International Journal of Interactive Multimedia and Artificial Intelligence

September 2016, Vol IV, Number 1, ISSN: 1989-1660

By far the greatest danger of Artificial Intelligence is that people conclude too early that they understand it.

Eliezer Yudkowsky (1979 -)

Special Issue on Artificial Intelligence Underpinning

http://www.ijimai.org

INTERNATIONAL JOURNAL OF INTERACTIVE MULTIMEDIA AND ARTIFICIAL INTELLIGENCE

ISSN: 1989-1660–VOL. IV, NUMBER 1

Editor's Note

The International Journal of Interactive Multimedia and Artificial Intelligence provides an interdisciplinary forum in which scientists and professionals can share their research results and report new advances on Artificial Intelligence and Interactive Multimedia techniques.

The research works presented in this issue are based on various topics of interest, among which are included: Pattern Recognition, Multimedia Information Retrieval, Knowledge extraction and knowledge mining, Data mining, Intelligent Systems & Artificial Intelligence, Wireless Technology, Network Telecommunication, Security & Network Management, Advanced Network Technologies.

Khaissidi, G. Et al. [1] proposes an unsupervised segmentation-free method for spotting and searching query, especially, for images documents in handwritten Arabic. The segmentation free approach performs spotting and segmentation concurrently using a sliding window. The features vectors are then specifically modeled by a Histograms of Oriented Gradients (HOGs), and learnt by a support vector machine classifier (SVM) to produce a better representation of the query.

Elgarrai, I. Et al. [2] presents a novel and uniform framework for both face identification and verification. The framework is based on a combination of Gabor filters and Fisher's Discriminant Analysis, and can be considered appearance based in that features are extracted from the whole face image. The design of Gabor filters for facial feature extraction is extensively discussed; the resultant descriptors are injected in a 1D- HMM to achieve the training and recognition steps. The method has been tested extensively for both identification and verification applications.

Sayoti, F. Et al. [3] is focused on one recently technique, the called Golden Ball (GB) for the flow shop scheduling problem (FSSP).This paper argues a new adaptation of the GB algorithm as a multiple-population meta-heuristic based on soccer concepts to find the optimal schedule with a minimal makespan. The proposed approach is very fast and achieves state-of-the-art results on solving the continuous Flow-Shop Scheduling problem by Metaheuristics.

El batteoui, I. Et al. [4] describes an interesting method of camera self-calibration with varying intrinsic parameters from an unknown planar scene. It then demonstrates that the relationship between two matches which have a best correlation score ZNCC. This approach permits to construct a non-linear cost function; its resolution provides the intrinsic parameters of the camera to find the optimum solution according to the intrinsic parameters in the two images of the scene. To evaluate the method and ideas proposed in this paper, a large number of experiments have been conducted to demonstrate their effectiveness

Ramchoun, H. Et al. [5] this paper presents a new approach for the definition and optimization of MLP Neural Network Architectures and Weights. It studies a genetic algorithm for the training and construction of a multilayer perceptron. It is a really interesting demonstration of how the genetic algorithm works on a layer-by-layer basis and (MLP) classifier has one or more hidden layers in between the input and the output layer. They explore MLP as a supervised learning technique called Backpropagation. Finally the proposed technique presented is expected to provide better learning scheme for a classifier

Boulid , Y. Et al. [6] This article provides a comprehensive survey of recent developments in Arabic handwriting recognition.They propose a novel algorithm for extracting text lines of handwritten Arabic manuscripts using Markov Decision Processes, and they also utilize a neighborhood connected component analysis, which provide information about the location of potential characters in the processed documents. The method requires binarization.

Harrati, Y. Et al. [7] investigates both the problems of scheduling and buffer management in delay tolerant networks. Delay Tolerant Networks (DTN) are wireless networks based on the store-carry-and-forward protocols where disconnections may occur frequently. In order to achieve data delivery in DTNs, there, a node may store a message in its buffer and carry it along for long periods of time. This work proposes an efficient joint scheduling and new drop policy MaxHopCount that can optimize different performance metric such as the average delivery rate and the average delivery delay.

Benalla, M. Et al. [8] addresses a multi-agent system for traffic fluidization for emergency convoy. It combines the benefits of a multi-agent system and intelligent traffic forecasts. They represent communications within a multi-agent system and propose a system that consists of three hierarchical levels: the Central Controller (decision-maker); the Communication and Coordination Controller (information provider); the Traffic Lights Controller (executor decision). They also provide a time and space optimal implementation of Dijksta's algorithm.

Settouti, N. Et al. [9] talks about the top 10 data mining algorithms identified by the IEEE International Conference on Data Mining (ICDM) in December 2006. With each algorithm, they provide a description of the algorithm, discuss the impact of the algorithm, and review current and further research on the algorithm. The work establishes a complete guideline for the use of nonparametric statistical procedures for performing multiple comparisons of several classifiers. Experimental studies of non-parametric statistical tests and post-hoc procedures devised to perform multiple comparisons of classification algorithms over medical and biological benchmark data sets.

Chebli,S. Et al. [10] proposes a new approach to compute the stability region for first order delay system controlled by proportional-integral (PI) for improving Active Queue Management (AQM) in the Internet. This result is based on an extension of the Hermite-Biehler Theorem to quasipolynomials. The performance of the closed-loop system with the new PI controller is tested.

Haddouch, K. Et al. [11] describes a new model of the binary Weighted Constraint Satisfaction Problems. This work also explores the techniques for solving WCSP using continuous Hopfield network. Overall, this paper shows a compelling demonstration that CHN is potentially capable of solving A Weighted Constraint Satisfaction problems. Computational results are implemented.

Magdin, M, et al. [12] presents a designed software wichh works by the use neural networks in real time which enable to apply the software into various fields of human lives and thus actively influence its quality. Validation of face emotion recognition software was annotated by using various experts. These expert findings were contrasted with the software results. By evaluating the emotional state, there is an attempt to overcome the barrier between man and non-emotional machine.

Mishra, P. and Shrawankar, U. [13] explains how nowadays game engines are imperative for building 3D applications and games. This is for the reason that the engines appreciably reduce resources for employing obligatory but intricate utilities. The paper elucidates about a game engine, popular games developed by these engines and its foremost elements. It portrays a number of special kinds of contemporary game developed by engines in the way of their aspects, procedure and deliberates their stipulations with comparison.

Bivde, V.S. et al [14] investigates a correlation between the values

DOI: 10.9781/ijimai.2016.410

of coupling metrics and the number of classes in the multimedia Java code. A case study of a banking multimedia Java project with its forty different versions is conducted to comments on this correlation. The analysis of the results shows that, if the input source code is with a large number of classes then it results in high coupling values.

Dr. Mohamed Bahaj

References

[1] Khaissidi, G. ET al. "Segmentation-free Word Spotting for Handwritten Arabic Documents", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 1, pp. 6-10, 2016.

[2] Elgarrai, Z. Et al. "Offline Face Recognition System Based on Gabor-Fisher Descriptors and Hidden. Markov Models", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 1, pp. 11-14, 2016.

[3] Sayoti, F. Et al. "Golden Ball Algorithm for solving Flow Shop Scheduling Problem" ", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 1, pp. 15-18, 2016.

[4] El batteoui, I. Et al. "A Self-Calibration Method of Zooming Camera ", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 1, pp. 19-25, 2016.

[5] Ramchoun, H. Et al." Multilayer Perceptron: Architecture Optimization and Training", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 1, pp. 26-30, 2016.

[6] Boulid, Y. Et al." Detection of Text Lines of Handwritten Arabic Manuscripts using Markov Decision Processes", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 1, pp. 31-36, 2016.

[7] Harrati, Y. Et al. "MaxHopCount: A New Drop Policy to Optimize Messages Delivery Rate in Delay Tolerant Networks", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 1, pp. 37-41, 2016.

[8] Benalla, M. Et al. "A Distributed Intelligent System for Emergency Convoy", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 1, pp. 42-45, 2016.

[9] Settouti, N. Et al." Statistical Comparisons of the Top 10 Algorithms in Data Mining for Classification Task", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 1, pp. 46-51, 2016.

[10] Chebli, S. Et al. "PI Stabilization for Congestion Control of AQM Routers with Tuning Parameter Optimization", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 1, pp. 52-55, 2016.

[11] Haddouch, K. Et al. "Solving the Weighted Constraint Satisfaction Problems Via the Neural Network Approach", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 1, pp. 56-60, 2016.

[12] Magdin, M, et al. "Evaluating the Emotional State of a User Using a Webcam", International Journal of Artificial Intelligence and Interactive Multimedia, vol. 4, no. 1, pp. 61-68, 2016.

[13] Mishra, P., Shrawankar, U., "Comparison between Famous Game Engines and Eminent Games", International Journal of Artificial Intelligence and Interactive Multimedia, vol. 4, no. 1, pp. 69-77, 2016.

[14] Bivde, V.S., Sararu, P., "Correlation Between Coupling Metrics Values and Number of Classes in Multimedia Java Projects: A Case Study", International Journal of Artificial Intelligence and Interactive Multimedia, vol. 4, no. 1, pp. 78-82, 2016.

TABLE OF CONTENTS

OPEN ACCESS JOURNAL

ISSN: 1989-1660

COPYRIGHT NOTICE

Segmentation-free Word Spotting for Handwritten Arabic Documents

G. Khaissidi[1], Y. Elfakir[1], M. Mrabti[1], Z. Lakhliai[1], D. Chenouni[1], M. El yacoubi[2]

[1] LIPI, ENS, FES, Maroc
[2] SAMOVAR, Télécom SudParis, CNRS, Université Paris-Saclay, France

***Abstract* — In this paper we present an unsupervised segmentation-free method for spotting and searching query, especially, for images documents in handwritten Arabic, for this, Histograms of Oriented Gradients (HOGs) are used as the feature vectors to represent the query and documents image. Then, we compress the descriptors with the product quantization method. Finally, a better representation of the query is obtained by using the Support Vector Machines (SVM).**

***Keywords* — Word Spotting, Arabic Handwritten Documents, Histograms Of Oriented Gradients (HOG), Image Recognition, Classification.**

I. Introduction

The search for information in Arabic manuscripts is not a simple process, for this, the conception of recognition system knows today a great expansion and seems as a necessity so as to exploit the wealth of information contained in ancient manuscripts. The manual manipulation repetitive of fragile documents could destroy them, for this, many digitization projects have been developed such as DEBORA (Digital access to Books Of Renaissance) [1], EAMMS (Electronic Access to Medieval Manuscripts) [2], Better Access to Manuscripts and Browsing of Images (BAMBI) [3] and Manuscript Access Through Standards for Electronic Records(MASTER)[4]…treat Latin scripts. For this reason above, and in order to develop a complete system for recognition Arabic handwriting, the first step for the creation of this system is presented in this article. In the survey of literature, it is found that many researchers of keyword spotting methods are inspired by one of two following categories:

- Learning-based methods use supervised machine learning techniques to train models of the words that the user wants to spot.
- Example-based methods, receive as input an instance of the word that the user wants to retrieve.

