a majority of jobs involves working with knowledge. "Contemporary society may be described as a knowledge society based on the extensive penetration of all its spheres of life and institutions by scientific and technological knowledge" [23] [24] [25] [26].

For Stehr knowledge is a capacity for social action. Science would become an immediate productive force, knowledge would no longer be primarily embodied in machines, but already appropriated nature that represents knowledge would be rearranged according to certain designs and programs. The economy of a knowledge society is largely driven not by material inputs, but by symbolic or knowledge-based inputs, there would be a large number of professions that involve working with knowledge, and a declining number of jobs that demand low cognitive skills as well as in manufacturing.

Also Alvin Toffler argues that knowledge is the central resource in the economy of the information society: "In a Third Wave economy, the central resource – a single word broadly encompassing data, information, images, symbols, culture, ideology, and values – is actionable knowledge".

In recent years the concept of the network society has gained importance in information society theory. For Manuel Castells network logic is besides information, pervasiveness, flexibility, and convergence a central feature of the information technology paradigm. "One of the key features of informational society is the networking logic of its basic structure, which explains the use of the concept of 'network society'" As a historical trend, dominant functions and processes in the Information Age are increasingly organized around networks. Networks constitute the new social morphology of our societies, and the diffusion of networking logic substantially modifies the operation and outcomes in processes of production, experience, power, and culture. For Castells the network society is the result of informationalism, a new technological paradigm [4] [5].

Jan Van Dijk defines the network society as a "social formation with an infrastructure of social and media networks enabling its prime mode of organization at all levels (individual, group/organizational and societal) [6]. Increasingly, these networks link all units or parts of this formation (individuals, groups and organizations)". According to Van Dijk networks have become the nervous system of society, whereas Castells links the concept of the network society to capitalist transformation, Van Dijk sees it as the logical result of the increasing widening and thickening of networks in nature and society.

The major critique of concepts such as information society, knowledge society, network society, postmodern society, postindustrial society, etc. that has mainly been voiced by critical scholars is that they create the impression that we have entered a completely new type of society. If there is just more information then it is hard to understand why anyone should suggest that we have before us something radically new.

Such neomarxist critics as Frank Webster argue that these approaches stress discontinuity, as if contemporary society had nothing in common with society as it was 100 or 150 years ago [30]. Such assumptions would have ideological character because they would fit with the view that we can do nothing about change and have to adopt to existing political realities. These critics argue that contemporary society first of all is still a capitalist society oriented towards accumulating economic, political, and cultural capital. They acknowledge that information society theories stress some important new qualities of society (notably globalization and informatization), but charge that they fail to show that these are attributes of overall capitalist structures. If there were a discourse on continuity and discontinuity, capitalism would enter into a new development stage.

Concepts such as knowledge society, information society, network society, informational capitalism, postindustrial society, transnational network capitalism, postmodern society, etc. show that there is a vivid discussion in contemporary sociology on the character of contemporary society and the role that technologies, information, communication, and co-operation play in it. Information society theory discusses the role of information and information technology in society, the question which key concepts shall be used for characterizing contemporary society, and how to define such concepts. It has become a specific branch of contemporary sociology.

V. Legal regulations of information society

The legal material concerning information society is interwoven into our legal system horizontally. The rules related to information society are enshrined to a greater or lesser extent in the several areas of law. As in any regulatory domain, the legal content concerning information society can be grouped according to the system of law. There are two distinct groups: the laws organising legal relations between the state and its citizens, and between the various state or public organizations (called public law), and the laws organising legal relations between citizens and partnerships, and between members of civil society (civil law) [11]. Differentiation is based on the relationship between those involved. While in the first case we can speak of an unequal legal relation based on subordination and superiority, in civil law the typical legal relation is one of equality and coordination.

In the continental legal system, we can distinguish between four main categories:

1. civil law,
2. criminal law,
3. administrative law,
4. constitutional law.

Civil law regulates the property personal and family relations of natural and legal persons in cases where the partners are equal and state intervention, except for legislation, occurs only in the event of a legal dispute. The most important areas affecting information society are as follows:

- e-commerce,
- digital signature,
- content regulation,
- protection of copyright and industrial property rights,
- media law,
- competition law.

Criminal law regulates acts that are a danger to society. We can group all those acts committed with or against IT technology which are dangerous for society and for which the law orders the sanction of punishment. Legal regulation of information society is primarily concerned with the following categories of crime:

- misuse of personal data,
- content-related crimes (e.g. distribution of child pornography, hate speech, etc.),
- crimes against computer systems and data,
- infringement of copyright.

Administrative law is the regulatory system of state functions. State administration extends beyond central government and local government to larger systems; for example the operation of transport, security, military and information systems. The following functions essential to information society belong to this group:

- electronic administration,
- electronic register of companies,
- administrative procedure,
- electronic public procurement.

The fourth field is constitutional law, which arose out of continental legal development. The object of regulation is to structure relations between the citizens and the state and the organizational structure of the state. The constitution is the document describing basic rights, responsibilities and procedures thus creating the basis for the process governing political, economic and social life. Areas of constitutional law related to the information society are as follows:

- electronic freedom of information,
- personal data protection,
- freedom of the press and freedom of expression.

VI. The effects of technology and innovation on society

Technique can be defined as the application of some devices or knowledge in order to accomplish a specific task or fulfill a purpose. These purposes may range from industrial use to social needs, improving working conditions or raising the standard of living. For humans, technique is an acquired way of using the surrounding environment for satisfying their own instinctive goals and cultural desires. It is the knowledge to create something new.

Under the term 'technology' I mean all the procedures and knowledge of procedures that are needed to perform a specific task.

Studies considering science and technology as an inseparable and organic part of society, like information society studies, do not have a unified conceptional and methodological apparatus, nor a comprehensive and prevailing scientific paradigm. We can talk about a variety of multidisciplinary and interdisciplinary studies, schools, theories and approaches interacting with each other and comprising works of scholars from various traditional sciences like history, economics, sociology or anthropology. The great number of diverse approaches makes it impossible to review them completely, so we have to forget about introducing schools like the technology theories of evolutionary economics in detail. On the whole, the goal of this chapter can be nothing more than to provide an ”intellectual crutch” for discussing and interpreting information communication technologies by reviewing the most relevant and important theories, concepts, models and notions of the topic.

Technological determinism argues that technology is the principal driving force of society determining its mode of operation, development, course of history, structure and values in a decisive manner. Converse effects are taken into account to a limited extent, fully disregarded or disclaimed. Technological development is thought to be propelled by the logic of science alone.

Most scientific concepts explicitly reject technological determinism; yet they assist its survival by studying only technology’s influence on society. This is more symptomatic of ICT related researches.

The beginning of Science, Technology and Society studies dates back to the early 1970s, when the first studies were published. The novelty in the pioneering works, which lends them their special character even today, was that they stressed – contrary to technological determinism – society’s crucial role in the development of science and technology, framing the three intermingling domains in complex theoretical systems. The works of philosophers, historians and sociologists were collected in two books in the mid-eighties, which have become the most cited publications of this school. Some of these approaches have developed into theories, generating further discourses and STS has been crystallized into an interdisciplinary field of research with both common research areas and methodology.

The STS school is far from being the dominant scientific paradigm of this area of knowledge, but has several advantages that make it indispensable when examining information society and ICT. These are its strong empirical basis and complex approach to analyzing interaction between technology and society, their manifold co-dependence, and complex co-development. Within the several concepts of STS, many schools exist criticizing and complementing each other.

A. *Studies of the interactions between science, technology and society*

The foundations of STS were laid down in the 1980s by the “Social Construction of Technology” school, which focuses on the development phase of technologies at the micro level, and pinpoints that technology (and natural scientific developments) are basically shaped by social processes. [14]

Any given technology stabilizes when debates are settled. This is the phase of ‘closure and stabilization’. Closure, however, does not mean finalizing: newly joined user groups can reopen the debates which can lead to new modifications to or variations of the existing technology.

Using the terminology of evolutionary approaches, we can say variations, mutations and hybrids are brought to life during the diffusion of a certain technology, which is chiefly true for ICT. Take the different variations of computers (desktop PC, portable notebook, PDA, etc.) or the convergence of mobile phones with other electronic devices (such as PDAs, digital cameras, mp3-players, game consoles, or GPS devices) which are typical hybrids.

Bijker and Pinch emphasize that the meanings assigned to technologies are determined by the norms and values of social groups which draw the “wider context” of socio-cultural and political environment into the set of determining factors. Drawing on the wider context concept, Laudan R. argues that changing social values can bring new technological constructs or their complete generation to life. The heterogeneous and hierarchical community of technological development functions as a mediator of social values and forces value orientation in society to change.

Capital mobility has increased incredibly, the economy has shifted to the service sector, and innovation has become the primary source of productivity growth in relation to engineering, organizations, institutions as well as individual workers. The "technical construction of society" has become a major issue, that is social processes are mainly mediated by technological development. The society's level of being informed, with exploiting the opportunities provided by information and communication technology, has been increased dramatically. This new technology, together with biotechnology opens new perspectives in the fields of industry, trade and education.The nature of economic competition has been undergoing huge changes, as more and more people think that there have been profound changes in the relation between economy and society and innovation requirements. The continuous and self-accelerating innovation processes characterized by the intense competition has brought about some changes in time relations. People start moving on a different time scale, time has been speeded up.

Space has become globalized, by turning into more unified and more complex at the same time. Socio-economic processes create new virtual spaces or even real spaces are modified: the processes are arranged in new ways in the interacting local, regional, national and supranational spaces. While integration processes are considered to be a general tendency, clear attempts for isolation also appear repeatedly. Knowledge has become the main economic source, and learning abilities and skills have become a criterion of adaptation at the levels of individuals, companies, local communities, nations, supranational organizations and the world taken as a global system.

Actor-Network-Theory is another school of STS studies, which is more and more widely used. It is a new branch of the sociology of science and technology, the basis of which was elaborated by Michel Callon, Bruno Latour and John Law in the 1980s. They – along with other scholars – developed their concepts into a theory.

A basic statement of ANT is that technological objects along with their socio-political context co-develop and shape each other mutually into socio-technical entities through constant interactions. The objects and their context form heterogeneous networks made up of human and non-human components which are connected to each other dynamically. These heterogeneous components can be objects, techniques, institutions, organizational solutions, human abilities or cognitive structures.

Human components as network builders are constantly formed and constituted by the networks they are part of. Actors in this network are connected by intermediaries, which in many cases, have social meanings. Texts, technical artifacts, currencies or human skills can function as intermediaries. One of ANT's – much debated – theorems is that the natural state of society is disorder. Order is achieved through the constant and endless efforts made by the actors to build networks.

Callon argues that an actor-network cannot be derived either from the actor or the network. The actions and the will of actors are inseparable from the network, and their effect runs through the whole network. This leads us to one of ANT's radical novelties: the boundaries between the actors disappear and even actions cannot be interpreted in the traditional way.

In the literature, the constant shifting of power between technology and society is called translation: as a result of this process, networks are formed progressively, in which certain entities gain control over other entities.

B. Diffusion of innovations

Innovation has become a key activity of information societies. It is the cornerstone of economic competitiveness. National and regional (such as European) administrations develop high level strategies to promote innovative activities in the economy.

Innovation can be defined as basically novel inventions or concepts – arising from either professional research or ideas by amateurs – translated into practice. An innovation can be a technological object, a new organizational solution or an idea. Innovations become market goods through product development and/or technology transfer. The product cycle consists of the following stages: introduction (to the market), growth, maturity and stabilization, and decline. The life cycle of common goods (e.g. road infrastructure) and public goods (e.g. public safety) go through the same stages. Rogers' theory applies to the life cycle of innovations as far as the maturity phase and at the level of communities and societies.

Rogers explains the diffusion of innovations as basically communicative: diffusion is the process by which an innovation is communicated through certain channels over time among the members of a social system. Diffusion is determined by the above mentioned four factors (innovation, communication channels, time and social systems). It is a process of decision making, in the stages of which different types of information and knowledge transferring mechanisms play crucial roles.

The diffusion of innovations – thus, of technologies too – takes place within social networks, so called diffusion networks. The ability of individuals to adapt depends on the cohesion of these networks, in other words, to the extent of its homophile (similar socio-economic status, qualifications, attitudes); on structural equivalence (on the individual's position in the network); and on the threshold of other users which makes it worthwhile for a group member to adopt the given technology.

Innovators play a crucial role in diffusing an innovation between homophile diffusion networks. They tend to use the technology first, and usually possess heterophile social relations (they maintain regular relationships with several social groups and through them, several networks of diffusion). Chronologically, the second group to adopt an innovation is called the early adopters; these are followed by the early majority, then the late majority, and lastly, the laggards. Each of these ideal-typical groups is characterized by specific socioeconomic factors, personality values and communication behavior. For example laggards are the most disadvantaged group along the socio-economic scale.

When studying the diffusion of ICT, at least one more category must be added: the refusers, who consciously resist usage throughout their lives (also known as diehards). The existence of this group indicates that no technology ever penetrates a society fully. To reach 100% diffusion both society and technology need to change as compared to their initial status when the innovation was introduced.

The process of diffusion is broken down into different stages from the individual user's point of view. First, one typically acquires information regarding innovation through mass media channels (or cosmopolitan communication channels). The following three phases are dominated by interpersonal channels (or local channels). In the second phase, persuasion and opinion forming take place, followed by deciding on the adaptation, finally evaluation and confirmation of the usage. Of course, refusing the implementation (even several times) is an option too, but it can be followed by acceptance, and vice versa, the evaluation of implementation can lead to discontinuing usage.

Rogers analyses the characteristics of an innovation affecting its own diffusion (such as relative advantage, compatibility, complexity, trial ability and observability), but gives little attention to their socially constructed nature [22].

The main advantage of Rogers' theory is that a key role is ascribed to communicative processes. This momentum makes the theory a close relative to other approaches such as SCOT and ANT. Rogers' theory can be drawn upon in the analyses of such information society related issues as the digital divide or e-inclusion.

VII. Conclusion

The application of these theories and schools on ICT is problematic in many respects. First, as we stated above, there is not a single, widely used paradigm which has synthesized the various schools and theories dealing with technology and society. Second, these fragmented approaches do not have a fully-fledged mode of application to the relationship of Information Control Technology (ICT) and (information) society [27].

Third, SCOT, ANT, the evolutionary- or the systems approach to the history of technology – when dealing with information society – does not take into account the results of approaches (such as information science or information systems literature or social informatics, information management and knowledge management, communication and media studies) studying the very essence of the information age: information, communication and knowledge. The list of unnoticed or partially incorporated sciences, which focuses on the role of ICT in human information processing and other cognitive activities, is much longer [31].

These, though, miss the approach of STS and evolutionary schools, particularly the concept of technology and society as a seamless web. Merging the two modes of understanding information society is in its infancy, though studying ICT systems cannot be completed without them both [29].

References

[1] D. Bell, "Az információs társadalom társas keretrendszere (The social framework of the information society) ", Információs Társadalom 2001/1

[2] D. Bell, "The Coming of the Post-Industrial Society, A Venture in Social Forecasting", London: Heinemann, 1976

[3] M. Buckland, "Information as thing. Journal of the American Society for Information Science", 1991

[4] M. Castells, "Az identitás hatalma (The Power of Identity) ", Budapest, Gondolat Kiadó, 2000

[5] M. Castells, "Az információ kora (The Information Age) ", Gondolat-Infonia, Az információs társadalom klasszikusai, Az információ kora, Gazdaság, társadalom és kultúra, II. kötet, 2000

[6] J. V. Dijk, "The Network Society", London: Sage. Second Edition, 2006

[7] P. Drucker, "The Age of Discontinuity: Guidelines to Our Changing Society", Harper & Row, New York, 1969

[8] P. Drucker, "A profit művészete (The Art Of Profit) ", Budapest, Gondolat Kiadó, 2003

[9] P. Drucker, "The Coming of the New Organization", Harvard Business Review, 1988

[10] C. Fuchs, "Internet and Society. Social Theory in the Information Age", NewYork: Routledge. Routledge Research in Information Technology and Society Series Number 8. ISBN 0415961327. 408 P.