Most of the authors prefer a learning-based approaches for applications where the keywords to spot are a priori known and fixed. In same handwritten Arabic documents, the segmentation step is not usually easy "Fig.1", any segmentation errors affect the subsequent word representations and matching steps, this dependence motivated the researchers to move towards complete segmentation-free methods.

From literature survey of handwritten word spotting techniques, we found that some researchers has done work on the handwritten Arabic documents where a million documents are written in various disciplines between the seventh and fourteenth centuries, whereas many works treat a Latin's manuscripts documents. For a given query image, Y. Leydier and al. [5] extract and encode interest points by a simple descriptor based on gradient information. The word spotting is then performed by trying to locate zones of the document images with similar interest points. Only the ones sharing the same spatial configuration than the query model are returned.

Kamble and Hegadi [6] extract the features of handwritten Marathi

Fig. 1. Process of the proposed system

DOI: 10.9781/ijimai.2016.411

characters using Rectangle Histogram Oriented Gradient, in this work; Feed-Forward Artificial Neural Network is used for classification step. For a query image, Rath et al. [7] extract discrete feature vectors that describe query, which are then used to estimate similarity after training step of the probabilistic classifier. Jon et al. [8] represent the document with a grid of HOG descriptors and use a sliding-window approach for word spotting in document images. A similar approach is used in [9] which the retrieval step is performed by using an exemplar support vector machine framework.

The approach used in [10] combines three slanted windows, vertical, left and right. They proposed this method to remedy with the problem of writing inclination, overlapping and diacritical marks. HMM-based classifier is used at the decision step. Kessentini et al. propose to use Multi-stream hidden Markov Models for off-line handwritten Arabic word recognition [11]. The approach used combines density based features and contour based features of sliding windows. In [12], each character is a state in Variable Duration Hidden Markov Models and has a variable duration to model a character model of multiple segments.

The remainder of this paper is organized as follows: Section 2 present the indexation system process. Section 3 describes HOG feature extraction approach for word-spotting and product quantization method. Afterwards, in section 4, classification using SVM classifier is presented. Section 5 discuss experimental results. Finally, conclusions are summarized in section 6.

II. Present Work

In the present work, the document images have been preprocessed to enhance them after scanning the collected datasheets. For this purpose, a model for the restoration of the degradations [13], which uses a series of multi-level classifiers [14] applied to document images. Then, the window slid on the image document Histograms of Oriented Gradients (HOG) features are then extracted from each word image to represent and to compare the query with the region of the document. The application of Product Quantization [15] to encode the HOG descriptors reduce the size of the descriptors provides better performance in time of descriptor computation. Finally, Support Vector Machines is used to produce a better representation of the query and to classify feature vectors. Identical positive set is produced by slightly the window around the query and sample negative set is obtained by taking a sampler random regions. The regions with high similarity will be used in reranking step. The general process of the proposed system is shown in "Fig.2".

The proposed indexation system is achieved in the following steps:

- Image preprocessing
- Feature Extraction
- Product Quantization
- SVM training set
- SVM classifying set
- Reranking

III. Hog Feature Extraction

Histogram of oriented gradients feature descriptors are used in computer vision and image processing for the object recognition purpose. The main idea behind the HOG descriptors is that local object appearance and shape within an image can be described by the distribution of intensity gradients or edge directions. The implementation of these descriptors can be achieved by dividing the image into small connected regions, called cells, and for each cell computing a histogram of gradient directions; histograms are also normalized based on their energy (regularized L2 norm). The

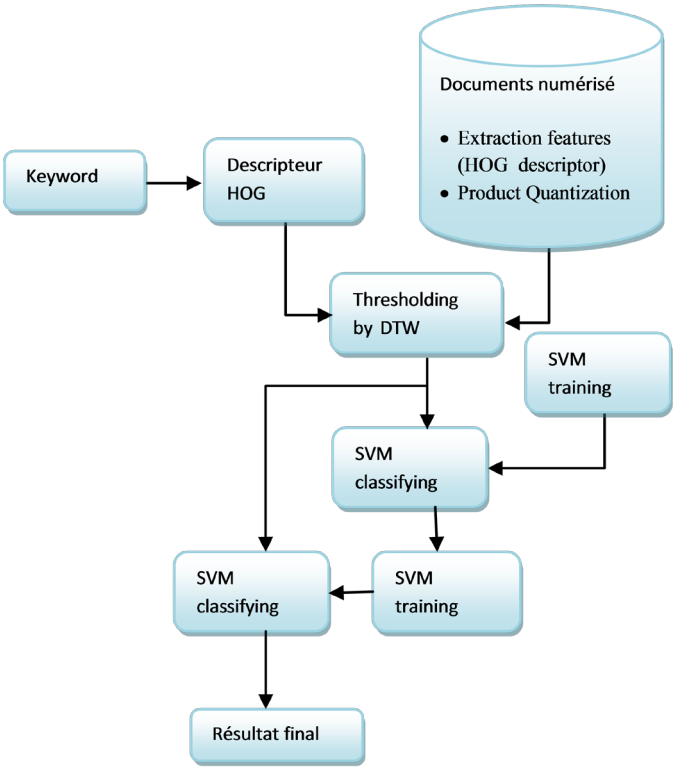


Fig. 2. Process of the proposed system.

combination of these histograms represents the descriptor.

A. Feature extraction

The application of HOG features descriptors and Product Quantization for word spotting in handwritten Arabic documents is the main contribution of this paper. The feature extraction is a most important part of word spotting system, that mean transforming the input query into the set of features. Histogram Oriented Gradient is used to detect and extract feature of Arabic handwritten documents "Fig. 3". Initially, we remove noise for sliding-window and region of documents with Gaussian filter. After smoothing, the Sobel kernel is used to calculate the horizontal and vertical components of the gradients. Let, is the smoothed image and the horizontal and vertical

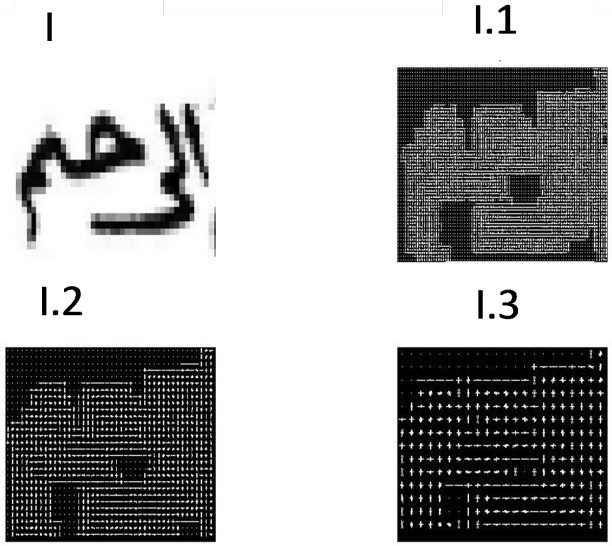


Fig. 3. Rectangle HOG of 2*2, 4*4 and 8*8 block size of Ibn Sina (I1, I2, I3) dataset

components of image gradient is Ix(x,y) and Iy(x,y) respectively.

$$I_x(x,y) = I_s * \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \quad \text{and} \quad I_y(x,y) = I_s * \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

The magnitude M(x, y) and direction D(x, y) of the gradient at pixel (x, y) in the smoothed image are computed as follows:

$$M(x,y) = \sqrt{I_x^2(x,y) + I_y^2(x,y)}$$

$$D(x,y) = \tan^{-1}\left(\frac{I_x(x,y)}{I_y(x,y)}\right)$$

Then, histogram of all blocks can be computed using the block size of character; each pixel is assigned in certain category according to its gradient direction, the feature extraction process is achieved in the following steps:

1. Query
2. Gradients in X and Y direction
3. Dividing query into Blocks
4. Feature extraction
5. Normalize the values
6. Feature concatenation form all blocks
7. Descriptor

B. Product Quantization

The main drawback of sliding-window based methods is the cost of re-computing the descriptors of every image with every new query. However, it's possible to be pre-computing and storing the HOG descriptors, that necessary a large amount of memory to keep them. For this, we propose to encode the HOG descriptors by means of Product Quantization (PQ) proposed by Jégou et al. in [15]. This method both to reduce the amount of memory needed to store the descriptors and reduce the computational cost of searching of the descriptor. This technique has shown excellent results on approximate nearest neighbor tasks. The idea is to decompose the space into a Cartesian product of low-dimensional subspaces and to quantize each subspace separately.

IV. Classification

The main objective of query recognition system is to achieve robust performance to identification of query. In our case we have used SVM classifier with linear function for the recognition. The Width of the margin between the classes is the major optimization criterion, the empty area around the decision boundary, defined by the distance to the nearest training pattern. These patterns called support vectors, which finally define the function for classification.

The kernel linear function used in SVM classification is:

$$K(x_i, x_j) = x_i^T \cdot x_j$$

In a linear model, separating hyper-plane has equation

$$w^T \cdot x + b = 0$$

Considering a binary classification problem with training data:

$\{(x_1, y_1).....(x_i, y_i)\}$ where $x_i \in (\mathrm{N}, \mathrm{P})$ and $y_i \in \{+1, -1\}$

The SVM attempts to find the hyper-plane < w, b > that maximizes the margin.

The positive and negative support vectors respectively is :

$w^T \cdot x_+ + b = +1$ and $w^T \cdot x_- + b = -1$ so

$$\frac{w}{\|w\|} \cdot (x_+ - x_-) = \frac{w^T (x_+ - x_-)}{\|w\|} = \frac{2}{\|w\|}$$

So we can deduce that maximize the margin amounts to minimizing. This can be casted as an optimization problem as:

$$\arg_w \min \frac{1}{2}\|w\|^2 + c_1 \sum_{(x_+, y_+) \in \mathrm{P}} L(y_+ w^T x_+) + c_2 \sum_{(x_-, y_-) \in \mathrm{N}} L(y_- w^T x_-) \; c_{1,2}$$

Is a regularization parameter, $y_+ = +1$ and $y_- = -1$

$x_+ \in \mathrm{P}$ is constructed by deforming the query, To produce the negative set , the sample random regions over all the documents $x_- \in \mathrm{N}.$

V. Experiments

In this section, we present the result of the approach proposed for searching query on Ibn Sina handwritten manuscripts datasets. MATLAB is used to measure all score and running times of the different sections (computing the HOG descriptors, calculating the scores with query and training the SVM). "Table 1" shows the mean average precision of the approach proposed and the time of descriptor computation per image document.