[11] L. Z. Karvalics, and T. Dessewffy, "Internet.hu, A magyar társadalom digitális gyorsfényképe (Internet.hu, A digital snapshot of the Hungarian society) ", Infonia-Aula, 2003

[12] L. Z. Karvalics, "Információ, társadalom, történelem, Válogatott írások, (Information, society, history, Selected works)", Typotex Kiadó, 2003

[13] L. Z. Karvalics, "Információs társadalom – mi az? Egy kifejezés jelentése, története és fogalomkörnyezete (Information Society – what is it exactly? The meaning, history and conceptual framework of an expression), Az információs társadalom, Az elmélettől a politikai gyakorlatig (The Information Society, From theory to political practice)", Gondolat – Új Mandátum, Budapest, 29-47, 2007

[14] A. Kincsei, "Technology and society in the information age", Budapest, 2007

[15] F. Machlup, "The Production and Distribution of Knowledge in the United States", Princeton UP, 1962

[16] Y. Masuda, "Az információs társadalom (The Information Society) ", OMIKK, Budapest, 1988

[17] Pallas Nagy Lexikon IX. kötet, (The Great Pallas Encyclopedia, Vol. 9)

[18] A. Payne, "CRM-kézikönyv, Ügyfélkezelés felsőfokon", HVG Kiadó Zrt., 2006

[19] M. Porat, "Communication Policy in an Information Society", 1977

[20] M. Porat, "The information economy: definition and measurement", Washington DC, Government Printing Office, 1977

[21] R. Richta, "The Scientific and Technological Revolution and the Prospects of Social Development, Scientific-Technological Revolution". Social Aspects, London: Sage. pp. 25–72., 1977

[22] E. M. Rogers, "Diffusion of Innovations", The Free Press, 1995

[23] N. Stehr, "Arbeit, Eigentum und Wissen", Frankfurt/Main: Suhrkamp, 1994

[24] N. Stehr, A World Made of Knowledge. Lecture at the Conference, "New Knowledge and New Consciousness in the Era of the Knowledge Society", Budapest, 2002

[25] N. Stehr, "Knowledge & Economic Conduct", Toronto: University of Toronto Press, 2002

[26] N. Stehr, "A modern társadalmak törékenysége (The fragility of modern societies) ", Gondolat – Infonia, 2007

[27] K. Szabo, and B. Hamori, "Információgazdaság (Information Economy)", Akadémiai Kiadó, Budapest, 2006

[28] A. Toffler, "The Third Wave", Budapest, Typotex, 2001

[29] A. Toffler, "A harmadik hullám (The Third Wave) ", Typotex, 2000

[30] A. Touraine, "Return of the Actor". Minneapolis, University of Minnesota Press, 1988

[31] T. Umesao, "Japanese Civilization in the Modern World", 2003

[32] F. Webster, "Theories of the Information Society", Routledge, 1995

[33] F. Webster, "The Information Society Revisited", In: Lievrouw, Leah A./Livingstone, Handbook of New Media, London: Sage., 2002, pp. 255-266.

[34] C. F. V. Weizsacker, "A német titánizmus (German Titanism) ", Európa Kiadó, 1989

Author Profile

Dr. Peter Sasvari received his MSc in Mechanical Engineering, MSc in Economics and his PhD in Business and Organisation Sciences at the University of Miskolc. Now he is an associate professor at the Institute of Business Science, Faculty of Economics, University of Miskolc, Hungary. His current research interests include different aspects of Economics, ICT and the information society.

Neural Network Based Hausa Language Speech Recognition

Matthew K. Luka
Department of Electrical and Information Engineering, Covenant University, Ota, Nigeria

Ibikunle A. Frank
Department of Electrical and Information Engineering, Covenant University, Ota, Nigeria

Gregory Onwodi
National Open University of Nigeria
Lagos, Nigeria

***Abstract*--Speech recognition is a key element of diverse applications in communication systems, medical transcription systems, security systems etc. However, there has been very little research in the domain of speech processing for African languages, thus, the need to extend the frontier of research in order to port in, the diverse applications based on speech recognition. Hausa language is an important indigenous lingua franca in west and central Africa, spoken as a first or second language by about fifty million people. Speech recognition of Hausa Language is presented in this paper. A pattern recognition neural network was used for developing the system.**

Keywords--Artificial neural network; Hausa language; pattern recognition; Speech recognition; speech processing

I. Introduction

Speech is a natural mode of communication for people because it is the most efficient modality for communication. This is coupled with the fact that the human brain has an impressive superiority at speech recognition as with other cognitive skills. Automatic speech recognition (ASR) is a process by which a machine identifies speech [1]. Speech recognition is used in different domain of life such as in telecommunications, mobile telephony, multimedia interaction, transcription, video games, home automation [2, 3].

The accuracy of a speech recognition system is affected by a number of factors. Generally, the error associated with discriminating words increases as the vocabulary size grows. Even a small vocabulary can be difficult to recognize if it contains confusable words. Also, speaker independence is difficult to achieve because system's parameters become tuned to the speaker(s) that it was trained on, and these parameters tend to be highly speaker-specific [4]. Other factors of interest are: continuity of speech, task and language constraint, and adverse conditions.

Generally, there are three methods widely used in speech recognition: Dynamic Time Warping (DWT), Hidden Markov Model (HMM) and ANNs [5]. Dynamic Time Warping algorithm is used to recognize an isolated word sample by comparing it against a number of stored word templates to determine the one that best matches it. This goal is complicated by a number of factors.

First, different samples of a given word will have somewhat different durations. This problem can be eliminated by simply normalizing the templates and the unknown speech so that they all have an equal duration. Dynamic Time Warping (DTW) is an efficient method for finding optimal nonlinear alignment between a template and the speech sample. The main problem of systems based on DTW is the little amount of learning words, high computation rate and large memory requirements [6]. A Hidden Markov Model is a statistical Markov Model in which the system being modelled is assumed to be a Markov process with unidentified (hidden) states. For speech recognition, the use of HMM is subject to the following constraint:

1) must be based on a first order Markov chain; 2) must have stationary states transitions; 3) observations-independence and 4) probability constraints.

Because speech recognition is basically a pattern recognition problem [4], neural networks, which are good at pattern recognition, can be used for speech recognition. Many early researchers naturally tried applying neural networks to speech recognition. The earliest attempts involved highly simplified tasks, e.g., classifying speech segments as voiced/unvoiced, or nasal/fricative/plosive. Success in these experiments encouraged researchers to move on to phoneme classification; this task became a proving ground for neural networks as they quickly achieved world-class results.

Speech recognition has been applied to most western and Asian languages. However, there is limited work reported on African languages. This research paper provides the background to the application of ANNs for speech processing of Hausa language. A pattern recognition network; which is a feed-forward network that can be trained to classify inputs according to target classes was used for recognition of Hausa language isolated words.

ANN is an information processing system that modelled on the performance characteristics of biological neural networks. ANNs are developed as generalizations of mathematical models of neural biology or human cognition based on the following assumptions [7]:

- Processing of Information occurs at many simple nodes (nodes are also called cells, units or neurons).
- Connection links are used for passing signals between the nodes.
- Each connection link is associated with a weight, which is a number that multiplies the signals.

- Each cell applies an activation function (usually nonlinear) to the weighted sum of its input to produce an output.

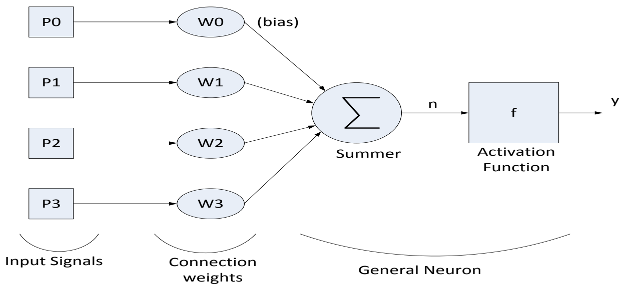


Figure 1. Mathematical Model of a basic neural network

The mathematical model of a three-input neuron is given in Fig. 1. The inputs are modified by the weights (synapses in biological neural network). A positive weight represents an excitatory connection, which a negative weight designates an inhibitory connection. A fictitious input known as the bias is normally used for training. The weighted inputs are summed in a linear fashion. Lastly, an activation function is applied to the weighted sum to determine the range and characteristic of the output. Mathematically,

$$n = \left(\sum_{i=1}^{3} w_j p_i + w_o p_o \right) \quad (1)$$

$$y = f(n) \quad (2)$$

There are several activation functions available. The four most commonly used ones include: (1) the threshold function, (2) the piecewise-linear function, (3) the sigmoid function and (4) the Gaussian (bell shaped) activation function. Based on architecture (connection patterns), artificial neural networks can be grouped into two groups; feed-forward networks and recurrent (feedback) networks. Feed-forward networks involve the one way flow of data from the input to the output. Examples of feed-forward networks include: single layer perceptron, multilayer perceptron and radial basis function networks. On the other hand, recurrent networks contain feedback connections which make their dynamical network properties important. Classical examples of feedback networks include: Competitive networks, Kohonen's Self-Organizing Maps, Hopfield networks and Adaptive Resonance Theory Models. Learning (or training) in the artificial neural network context is the task of updating network architecture and connection weights so that a network can efficiently perform a function [8]. The major difference between learning in ANN and experts systems is that in the latter, learning is based on a set of rules specified by human experts whereas in ANN, learning process is automatic (e.g. from input-output relationships). There are three main learning paradigms: Supervised, Unsupervised and Hybrid. Also, there are four fundamental learning rules (how connections are updated): Hebbian, error-correction, Boltzmann and competitive learning. The manner in which this rules are used for training a specific architecture is referred to as the learning algorithm.

II. Overview of Speech Recognition

Speech recognition deals with the analysis of the linguistic content of a speech signal. Speech recognition is one of the areas of research in speech processing and it is the study of speech signals and processing methods. Other aspects of speech processing include: speaker recognition, speech enhancement, speech coding, and voice analysis and speech synthesis. Based on the type of utterance to be recognized, speech recognition systems can be classified into four categories [9]: Isolated words system, connected words, Continuous speech and spontaneous speech systems. Human speech recognition starts with receiving speech waveform through the ear. It is based on "Partial recognition" of information across frequency, probably in the form of speech features that are local in frequency (e.g formants) [10]. The partial recognition (extracted features) is then integrated into sound units (phonemes), and the phonemes are then grouped into syllables, then words and so forth.

In order to analyse linguistic content of speech, the acoustic signal is first converted into an analogue signal that can be processed by an electronic device. The analogue signal is sampled and digitized for storage and further processing. In this work, PRAAT- a flexible program was used for the analysis and reconstruction of acoustic speech signals. PRAAT offers a wide range of procedures, including spectrographic analysis, pitch analysis, intensity analysis, articular synthesis and neural networks.

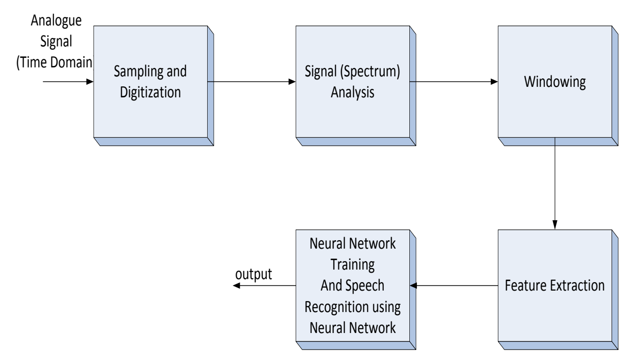


Figure 2. Connectionist based speech recognition model

Speech recognition may be considered to involve five important processes which are:

- Input Acquisition
- pre-processing (pre-emphasis and windowing)
- Feature Extraction
- Modelling
- Model Validation (testing)

A. Input Acquisition

The acoustic signals were captured in a low noise environment using a good quality microphone. The analogue output signal of the microphone is recorded using PRAAT at a sampling rate of 12 KHz. To facilitate ease of editing and manipulation, the digital data is saved as a .wav file.

B. Pre-Processing

Pre-Processing is the separation of the voiced region from the silence/unvoiced portion of the captured signal [11]. Pre-processing is necessary because most of the speaker or speech specific information are present in the voiced part of the speech signal [12]. Pre-processing also helps significantly in reducing the computational complexity of later stage [13]. Two vital pre-processing steps are pre-emphasis and windowing.

Pre-emphasis: In general, the digitized speech waveform has a high dynamic range and suffers from additive noise to reduce this range pre-emphasis is applied. By pre-emphasis, we imply the application of a high pass filter, which is usually a first-order FIR of the form:

$$H(z) = \sum_{k=0}^{N} \alpha(k) z^{-k} \qquad (3)$$

Normally, a single coefficient filter digital filter known as pre-emphasis filter is used:

$$H(z) = 1 - az^{-1} \qquad (4)$$

Where the pre-emphasis factor α is computed as:

$$\alpha = \exp(-2\pi F \Delta t) \qquad (5)$$

Where F is the spectral slope will increase by 6dB/octave and Δt is the sampling period of the sound. The pre-emphasis factor is chosen as a trade-off between vowel and consonants discrimination capability [14].The new sound y is then computed as:

$$y_i = x_i - \alpha x_{i-1} \qquad (6)$$

Hearing is more sensitive above the 1 kHz. The pre-emphasis filter amplifies this region of the spectrum. This assists the spectral analysis algorithm in modelling the perceptually important aspects of the speech spectrum [15]. In this work, a pre-emphasis factor of 0.9742 was used.

Windowing: To minimize the discontinuity of signal at the beginning and end of each frame, the signal should be tapered to zero or near zero, and hence reduce the mismatch. This can be achieved by windowing each frame to increase the correlation of the Mel Frequency Cepstrum Coefficients (MFCC), Spectral estimates between consecutive frames [16]. To the chosen 12 Mel-Frequency coefficients, and for time 0.005 seconds, a window length of 0.015 is selected using the PRAAT Object software tool.

A. Feature Extraction

Feature extraction involves computing representations of the speech signal that are robust to acoustic variation but sensitive to linguistic content [17]. There are different ways of representing speech. Some feature extraction techniques include: Linear Discriminate Analysis, Linear Predictive Coding, Cepstral Analysis, Mel-frequency scale analysis, principal component analysis and Mel-frequency Cepstrum.

The feature extracted depends on the situation and the kind of speech information to be represented. A waveform, which is the variation of speech amplitude in time, is the most general way to represent a signal. However, a waveform contains too much irrelevant data to use it directly for pattern recognition. The spectrogram offers a better three dimensional representation of speech signals. The patterns represented by spectrogram tend to vary significantly, which makes it unsuitable for pattern recognition.

The Mel Filter Bank is a set of triangular filter banks used to approximate the frequency resolution of the human ear. The filter function depends on three parameters: the lower frequency, the central frequency and higher frequency. On a Mel scale the distances between the lower and the central frequencies, and that of the higher and the central frequencies are equal. The filter functions are:

$$H(f) = 0 \;\forall\; f \le f_l \; and \; f \ge f_h \qquad (7)$$

$$H(f) = \frac{f - f_l}{f_c - f_l} \quad \forall\; f_l \le f \le f_c \qquad (8)$$

$$\mathrm{H(f)} = \frac{\mathrm{f_h - f}}{\mathrm{f_h - f_c}} \quad \forall\; \mathrm{f_c} \le \mathrm{f} \le \mathrm{f_h} \qquad (9)$$

The Mel frequency cepstral coefficients are found from the Discrete Cosine Transform of the Filter bank spectrum by using the formula given by Davis and Mermelstein [18].

$$C_i = \sum_{j=1}^{N} P_j \cos(i\pi/N(j - 0.5)) \qquad (10)$$

Where $P_j$denotes the power in dB in the jth filter and Ndenotes number of samples. 12 Mel frequency coefficients are considered for windowing. Mel-Frequency analysis of speech is based on human perception experiments.

The signal is sampled at 12 kHz. The sampled speech data is applied to the Mel filter and the filtered signal is trained. The number of frames for each utterance is obtained from the frequency coefficients by using PRAAT object software tool.

III. Implementation of the System

The number of frame for each word is used as inputs for the neural network.The vocabulary set is composed of ten Hausa words. Each word is spoken eight times by four speakers (two male and two female).

Thus the database is composed of 320 words. Frames obtained for each utterance of the speaker form Mel-Frequency Cepstral Coefficients are as shown in Table 1.

The database was divided randomly into training, testing and validation sets. A respective division ratio of 70%, 15% and 15% was adopted in splitting the data. The training subset is used for computing the gradient and updating the network weights and biases. The validation subset is used to prevent model overfitting.