As we see in "Fig. 5" and "Fig. 6", the precision and recall mean recall results depending on the query's size.

$$\text{Precision} = \frac{|(\text{relevant document}) \cap (\text{retrieved document})|}{|(\text{retrieved document})|}$$

$$\text{Recall} = \frac{|(\text{relevant document}) \cap (\text{retrieved document})|}{|(\text{relevant document})|}$$

$$\text{mAP} = \frac{\sum_{q=1}^{Q} P(q)}{Q}$$

TABLE I
TIME OF DESCRIPTOR COMPUTATION AND MEAN AVERAGE PRECISION

Dataset	Time computation	Mean average precision
Ibn Sina	125s	68.4%

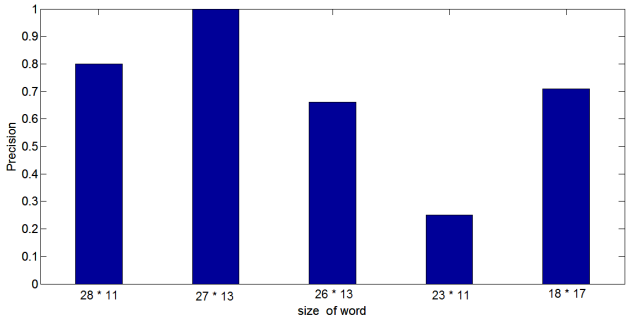


Fig. 4. Precision for different size of word .

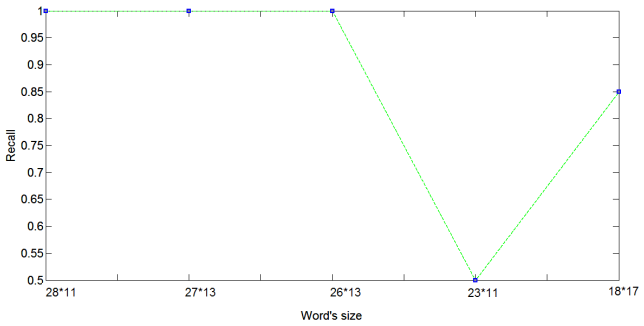


Fig. 5. Recall for different size of word.

VI. Conclusion

In this paper we have presented segmentation-free word spotting method for Handwritten Arabic documents using HOG descriptor. This method has been evaluated on a large amount of handwritten Ibn Sina. The experimental results "Table 1" and "Fig. 5" show that the used of HOG based feature extraction method and SVM classifier with linear function provides good results in mAP and the time of descriptor computation. In future, this work can extend to enhance the performance and reducing the time of descriptor by adding some more relevant features.

References

[1] DEBORA: projet européen n°. LB 5608 A. Coordinateur R. Bouché, juin 2000.179 pages

[2] http://www.hmml.org/eamms/index.html.

[3] CALABRETTO, Sylvie; BOZZI, Andrea; PINON, Jean-Marie, décembre 1999. "Numérisation des manuscrits médiévaux", le projet européen BAMBI, in: Actes du colloque Vers une nouvelle érudition: numérisation et recherche en histoire du livre, Rencontres Jacques Cartier, Lyon.

[4] BURNARD, Lou. ; ROBINSON, PETER. "Vers un standars européen de description des manuscrits", le projet Master. Document numérique. 1999, vol 3, no 1-2, p.151-169.

[5] Y. Leydier, A. Ouji, F. Lebourgeois, H. Emptoz, 2009. "Towards an omnilingual word retrieval system for ancient manuscripts", Pattern Recognit. 42 (2009) 2089–2105.

[6] M.Kamble, S.Hegadi, 2015. "Handwritten Marathi character recognition using R-HOG Feature", in: International Conference on Advanced Computing Technologies and Applications (ICACTA), Procedia Computer Science 45 (2015) 266 – 274015.

[7] T. Rath, V. Lavrenko, and R. Manmatha, 2003. "Retrieving historical manuscripts using shape", Technical Report, Center for Intelligent Information Retrieval Univ. of Massachusetts, Amherst.

[8] J. Almazán, A. Gordo, A. Fornés, E. Valveny, 2014. "Segmentation-free word spotting with exemplar SVMs", Pattern Recognition, 47 (12), pp. 3967–3978.

[9] Y. Elfakir, G. Khaissidi, M. Mrabti, D. Chenouni, "Handwritten Arabic Documents Indexation using HOG Feature," International Journal of Computer Applications", vol. 126, no. 9, pp. 14–18, Sep. 2015

[10] R.A. Mohamad, L. Likforman-Sulem, C. Mokbel, 2009. "Combining slanted-frame classifiers for improved HMM-based Arabic handwriting recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence 31 (7) (2009) 1165–1177.

[11] Y. Kessentini, T. Paquet, A. Ben Hamadou, "Off-line handwritten word recognition using multi-stream hidden Markov models", Pattern Recognition Letters 31 (1) (2010) 60–70.

[12] A. Kundu, T. Hines, J. Phillips, B. Huyck, L. Van Guilder, 2007. "Arabic handwriting recognition using variable duration HMM", in: 9th International Conference on Document Analysis and Recognition (ICDAR), pp. 644–648.

[13] M .Cheriet, R. Farrahi Moghaddam, R. Hedjam, 2013. "A learning framework for automation and optimization of document binarization methods", Computer Vision and Image Understanding 117(3): 269-280.

[14] Y.Elfakir, G. Khaissidi, M. Mrabti, "Traitement des documents anciens par les classificateurs multi-niveaux", Colloque International sur le Monitoring des Systèmes Industriels, ENSA Marrakech, CIMSI14.

[15] H. Jégou, M. Douze, C. Schmid, "Product quantization for nearest neighbor search", IEEE Trans. Pattern Anal. Mach. Intell. 33 (1) (2011) 117–128.

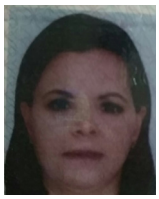
Ghizlane KHAISSIDI, National PhD holder in 2009 of the Sidi Mohamed Ben Abdellah University in Fez in Image Processing and Computer. Currently Professor at the National School of Applied Sciences (ENSA), University USMBA Fez (Morocco) and member of Lab of computering and interdisciplinary physics (L.I.P.I) at (E.N.S.F), Her research activities concern the image processing and its applications in medicine, heritage preservation (indexing of old manuscripts), societal dimension of applications (applications for the blind and visually impaired), handwriting and machine print recognition.

Youssef Elfakir received his Master degrees in 2013, both from the Department of Physics of the Faculty of Sciences Dhars Elmehraz, Fes University Marocco. He is currently a PhD student in Computer Science Engineering at the National School of Applied Sciences. His main research interests are image preprocessing, image analysis of visual manuscript and word spotting.

Mostafa MRABTI, state PhD holder in 1996 of the Sidi Mohamed Ben Abdellah University in Fez Automatic and Signal Processing. Currently Professor at the National School of Applied Sciences (ENSA), University USMBA Fez (Morocco) and member of Lab of computering and interdisciplinary physics (L.I.P.I) at (E.N.S.F),. His research activities concern the error correcting codes, implementation of signal processing algorithms on dedicated circuits, image processing, indexation of old manuscripts, handwriting and machine print recognition..

Mounîm A. El Yacoubi obtained his PhD in Signal Processing and Telecommunications from the University of Rennes I, France, in 1996. During his PhD, he was with the Service de Recherche Technique de la Poste (SRTP) at Nantes, France where he developed software for Handwritten Address Recognition that is still used for real-life Automatic French mail sorting. He was a visiting scientist for 18 months at the Centre for Pattern Recognition and Machine Intelligence (CENPARMI) in Montréal, Canada. He then became an associated professor (1998-2000) at the Catholic University of Parana (PUC-PR) in Curitiba, Brazil. From 2001 to 2008, he was a Senior Software Engineer at Parascript, Boulder (Colorado, USA), a world leader company in automatic processing of handwritten and printed documents (mail, checks, forms). Since June 2008, he is a Directeur d'Etudes / Professor at the University of Paris Saclay, Telecom SudParis within the Intermedia team. His main interests include Machine Learning, Statistical Pattern Recognition, Video, Image and Signal Processing, Human Gesture and Activity recognition, Human Robot (Computer) Interaction, e-Health, Video Surveillance, Biometrics, Smart Cities, Information Retrieval, and Handwriting and Machine Print Recognition.

Driss Chenouni received the Ph.D. degree in physics from the University of Montpellier II, France, in 1989, and the State Doctor's degree from the University of Fes, Morocco, in 1996. He is currently a Lab director of computering and interdisciplinary physics (L.I.P.I), at (E.N.S.F), and a Director of the Ecole Normale Supérieur at Sidi Mohammed Ben Abdellah University (USMBA), Fez, Morocco. His current research interests include Multi-Agent systems, Enterprise Architecture, Modeling, Web services, Autonomic computing, image processing and indexation of old manuscripts.



Zakia Lakhliai received the Ph.D. degree in physics from the University of Montpellier II, France, in 1987, and the State Doctor's degree from the University of Fes, Morocco, in 1996. She is currently a member of Lab of computering and interdisciplinary physics (L.I.P.I), at (E.N.S.F), and a Professor in the Superior School of technology (E.S.T.F.) at Sidi Mohammed Ben Abdellah University (USMBA), Fez, Morocco. Her current research interests signal and image processing, indexation of old manuscripts.

Offline Face Recognition System Based on Gabor-Fisher Descriptors and Hidden Markov Models

Zineb Elgarrai[1], Othmane Elmeslouhi[2], Mustapha Kardouchi[3], Hakim Allali[1], Sid-Ahmed Selouani[4]

[1] FST of Hassan 1st University Settat /LAVETTE Laboratory,
[2] FPO of Ibnou Zohr University /LabSIE Laboratory
[3] Université de Moncton /Département d'Informatique,
[4] Université de Moncton/Département de Gestion de l'Information

Abstract — **This paper presents a new offline face recognition system. The proposed system is built on one dimensional left-to-right Hidden Markov Models (1D-HMMs). Facial image features are extracted using Gabor wavelets. The dimensionality of these features is reduced using the Fisher's Discriminant Analysis method to keep only the most relevant information. Unlike existing techniques using 1D-HMMs, in classification step, the proposed system employs 1D-HMMs to find the relationship between reduced features components directly without any additional segmentation step of interest regions in the face image. The performance evaluation of the proposed method was performed with AR database and the proposed method showed a high recognition rate for this database.**

Keywords— **Face Recognition, Hidden Markov Models, Gabor wavelets, Fisher's Discriminant Analysis**

I. Introduction

Face recognition from still images and video sequences has become an important part of user authentication and security infrastructure in recent years. Face recognition (FR) system consists of two major tasks: face feature extraction and face classification (Fig 1). Both of them, have an important impact on the performances of recognition method.

Feature extraction approaches are classified usually into two main categories: feature-based methods where features are extracted from local facial features and holistic methods where features are extracted from the whole face image [1] [2] [3].

In this paper we used a local face descriptor based on Gabor features [4] [5]. Gabor features have been widely used in face identification because of their good performances in illumination and facial expressions change. This robustness is due to the fact that Gabor kernels captures salient visual properties such as spatial localization, orientation selectivity and spatial frequency characteristic [6][7]. However, Gabor magnitude features have a very high dimensionality which needs long computational time. To overcome this dimensionality issue, the Gabor features magnitude are projected to a subspace using Fisher's Discriminant Analysis method (also known as Linear Discriminant Analysis (LDA)) to select the most discriminative features representing the most important information [5].

To deal with the classification task, many methods have been proposed in this topic. Some approaches are based on neural network [8], SVM [9] and HMMs [10] [11] [12] [13] [14].