The error on the validation set is monitored during the training process. Normally, the validation error decrease during the starting phase of the training as does the training set error.

However, the validation set will typically begin to rise when the network begins to over fit. If the validation error continues to increase for a given number of epochs, the training will be stopped. The testing subset is not used during training, but it is used to compare different models.

A Multilayered Neural Network is used for speech recognition task. The network is made up of eight (8) inputs corresponding to an utterance of each word for every speaker. A hidden layer consisting of ten neurons and an output layer of ten neurons was used as shon in Figure 3. The scaled conjugate gradient back-propagation algorithm was used for training because it converges faster which helps improve generalization.

TABLE I. FRAMES OBTAINED FOR MFCC FEATURE

	Words (utterance)									
	Daya	*Biyu*	*Uku*	*Hudu*	*Biyar*	*Shida*	*Bakwai*	*takwas*	*tara*	*Goma*
Speaker1										
	353	438	379	344	353	353	430	353	404	327
	319	438	464	336	344	361	319	361	310	361
	285	430	344	387	336	344	336	421	361	370
	276	575	404	310	361	319	353	353	344	319
	285	430	438	370	413	319	336	327	344	310
	268	472	387	336	336	336	353	361	370	310
	225	430	370	327	327	344	336	413	404	344
	302	558	396	319	310	387	344	566	302	310
Speaker2										
	413	447	344	379	327	361	387	310	353	336
	370	404	268	353	319	293	276	327	336	293
	455	489	507	327	319	302	379	293	302	319
	361	336	421	413	319	276	276	336	455	302
	361	319	361	293	370	293	379	319	276	285
	498	438	626	310	302	310	396	396	259	310
	353	336	421	566	507	344	327	285	310	310
	319	336	327	327	455	293	327	327	404	319
Speaker3										
	430	370	379	319	379	387	524	379	396	396
	438	361	327	361	379	353	319	319	353	361
	498	353	336	319	387	396	344	549	370	361
	447	447	413	396	336	336	396	438	481	353
	310	379	361	327	310	438	310	464	370	327
	319	387	361	379	353	344	319	379	438	455
	327	319	361	481	387	327	327	455	353	327
	455	600	361	379	353	293	336	464	353	387
Speaker4										
	293	293	379	370	302	293	353	344	319	293
	327	276	285	524	276	233	319	404	507	327
	276	293	319	310	293	302	302	285	293	276
	310	242	319	404	285	464	302	276	336	268
	293	268	293	344	293	225	293	310	507	268
	276	242	285	370	251	251	327	285	276	293
	285	276	285	259	336	293	242	370	268	302
	300	635	276	498	276	327	293	319	302	319

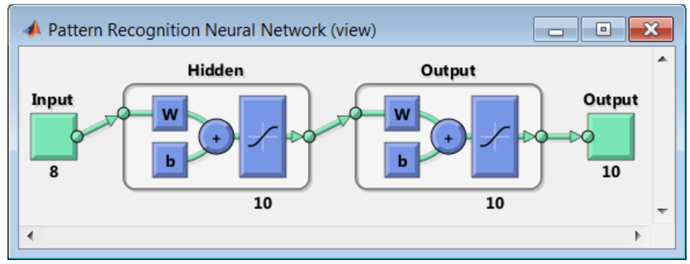


Figure 3. Multi-layered Pattern Recognition Neural Network

IV. SIMULATION RESULTS AND DISCUSSION

The experimental results for the four speakers are presented in Figures 4-7. The performance of the system depends on a number of factors. Adequate pre-processing is a very vital element to ensuring good results in a speech recognition system.

The choice of adequate training parameters helps in arriving at good generalization. From the results, the speech model performance for speaker 2 and speaker 4 was better than that of speakers 1 and 3. The results obtained for the overall data of all the speakers combined is given in Figure 8. When the data was combined, a better generalization was obtained than that of individual speakers.

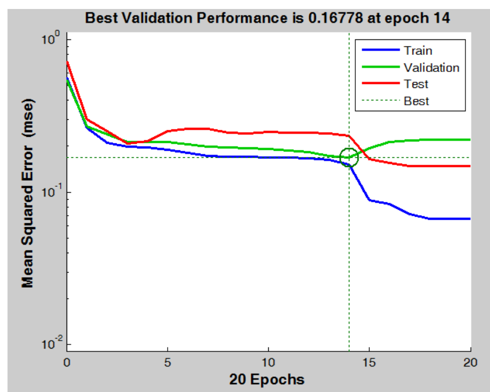


Figure 4. Simulation results for speaker 1 speech recognition

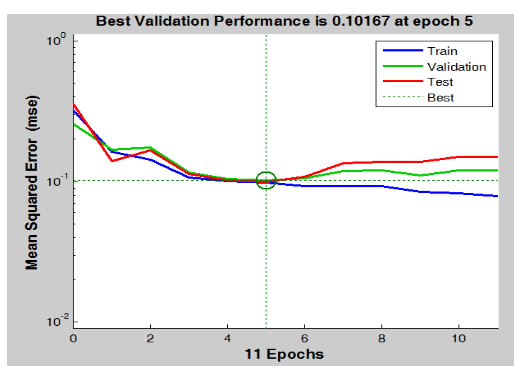


Figure 5. Simulation results for speaker 2 speech recognition

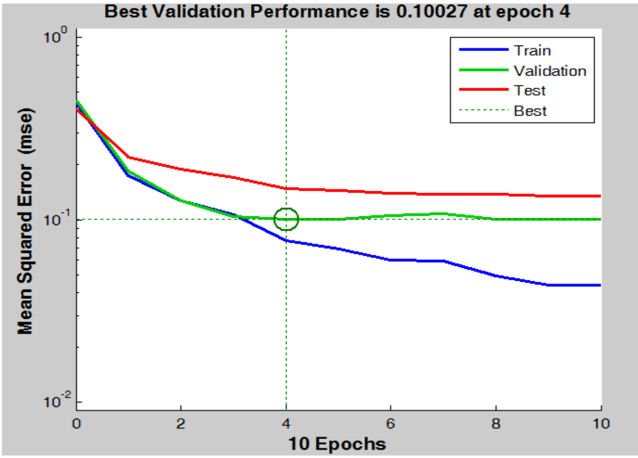


Figure 6. Simulation results for speaker 3 speech recognition

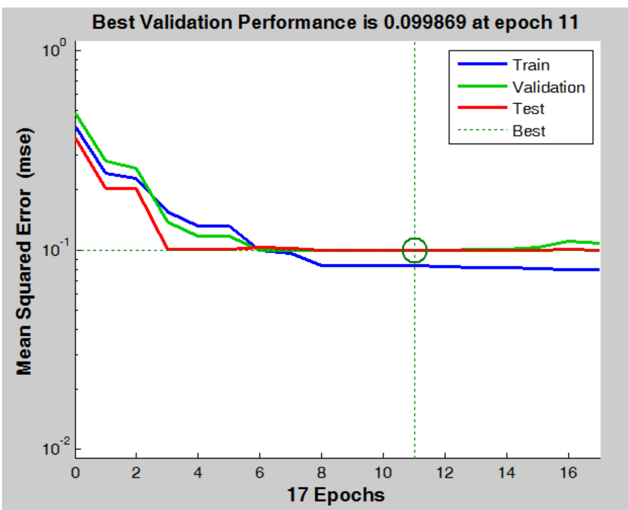


Figure 7. Simulation results for speaker 4 speech recognition

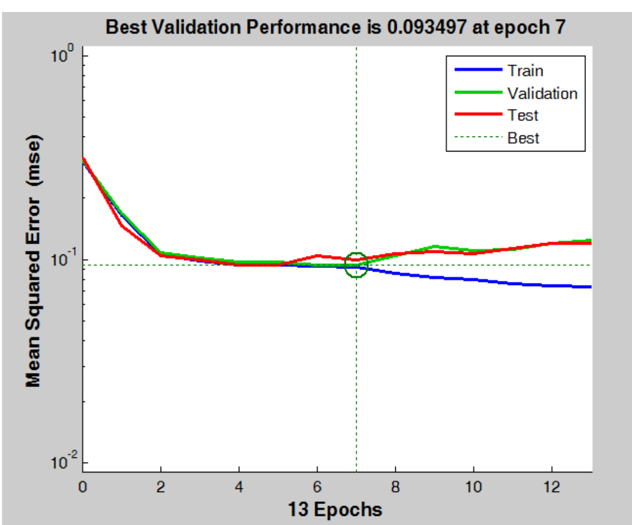


Figure 8. Hausa speech recognition for all the speakers.

V. CONCLUSION

This paper presents speech recognition of Hausa language using pattern recognition artificial neural network. The captured data was first pre-processed then the Mel-Frequency Cepstral Coefficients (MFCC) was extracted. These extracted features were used to train the neural network. The choice of pre-emphasis filter factor was carefully chosen because Hausa is a tone language, in which syllable-based pitch differences add as much to the meaning of words as do consonants and vowels. A better generalization was achieved by adopting early stopping achieved using model validation. The generalization power of the neural network increase when the size of the database increases. This work is a primer for more extensive research in speech processing of African languages (Hausa language in particular).

Future work will focus on increasing the database and expanding the loci of research to other domains of speech processing such as speech enhancement, speech coding, and voice analysis and speech synthesis.

REFERENCES

[1] R.L.K.Venkateswarlu, R. Vasantha Kumari, G.VaniJayaSri, "Speech Recognition by Using Recurrent NeuralNetworks", International Journal of Scientific & Engineering Research Volume 2, Issue 6, June-2011 1ISSN 2229-5518

[2] Lawrence R. Rabiner, "Applications of Speech Recognition in the Area of Telecommunications", retrieved on March 20th, 2012, from http://www.ece.ucsb.edu/Faculty/Rabiner/ece259/Reprints/341_telecom%20applications.pdf

[3] Wikipedia Encyclopaedia, "Speech Recognition", Retrieved March 20th, 2012, from http://en.wikipedia.org/wiki/speech

[4] DanielRios, "Speech Recognition ", *NeuroAI: Algorithms and applications*, Retrieved March 20th , 2012, from www.learnartificialneuralnetworks.com/speechrecognition.htm

[5] Song Yang, MengJooEr and Yang Gao. " A High Performance Neural-Network-Based Speech Recognition System". IEEE trans. PP 1527, 2001

[6] Meysam Mohamad Pour and Fardad Farokhi." An Advanced Method for Speech Recognition". World Academy of Science and Technology, pp. 995-1000, 2009

[7] Laurene Fausett, "Fundamentals of Neural Networks: Architectures, Algorithms and Applications", 1st edit., (pg. 3) Prentice hall, 1993

[8] Anil K. Jain and Jianchang Mao, "Artificial Neural Networks: A Tutorial" *IEEE Computer* (pg.31-44), 1996.

[9] SantoshK.Gaikwad, BhartiW.Gawali and PravinYannawar, "A Review on Speech Recognition Technique", International Journal of Computer Applications (0975 – 8887), Vol., 10, No.3, Nov. 2010

[10] Jont B. Allen. "How do Humans Process and Recognize Speech". *IEEE trans.*, pp. 567-577, Vol.2 No.4, Oct., 1994

[11] AyazKeerio et al, "On Pre-processing of Speech Signals", *World Academy of Science, Engineering and Technology* 47 (pg. 317-323) 2008.

[12] Saha. G., Chakroborty. S., and Senapati. S, "A new SilenceRemoval and End Point Detection Algorithm for Speech andSpeaker Recognition Applications", in Proc. of Eleventh National Conference on Communications (NCC), IITKharagpur,India, January 28-30, 2005, pp. 291-295.

[13] Mitra. A., Chatterjee. B., Mitra. B. K., "Identification ofPrimitive Speech Signals using TMS320C54X DSP Processor",in Proc. of Eleventh National Conference on Communications (NCC), IIT-Kharagpur, India, January 28-30, 2005, pp. 286-290.

[14] Kuldip K. Paliwal, "Effect of Pre-emphasis on Vowel Recognition Performance", *speech communication 3* (pg.101-106), North-Holland, 1984.

[15] J. W. Picone. "Signal modelling technique in speech recognition". IEEE Proc., vol. 81, no.9, pp. 1215-1247, Sep. 1993.

[16] Picton, P, "Neural Networks," Palgrave, NY 2000.

[17] R.L.K. Venkateswarlu and R. Vasanthakumari. "Neuro Based Approach For Speech Recognition By Using Mel-Frequency Cepstral Coefficients" IJSC, Vol. 2, No. 1, January-June 2011, pp. 53– 57

[18] Davis S. B and Mermelstein P., "Comparison of Parametric Representations for Mono-symbolic word Recognition in continuously spoken sentences", *IEEE trans. on Acoustic Speech Signal Process.*, Vol. ASSP-28, No. 4, 1980, pg. 357-366

AUTHORS PROFILE

Matthew K. Luka is a postgraduate student of Covenant University. His research Interest include: Artificial Intelligence, Wireless Communications.

Frank A. Ibikunle (PhD) is a senior lecturer with Covenant University, Ota. He is a seasoned lecturer with extensive industry experience in wireless communications and internet systems. He also has novel publications on various application of artificial intelligence.

Gregory Onwodi is with the National Open University of Nigeria

Improving Performance Analysis Of Routing Efficiency In Wireless Sensor Networks Using Greedy Algorithm

Dr. VUDA SREENIVASARAO[1]
Professor, Dept. of CIT
Defence University College,
Deberezeit. Ethiopia.

CHANDRA SRINIVAS POTLURI[2]
Asst Professor, Dept. of CSE
Aurora Sci Tech& Research Academy
Hyderabad, India.

SREEDEVI KADIYALA[3]
Lecturer, Dept. of CSE
Jigjiga University, Jigjiga.
Ethiopia

Capt. GENETU YOHANNES[4]
HOD, Dept. of CIT
Defence University College,
Deberezeit,Ethiopia.

***Abstract*: The void problem causing the routing failure is the main challenge of the greedy routing in the wireless sensor networks. The current research work still cannot fully deal with the void problem since the excessive control overheads should be consumed so as to guarantee the delivery of packets. In this paper the solution to the void problem is taken up as the issue. This situation exists in the currently existing greedy routing algorithms in wireless sensor networks. The GAR protocol is a new protocol proposed here to guarantee the delivery of packets and excessive consumption of control overheads is resolved.**

Keywords: Greedy routing; void problem; localized algorithm; wireless sensor network.

I. Introduction

Wireless sensor networks consist of a large number of inexpensive sensor nodes distributed in environment uniformly, having limited energy, therefore, in the most cases, nodes communicate with central node via their neighbours. On the other hand, an optimal route must be selected because there are different routes to central node from any other nodes. On the other hand, frequent use of one route results in reduction of energy of sensors located on that route and, ultimately, in sensors destruction. For solving this problem, we can consider a wireless sensor network as a graph in the nodes (hosts) which are the sensors and edges show the links between sensors.

Wireless sensor networks are currently using in different fields of research in both academic and industry. A wireless sensor networks is composed of a large number of nodes that are choose randomly dispersed over some of the interest area. Here all nodes in a wireless networks not communicate directly, so a multi hop routing protocol concept is required. Most of the routing protocols in wireless networks have been designed for networks of some number of hundreds of nodes and do not scale to networks with thousands of nodes.

Sensors combined in to machinery, environment, structures, coupled with the effective delivery of sensed information, could provide best benefits to society. The wireless sensor is networked and scalable, consumes very little power, it is software programmable and smart, it is good capable of speed data acquisition, reliable and accurate over the long term, costs little to purchase and install, and requires no real maintenance. This capability is enabling networks of very low cost sensors that are able to communicate with each other using low power wireless data routing protocols.

A wireless sensor network generally consists of a base station can be communicate with a number of wireless sensors via a radio link. Data is collected at the wireless sensor node, compressed and transmitted to the gateway directly or, if required, uses other wireless sensor nodes to forward data to the gateway. The transmitted data is presented to the system by the base station connection. A wireless sensor network (WSN) consists of sensor nodes (SNs) with wireless communication capabilities for specific sensing tasks.

Due to the limited available resources, efficient design of localized multi hope routing protocols becomes a crucial subject within the WSNs. How to guarantee delivery of packets is considered an important issue for the localized routing algorithms. The well-known greedy forwarding (GF) algorithm is considered a superior scheme with its low routing overheads. However, the void problem, which makes the GF technique unable to find its next closer hop to the destination, will cause the GF algorithm failing to guarantee the delivery of data packets.