The methods based on HMMs can be classified roughly into three categories: 1D-HMMS [10] [14], pseudo 2D- HMMs [13] and full 2D-HMMs [11] [15]. In this paper we propose a new way based on 1D-HMMs for the face classification step. The proposed approach is somewhat different from the conventional 1D-HMMs methods, by the fact that it works well without the need to previously localize significant facial regions like eyes, nose, mouth, etc. Thereby, the method can be used even if there are no frontal images in the database. The rest of the paper is organized as follows. In Section 2, Gabor-Fisher features are briefly resumed, and, in Section 3, the proposed HMM approach is described. Experimental results are presented in Section 4. Finally, in Section 5, conclusions are drawn.

II. Features Extraction

Gabor wavelets (also called Gabor kernels or filters) have proven themselves to be a powerful tool for facial image feature extraction and recognition. The 2D Gabor wavelet can be represented by two components: a real and an imaginary components [4]:

-real component:

$$g(z,\lambda,\theta,\varphi,\sigma,\nu)=\exp(\frac{x'^2+\nu^2y'^2}{2\sigma^2})\cos(2\pi\frac{x'}{\lambda}+\varphi) \quad (1)$$

-imaginary component:

$$g(z,\lambda,\theta,\varphi,\sigma,\nu)=\exp(\frac{x'^2+\nu^2y'^2}{2\sigma^2})\sin(2\pi\frac{x'}{\lambda}+\varphi) \quad (2)$$

where x'= x cos θ + y sin θ and y' = x sin θ + y cos θ and (x , y) is the coordinate of a pixel in the image plan .and λ represents the wavelength of the sinusoidal factor, θ represents the orientation of a Gabor function, φ is the phase offset, σ is the standard deviation of the Gaussian envelope and is the spatial aspect ratio.

The Fig. 2 shows the result of the convolution of a facial image with 40 banks of Gabor kernel (8 orientations and 5 scales). However, for a given face image, even for a small face image of, for example, 64 × 64 pixels, the 40 magnitude responses reside in a 163840 (64 ×64 × 40) dimensional space, which is far too extensive for efficient processing and storage. To overcome this dimensionality issue, we use the fisher's discriminant analysis (also called LDA) dimensionality reduction technique to project the Gabor magnitude feature vectors into a subspace where between-class variations of the projected patterns are maximized while within-class variations are minimized [16]. To avoid singularity issues, when computing the inverse of the within-class scatter matrix, the LDA reduction method is implemented in the Principal Component Analysis PCA subspace as suggested in [17]. The resultant descriptors are called: Gabor-Fisher descriptors. We notice that for the LDA, the best features dimension is the number of

DOI: 10.9781/ijimai.2016.412

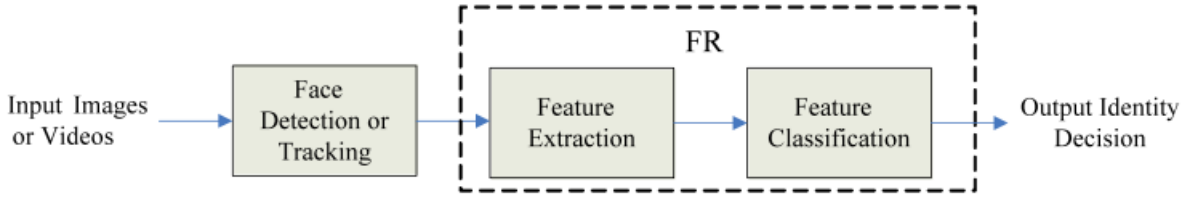


Fig. 1. Face recognition system architecture.

subject minus one [18]. Finally, each Gabor-Fisher feature obtained is sampled to small sequences with a size L. By this way, we avoid any supplementary segmentation of face regions as depicted in previous face recognition methods based on 1D-HMMs.

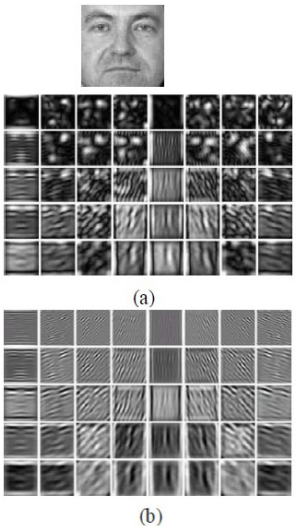


Fig. 2. An example of the Gabor magnitude and phase output: (a) the magnitude and (b) the phase output of the filtering operation with a bank of 40 Gabor filters

III. Features Classification

An HMM is a Markov chain with a finite number of unobservable states [25]. Although the Markov states are not directly observable, each state has a probability distribution associated with the set of possible observations. As mentioned above, 1D-HMMS approach is used in the recognition step. Generally, the elements of 1D-HMMs are defined by the triplet $\lambda = \{A, B, \Pi\}$ where, A is an $N \times N$ state transition matrix that gives the state transition probabilities between N states, B is an $k \times N$ emission probability matrix while being in a particular state, and Π is an $1 \times N$ matrix, called initial state probability matrix, and it gives the probability of being in a particular state at the start of the process. To use the proposed system we need three steps:

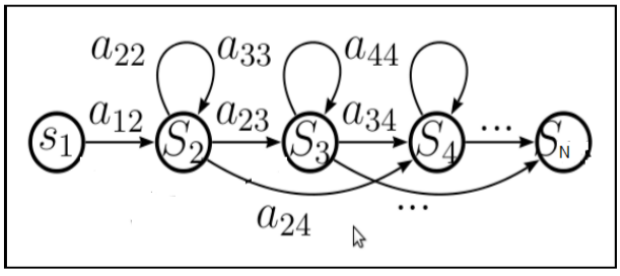


Fig. 3. An example of a Left-Right 1D-HMM with N states

A. System configuration

As mentioned above, we use a 1D left to right Hidden Markov Model (see Fig.3). In this model, the number of states N is equal to the number of classes in dataset minus one divided by L (size of sampled features obtained in the features extraction step). The entry state 1 and the exit state N are non-emitting. The transition matrix A will have N rows and N columns. We assume that each observation probability distribution is represented with single Gaussian distributions. In this case the density of the observation u_t in state i is thus:

$$b_i(u_t) = \frac{1}{\sqrt{(2\pi)^k \det(\sum_i)}} \exp\left(\frac{-1}{2}(u_t - u_i)^t \sum_i^{-1} (u_t - u_i)\right) \quad (3)$$

Where k is the dimension of u_t, and where u_i and $\sum_i$ are the mean vector and covariance matrix, respectively.

Finally, the Gabor-Fisher features will be divided into two sets: training and testing sets.

B. Training step

During this step, the Gabor-Fisher features training set will be injected in the system to estimate the face model parameters for each subject. These parameters can be estimated using an iterative procedure, known as the Baum–Welch Algorithm (BWA) [19].

C. Recognition step

It follows the training step. In this step, the proposed system utilizes the Viterbi algorithm [19] to find the highest likelihood score between features in the testing sets and models obtained in the training step.

IV. Expirimental Result

In this section, some experiment results will be given to evaluate the proposed method. The cropped AR face database [20] is used for this purpose. The AR face database contains over 3,200 frontal color face images of 126 subjects (26 different images for each person), including different facial expressions, with various occlusions and under different lighting conditions. Most of the pictures were recorded in two sessions (separated by two weeks). All images were taken by the same camera under tightly controlled conditions of illumination and viewpoint.

For our experiments, like in the work of [20], in the first time, 100 different subjects (50 males and 50 females) were randomly selected from this database, then, all selected images were segmented using an oval-shaped mask and finally all color images are transformed into gray images. Fig. 4 shows some sample images extracted from the obtained face database.

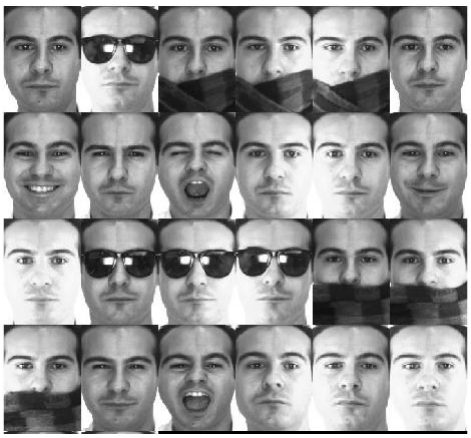


Fig. 4. Sample images from the AR face database

In this experiment part, we use Hidden Markov Models Toolkit (HTK) [21]. The HTK toolkit is builded initially for speech recognition system. We followed the same steps as in [22] to adapt this toolkit to the face recognition domain.

The first set of experiments was performed to discover the optimal number of states per model. To find the best value of this parameter, tables 1 show recognition rates obtained when using different number of states. Obtained results show that the best recognition rate is obtained when the number of states equals 18 for AR database and it shows high sensitivity when moving from 18 states to 19 states.

Furthermore, to prove the effectiveness of the proposed approach we vary sizes of testing and training sets. Table 2 illustrates the recognition rates for various training sizes.

From the above results we can see that for the AR database the recognition rate reaches the best value (i.e 100%) when the size of the training sets is equal to 15.

Finally, we make a comparison between the gabor descriptor with others descriptors like: HOG [23] and LBP [24]. We notice that the size of features of those three descriptors is reduced using Fisher's Discriminant Analysis method. The comparison results are shown in table 3. From this table, we notice that the results obtained by the gabor filters outperform those obtained using HOG and LBP descriptors in the AR datasets.

TABLE I
RECOGNITION RATE FOR DIFFERENT NUMBER OF STATES FOR AR DATABASE.

Number of states	RECOGNITION RATE (%)
13	66.55
15	79.45
17	96.91
18	100
19	none

TABLE II
RECOGNITION RATE VS NO. OF TRAINING SAMPLES

No. of Training Samples Recognition Rate	No. of Training Samples Recognition Rate
9	94,35%
12	98,14%
14	99,83%
15	100%
9	94,35%

TABLE III
FACE RECOGNITION RATES USING DIFFERENTS DESCRIPTORS WITH 15 TRAINING SAMPLES FOR AR DATASETS.

Descriptors Recognition Rate	Descriptors Recognition Rate
LBP	94 ,18%
HOG	98 ,12%
Gabor-Fisher	100%

V. Conclusion

A new face recognition system based on 1D-HMM was presented in this paper. In the first, the system extract facial features using Gabor filters and reduce their dimensionality using Fisher's Discriminant Analysis method. Secondly the resultant descriptors are resampled and injected in a 1D- HMM to achieve the training and recognition steps. The AR standard database is used to evaluate the proposed system. The recognition rate reaches 100%.

Our future research will be focused on testing the performances of the proposed system in low resolution facial images in order to integrate it in real applications such as surveillance.

References

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," ACM Comput. Surv., vol. 35, no. 4, pp. 399–458, Dec. 2003.

[2] K. W. Bowyer, K. Chang, and P. Flynn, "A survey of approaches and challenges in 3d and multi-modal 3d + 2d face recognition," Comput. Vis. Image Underst., vol. 101, no. 1, pp.

[3] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2d and 3d face recognition: A survey," Pattern Recogn. Lett., vol. 28, no. 14, pp. 1885–1906, Oct. 2007.

[4] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," Trans. Img. Proc., vol. 11, no. 4, pp. 467–476, Apr. 2002.

[5] V.Struc and N. Pavesic, "The complete gabor-fisher classifier for robust face recognition," EURASIP Advances in Signal Processing, vol. 2010, p. 26, 2010.

[6] U. Park, Y. Tong, and A. K. Jain, "Age-invariant face recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 5, pp. 947–954, May 2010.

[7] J. Sung and D. Kim, "Pose-robust facial expression recognition using view-based 2d + 3d AAM," IEEE Transactions on Systems, Man, and Cybernetics, Part A, vol. 38, no. 4, pp. 852–866, 2008.

[8] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: A convolutional neural-network approach," Trans. Neur. Netw., vol. 8, no. 1, Jan. 1997.

[9] G. Guo, S. Z. Li, and K. L. Chan, "Face recognition by support vector machines," in Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on. IEEE, 2000, pp. 196–201.

[10] F. Samaria, "Face recognition using hidden markov models," Phd thesis, Cambridge University, 1994.

[11] H. Otluman and T. Aboulnasr, "Low complexity 2-d hidden markov model for face recognition," in Circuits and Systems, 2000. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium on, vol. 5. IEEE, 2000, pp. 33–36.

[12] V. Kohir and U. Desai, "Face recognition using dct-hmm approach," in Workshop on Advances in Facial Image Analysis and Recognition Technology (AFIART) Germany, 1998.

[13] A. V. Nefian, M. H. Hayes, and III, "An embedded hmm-based approach for face detection and recognition," in In Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, 1999, pp. 3553–3556.

[14] M. Bicego, U. Castellani, and V. Murino, "Using hidden markov models and wavelets for face recognition," in Proceedings of IEEE International Conference on Image Analysis and Processing (ICIAP03, 2003, pp. 52–56.

[15] H. Othman and T. Aboulnasr, "A simplified second-order HMM with application to face recognition," in International Symposium on Circuits and Systems (ISCAS 2001), 6-9 May 2001, Sydney, Australia, 2001, pp. 161–164.

[16] T. Savic and N. Pavesic, "Personal recognition based on an image of the palmar surface of the hand," Pattern Recogn.,vol. 40, no. 11, pp. 3152–3163, 2007.

[17] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman,"Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," IEEE Transactions On Pattern Analysis and Machine Intelligence, vol. 19, pp. 711–720, 1996.

[18] M. Loog, B. van Ginneken, and R. Duin, "Dimensionality reduction by canonical contextual correlation projections," in Computer Vision - ECCV 2004, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2004, vol. 3021.

[19] L. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," Proceedings of the IEEE,vol. 77, no. 2, pp. 257–286, 1989.

[20] A. M. Martınez and A. C. Kak, "Pca versus lda," IEEE Trans. Pattern Anal. Mach. Intell., pp. 228–233, 2001.

[21] S. Young, G. Evermann, M. Gales, T. Hain, D. Kershaw, X. Liu, G. Moore, J. Odell, D. Ollason, D. Povey et al.,"The htk book (for htk version 3.4)," Cambridge university engineering department, vol. 2, no. 2, pp. 2–3, 2006.

[22] Z. Elgarrai, O. E. Meslouhi, H. Allali, M. Kardouchi, and S.-A. Selouani, "Face recognition system using gabor features and htk toolkit," in Proceedings of the 2014 Tenth International Conference on Signal-Image Technology and Internet-Based Systems, ser. SITIS '14, 2014, pp. 32–36.

[23] O. Deniz, G. Bueno, J. Salido, and F. De la Torre, "Face recognition using histograms of oriented gradients," Pattern Recogn. Lett., vol. 32, no. 12, pp. 1598–1603, 2011.

[24] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 7, 2002.

[25] H. Bolivar, M. Martínez, R. González-Crespo, O. Sanjuán. "Complexity analysis of a matchmaker based on hidden Markov model for decentralized grid scheduling", vol. 5, no. 3, pp. 190-197., 2014.

Zineb Elgarrai received her engineer degree in Computer science in 2009 at the National School of computer science and systems analysis (ENSIAS)
School, Rabat, Morocco. In 2012 she joined the LAVETE Laboratory of FST of Hassan 1st University, Settat, Morocco. Her actual main research interests concern Face Recognition and Computer Vision. E-mail: elgarrai@gmail.com

Othmane Elmeslouhi is an Associate Professor in the Mathematics and Computer Science Department of the Polydisciplinary Faculty Ouarzazate of Ibnou University.
He has a PhD from the University of 1st University Settat in computer engineering. His research interests include Computer Vision, Pattern Recognition and Medical Imaging. E-mail: o.elmeslouhi@uiz.ac.ma

Mustapha Kardouchi is a full professor at the Computer Science Departement, Moncton University, Canada. His research interests are primarily in the field of Computer Vision with a focus on Image Recognition and Video Analysis.
E-mail : mustapha.kardouchi@umoncton.ca

Sid Ahmed Selouani is a full professor at the Information Management Departement, Moncton University, Canada. His research interests are Speech Recognition and their applications.
E-mail: sid-ahmed.selouani@umoncton.ca

Hakim Allali was born in Morocco on 1966. He received the Ph.D degree from Claude Bernard Lyon I University (France) in 1993 and the "Docteur d'Etat" degree from Hassan II-Mohamedia University, Casablanca (Morocco) in 1997. He is currently Professor at Faculty of Sciences and Technologies of Hassan 1st University of Settat (Morocco) and director of LAVETE Laboratory.
He is executive manager and founder of IT Learning Campus. His research interests include technology enhanced learning, modeling, image processing, computer networking and GIS.E-mail: hakim-allali@hotmail.fr

Golden Ball Algorithm for solving Flow Shop Scheduling Problem

Fatima Sayoti, Mohammed Essaid Riffi

LAROSERI Laboratory, Dept. of Computer Science, Faculty of Sciences, University of Chouaib Doukkali, El Jadida, Morocco

***Abstract* — The Flow Shop Scheduling Problem (FSSP) is notoriously NP-hard combinatorial optimization problem. The goal is to find a schedule that minimizes the makespan. This paper proposes an adaptation of a new approach called Golden Ball Algorithm (GBA). The proposed algorithm has been never tested with FSSP; it's based on soccer concept to obtain the optimal solution. Numerical results are presented for 22 instances of OR-Library. The computational results indicate that this approach is practical for small OR-Library instances.**

***Keywords* — Combinatorial Optimization, Flow Shop Scheduling Problem, Golden Ball Metaheuristic, Makespan Time, Metaheuristics.**

I. Introduction

THE flow shop scheduling problem (FSSP) [1] is to schedule a set of n jobs on a set of m machines. The both parameters n and m are given. Each job consists of a chain of operations. All machines should process the jobs in the same order of jobs.

The objective is to schedule jobs in such a way as to minimize the maximum of the completion time of all the jobs (makespan). In the FSSP some constraints must be satisfied such as:

- All jobs are independent and ready for processing at time zero
- No n jobs may be processed at the same time on the same machine
- No preemption of a given job
- The precedence relations have to be respected
- The makespan of all jobs should be minimized

The FSSP is one of the most difficult combinatorial optimization problem and it belongs to NP-hard problems [2].Finding a good solutions to the FSSP is of great interest to the industrial sector.

Several evolutionary methods are used to solve the flow shop scheduling problem such as: genetic algorithms [3]-[4], ant colony optimization [5], partial swarm optimization [6]-[7], Tabu search method [8]-[3] etc.

In this paper we present an adaptation of the Golden Ball algorithm (GBA) to the flow shop scheduling problem (FSSP). This algorithm is based on the soccer concepts. The objective of this approach is to produce satisfactory results more near optimal and in less time.

The rest of this paper is organized as follows: In section II, flow shop scheduling problem formulation. In section III, the golden ball metaheuristic. In section IV, the golden ball adaptation. In section V, experimental results on 22 OR-Library instances [9] and finally a conclusion.

II. Flow Shop Scheduling Problem Formulation

A flow shop scheduling problem (FSSP) can be defined by a set J of n jobs J= {J1,…, Jn} which have to be processed on a set M of m machines M= {M1,…., Mm}. Each job consists of a series of operations Oik, where k indicates the machine Mk on which the operation must be processed and i defines the job to which the operation belongs. Each operation needs to be processed during an uninterrupted period of time on a given machine pik. Every machine processes the jobs in a fixed order of a given jobs. Each machine may process only one operation during the period of time. The objective of FSSP is to minimize the makespan Cmax of the whole process and to find the optimal schedule. Cmax is the completion time of all jobs.

We consider an example of FSSP with 3 machines and 4 jobs where: M= {M1, M2, M3}, J= {J1, J2, J3, J4},

J1= {O11, O12, O13}

J2= {O21, O22, O23}

J3= {O31, O32, O33}

J4= {O41, O42, O43},

Each machine processes the jobs in this order: {J1, J2, J4, J3}

TABLE I shows the processing time of all operations:

TABLE I
OPERATIONS PROCESSING TIMES

JOBS	OPERATION TIMES		
J1(O11, O12, O13)	6	2	5
J2(O21, O22, O23)	2	3	2
J3(O31, O32, O33)	4	7	3
J4(O41, O42, O43)	1	3	1

The makespan is calculated using the Gantt chart representation (Fig.1):

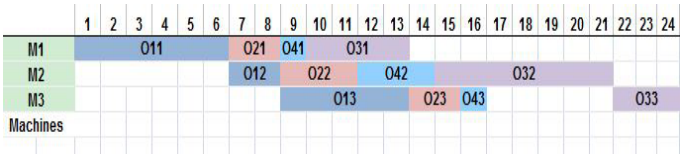


Fig. 1. Gantt chart representation for the schedule {J1, J2, J4, J3}

In this example the best schedule obtained is:

{J1, J4, J3, J2}

with a minimal makespan Cmax=23.

In this section we describe the operation of a new metaheuristic recently proposed called Golden Ball (GB). This method is based on several concepts of soccer to find the optimal solution. It was proposed firstly in 2013 by E.Osaba et al. [10].The recent version of GB was published in 2014 by the same authors [11]. This technique divides the initial solutions into groups. Each group represents a team. Each team works independently and competes with other teams to get the best solution.

Golden Ball algorithm is based on 4 main phases (Fig.2): Initialization phase, Training phase, Competition phase and Transfer phase.

DOI: 10.9781/ijimai.2016.413

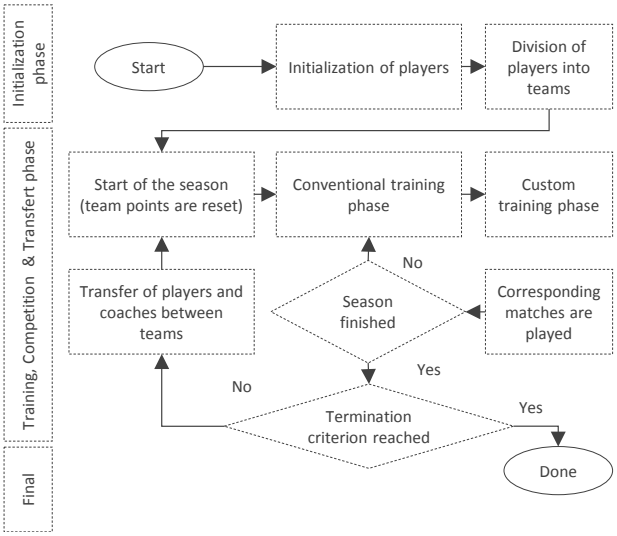


Fig. 2. Flowchart of GB metaheuristic

III. The Golden Ball Adaptation

In this section we present the equivalence of main terms used to solve the FSSP with the GBA.