Next generation environments represent the next evolutionary development step in utilities, building, home, industrial, shipboard, and transportation systems automation. Sensor network data comes from multiple sensors of different modalities in distributed in different locations.

Routing algorithms are proposed to either resolve or reduce the void problem, which can be classified into non-graph-based and graph-based schemes. In the non-graph- based algorithms, the intuitive schemes as proposed in construct a two-hop neighbor table for implementing the GF algorithm. The network flooding mechanism is adopted within the GRA and PSR schemes while the void problem occurs. There also exist routing protocols that adopt the backtracking method at the occurrence of the network holes (such as GEDIR, , DFS , and SPEED). The routing schemes as proposed by ARP and LFR memorize the routing path after the void problem takes place. Moreover, other routing protocols (such as PAGER NEAR and YAGR) propagate and update the information of the observed void node in order to reduce the probability of encountering the void problem. By exploiting these routing algorithms, however, the void problem can only be either 1) partially alleviated or 2) resolved with considerable routing overheads and significant converging time.

Several routing algorithms are proposed to either resolve or reduce the void problem, which can be classified into non-graph-based and graph-based schemes. In the non-graph- based algorithms [3], [4], [5], [6], [7], [8], [9], [10], [11], the intuitive schemes as proposed in [3] construct a two-hop neighbor table for implementing the GF algorithm. The network flooding mechanism is adopted within the GRA [4] and PSR schemes while the void problem occurs. There also exist routing protocols that adopt the backtracking method at the occurrence of the network holes (such as GEDIR, [3], DFS [5], and SPEED [6]). The routing schemes as proposed by ARP and LFR memorize the routing path after the void problem takes place. Moreover, other routing protocols (such as PAGER [7], NEAR [8], DUA [9], INF [10], and YAGR [11]) propagate and update the information of the observed void node in order to reduce the probability of encountering the void problem. By exploiting these routing algorithms, however, the void problem can only be either 1) partially alleviated or 2) resolved with considerable routing overheads and significant converging time. The Gabriel graph (GG) and the relative neighborhood graph (RNG) are the two commonly used localized planarization techniques that abandon some communication links from the UDG for achieving the planar graph. Nevertheless, the usage of the GG and RNG a graph has significant pitfalls due to the removal of critical communication links, leading to longer routing paths to the destination.

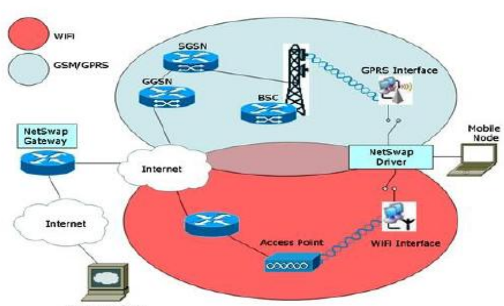


Figure 1. GPRS and Wi-Fi Interface

In figure 1 the working mechanism for GPSR and Wi-Fi interface has been shown. The representative planar graph-based GPSR scheme cannot forward the packets from NV to NA directly since both the GG and the RNG planarization rules abandon the communication link from NV to NA. Considering the GG planarization rule for example, the communication link from NV.

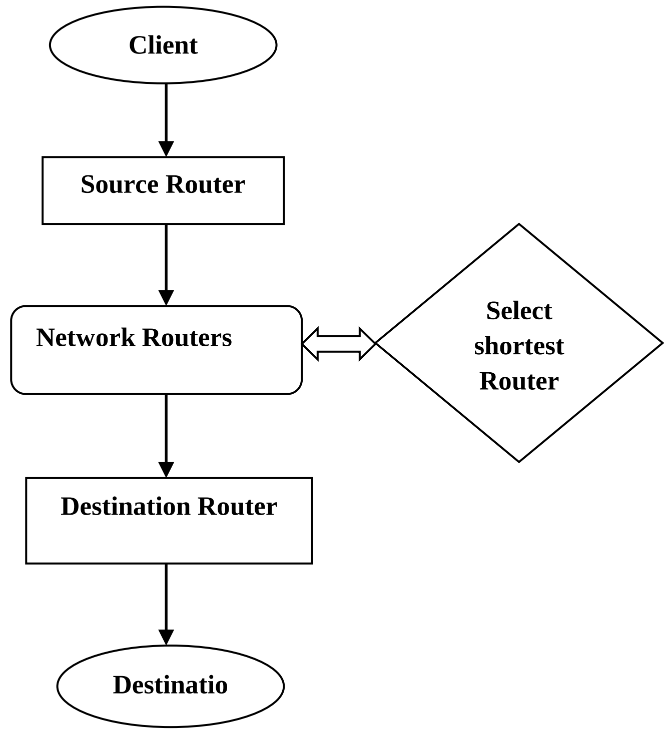


Figure 2. process of routers.

II. RELATED WORK

A. Routing Algorithm:

Let N is the number of nodes in the ad hoc network under consideration, Let A be the source node and Z be the destination node, L[X] be the neighbor list of some arbitrary node X.

E num power status = {very low, low, medium, high, very high};

hop count= 0;

current node=A

path= ‘A’;

do

{

If Z Î L [current node] /* If Z is the immediate neighbor */

{ /*Route found*/

Path= Path + ’Z’;

Current node= Z; /* destination reached*/

hop_ count ++;

route=path;

Return route and hop_count;

Exit (0); /*Come out of the program*/

}

else /*Z is not the immediate neighbour*/

{

T=0 ; /* Number of neighbors whose Z is a immediateneighbor */

neighb_list=L[Current node]

while (neighb_list != empty)

{ Remove first element of neighb_list, say X.

Send unicast message to X to enquire about number of its neighbors N, Power Status

P and Ans to know if Z is a immediate neighbor of Current node;

Read and Record (node_id, P, N, Ans);

If (Ans=”Y”) T++;

};

If (T==0) K=1; /* no one has information about the destination node Z*/

if (T==1)K=2; /* only one node has information about the destination node Z*/

if (T >=1)K=3; /* more then one node has information about the dhestination node Z*/

Switch (K)

{

Case 1: /* there is no neighbor node with Z as its immediate neighbor */

For each neighbor of current node

Check (node id, P, N, ANS)

If (P= medium || P= High || P= Very high)

{

Add node to the probable list;

For every node i belonging to probable list, find

suitability index Su[i].

Let X be node with highest suitability index.

Su=ord(power status) – N/C;

Choose X as the next node in route.

hop_count ++;

current node = X;

path =path + ‘X’;

}

Break;

Case 2: /* there is only one neighbor with Z as the immediate neighbor */

find record with (Ans= “Y”);

Let this node be X;

path=path +’X-Z’;

Route = path;

hop_count = hop_count + 2;

Return route and hop_count;

Exit (0);

Break;

Case 3: /* there is more than one neighbours with Z as the immediate neighbour” */

find records with (Ans= “Y”);

Let these node be X1, X2, ……….Xk;

For each node Xi to Xk find suitability index Su.

Su=ord(power status) – N/C;

Choose the node with highest suitability index.

Let the node be Y.

path=path +’Y-Z’;

route = path;

hop count=hop_count +2;

Return route and hop_count;;

Exit (0);

Break;

}

III. RESULTS

Giving the Destination address in the source form:

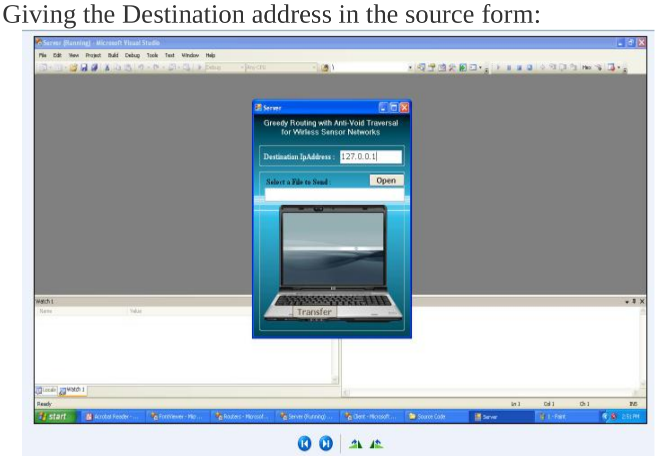


Figure 3. specify the destination IP address.

Server with selected file and destination address

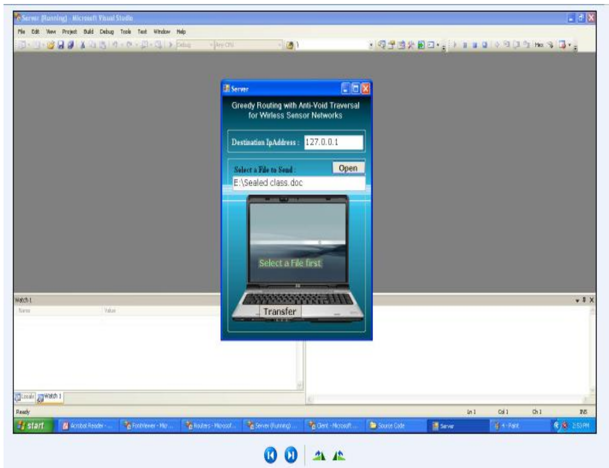


Figure 4. Destination IP address and a file to send.

Router Form to select the destination node

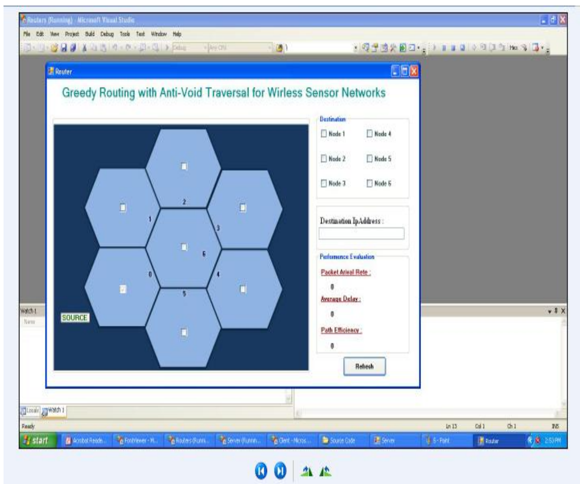


Figure 5. seven nodes selecting '0' as Source node.

Selecting destination node

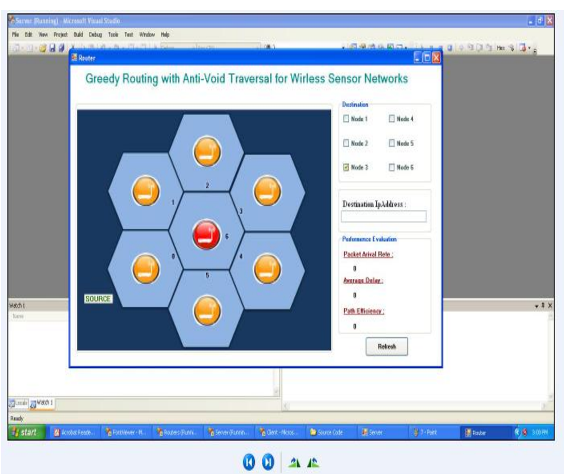


Figure 6. the node 3 as destination node.

Screen showing the data transfer through routing path

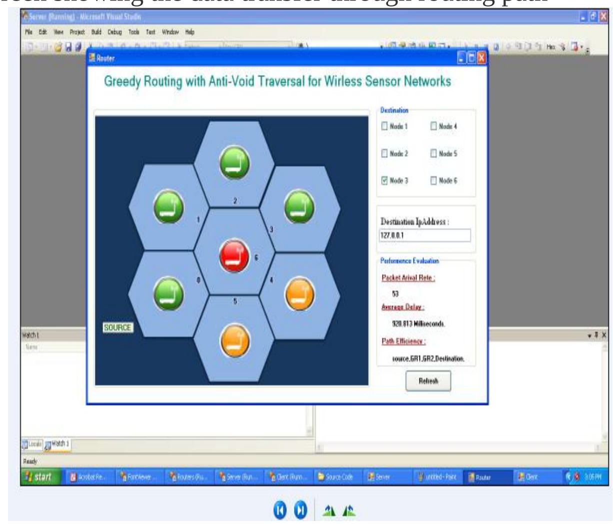


Figure 7. the transmission of file from one node to another node with green color. i.e. node 3.

The file received by the destination

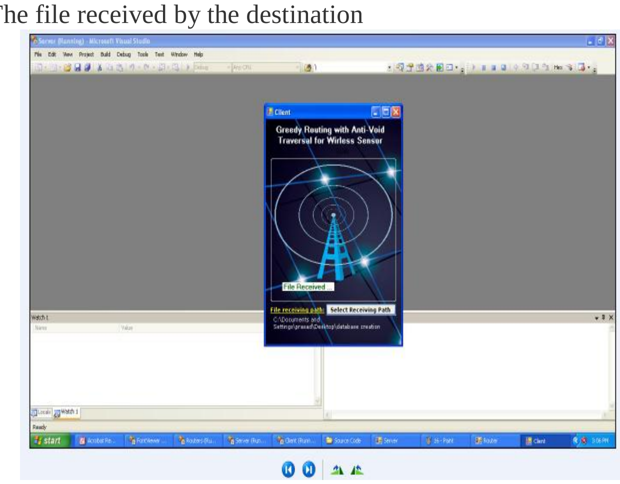


Figure 8. The successful transfer of file can be seen by observing "File Received" tab.

IV. CONCLUSION

In this paper, a UDG based GAR protocol is proposed to solve the some void problems by using conventional GF algorithm. The RUT scheme is adopted with in the GAR protocol solve by using the boundary finding problem, results can be delivery of data packets under UDG networks. The correctness of the RUT scheme and the GAR algorithm is properly proven.

The performance of GAR Protocol is evaluated and compared with existing localized routing algorithms via simulations. The simulation study shows that the proposed GAR algorithm can guarantee the delivery of data packets with reduced communication overhead under different network scenarios.

REFERENCES

[1] An efficient routing method for Ad hoc Networks by Mohinder Gupta.

[2] Realization of Greedy Anti-Void Routing Protocol for Wireless Sensor Networks by Wen-Jiunn Liu and Kai-Ten Feng.

[3] B.Chella prabha, B.Anuradha,V.Kanimozhi, M.Dhivya "Enhanced Greedy Routing with Anti-Void Traversal for Wireless Sensor Networks" International journal of computing Applications in Engineering Sciences, Vol 1, Issue 1,March2011,pp- 18-22.

[4] Next century challenges: Scalable coordination in sensor Networks—D.Estrin, R.Govindan, J.Heidemann, and S.Kumar, ACM-PP-263-270,Aug 1999.

[5] Data Communications and Networking, by Behrouz A Forouzan.

[6] Computer Networking: A Top-Down Approach, by James F. Kurose.

[7] Operating System Concepts, by Abraham Silberschatz.

[8] A Novel Routing Strategy For Ad Hoc Networks With Selfish Nodes by Shailender Gupta Chander Kumar,IJEECS.

[9] B.Chella prabha, B.Anuradha, V.Kanimozhi, M.Dhivya, "Enhanced Greedy Routing with Anti-Void Traversal for Wireless Sensor Networks", Journal of Computer Applications in Engineering Sciences (IJCAES) VOL I, ISSUE I, MARCH 2011.

AUTHORS BIBLIOGRAPHY

Dr. VUDASREENIVASARAO received his M.Tech degree in computer science and engineering from Sathyabama University from 2007.He received PhD degree in computer science and engineering from Singhania University from 2010. Currently working as Professor in the department of Computer and Information Technology, Defence University College, Deberezeit, Ethiopia. His main research interests are Data mining, Fuzzy logic and Network Security. He has got 11 years of teaching experience. He has published 26 research papers in various international journals and one Springer international conference paper. He has 38 Editorial Board / Reviewers memberships in various international journals. He is a life member of various professional societies like IEEE, ACM, MAIRCC, MCSI, SMIACSIT, MIAENG, MCSTA and MISTE.

CHANDRA SRINIVAS POTLURI received his M.Tech degree in Computer science from JNT University. Currently working as Assistant Professor in Aurora Scientific Technological and Research Academy, Hyderabad, India. His main research interests are Data Mining & Mobile Computing. He has got 5 years of teaching experience & 4 years of Industrial experience worked in Satyam, Hyderabad, India.