Player: Schedule

Team: Group of schedules

NT: Number of groups of schedules

NP: Number of schedules per group

Quality: Completion time of schedule (Cmax)

Strength value: Average completion time of each group, it is equal to the sum of all Cmax divided by NP

Coach: Training function

Captain: Best schedule of the group

There are two kinds of training function: Conventional trainings and Custom trainings.

We use the following methods as conventional training functions:

2-opt [12]-[13]: a local search operator highly used by several researches to solve TSP. Its goal is to improve the path by replacing in each step of 2-opt two edges by two other edges.

Insertion method [14]: choose and extract a random operation, insert it in position k between i and j in a way that the whole schedule is improved.

Swapping mechanism [15]:

- First possible swap: select and swap two random operations.
- Second possible swap: reverse the operations order between two random positions.
- Third possible swap: divide the path into three parts. Copy the last part into the first part of the new schedule. Copy the second part into the second part of the new schedule and copy the inverse of the first part into the last part of the new schedule.

As a custom training function the proposed adaptation uses the **Ordered Crossover** (OX) [16].

Step1: select randomly two positions pos1 and pos2 to select the substring to be copied.

Step2: copy the substring from the parent into the corresponding position of the child schedule.

Step3: starting with the pos2, select all operations which are not already in the substring from the second parent

Step4: place these operations into the positions of the child schedule from the left to the right according to the order of the sequence.

In the competition phase we order the schedules of each group according to their makespan in ascending order.

In the transfer phase especially in the season transfer all groups of schedules must be sorted according to their average cost in ascending order.

GB ALGORITHM STEPS

1. Determine the value of NT and NP.
2. Generate NT*NP random schedules
3. To each group assign randomly NP schedules and a training function.
4. For each group find its best schedule and calculate its average completion time.
5. Start of the season and points initialized to 0.
6. Start the training session for each group.
7. If the session is finished, sort schedules of each group according to their Cmax in ascending order. Else, start the transfer phase if it is necessary and go to step 6.
8. Compare each schedule of the group with another existing in other group chosen randomly. The group who has the better schedule receives 3 points. If the two schedules are equal the both groups receive 1 point.
9. If the season is not finished, go to step 4.

   Else,

   If the optimal solution is found, stop the program.

   Else,

   - sort all groups according to their points and average completion time in ascending order.
   - exchange schedules and training functions between groups.
   - go to step 4.

IV. Numerical Results

The program is tested on different instances of OR-library. The GB algorithm was implemented in C language and compiled using Microsoft Visual Studio 2008, the program code was executed in computer with Genuine Intel(R) 575 @ 2.00 GHz 2.00 GHz RAM 2,00 Go.

The program uses four training functions and two parameters: NT (number of groups), NP (number of schedules per group. NP is assumed as constant. The effect of NT parameter is evaluated in the test problems with 6 various values.

The parameters values in the TABLE II produce better results during the algorithm run. The results obtained for the different values of NT parameter are shown in the TABLE III. The application is run five times for each instance. The program stops after 60s.

TABLE II
PARAMETERS VALUES

NT	7
NP	4
Maximum execution time of the program	500 seconds

TABLE III
RESULTS OBTAINED FOE EACH INSTANCE

NT	CAR1	CAR2	CAR3	REC01	REC03	REC05	REC07
2	7038	7166	7347	1259,4	1110,8	1256,8	1600,6
3	7038	7166	7331,6	1258,6	1112,6	1259,4	1611,2
4	7038	7166	7312	1264,8	1113,4	1253,8	1596
5	7038	7166	7333,6	1256,8	1113,4	1251,8	1596,8
6	7038	7166	7324	1250,4	1111,4	1253,4	1590,2
7	7038	7166	7322,8	1251,8	1111,2	1247	1588,4

TABLE IV
COMPUTATIONAL RESULTS OF OR-LIBRARY INSTANCES

Problem			**Golden Ball Algorithm**					
Instance	n × m	BKS	Best	Nbest	Worst	Average	%Error	Time
Car1	11× 5	7038	7038	10	7038	7038	0,0000	0
Car2	13× 4	7166	7166	10	7166	7166	0,0000	0
Car3	12× 5	7312	7312	10	7312	7312	0,0000	0
Car4	14× 4	8003	8003	10	8003	8003	0,0000	0
Car5	10× 6	7720	7720	10	7720	7720	0,0000	0
Car6	8× 9	8505	8505	10	8505	8505	0,0000	0
Car7	7× 7	6590	6590	10	6590	6590	0,0000	0
Car8	8× 8	8366	8366	10	8366	8366	0,0000	0
Hel1	100× 10	516	525	01	537	531,1	2,9263	500
Hel2	20× 10	136	137	02	140	138,4	1,7647	500
Rec01	20× 5	1247	1249	09	1251	1249,2	0,1764	13
Rec03	20× 5	1109	1110	02	1117	1111,8	0,2524	11
Rec05	20× 5	1242	1245	09	1253	1245,8	0,3059	19
Rec07	20× 10	1566	1572	02	1584	1579,8	0,8812	500
Rec09	20× 10	1537	1553	01	1583	1567,9	2,0104	500
Rec11	20× 10	1431	1438	01	1472	1460,5	2,0614	500
Rec13	20× 15	1930	1951	01	1993	1969,3	2,0362	500
Rec15	20× 15	1950	1982	01	2013	1995,1	2,3128	500
Rec17	20× 15	1902	1931	01	1964	1946,8	2,3554	186
Rec19	30× 10	2093	2123	01	2187	2153,4	2,8858	500
Rec21	30× 10	2017	2054	01	2081	2070,3	2,6425	500
Rec23	30× 10	2011	2058	01	2101	2071,9	3,0283	500

The table above (TABLE IV) represents the following information:

BKS: Best known Solution

Best: Best schedule

N_{Best}: The number of times the algorithm reaches the best schedule

Worst: The worst schedule

Average: The average cost

%Error: The percent error is calculated as follows:

$$\%Error = \frac{\text{Average} - \text{BKS}}{\text{BKS}} \times 100$$

The application is run ten times for each test instance.

The program stops when the best solution is reached. The maximum execution time of the application is 500s.

The following table (Table V) compares the proposed approach with other existed algorithms in the literature of flow shop scheduling problem such as Palmer [17] CDS [18] and NEH [19].

TABLE V
COMPARISON OF GB ALGORITHM WITH ALGORITHMS IN THE LITERATURE OF FSSP

Problem			***C_{MAX}***			
Instance	***n × m***	***BKS***	***GBA***	***Palmer***	***CDS***	***NEH***
Car1	11× 5	**7038**	**7038**	7472	7202	7038
Car2	13× 4	**7166**	**7166**	7940	7410	7376
Car3	12× 5	**7312**	**7312**	7725	7399	7443
Car4	14× 4	**8003**	**8003**	8423	8423	8003
Car5	10× 6	**7720**	**7720**	8520	8627	8090
Car6	8× 9	**8505**	**8505**	9487	9553	9079
Car7	7× 7	**6590**	**6590**	7639	6829	7468
Car8	8× 8	**8366**	**8366**	9023	8903	8967
Rec01	20× 5	**1247**	**1249**	1391	1399	1334
Rec03	20× 5	**1109**	**1110**	1223	1273	1136
Rec05	20× 5	**1242**	**1245**	1290	1338	1294
Rec07	20× 10	**1566**	**1572**	1715	1697	1637
Rec09	20× 10	**1537**	**1553**	1915	1639	1692
Rec11	20× 10	**1431**	**1438**	1685	1597	1635

As the results show, the GBA is an effective algorithm for the flow shop scheduling problem.

V. Conclusion

This paper presents a new approach called Golden Ball algorithm (GBA) for the flow shop scheduling problem (FSSP). This proposed technique is based on soccer concept to find the optimal schedule with a minimal makespan. GBA is able to very quickly find the optimal schedule for the small flow shop schedule problem. For the other OR-Library instances, the algorithm produces results near optimal. The future work may be to increase the performance and the quality of the proposed adaptation for the large instances.

References

[1] J.ND.Gupta, et E.F.Stafford. "Flowshop scheduling research after five decades". European Journal of Operational Research, vol. 169, no 3, p. 699-711, 2006.

[2] S.Sahni, et T.Gonzalez. "P-complete approximation problems". Journal of the ACM (JACM), vol. 23, no 3, p. 555-565, 1976.

[3] J.Knox, et F.Glover. "Comparative Testing of Traveling Salesman Heuristics Derived from TABU Search, Genetic Algorithms, and Simulated Annealing". publisher not identified, 1989.

[4] John Henry Holland. "Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence". MIT press, 1992.

[5] A.Colorni, M.Dorigo, V.Maniezzo, et al. "Distributed optimization by ant colonies". In : Proceedings of the first European conference on artificial life. p. 134-142. 1991.

[6] B.Jarboui, S.Ibrahim, P.Siarry, et al. "A combinatorial particle swarm optimisation for solving permutation flowshop problems". Computers & Industrial Engineering, vol. 54, no 3, p. 526-538, 2008.

[7] M.F.Tasgetiren, M.Sevkli, Y.C.Liang, et al. "Particle swarm optimization algorithm for permutation flowshop sequencing problem". In : Ant Colony Optimization and Swarm Intelligence. Springer Berlin Heidelberg. p. 382-389, 2004.

[8] F.Glover. "Future paths for integer programming and links to artificial intelligence". Computers & operations research, vol. 13, no 5, p. 533-549, 1986.

[9] J.E.Beasley. "OR-Library: distributing test problems by electronic mail". Journal of the operational research society, p. 1069-1072, 1990.

[10] E.Osaba, F.Diaz, et E.Onieva. "A novel meta-heuristic based on soccer concepts to solve routing problems". In : Proceedings of the 15th annual conference companion on Genetic and evolutionary computation. ACM. p. 1743-1744, 2013.

[11] E.Osaba, F.Diaz, R.Carballedo, et al. "Focusing on the Golden Ball Metaheuristic: An Extended Study on a Wider Set of Problems". The Scientific World Journal, vol. 2014, 2014.

[12] G. Croes, "A method for solving traveling-salesman problems. Operations Research", Vol. 6, no 6, pp. 791-812, 1958.

[13] S.LIN. "Computer solutions of the traveling salesman problem. Bell System Technical Journal, The, vol. 44, no 10, p. 2245-2269, 1965.

[14] M.Fischetti, J.J.Salazar Gonzalez, et P.Toth, "A branch-and-cut algorithm for the symmetric generalized traveling salesman problem". Operations Research, Vol. 45, no 3, pp. 378-394, 1997.