SREEDEVI KADIYALA received her M.Tech degree in Computer Science and Engineering from JNT University and M.Phil degree in Computer Science and Engineering from Bharathiar University. Currently working as a Lecturer in the department of CSE at Jigjiga University , Jigjiga, Ethiopia. Her main research interests are Data Mining and Data Ware housing, Network Security. She has got 9 years of teaching experience. She has published two papers in IJAST and IJCIS.

Capt. GENETU YOHANNES received his M.Tech degree in Computer Science & Engineering from Defence Institute of Advanced Technology, Pune, in Year 2011. Currently working as HOD in the Department of Computer and Information Technology, Defence University College, Deberezeit, Ethiopia. He has got 7 years of teaching experience. His main research interests are Data Mining and Wireless Networks.

Automated Marble Plate Classification System Based On Different Neural Network Input Training Sets and PLC Implementation

Irina Topalova
Department of Automation of Discrete Production Engineering, Faculty of Mechanical Engineering
Technical University of Sofia, Bulgaria
Sofia, Bulgaria

***Abstract*— The process of sorting marble plates according to their surface texture is an important task in the automated marble plate production. Nowadays some inspection systems in marble industry that automate the classification tasks are too expensive and are compatible only with specific technological equipment in the plant. In this paper a new approach to the design of an Automated Marble Plate Classification System (AMPCS),based on different neural network input training sets is proposed, aiming at high classification accuracy using simple processing and application of only standard devices. It is based on training a classification MLP neural network with three different input training sets: extracted texture histograms, Discrete Cosine and Wavelet Transform over the histograms. The algorithm is implemented in a PLC for real-time operation. The performance of the system is assessed with each one of the input training sets. The experimental test results regarding classification accuracy and quick operation are represented and discussed.**

Keywords- Automated classification; DCT; DWT; Neural network; PLC

I. INTRODUCTION

Many modern systems for automated vision control in stone production industry are developed for inspection and classification of the surface texture of marble slabs and plates after cutting. . Many of them have high market price and are compatible only with the proprietarily company production equipment. The existing software products for texture recognition are not intended for implementation in Programmable Logic Controllers (PLC) widely used for control of technological processes. A specific task in marble plate production is the sorting of tiles with identical surface textures. The high accuracy classification in this case is difficult because the parametrical descriptions of the textures are high correlated. In addition the textures have to be recognized during movement (transportation on a conveyer belt) and with some changes in the working area illumination corresponding to the specifics of the production process. That is the reason to develop effective methods and algorithms aiming at high recognition accuracy for different kinds of similar textures when evaluating them in real production environment conditions. On the other hand optimal (considering accuracy, quick-operating and cost) software and hardware system solutions have to be sought – suitable for implementation in PLCs, because these devices are preferable and widely used in up-to-date automated production systems. In this paper, a new approach for design of an Automated Marble Plate Classification System (AMPCS) based on different neural network input training sets is proposed, aiming at high classification accuracy using simple processing and application of only standard devices and communication protocols. It is based on training a classification Multi-Layer-Perceptron Neural Network (MLP NN) with three different input training sets: extracted texture histograms, Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT) over the histograms. The algorithm is implemented in a PLC for real-time operation. The system performance is assessed with each one of the input training sets. The modeling technique performance is assessed with different training and test sets. The experimental test results regarding classification accuracy and quick operation are represented and discussed.

II. EXISTING METHODS FOR TEXTURE CLASSIFICATION

There are many contemporary methods and algorithms for texture analysis and classification. The task of classification is to associate an undefined texture with a preliminary known class. Tuceryan and Jain [1] divide the classification methods in four basic categories: statistical, geometrical, model based and filter based methods. Statistical first order data as the mean value and the dispersion are not appropriate for description of the pixel cross correlation. The second order data given by the differential statistics and the auto-correlation functions are more appropriate for texture classification [2]. The geometrical method is based on extraction of primitives, detects all the particles in the image applying operators as Sobel, Prewitt, Roberts etc. and after morphological analysis builds a detailed report on the geometrical parameters of each particle [3]. The application of statistical and filter based approaches [4] is due to the fact, that there are many software products, offering an opportunity for easy simulation and test of the results. Context analysis based on "direct local neighborhood" is also lately applied and it is a variant of a method based only on filtration. The model based analysis aims construction of a base model of the image. On the base of the constructed model a description and synthesis of the texture is feasible. The main disadvantage of the existing methods is the very high computational complexity and relatively low recognition accuracy. All of these methods are not efficient when the textures are very similar and have overlapping parametrical descriptions. When investigating recognition and classification of a preliminary known texture classes, more suitable is to apply an adaptive recognition method and a supervised learning scheme [5], since this method gives the more accurate results. In this instance the best variant is to choose neural networks (NN) because of the good NN capabilities to adapt to changes in the input vector, to set precisely the boundaries between the classes therefore offering high recognition accuracy and fast computations in the recognition phase [6]. After choosing a NN for combination between a supervised learning scheme and utilization of the histogram parameters, the more important thing is the right choice of the input NN data. This choice is influential for the right and fast convergence of the NN, for the number of parameters in the input vector and the whole NN topology. The initial choice of variables is guided by intuition. Next the number of NN input parameters has to be optimized, developing a suitable method for their reduction aiming to preserve only the informative parameters without loose of any information useful for accurate class determination [7].

III. Structure and Functioning Modes of AMPCS

The AMPCS is based on training a classification MLP neural network with three different input training sets: extracted texture histograms, Discrete Cosine Transform (DCT) over the extracted histogram and Discrete Wavelet Transform (DWT) coefficients over the histogram. Many histograms of each investigated grey scale texture image, corresponding to different movement speeds of the plate on the conveyer belt and different illuminations are extracted, sampled and stored for further processing. Next DCT and DWT over the extracted histograms are also calculated and added to the stored data. The histogram values, obtained DCT and DWT coefficients are used consecutively for training a MLP neural network, changing the NN topology and the NN training parameters in order to find the best case in classification phase. The functioning structure of the designed AMPCS is shown in Figure 1. After acquisition of the marble

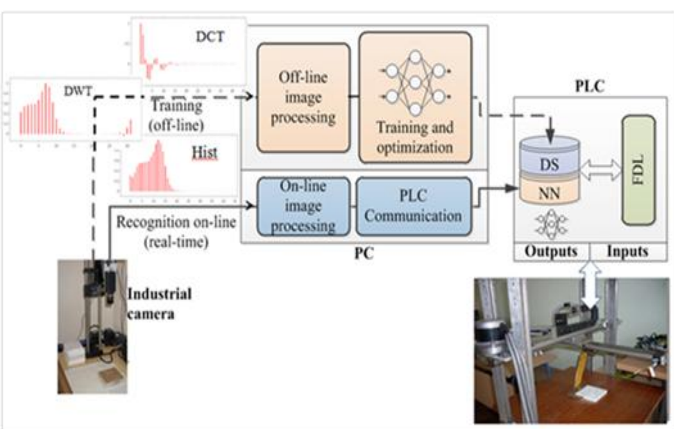


Figure 1. Structure of the designed AMPCS.

Figure 2. Exemplars of the tested marble textures.

plate images through a high resolution CCD industrial camera, they are transferred using Ethernet communication to a conventional PC [8]. The system is working in two modes – off-line or training and on-line or classification mode.

IV. Calculation of the Input Training Sets

A. *Histograms*

The image histogram is the most important statistical characteristic. It contains information about the image contrast and brightness. Textures can also be described through their histograms. The mathematical definition of a histogram is as follows:

$$n = \sum_{i=0}^{N-1} m_i \qquad (1)$$

where for an image with a number of k grey levels, m_i is each of $N = 2^k$ columns, containing the number of pixels with intensity value of i. Since the histogram values are used as NN input training vector, the values have to be fitted to the range of NN's transfer function argument - i.e. between 1 and -1. Therefore, each histogram value is divided by the maximum histogram value. To reduce NN's input layer neurons, the normalized histogram is reduced through sampling each sixth histogram value according to the method given in [9].

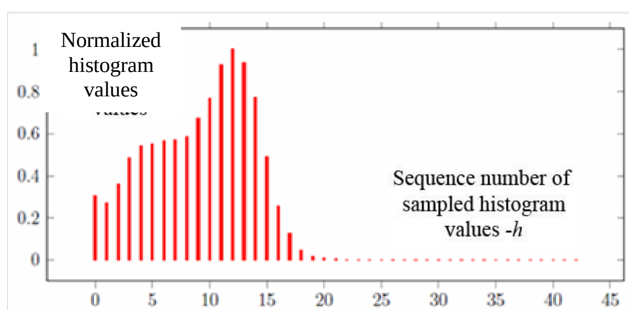


Figure 3. Normalized and sampled histogram of marble texture *Balmoral Red*

Figure 3 shows the normalized and sampled histogram for class *Balmoral red*.

B. Discrete Cosine Transform

Discrete Cosine Transform (DCT) represents the input signal as a sum of cosines with increasing frequencies. DCT over a signal length of N is equivalent to a Fourier transform of an even symmetrical time signal length of 2(N-1) with imaginary part equal to zero. Thus, the calculations are simplified which is a precondition for fast computations. DCT calculation over the sampled histogram h[t] is given in (2). The obtained DCT(w) is normalized with regard to the absolute maximum coefficient |DCTmax(w)|, Fig.4.

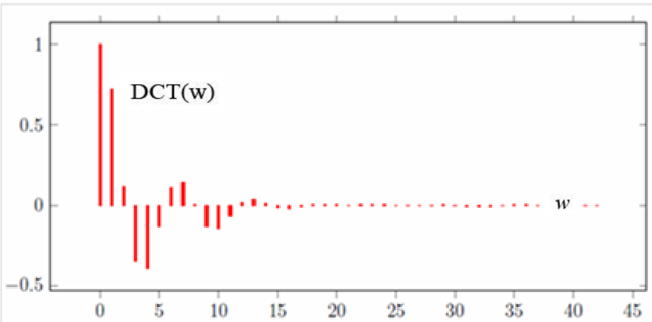


Figure 4. DCT over the histogram of the texture Balmoral Red given in Fig.3

$$DCT(w) = C_w \sum_{k=0}^{N} h[t] . cos\left(\frac{(2k+1)\pi w}{2n}\right) \quad (2)$$

$w = \{0, \dots, N-1\}$ and

$$C_w = \begin{cases} \frac{1}{\sqrt{N}}, for\ w = 0 \\ \sqrt{\frac{2}{N}}, for\ w > 0 \end{cases}$$

C. Discrete Wavelet Transform

Wavelet analysis consists of breaking up of signal into shifted and scaled versions of the original Wavelet ψ(t). Wavelets are defined by the original wavelet function ψ(t) (i.e. the mother wavelet) and scaling function Ø(t) (also called father wavelet) in the time domain amplitudes. The most popular and used mother wavelet functions are

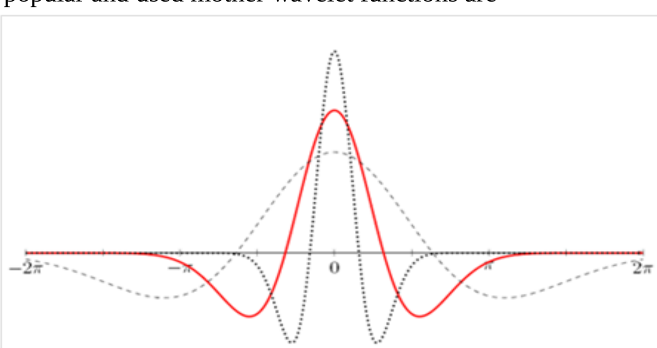


Figure 5. Three Mexican hat wavelets $\psi_{a,b}(t) = (\frac{1}{\sqrt{a}})\ \psi\ ((t-b)/a)$ at three dilations, a = 0.5, 1.0, 2.0 and all located at b = 0.

Gaussian, Morlet, b-spline, Shannon, Mexican hat, Cohen-Daubechies-Feauveau (CDF), [10,11] etc. Here we use Mexican hat wavelet [12], Fig. 5. In this case the normalised wavelet function is often written more compactly as:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad (3)$$

This dilation and contraction of the wavelet is governed by the dilation parameter a which, for the Mexican hat wavelet, is (helpfully) the distance between the center of the wavelet and its crossing of the time axis. The movement of the wavelet along the time axis is governed by the translation parameter b. A natural way to sample the parameters a and b is to use a logarithmic discretization of the a scale and link this, in turn to the size of steps taken between b locations. To link b to a we move in discrete steps to each location b which are proportional to the a scale [12]. This kind of discretization of the wavelet has the form:

$$\psi_{m,n}(t) = \frac{1}{\sqrt{a_0^m}}\psi\left(\frac{t-nb_0a_0^m}{a_0^m}\right) \quad (4)$$

where the integers m and n control the wavelet dilation and translation respectively; a0 is a specified dilation step parameter set at a value greater than 1, and b0 is the location parameter which must be greater than zero[12]. Substituting a0 = 2 and b0 = 1 into equation (4) we see that the dyadic grid wavelet can be written in a very compact form as:

$$\psi_{m,n}(t) = 2^{-m/2}\psi\left(2^{-m}t - n\right) \quad (5)$$

Using the dyadic grid wavelet, the discrete wavelet transform (DWT) can be written using wavelet (detail) T m,n coefficients as:

$$T_{m,n} = \int_{-\infty}^{\infty} h(t)\ \psi_{m,n}(t)dt \quad (6)$$

The scaling function can be convolved with the signal to produce approximation coefficients as follows:

$$S_{m,n} = \int_{-\infty}^{\infty} h(t)\ \emptyset_{m,n}(t)dt \quad (7)$$

In our case the discrete input signal S0,n = h[t] is the gray scale image histogram of finite length N=256 and M=8, which is an integer power of 2: N = 2M . Thus, the range of scales we can investigate is 0 < m < M. We can represent the histogram signal h[t] using a combined series expansion using both the approximation coefficients Sm,n and the wavelet (detail) Tm,n coefficients as follows [12]:

$$h[t] = \sum_{n=-\infty}^{\infty} S_{m_{0},n}\ \emptyset_{m_{0},n}(t) + \sum_{m=-\infty}^{m_0}\sum_{n=-\infty}^{\infty} T_{m,n}\ \psi_{m,n}(t) \quad (8)$$

D. Applied Wavelet multi-resolution calculation algorithm

Once we have our discrete input signal S0,n, we can compute S m,n and T m,n. This can be done for scale indices m > 0, up to a maximum scale determined by the length of the input signal. To do this, we use an iterative procedure as follows.

First we compute S1,n and T1,n from the input coefficients S0,n, i.e.

$$S_{1,n} = \frac{1}{\sqrt{2}} \sum_k C_k \, S_{o,2n+k} \quad (9)$$

$$T_{1,n} = \frac{1}{\sqrt{2}} \sum_k b_k \, S_{o,2n+k} \quad (10)$$

Here the scaling equation (or dilation equation) describes the scaling function Ø(t) in terms of contracted and shifted versions of itself as follows:

$$\emptyset(t) = \sum_k C_k \, \emptyset(2t-k) \quad (11)$$

where Ø (2t - k) is a contracted version of Ø (t) shifted along the time axis by an integer step k and factored by an associated scaling coefficient, ck.

Integrating both sides of the above equation, we can show that the scaling coefficient must satisfy the following constraint:

$$\sum_k C_k = 2 \quad (12)$$

In addition, in order to create an orthogonal system we require that

$$\sum_k C_k C_{k+2k'} = \begin{cases} 2 \; if \; k' = 0 \\ 0 \;\; otherwise \end{cases} \quad (13)$$

The reconfigured coefficients used for the wavelet function are written more compactly as:

$$b_k = (-1)^k C_{N_k - 1 - k} \quad (14)$$

Next, applying the Haar filtering algorithm [10,11,12,13,14] we can find S2,n and T2,n from the approximation coefficients S 1,n and so on, up to those coefficients at scale index M, where only one approximation and one detail coefficient is computed: SM,0 and TM,0. At scale index M we have performed a full decomposition of the finite-length, discrete input signal. We are left with an array of coefficients: a single approximation coefficient value, SM,0, plus the detail coefficients, Tm,n, corresponding to discrete waves at scale a = 2m and location b = 2mn. The finite time series is of length N = 2M.

This gives the ranges of m and n for the detail coefficients as respectively 1 < m < M and 0 < n < 2 M-m - 1[12]. At the smallest wavelet scale, index m = 1, 2M=21 = N/2 coefficients are computed, at the next scale 2M=22 = N/4 are computed and so on, at larger and larger scales, until the largest scale (m = M) where only one (= 2M=2M) coefficient is computed. The total number of detail coefficients for a discrete time series of length N = 2M is then, 1 + 2 + 4 + 8 + … + 2 M-1 = N - 1. In addition to the detail coefficients, the single approximation coefficient SM,0 remains.

As the histogram S0,n = h[t] input signal vector contains N=256 components, the decomposition of the signal in our case corresponds to a number of 32 approximation coefficients Sm,n when stopped at subsequent level m=3.

The obtained S3,n approximation coefficients for the texture Balmoral Red are illustrated in Fig.6.

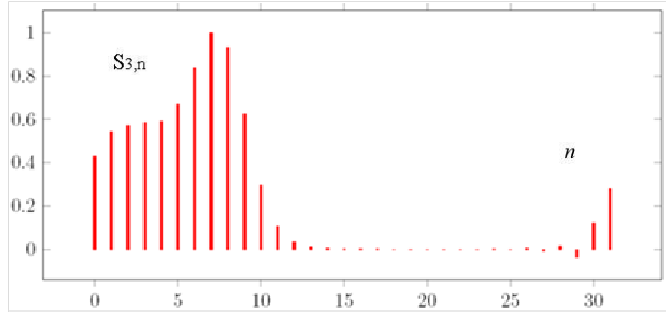


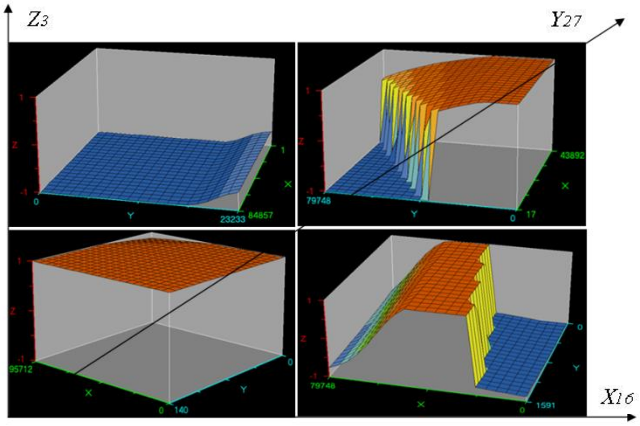
Figure 6. Three-stage (m=3) approximation DWT coefficients $S_{3,n}$ for texture *Balmoral Red*

V. EXPERIMENTS AND RESULTS

The experiments were carried out with the ten classes of marble textures shown in Figure 2. The image acquisition was accomplished using a CCD camera Basler scA1000-20gm with exposure time of 120µs to 29.5 ms and resolution 1034 x 779 pixels. The value of 9Pix Motion Blur (MB), corresponding to an image resolution of 300 dpi or 118 Pix/cm, 25 m/min linear velocity of the conveyer belt, and 1/500 sec camera exposure time was applied to the images. Also 10% brightness variations were added to some of the texture samples. The MLP NN structure was trained in off-line mode of AMPCS, consecutively with sampled histogram values (each 8th value was sampled) of the ten textures, with their DCTs, sampled on the same way, and with DWTs approximating coefficients S3,n. Thus, all MLP NN input vectors are composed of 32 components. The right choice of MLP NN topology (number of layers, number of neurons in the layers) and NN parameters (activation function, MSE - mean square error, input data) is decisive for optimal training the NN [15]. Initially we use three - layered 32-50-10 MLP NN topology, where 10 is the number of output layer neurons, corresponding to the number of tested texture classes. With the training of MLP NN we want to obtain "softer" transitions or larger regions, where the output stays near to "1" or "-1" (using tangent hyperbolic as activation function).

With 3D graphic presentation it is possible to view the input/output characteristic of two inputs and one output of the neural network in spatial representation, Fig.7. After fixing the number of input values, the number of neurons in the hidden layer and applying the method given in [16], the training is repeated with reduced MSE. These reduction aims at establishing the boundaries between the classes more precisely. After that the 3D-surfaces are observed, repeating the step and reducing MSE again until the 3D-surfaces of almost all outputs have areas, where the output have regions near to "1" or "-1" as on Figure 5/b.

At this point the train phase is completed. The number of hidden neurons in MLP NN best final topology varies between 25 and 50 with MSE between 0.01 and 0.31. It was trained after 32000 iterations on the average for all 10 classes with 300 training samples per class for each of the calculated three input sets. Finally the trained NN for the three investigated input training sets is downloaded in the Programmable Logic Controller - SIMATIC S7-317 DP [17] for on-line mode operation.



a/ b/

Figure 7. 3D surface representation of MLP NN output 3 (Z_3) depending on inputs № 16 (X_{16}) and № 27 (Y_{27}) of the corresponding DCTs in training phase – a/ bad training, b/ good training

These reduction aims at establishing the boundaries between the classes more precisely. After that the 3D-surfaces are observed, repeating the step and reducing MSE again until the 3D-surfaces of almost all outputs have areas, where the output have regions near to "1" or "-1" as on Fig.7/b. At this point the train phase is completed. The number of hidden neurons in MLP NN best final topology varies between 25 and 50 with MSE between 0.01 and 0.31. It was trained after 32000 iterations on the average for all 10 classes with 300 training samples per class for each of the calculated three input sets. Finally the trained NN for the three investigated input training sets is downloaded in the Programmable Logic Controller - SIMATIC S7-317 DP [18] for on-line mode operation.

VI. CLASSIFICATION PERFORMANCE

The algorithms for a particular neuro - application are calculated in the processor of the SIMATIC S7-317 DP and more precisely, after an unconditional call by a user program or cyclically at time-controlled intervals. MLP NN in the PLC is represented with a pair of a Function Block (reserved for NN is FB101) and an associated instance Data Block (DB101). The results of the trained NN - i.e. the obtained matrix of adapted weight coefficients are stored in the DB101. When calling FB101, NN inputs are provided with the calculated histogram, DCT or DWT values over the current accepted texture. Next a FDL-Final Decision Logic procedure for automated evaluation and finding the MLP output with maximal value (corresponding to the recognized texture class) was designed, using ladder diagram method in the PLC.

At that point we use FB1 and instance data block DB1. FB1 receives as inputs the output values of MLP NN (Fig.9). FB1 contains the developed algorithm for finding the maximum value of NN outputs. Fig.10 shows a single compare logic element (CMP), which compares one of the NN outputs with the current memorized maximum value MD201. MLP outputs together with FDL results were applied to the physical PLC outputs and to other FBs in the PLC for direct control of the sorting mechanical devices.

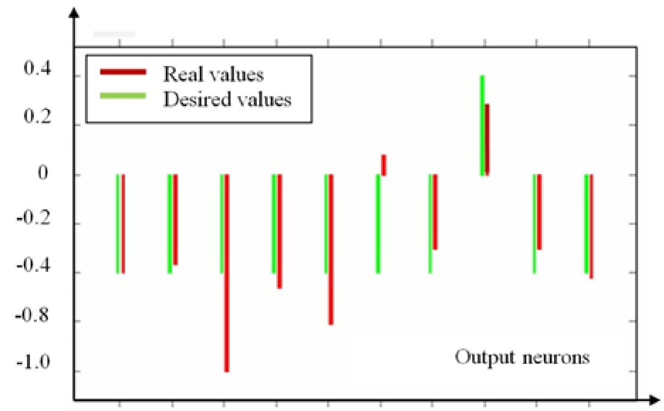


Figure 8. Output neuron values for classification of a sample of class8 – *Santa Cecilia Dourado*

Each class was tested with 40 samples of textures not included in the training set. The tests were provided with the defined three different input sets. Fig.8 shows the output neuron values for classification of a sample of class8 – Cecilia Dourado. The obtained classification accuracy, calculated as the overall number of correct classifications, divided by the number of instances in the data testset given on the average for the ten classes is presented in Table1.

The maximum value of the time needed for image acquisition, histogram/DCT/DWT calculation, giving the values to the PLC NN Function Block, classification through the trained NN and giving the result to the physical PLC outputs is calculated for different number of hidden neurons with different MSE. Fig.11 shows MLP NN outputs for correct classification of 40 test samples of class 4 - Gris Mondaris.

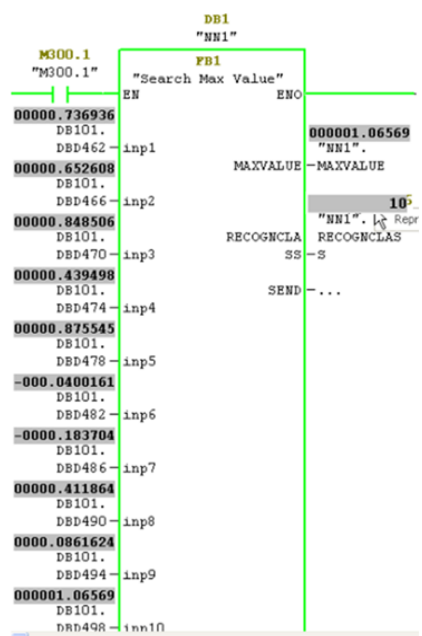


Figure 9. FDL components: Function Block of FDL

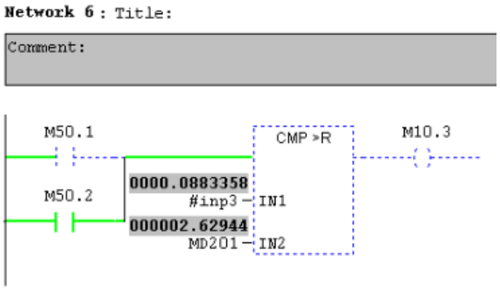


Figure 10. Compare Logic Element (CMP) in FB1

The results given in Table 1 show that the highest classification accuracy is obtained when DWT coefficients are used in AMPCS - i.e. 100% (calculated as the number of correct classified/common number of tested samples) with Mexican hat is used as mother wavelet and only 32 approximated DWT coefficients are used. When DCT coefficients are used as input set, an accuracy of 87.5 - 95% is achieved by the same length of the input vector. When a histogram with 32 sampled values is used, the accuracy is between 80% - 85%.

The best accuracy is obtained when NN has 45 hidden neurons applying the histogram method, 50 neurons – for the DCT method, and only 25 neurons for the DWT method. It is a good precondition for further minimizing the number of hidden neurons without affecting the accuracy and at the same time for reducing the real-time operation. Another good research continuation would be the testing of the accuracy and computational time applying DWT approximation coefficients with m>3 - i.e. reducing the input vector components to n=16 and even n=8. Because the faster real-time operation is achieved for the histogram method and the slowest – for the DWT method, this would compensate the relatively complicated calculations.

Finally, it is recommended to investigate the cross-correlation between texture histograms before choosing one of the proposed in this paper input sets [17]. When the cross-correlation coefficient is greater than 0.6 - 0.7, it is recommended to apply DWT method because Wavelets are building blocks that can quickly de-correlate data. In the case of slight correlated histograms (i.e. not very similar textures), it is better to apply histogram or DCT method because according to Table1 they ensure faster computation and sufficient accuracy.

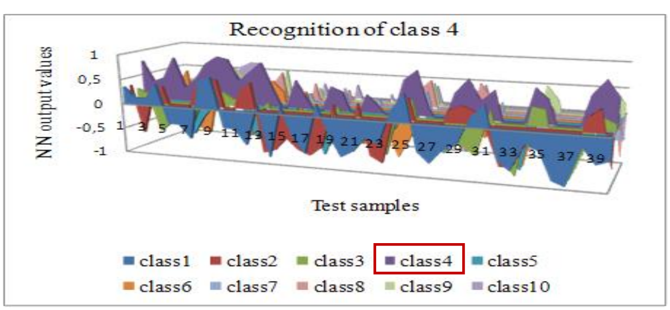


Figure 11. Classification of 40 test samples of class 4 Gris Mondaris

Method	Number of hidden layer neurons	Mean Square Error [%]	Classification accuracy	Real-time execution [ms]
Histogram	50	0.16	34/40	578
	45	0.18	34/40	553
	40	0.21	33/40	538
	35	0.24	33/40	519
	30	0.27	33/40	487
	25	0.31	32/40	461
DCT	50	0.01	38/40	638
	45	0.01	37/40	626
	40	0.01	37/40	611
	35	0.01	37/40	595
	30	0.01	35/40	578
	25	0.01	35/40	561
DWT	50	0.09	40/40	768
	45	0.10	40/40	743
	40	0.11	40/40	721
	35	0.12	40/40	698
	30	0.13	40/40	672
	25	0.16	40/40	649

TABLE I. SYSTEM PERFORMANCE ACCORDING TO THE NUMBER OF HIDEN LAYER NEURONS

II. CONCLUSION

An AMPCS using adaptive classification method was trained and tested with various marble texture images, added MB and brightness to simulate the real production conditions. It is based on training a classification MLP NN with three different input training sets: extracted texture histograms, DCT and DWT over the histograms. The developed method applying training with DWT approximation coefficients shows higher recognition accuracy compared to the other two methods and offers more resources for further reduction of the computational time while retaining the accuracy.

The proposed adaptive algorithm is implemented and proved for real-time operation on standard PLC SIMATIC S7-317 DP instead of on other embedded systems, because PLCs are widely used as control devices in automated production. Therefore, there is no need to develop additional equipment to the PLCs and this is a good opportunity for replication of the proposed AMPCS results in similar processes. The achieved short maximum system reaction time and estimated high recognition accuracy is a good precondition for proper system operation even under worse production conditions.

REFERENCES

[1] Tuceryan, M., Jain, A., Texture Analysis, Handbook of Pattern Recognition and Computer Vision, 2nd ed., World Scientific Publishing Company, pp. 207–248, 1998.

[2] Baraldi, A. & Parmiggiani, F., An investigation of the Textural Characteristics Associated with Gray Level Cooccurrence Matrix Statistical Parameters, IEEE Trans. On Geoscience and Remote Sensing, ISSN: 0196-2892, vol. 33, no. 2, pp. 293-304, 1995.

[3] Mallick-Goswami, B., Datta, A., Detecting defects in fabric with laser-based morphological image processing, Textile Research Journal, ISSN: 0040-5175, 70: pp.758–762, 2000.

[4] Sousa, C. & Pinto, C., Comparison of Intelligent Techniques Applied to Marble Classification, ICIAR 2004, ISSN 0302-9743, pp. 802–809, Porto, Portugal, 2004.

[5] Neural Networks for Instrumentation, Measurement and Related Industrial Applications, NATO Science Series, Series III: Computer and Systems Sciences – Vol.185, 2003.

[6] M. Sartori, P Antsaklis, A simple method to derive bounds on the size and to train multi-layer neural networks, IEEE Trans. Neural Networks, vol. 2, no. 4, pp. 467-471, 1991.

[7] Camargo, L. S., Yoneyama, T., Specification of training sets and the number of hidden neurons for multilayer perceptrons, Neural Computing, vol. 13, pp. 2673-2680, 2001.

[8] Topalova, I. & Tzokev, A., Automated Classification of Marble Plate Textures with MLP-PLC Implementation, International Conference on Engineering and Meta-Engineering: ICEME'2010, pp.115-119, 6-9 April, Orlando, USA, 2010.

[9] Paul S. Addison, The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance, Institute of Physics Publishing, Bristol, UK, 2002

[10] Stéphane Mallat, A Wavelet Tour of Signal Processing, 3rd ed., Academic Press, 2009.

[11] Alexander G. and Jean M., Decomposition of Hardy functions into square integrable wavelets, 723–736,1984.

[12] http://www.scribd.com/doc/36971083/Addison-EEG-Review, last access on 14.04.2012.

[13] Stéphane Mallat, A Wavelet Tour of Signal Processing, 3rd ed., Academic Press, 2009. http://etd.lib.fsu.edu/theses/available/etd-11242003-185039/unrestricted/09_ds_chapter2.pdf

[14] Widrow, B. & Stearns, S., Adaptive Signal Processing, Prentice-Hall, Inc. Englewood Cliffs, N.J. 07632, pp.36-40, 2004.