[15] C.D.TARANTILIS." Solving the vehicle routing problem with adaptive memory programming methodology". Computers & Operations Research, vol. 32, no 9, p. 2309-2327, 2005.

[16] L.DAVIS. "Applying adaptive algorithms to epistatic domains". In : IJCAI. p. 162-164. 1985.

[17] D.PALMER. "Sequencing jobs through a multi-stage process in the minimum total time--a quick method of obtaining a near optimum". OR , p. 101-107, 1965.

[18] H.G.CAMPBELL, R.A. DUDEK, et M.L.SMITH. "A heuristic algorithm for the n job, m machine sequencing problem. Management science", vol. 16, no 10, p. B-630-B-637, 1970.

[19] M.NAWAZ, E.E.ENSCORE, et I.HAM. "A heuristic algorithm for the m-machine, n-job flow-shop sequencing problem". Omega, vol. 11, no 1, p. 91-95, 1983.

Ms. Fatima Sayoti (15-02-1988) received a master's degree of software quality from University of Chouaib Doukkali, El Jadida, Morocco in 2011. She is currently pursuing his Ph.D. degree (Computer Science) at the University of Chouaib Doukkali, Faculty of Sciences, El Jadida, Morocco. She is also a temporary teacher at Polydisciplinary Faculty, El Jadida, Morocco.

A Self-Calibration Method of Zooming Camera

Ismail El batteoui[1], Abderrahim Saaidi[1,2], and Khalid Satori[1]

[1]LIIAN, Department of Mathematics and Informatics, Faculty of sciences Dhar-Mahraz B.P1796 Atlas Fez, University of Sidi Mohamed Ben Abdellah- Morocco
[2]LSI, Department of Mathematics, Physics and Informatics, Polydisciplinary faculty. B.P 1223, Taza University of Sidi Mohamed Ben Abdellah - Morocco

***Abstract* — In this article we proposed a novel approach to self-calibrate a camera with variable focal length. We show that the estimation of camera's intrinsic parameters is possible from only two points of an unknown planar scene. The projection of these points by using the projection matrices in two images only permit us to obtain a system of equations according to the camera's intrinsic parameters . From this system we formulated a nonlinear cost function which its minimization allows us to estimate the camera's intrinsic parameters in each view. The results on synthetic and real data justify the robustness of our method in term of reliability and convergence.**

***Keywords* — Variable Focal Length, Unknown Planar Scene, Projection Matrices, Nonlinear Cost Function.**

I. Introduction

The self-calibration of camera [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16] is a fundamental area in computer vision, it consist of estimation of camera's parameters with no knowledge about the observed scene. The principle of self-calibration consist to formulate equations based on the camera's intrinsic parameters by using invariant primitives (circular points, absolute conic...) in the scene space. These primitives describe the euclidean structure of the scene.

Our self-calibration method presented in this paper consist of calibrate a camera from a unknown planar scene and by considering that the camera's intrinsic parameters are constants except to the focal length which may be variable from one view to another which is suitable with the current applications of computer vision. The highlight of our idea lies partly in the use of only two points from an unknown plananar scene and the projection of this scene in only two image planes and secondly in the implementation of a relationship between images of the absolute conic in the both views and the projection matrices, to determinate, of the two points used in the scene. The procedure of our method of self-calibration described in this work can be presented in the form of the following steps:

- Acquisition of two images of an unknown planar scene.
- Detection of interest points in the two images by using the Harris detector.
- The matching of the interest points by using the ZNCC correlation measure
- Estimation of the projection matrices of the two points of the scene
- Development of a relationship between the projection matrices estimated in the previous step and the images of the absolute conic in the two images of the scene.
- Formulation of a non-linear cost function on the basis of the relation found in the previous step.
- Optimization of the cost function by the Levenberg-Marquardt [17] through two phases:
  - Initialization phase providing an initial solution by applying constraints on the two images and some intrinsic parameters of the camera.
  - Minimizing of the cost function to find the optimum solution according to the intrinsic parameters of our camera in the two images of the scene.

The present article is organized as follows: the second part has a review of previous works dealing with the problem of self-calibration. The self-calibration tools used in this article are discussed in the third section then our method of self-calibration is addressed in the fourth section. Our experimental results are presented in the fifth section and the last part contains a conclusion of our work.

II. Related works

In this section we present with a non-exhaustive manner some works done in the field of the camera self-calibration. In [18] the singular value decomposition of the fundamental matrix gives a simple form of Kruppa equations that solves the problem of self-calibration of cameras with varying intrinsic parameters. A new method for camera self-calibration performing a sequence of critical motions is processed in [19]. In this method, the intrinsic parameters are assumed to be known except to the focal length which may vary freely from one view to another. The determination of the intrinsic parameters is based on the use of the image of the absolute conic in the different views. In [20], the authors are proposed a new method of self-calibration from a sequence of images of an unknown 3D scene. The projection of the scene points in the image planes is used with the fundamental matrix to determine the projection matrices of scene points. These projection matrices are used to formulate a nonlinear cost function which resolution can estimate the intrinsic parameters in the different views. In [21], the resolution of three linear equations obtained from circles and their centers respectively determine the vanishing line. This vanishing line and the circular points are used to estimate the camera's intrinsic parameters. The problem addressed in [22] is the self-calibration of the cameras with varying intrinsic parameters, This method is based on a quasi-affine reconstruction (by localization the plane in infinity) which provides a system of equations based on the intrinsic parameters of the camera. The resolution of this system gives the intrinsic parameters in each view. In [23] the authors have proposed a linear approach to self-calibrate a camera performing a small rotation and general translation with the possibility of variation of the focal length. This method is based on using the homography induced by the plane at infinity and the absolute quadric. In [24], a method of camera self-calibration with varying intrinsic parameters is presented. It's based on the quasi-affine reconstruction, after thissreconstruction the homography of the plane at infinity is determined. This homography is used with constraints on the image of the absolute conic to estimate the intrinsic parameters of the camera [29].

DOI: 10.9781/ijimai.2016.414

III. Camera model and image of absolute conic

Our camera is modeled by a model called pinhole. Denote by $M = (X\ Y\ Z\ w)^T$ a point in homogeneous coordinates of the scene, $m = (x\ y\ z\ w)^T$ an image point. The projection of the M in its image m is denoted by:

$$m \sim \xi (R \quad t) M \tag{1}$$

With R is the rotation matrix and t the translation vector of the camera in space and ξ is the matrix of the intrinsic camera parameters defined by:

$$\xi = \begin{pmatrix} l & \gamma & \mu_0 \\ 0 & \eta l & \nu_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2}$$

l : Focal length.

η : Scale factor.

γ : Skew factor.

μ_0 and ν_0 : Coordinates of principal point.

Considering the space of our scene is a projective space. In this space the points satisfying $w = 0$ are called the points at infinity, these points form the plane at infinity. In the plane at infinity the points satisfying $M^T M = 0$ form the absolute conic denoted Ω_∞. The absolute conic is a complex conic of the plane at infinity, the image of absolute conic ω is given by:

$$\omega = (\xi \xi^T)^{-1} \tag{3}$$

From (3) we deduce that the image of the absolute conic depends only on the intrinsic parameters of the camera. Once ω is estimated we can easily estimate the intrinsic parameters of the camera using the Cholesky factorization.

IV. Proposed method

The main idea of our new self-calibration method is based on the projection of two points of an unknown planar scene in only two image planes using two projection matrices which must be estimated.

Let (*S*) an unknown planar scene and $A(x_A, y_A, z_A, 1)^T$ and $B(x_B, y_B, z_B, 1)^T$ two points in homogeneous coordinates of (*S*) and $a_k(u_{ak}, v_{ak}, 1)^T$ and $b_k(u_{bk}, v_{bk}, 1)^T$ for *k* = *i* or *j* its projections in the two images (see Fig. 1).

Assuming that the plane containing the segment $[AB]$ is in $Z = 0$ and $0 < \theta < \Pi/2$. From the Fig. 1 we can deduce: $x_A = \frac{\|AB\|}{2}\cos(\theta)$ and $y_A = \frac{\|AB\|}{2}\sin(\theta)$

$$x_B = -\frac{\|AB\|}{2}\cos(\theta) \quad \text{and} \quad y_B = -\frac{\|AB\|}{2}\sin(\theta)$$

From (1) we have for the image i :

$$a_i \sim \xi_i (r_{1i}\ r_{2i}\ r_{3i}\ t_i) A \tag{4}$$

with r_{qi} is the column number *q* of R_i .
The equation (4) can be rewritten as follows:

$$a_i \sim \xi_i (r_{1i}\ r_{2i}\ t_i) \begin{pmatrix} x_A \\ y_A \\ 1 \end{pmatrix} \tag{5}$$

Let:

$$H_i \sim \xi_i (r_{1i}\ r_{2i}\ t_i) \tag{6}$$

H_i is a 3x3 matrix called the homography matrix scene to image *i*.
The equation (5) becomes:

$$a_i \sim H_i \begin{pmatrix} x_A \\ y_A \\ 1 \end{pmatrix} \tag{7}$$

Let $H_i = (h_{1i}\ h_{2i}\ h_{3i})$ with h_{ki} are column vectors of H_i . The vectors r_{1i} and r_{2i} are orthogonal which permit us to write from (6) :

$$\begin{cases} h_{1i}^T \xi_i^{-T} \xi_i^{-1} h_{2i} = 0 \\ h_{1i}^T \xi_i^{-T} \xi_i^{-1} h_{1i} = h_{2i}^T \xi_i^{-T} \xi_i^{-1} h_{2i} \end{cases} \tag{8}$$

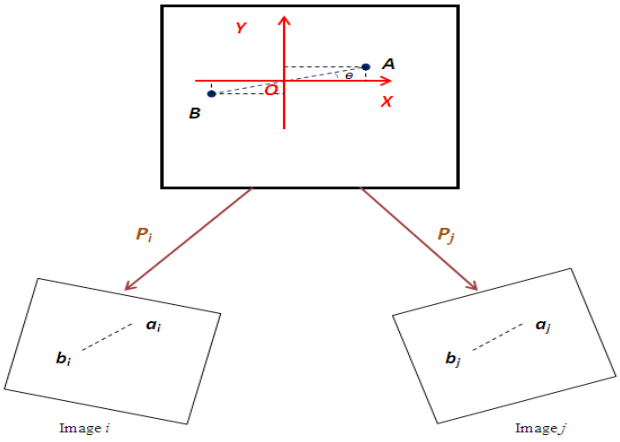


Fig. 1. Vision system used

Let:

$$A' = (x_A\ y_B\ 1)^T \quad \text{and} \quad B' = (x_B\ y_B\ 1)^T$$

From (7) we have for image *i*:

$$\begin{cases} a_i \sim H_i A' \\ b_i \sim H_i B' \end{cases} \quad (9)$$

For image *j*:

$$\begin{cases} a_j \sim H_j A' \\ b_j \sim H_j B' \end{cases} \quad (10)$$

From (9) and (10) we can write:

$$\begin{cases} a_i \sim H_i Q W \\ a_j \sim H_j Q W \end{cases} \quad (11)$$

And:

$$\begin{cases} b_i \sim H_i Q W' \\ b_j \sim H_j Q W' \end{cases} \quad (12)$$

with $Q = \begin{pmatrix} \frac{\|AB\|}{2}\cos(\theta) & 0 & 0 \\ 0 & \frac{\|AB\|}{2}\sin(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $W = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and $W' = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$.