[15] Topalova, I., Tzokev, A., Automated Texture Classification of Marble Shades with Real-Time PLC Neural Network Implementation, WCCI 2010 IEEE World Congress on Computational Intelligence July, 18-23, 2010 - CCIB, pp.1031-1038, 18-23 July, ISBN:978-1-4244-8126-2/10, Barcelona, Spain, 2010.

[16] Programmable Logic Controllers S7-300, Module Data, Reference Manual, Siemens, 2005.

[17] Baraldi, A. & Parmiggiani, F., An investigation of the Textural Characteristics Associated with Gray Level Cooccurrence Matrix Statistical Parameters, IEEE Trans. On Geoscience and Remote Sensing, ISSN: 0196-2892, vol. 33, no. 2, pp. 293-304, 1995.

Analysis, Design and Implementation of Human Fingerprint Patterns System "Towards Age & Gender Determination, Ridge Thickness To Valley Thickness Ratio (RTVTR) & Ridge Count On Gender Detection

[1] E.O. OMIDIORA, [2]O. OJO.
[1,4] Dept Of Computer Science & Engineering
14Ladoke Akintola University of Technology
14Ogbomosho Nigeria

[3]N.A. YEKINI, [4]T.O. TUBI
[2,3]Dept Of Computer Technology
Yaba College of Technology
Yaba Lagos Nigeria

Abstract-The aim of this research is to analyze humans fingerprint texture in order to determine their Age & Gender, and correlation of RTVTR and Ridge Count on gender detection. The study is to analyze the effectiveness of physical biometrics (thumbprint) in order to determine age and gender in humans. An application system was designed to capture the finger prints of sampled population through a fingerprint scanner device interfaced to the computer system via Universal Serial Bus (USB), and stored in Microsoft SQL Server database, while back-propagation neural network will be used to train the stored fingerprint. The specific Objectives of this research are to: Use fingerprint sensor to collect different individual fingerprint, alongside their age and gender, Formulate a model and develop a fingerprint based identification system to determine age and gender of individuals and evaluate the developed system.

Keywords: Age; Gender; Fingerprint; Ridges Count; RTVTR.

I. INTRODUCTION

A Fingerprint is the representation of the epidermis of a finger; it consists of a pattern of interleaved ridges and valleys. Fingertip ridges evolved over the years to allow humans to grasp and grip objects. Like everything in the human body, fingerprint ridges form through a combination of genetic and environmental factors. This is the reason why even the fingerprint of identical twins is different (Maltoni and Cappelli, 2006).

The concept of fingerprint pattern being studied has been of significant use over time, when scanning it involves the conversion of fingerprint by small portion of light solid-state devices into alphanumeric formula (Glaton, 1982).

The technology of using human body for identification purpose is known as "Biometrics", the word was derived from Greek word meaning Bio for "*life*" and Metric for "*Measurement*", the authentication of biometrics for personal identification is extremely and more reliable compared with something you know which can be forgotten (like password, registration numbers) or what you have (like identity card, physical lock, smart card) which can be misplaced compared with something you are or parts of your body. (Neil and Adnan 2004).

Biometrics measures the unique physical or behavioral characteristics of individual as a means to recognize or authenticate their identity. Common physical biometrics includes fingerprints, hand or palm geometry, and retina, iris or facial characteristics. Behavioral characteristics include signature, voice (which also has a physical component), keystroke pattern and gait. Although some technologies have gained access control and biometrics as a whole shows great potential for use in end user segments, such as airports, stadiums, defense installations and the industry and corporate workplaces where security and privacy are required (Jain et al 2003).

II. LITERATURE REVIEW

A. Overview of Fingerprinting

Human fingerprints have been discovered on a large number of archeological artifacts and histological items. Although these findings provide evidence to show that ancient people were aware of the individuality of fingerprints, it was not until the late sixteenth century that the modern scientific fingerprint technique was first initiated (Jain, et al, 2003). In 1686, Marcello Malpighi, a professor of anatomy at the University of Bologna noted in his writings the presence of ridges, spirals and loops in fingerprints. Since then, a large number of researchers have invested huge amounts of effort on fingerprint studies (Gu, Zhou, and Yang, 2006).

Henry Fauld, in 1880, was the first to scientifically suggest the individuality of fingerprints based on an empirical observation. At the same time, Herschel asserted that he had practiced fingerprint recognition for about 20 years. These findings established the foundation of modem fingerprint recognition. In the late nineteenth century, Sir Francis Galton conducted an extensive study on fingerprints; he introduced the minutiae features for fingerprint matching in 1888. An important advance in fingerprint recognition was made in 1899 by Edward Henry, who established the well-known

Henry system" of fingerprint classification (Lee and Gaensslen, 2001).

In the early twentieth century, fingerprint recognition was formally accepted as a "valid personal identification method and became a standard routine in forensics (Lee and Gaensslen, 2001). Fingerprint identification agencies were set up world-wide and criminal fingerprint databases were established (Lee and Gaensslen, 2001).

Various fingerprint recognition techniques, including latent fingerprint acquisition, fingerprint classification and fingerprint matching were developed. For example, the FBI fingerprint identification division was set up, in 1924, with a database of 810,000 fingerprint cards (Federal Bureau of Investigation, 1984).

European explorer Joao de Barros recorded the first known example of fingerprinting, which is a form of biometrics, in China during the 14th century. Chinese merchants used ink to take children's fingerprints for identification purposes. In 1890, Alphonse Bertillon studied body mechanics and measurements to help in identifying criminals. The police used his method, the Bertillon age method, until it falsely identified some subjects. The Bertillon age method was quickly abandoned in favor of fingerprinting, brought back into use by Richard Edward Henry of Scotland Yard. Karl Pearson, an applied mathematician studied biometric research early in the 20^{th} century at University College of London. He made important discoveries in the field of biometrics through studying statistical history and correlation, which he applied to animal evolution. His historical work included the method of moments, the Pearson system of curves, correlation and the chi-squared test.

In the 1960s and '70s, signature biometric authentication procedures were developed, but the biometric field remained fixed until the military and security agencies researched and developed biometric technology beyond fingerprinting (Caplan and Dermatol. 1990).

Since 700 AD, this science of fingerprint has been used for the purpose of identification. Chinese used fingerprints as official documents in 3000BC. The system was first used in India in 1858 by Sir William Herschel to prevent impersonation, but the credit is given to Sir Francis Galton for having it systematized for the identification of criminals. His system was officially adopted in England in 1894, and was further modified by Sir Edward Henry. Afterwards the studies have been conducted on fingerprint ridges mainly its types, classification, methods of lifting fingerprints, recording of fingerprints and materials used to develop fingerprint (Gungadin 2007).

Fingerprint identification and classification has been extensively researched in times past, however very few researchers have studied the fingerprint gender classification problem, (Acree 1999) used the ridge density, which he defined as the number of ridges in a certain space; it was shown that the females have higher ridge density (Acree 1999). (Kralik 2003) also showed that the males have higher ridge breadth,) which was defined as the distance between the centers of two adjacent valleys), than females.

Two studies showed that the males have higher ridge count than the females (Hall, Timura 2003) and (Cote, Earls, Lalumiere 2002). It was shown that both males and females have higher rightward directional asymmetry in the ridge count, with the asymmetry being higher in males than females (Hall and Kimura 2003), and higher incidence of leftward asymmetry in females. (Cote et al 2002). Female's fingerprints are significantly of lower quality than male fingerprints (Hicklin and Reedy. 2002). While the appearance of white lines and scars in fingerprint images are very common in housewives (Wendt 1955).

By the seventh month of natal development, the dermatoglyphic patterns of the fingers are complete and no further modifications can occur during maturation except in the case of accident. Since, during growth, the overall size of the palm increases, the fingerprint increase in size without adding new ridges and ridge breadth this is defined as the measurement from the center of one furrow across the ridge to the center of the next furrow. (Penrose 2001)

Within individual, the breadth of fingerprint ridge varies within the hand and between hands, but the difference is quite small, on the order of 0.05mm and less (Cummings 1994).

Although it has been found that males tend to have more ridge counts than females, inconsistent results have been obtained with regard to the statistical significance of this sex difference. It has also been shown that women tend to have a higher ridge density (ridge counts divided by the size of the corresponding fingertip area) than men but the sex determination accuracy of this feature does not seem to be very satisfactory (Acree 1999).

The female's fingerprint is characterized by a high Ridge Thickness Valley Thickness Ratio, while the male's fingerprint is characterized by low Ridge Thickness Valley Thickness Ratio, with the exception of small percentage of male's fingerprints having high Ridge Thickness Valley Thickness Ratio, and female's fingerprints having low Ridge Thickness Valley Thickness Ratio (Baldawi et al. 2008)

The female's fingerprint is characterized by high count of white lines, with the exception of small percentage having few or no white lines. The male's fingerprint is characterized by having no or few number of white lines, with the exception of small percentage having high count of white lines also that the male's ridge count is slightly higher than the female's, with high standard deviation in its distribution among both genders, ridge count for the females, with μ=13.6671, σ=4.9845, and the males, with μ=14.6914, σ=4.9336, with t-value =4.802, and ρ-value=1.685E-06 (Badawi et al. 2008)

The size of the fingertip has a strong relationship to the values of ridge counts and ridge density. If males have more ridge counts and smaller ridge densities than females, then the finger size difference between males and females should be more significant than the features of ridge count and ridge density (Wang et al 2007).

Gungadin (2007) opined that a ridge count of ≤13 ridges/25 mm2 is more likely to be of male origin and that of ≥14 ridges/25 mm2 is likely to be of female origin. The outcome of his study is that women tend to have more ridge density than men.

B. Fingerprint Classification

Fingerprint identification and verification are one of the most significant and reliable identification methods. It is virtually impossible that two people have the same fingerprint, having a probability 1 / 1.9 X 10^{15} (Hong et al. 2000). In fingerprint identification and verification applications world-wide, a large volume of fingerprints are collected and stored for a wide range of applications, including forensics, civilian, commercial and law-enforcement applications.

Automatic identification of humans based on fingerprints requires the input fingerprint to be matched with a large number of fingerprints in a database (for example, the FBI database contains approximately 70 million fingerprints). To reduce the search time and computational complexity, it is desirable to classify the database into accurate and consistent classes so that input fingerprint is matched only with a subset of the fingerprints in the database. The nature of each application will determine the degree of accuracy required. For example, a criminal investigation case may require higher degree match than access control case systems.

Automatic fingerprint classification methods, such as methods introduced in (Drets and Liljenstrom, 1999) rely on point patterns in fingerprints, which form ridge endings and bifurcation unique to each person.

Traditionally, activities to solve a pattern recognition task are twofold. First, a set of features has to be found describing the object(s) being classified. Second, after a set of features has been found, a classification mechanism is chosen and optimized. These two steps are highly interdependent, since the choice of features influences the conditions under which a classifier operates, and vice versa. (Mohamed, Nyongesa and Siddiqi 2002).

With the advent of neural networks however, more and more problems are solved by simply feeding large amounts of ’raw data’ (e.g. images, sound signals, stock market index ranges) to a neural network. This approach, however, is not feasible in fingerprint classification, which are highly susceptible to noise and elastic distortions. Therefore, it is desirable to extract features from the images that are invariant to such distortions. During training the classification network learns the association and significance of features. An attempt has been made previously to study fuzzy logic and artificial neural network techniques in fingerprint identification. (Mohamed, Nyongesa and Siddiqi 2002).

C. Fingerprint Feature Extraction (FFE)

The central problem in designing a fingerprint classification system is to determine what features, should be used and how categories are defined based on these features. There are, mainly two types of features that are useful for fingerprint recognition system:

(i) Local ridge and valley details (minutiae) which have different characteristics for different fingerprints, and
(ii) Global pattern configurations, which form special patterns of ridges and valleys in the central region of the fingerprint.

The first type of features carries for the information about the individuality of fingerprints and the second type of features carry information about the fingerprint class. Therefore, for fingerprint classification, the features derived from the global pattern configurations should be used. These features should be invariant to the translation and rotation of the input fingerprint images. Generally, global fingerprint features can be derived from the orientation field and the global ridge shape. The orientation field of a fingerprint consists of the ridge orientation tendency in local neighborhoods and forms an abstraction of the local ridge structures. It has been shown that the orientation field is highly structured and can be roughly approximated by the core and delta models (Monro and Sherlock, 1993) which are known as singular points details.

III. Methodology

A. Fingerprint Acquisition

The ten fingerprints of two hundred individual was collected, one hundred male and one hundred females for different age groups. The age group was divided in series of five years (i.e. 0-5, 6-10, 11-15, 16-20, 21-25…….). The captured fingerprint was stored in a Microsoft SQL Server database through an interfaced fingerprint reader.

B. CREATING THE APPLICATION

The application was created using Microsoft Visual Studio 2008 integrated development environment (IDE). *Visual Studio .NET* is Microsoft’s integrated development environment (IDE) for creating, documenting, running and debugging programs written in a variety of .NET programming languages. Visual Studio .NET also offers editing tools for manipulating several types of files. Visual Studio .NET is a powerful and sophisticated tool for creating business-critical and mission-critical applications. Figure 1 shows Microsoft Visual Studio .Net 2008 IDE environment. C# is the ideal language used for the development of the application.

Figure 2 shows the runtime environment of the application. The first page shown here is Data Collection/Identification page. Data collection involve collection of fingerprint, sex and age range of different people. “Save button” will be clicked in other to save the data into the database. This will then trigger the fingerprint reader to ask for thumb. Clear button can be used to clear the data in other to save another one.

Before identification, the network must be trained properly in other to give better result. Train Network button can be used to train the network, while Load Network button can be used to load the already trained network.

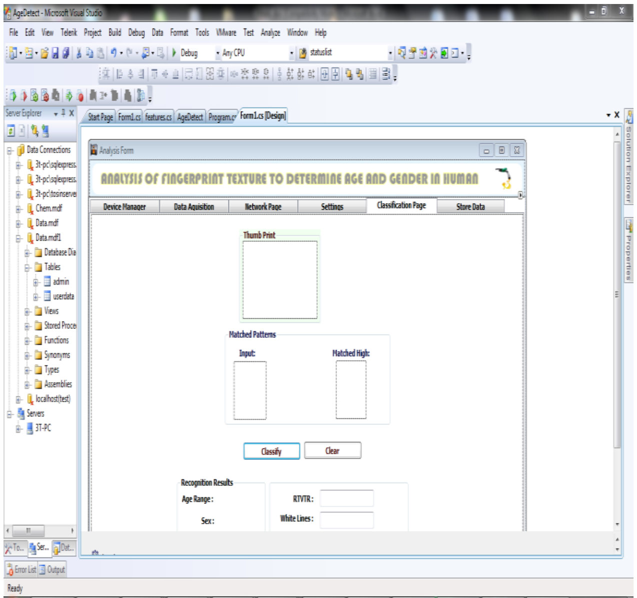


Figure 1: Design environment of the application.

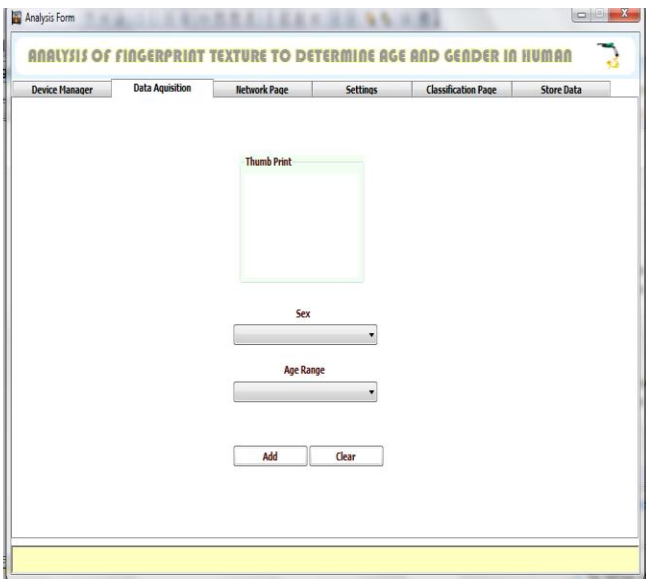


Figure 2: Runtime environment of the application.

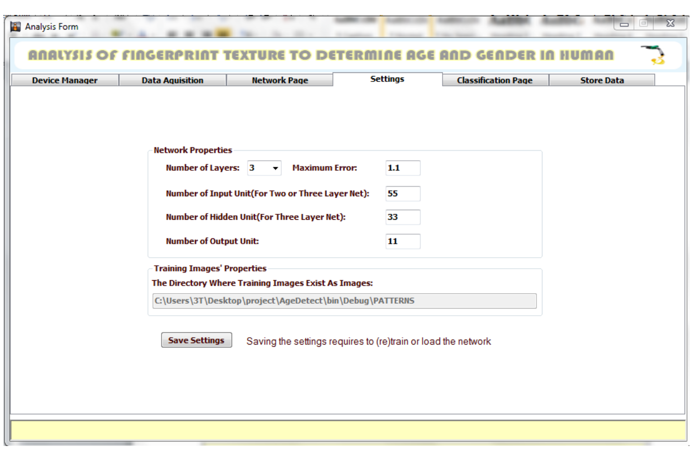
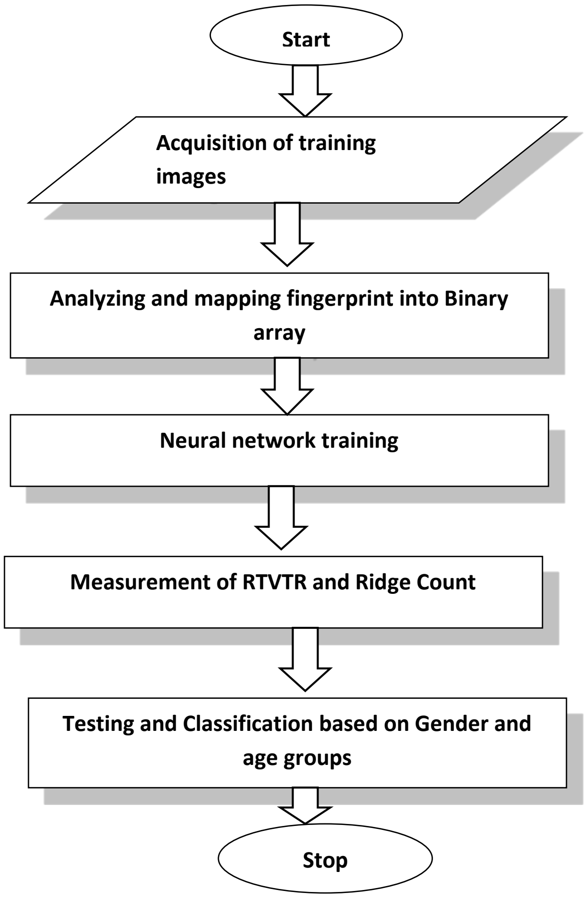


Figure 3: Network Settings Form.

Figure 3 shows the required network properties. The flowchart for the methodology is as shown below:



FLOWCHART FOR METHODOLOGY

IV. IMPLEMENTATION AND DISCUSSION

A. Analysis of Results

Having trained the network and used some samples of fingerprint to predict the gender and age range of some some individuals. The following results were obtained.

TABLE 1: RESULTS OF AGE AND GENDER PREDICTION.

S/N	AG	EG	AA	EA	GC	AC
1	M	M	21-25	26-30	YES	NO
2	M	F	26-30	26-30	NO	YES
3	F	F	26-30	26-30	YES	YES
4	M	M	21-25	26-30	YES	NO
5	M	M	26-30	26-30	YES	YES
6	F	M	26-30	26-30	NO	YES
7	M	M	21-25	26-30	YES	NO
8	F	F	21-25	16-20	YES	NO
9	F	F	26-30	26-30	YES	YES
10	M	M	26-30	16-20	YES	NO

11	F	F	16-20	26-30	YES	NO
12	M	F	21-25	26-30	NO	NO
13	F	F	21-25	26-30	YES	NO
14	M	M	26-30	16-20	YES	NO
15	M	M	21-25	26-30	YES	NO

TABLE 2: PERCENTAGE ACCURACY OF THE MODEL, WHERE AG = ACTUAL AGE, EG = EXPERIMENTAL AGE, AG = ACTUAL AGE, EA = EXPERIMENTAL AGE, GC = GENDER CORRELATION AND AC = AGE CORRELATION

GENDER ACCURACY	AGE ACCURACY
80.00%	33.33%

From the table above, the percentage accuracy of the model for age is 33.3% while that of gender is 80%.

B. *Ridge Thickness To Valley Thickness Ratio*

Measuring the Ridge thickness to valley thickness ratio (RTVTR), the following results were obtained for 15 randomly selected samples.

The result shows that the females have a higher RTVTR compared to the males.

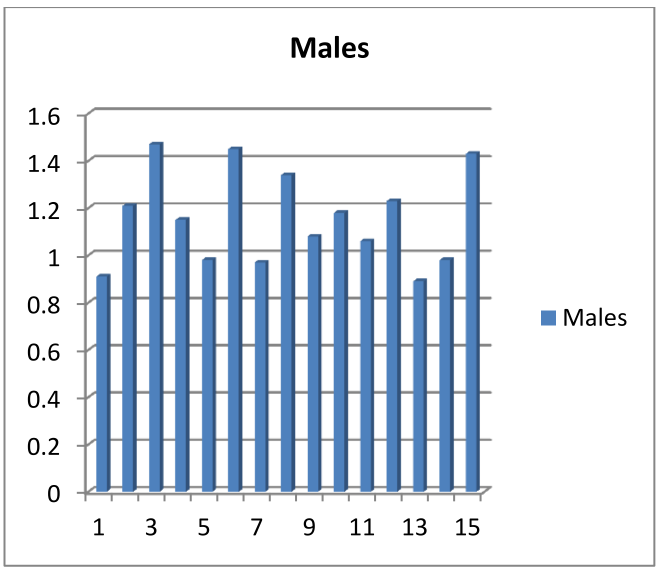


Figure 4: Histogram of the RTVTR obtained for males

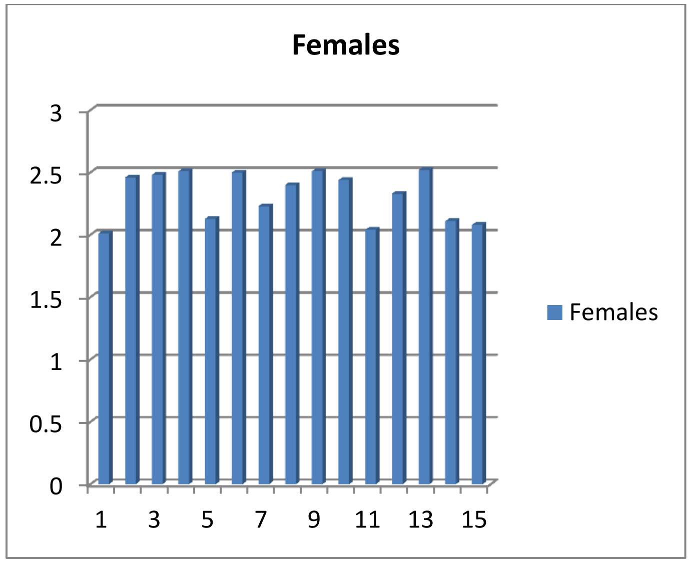


Figure 5: Histogram of thhe RTVTR obtained for females

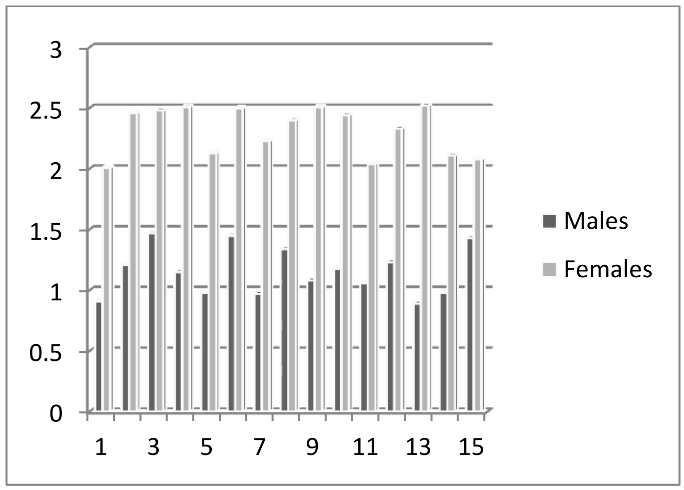


Figure 5: Histogram for the RTVTR obtained for females versus the males

TABLE 4: MEAN, STANDARD DEVIATION AND VARIANCES FOR THE RESULT OBTAINED

	Mean	S.D	Variance
Males	1.155	0.191	0.0353
Females	2.310	0.188	0.0363

TABLE 3: RTVTR RESULT

S/N	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Male	0.9	1.2	1.5	1.2	1.0	1.5	1	1.3	1.1	1.2	1.1	1.2	0.9	1.0	1.4
Female	2.0	2.5	2.5	2.5	2.1	2.5	2.2	2.4	2.5	2.4	2.0	2.3	2.5	2.1	2.1

C. Ridge Count

Ridge count is the number of ridges occured in a particular region of a particular section of the fingerprint. The result of the Ridge count is shown in the table below, and it shows that the males have a slightly higher ridge count than the females.

TABLE 5: RIGDE COUNTS

S/N	Males	Females
1	14.642	13.661
2	14.352	13.781
3	14.253	12.978
4	13.948	13.465
5	14.645	13.875
6	16.473	13.667
7	14.731	13.657
8	14.532	13.898
9	14.572	13.675
10	14.493	13.643
11	14.343	13.794
12	14.637	13.103
13	15.362	13.133
14	14.546	12.981
15	14.691	13.408

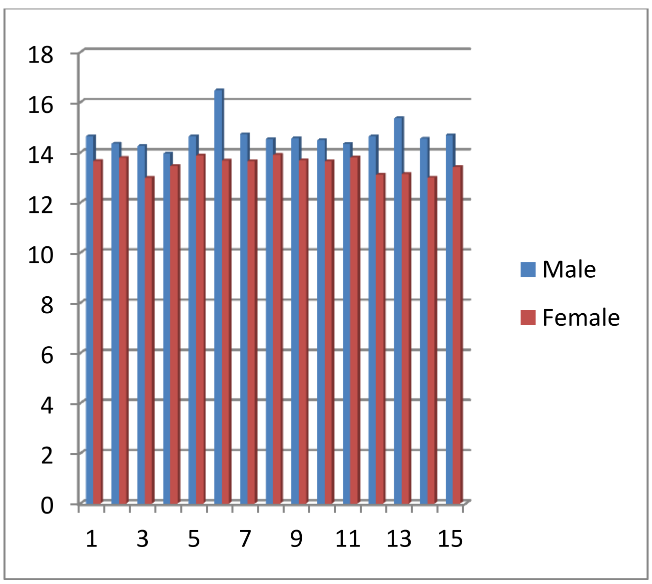


Figure 6: Histogram of the Ridge Count for both male and Female Gender

From the results above, three observations can rightly be made

1. The females have a higher ridge thickness to valley thickness ratio than the males.
2. The Males has a slightly higher ridge count than the females.
3. There is no particular relationship between the age of subjects and their fingerprint pattern, as it does not change (only as a result of accident or mutation)

The network has been trained and tested for a number of fingerprint image of different sex and age range. Since the implementation of the software is open and the program code is scalable, the inclusion of more number of fingerprint images from different people is straight forward.

The necessary steps are preparing the sequence of input symbol images in a single image file (*.bmp [bitmap] extension). The application will provide a file opener dialog for the user to locate the *.net network file and will load the corresponding training file by itself.

Fingerprint image size is 10* 10 in dimension. Use of any other size is also possible by preparing the input/desired output set as explained. The application can be operated with fingerprint images as small as 10*10 dimensions in size.

V. CONCLUSION AND RECOMMENDATION

This research works have presented a model towards the determination of gender through the fingerprint information. The technical approach followed in processing input images, detecting graphic symbols, analyzing and mapping the fingerprint and training the network for a set of desired corresponding image to the input images have been discussed. Multi-Layer Perceptron (MLP) model technique was used, and results were obtained for age & gender, Ridge Thickness To Valley Thickness Ratio & Ridge Count for trained finger prints sampled. We observed that the females have a higher ridge thickness to valley thickness ratio compare to males. The Males has a slightly higher ridge count compare to females, there is no particular relationship between the age of subjects and their fingerprint pattern, as it does not change (only as a result of accident or mutation)

RECOMMENDATION

Further research is necessary to develop feature extraction approaches that can reliably and consistently extract a set of features that provide rich information to age, paternalism e.t.c.

REFRENCES

[1] Acree M.A. (1999) Is there a gender difference in fingerprint ridge density? *Forensic Science International*, 102: 35–44.

[2] Badawi Ahmed, Mohamed Mahfouz, RimonTadross, Richard Jantz. (2008) Fingerprint-Based Gender Classification.

[3] Capelli R, Maltoni D. and Ferrara M. (2006) The quality of fingerprint scanners and the impact on the Accuracy of Fingerprint Recognition Algorithms *In Proceedings of Multimedia Content Representation, Classification and Security (MRSC 2006)*, pages 10,16.

[4] Caplan R.M and J.ADermatol, 1990. How fingerprint came into use for personal identification. Graw-Hill Press.

[5] Cote Karine,. Earls Christopher M, and . Lalumiere Martin L. (2002) "Birth Order, Birth Interval, and Deviant Sexual Preferences Among Sex Offenders." Sexual Abuse: A Journal of Research and Treatment, Vol. 14, No. 1.

[6] Drets. G and H. Liljenstrom, ''Fingerprint Sub-Classification, (1982) A Neural Network Approach,'' Intelligence Biometric Techniques In Fingerprint and Face Recognition, PP. 109-134, 1999. Galton F. *Finger Prints.* Macmillian London

[7] Gu J., Zhou J., and Yang C. (2006) Fingerprint Recognition by combining global structure and local clues. *IEEE Transactions on Image Processing, 15(7):1924{1964.*

[8] Gungadin, S. (2007) Sex Determination from Fingerprint Ridge Density. Internet Journal of Medical Update, 2: Jul-Dec 2(2).

[9] Gurney, K. (1997). *An Introduction to Neural Networks* London: Routledge.

[10] Hall, J. A. Y. and Kimura D.(1994) "Dermatoglyphic Asymmetry and Sexual Orientation in Men." ,Behavioral Neuroscience, 108, 1203-1206,

[11] Haykin, S. (1999) *Neural Networks: A Comprehensive Foundation*, Prentice Hall, ISBN 0-13-273350-1

[12] Hicklin R. Austin, Christopher L. Reedy, (2002) "Implications of the IDENT/IAFIS Image Quality Study for Visa Fingerprint Processing", Mitertek Systems (MTS).

[13] Hong L., Prabhakar. S., Jain. A. K., and Pankanti. S., ''Filterbank-Based Fingerprint Matching,'' IEEE Transactions of Image Processing, Vol. 9:5, PP. 846-859, 2000.

[14] Jain A.K. Maltoni D. Maio D. and Prabhakar S. (2003) *Handbook of Fingerprint Recognition.*Springer, New York.

[15] Kralik M., Novotny V. (2003) Epidermal Ridge Breadth: An Indicator of Age and Sex in Paleodermatoglyphics. *Variability and Evolution*. Vol-11; 5-30

[16] Lee C. Henry and Gaensslen E. Robert (2001) *Advances of Fingerprint Technology*. Elsevier Science.

[17] Mohamed.S., H. O. Nyongesa, and J Siddiqi (2000) ''Automatic Fingerprint Identification System Using Fuzzy Neural Techniques, 'Proceedings of the International Conference on Artificial Intelligence, Volume 2 PP. 859-865, CSREA Press, Las Vegas.

[18] Neil Yager, Adnan Amin (2004) "*Fingerprint classification: a review*," Springer-Verlag London.

[19] Penrose L.S. (2001). Memorandum on Dermatoglyphic Nomenclature. *Birth Defects Original Article Series*, 4: 1–13.

[20] The Federal Bureau of Investigation (1984). The Science of Fingerprints: Classification and Uses. Technical Reports U.S Government Publication, Washington,DC.

[21] Wang Y, Hu J, Phillips D. A (2007) Fingerprint Orientation Model Based on 2D Fourier Expansion (FOMFE) and its application to Singular-Point Detection and Fingerprint Indexing. *IEEE Trans Pattern Anal Mach Intell.*;29(4):573-85.

[22] Wendt GG (1955) "ÜberweißeLinienimAbdruck der men schlichenFingerbeere." Homo 6:180–188.