Let:

$$\begin{cases} P_i = H_i Q \\ P_j = H_j Q \end{cases} \quad (13)$$

The equations (11) and (12) become:

$$\begin{cases} a_i \sim P_i W \\ a_j \sim P_j W \end{cases} \quad (14)$$

And:

$$\begin{cases} b_i \sim P_i W' \\ b_j \sim P_j W' \end{cases} \quad (15)$$

The matrices P_i and P_j are the projection matrices of the points A and B in the both images *i* and *j*.

From (13) we can write:

$$P_j = H_{ij} P_i \quad (16)$$

With $H_{ij} = H_j H_i^{-1}$ is the homography matrix between the images *i* and *j*. From (14), (15) and (16) we can write:

$$\begin{cases} a_i \sim P_i W \\ a_j \sim H_{ij} P_i W \\ b_i \sim P_i W' \\ b_j \sim H_{ij} P_i W' \end{cases} \quad (17)$$

The matrix P_i consists of eight unknowns and the above system gives us twelve equations that allows us to estimate the matrix P_i. The matrix P_j can be estimated from (16).

Let $P_i = \begin{pmatrix} p_{1i} & p_{2i} & p_{3i} \end{pmatrix}$, with the p_{ki} are the column vectors of P_i. From (13) we can deduce:

$$\begin{cases} p_{1i} = \frac{\|AB\|}{2}\cos(\theta) h_{1i} \\ p_{2i} = \frac{\|AB\|}{2}\sin(\theta) h_{2i} \end{cases} \quad (18)$$

The system of equations (18) gives us:

$$\begin{cases} h_{1i} = \frac{2}{\|AB\|}\cos(\theta)^{-1} p_{1i} \\ h_{2i} = \frac{2}{\|AB\|}\sin(\theta)^{-1} p_{2i} \end{cases} \quad (19)$$

By replacing in (9) h_{1i} and h_{2i} by their expressions in (19) and by posing $\lambda_1 = \frac{2}{\|AB\|}\cos(\theta)^{-1}$ and $\lambda_2 = \frac{2}{\|AB\|}\sin(\theta)^{-1}$, we obtain for image *i* :

$$\begin{cases} \lambda_1 \lambda_2 p_{1i}^T \omega_i p_{2i} = 0 \\ \lambda_1^2 p_{1i}^T \omega_i p_{1i} = \lambda_2^2 p_{2i}^T \omega_i p_{2i} \end{cases}$$

Or:

$$\begin{cases} p_{1i}^T \omega_i p_{2i} = 0 \\ \frac{\lambda_1^2}{\lambda_2^2} = \frac{p_{2i}^T \omega_i p_{2i}}{p_{1i}^T \omega_i p_{1i}} \end{cases} \quad (20)$$

And for image *j* :

$$\begin{cases} \lambda_1 \lambda_2 p_{1j}^T \omega_j p_{2j} = 0 \\ \lambda_1^2 p_{1j}^T \omega_j p_{1j} = \lambda_2^2 p_{2j}^T \omega_j p_{2j} \end{cases}$$

Or:

$$\begin{cases} p_{1j}^T \omega_j p_{2j} = 0 \\ \frac{\lambda_1^2}{\lambda_2^2} = \frac{p_{2j}^T \omega_j p_{2j}}{p_{1j}^T \omega_j p_{1j}} \end{cases} \quad (21)$$

From (20) and (21) we can deduce that:

$$\frac{p_{2i}^T \omega_i p_{2i}}{p_{1i}^T \omega_i p_{1i}} = \frac{p_{2j}^T \omega_j p_{2j}}{p_{1j}^T \omega_j p_{1j}}$$

Or:

$$p_{2i}^T \omega_i p_{2i} p_{1j}^T \omega_j p_{1j} = p_{2j}^T \omega_j p_{2j} p_{1i}^T \omega_i p_{1i} \quad (22)$$

Let $H_{ij} = \begin{pmatrix} \overline{h}_{ij1} \\ \overline{h}_{ij2} \\ \overline{h}_{ij3} \end{pmatrix}$ with $\overline{h}_{ij1}$, $\overline{h}_{ij2}$ and $\overline{h}_{ij3}$ are the row vectors of H_{ij}.

From (16) we can deduce:

$$\begin{cases} p_{1j} = \overline{h}_{ij1} p_{1i} \\ p_{2j} = \overline{h}_{ij2} p_{2i} \\ p_{3j} = \overline{h}_{ij3} p_{3i} \end{cases} \quad (23)$$

By replacing p_{1j} and p_{2j} by its expressions in (22) we obtain:

$$p_{2i}^T \omega_i p_{2i} (\overline{h}_{ij1} p_{1i})^T \omega_j (\overline{h}_{ij1} p_{1i}) = (\overline{h}_{ij2} p_{2i})^T \omega_j (\overline{h}_{ij2} p_{2i}) p_{1i}^T \omega_i p_{1i} \quad (24)$$

By developing the first equation in (13) we obtain for image *i*:

$$P_i^T \omega_i P_i = \begin{pmatrix} Q'^T Q' & Q'^T R_i^T t_i \\ t_i^T R_i Q' & t_i^T t_i \end{pmatrix} \quad (25)$$

With $\omega_i = (\xi_i \xi_i^T)^{-1}$ is the image of absolute conic in image *i* and

$$Q' = \begin{pmatrix} \frac{\|AB\|}{2}\cos(\theta) & 0 \\ 0 & \frac{\|AB\|}{2}\cos(\theta) \\ 0 & 0 \end{pmatrix}$$

For the image *j*:

$$P_j^T \omega_j P_j = \begin{pmatrix} Q'^T Q' & Q'^T R_j^T t_j \\ t_j^T R_j Q' & t_j^T t_j \end{pmatrix} \quad (26)$$

Let:

$$P_i^T \omega_i P_i = \begin{pmatrix} e_{11i} & e_{12i} & e_{13i} \\ e_{21i} & e_{22i} & e_{23i} \\ e_{31i} & e_{32i} & e_{33i} \end{pmatrix} \quad (27)$$

and:

$$P_j^T \omega_j P_j = \begin{pmatrix} e_{11j} & e_{12j} & e_{13j} \\ e_{21j} & e_{22j} & e_{23j} \\ e_{31j} & e_{32j} & e_{33j} \end{pmatrix} \quad (28)$$

From (25) and (26) we can deduced that the first two rows and columns of the matrices $P_i^T \omega_i P_i$ and $P_j^T \omega_j P_j$ are identical, which permit us to write:

$$\begin{cases} e_{11i} e_{12j} - e_{11j} e_{12i} = 0 \\ e_{11i} e_{22j} - e_{11j} e_{22i} = 0 \end{cases} \quad (29)$$

From (20), (21), (22), (24) and (29) we can deduce the following system:

$$\begin{cases} p_{1i}^T \omega_i p_{2i} = 0 \\ p_{1j}^T \omega_j p_{2j} = 0 \\ p_{2i}^T \omega_i p_{2i} p_{1j}^T \omega_j p_{1j} = p_{2j}^T \omega_j p_{2j} p_{1i}^T \omega_i p_{1i} \\ p_{2i}^T \omega_i p_{2i} (\overline{h}_{ij1} p_{1i})^T \omega_j (\overline{h}_{ij1} p_{1i}) = (\overline{h}_{ij2} p_{2i})^T \omega_j (\overline{h}_{ij2} p_{2i}) p_{1i}^T \omega_i p_{1i} \\ e_{11i} e_{12j} - e_{11j} e_{12i} = 0 \\ e_{11i} e_{22j} - e_{11j} e_{22i} = 0 \end{cases} \quad (30)$$

In the present article we are assumed in the one hand that the intrinsic camera parameters are unknown and constant except for the focal length which can be vary from one view to another and in the second hand we are used two images only of an unknown planar scene which gives us six unknown, the system of equations (30) contains six equations where the possibility to estimate the intrinsic parameters of camera in the both views.

The equations of the system (30) are nonlinear therefore to solve it we minimize the cost function given by the following expression:

$$\min_{\Delta_{ij}} \sum_{j=i+1}^{n} \sum_{i=1}^{n-1} \left(\beta_{1i}^2 + \beta_{2j}^2 + \beta_{3ij}^2 + \beta_{4ij}^2 + \beta_{5ij}^2 + \beta_{6ij}^2 \right) \quad (31)$$

with:

$$\beta_{1i} = p_{1i}^T \omega_i p_{2i}$$

$$\beta_{2j} = p_{1j}^T \omega_j p_{2j}$$

$$\beta_{3ij} = p_{2i}^T \omega_i p_{2i} p_{1j}^T \omega_j p_{1j} - p_{2j}^T \omega_j p_{2j} p_{1i}^T \omega_i p_{1i}$$

$$\beta_{4ij} = p_{2i}^T \omega_i p_{2i} (\overline{h}_{ij1} p_{1i})^T \omega_j (\overline{h}_{ij1} p_{1i}) - (\overline{h}_{ij2} p_{2i})^T \omega_j (\overline{h}_{ij2} p_{2i}) p_{1i}^T \omega_i p_{1i}$$

$$\beta_{5ij} = e_{11i} e_{12j} - e_{11j} e_{12i}$$

$$\beta_{6ij} = e_{11i} e_{22j} - e_{11j} e_{22i}$$

And Δ_{ij} is the vector of the intrinsic parameters of the camera in the two images *i* and *j* and *n* is the number of images used.

The minimization of (31) is given by the Levenberg-Marquardt algorithm which requires an initialization step, for this purpose we assume the following assumptions:

- The main point is in the center of the image, so μ_{i0}, ν_{i0}, μ_{j0} and ν_{j0} are known.
- The pixels are square so $\eta_i = \eta_j = 1$ and $\gamma_i = \gamma_j = 0$.

By substituting these parameters in the system of equations (30) we can determine the focal lengths l_i and l_j for the both images *i* and *j*.

V. Experimental Results

To test the effectiveness of our method two types of data are used synthetic and real.

A. Simulation

In this simulation a sequence of ten 512x512 images of a calibration target is used to test the effectiveness of our self-calibration method presented in this paper. After the detection of interest points by the Harris detector [25] the corresponding points between images are determined using the correlation measure ZNCC [26]. The images are perturbed by adding a Gaussian noise of deviation σ to the pixels of images. The projection of the scene in image planes allows us to obtain linear equations whose its solution allows to determine the projection matrices. The self-calibration equations are then built to formulate a nonlinear cost function which its minimization by the Levenberg-Marquardt algorithm helps us to estimate the intrinsic parameters of our camera in all views. To evaluate the performance of our approach we compared the results obtained with those given by a well-known calibration method [6] published by Z. Zhang and those given by [27]. The figures (2) and (3) consecutively present the relative error of the focal length estimated by our method and those of Z. Zhang and C-R.Huang according to the number of images used and added noise.

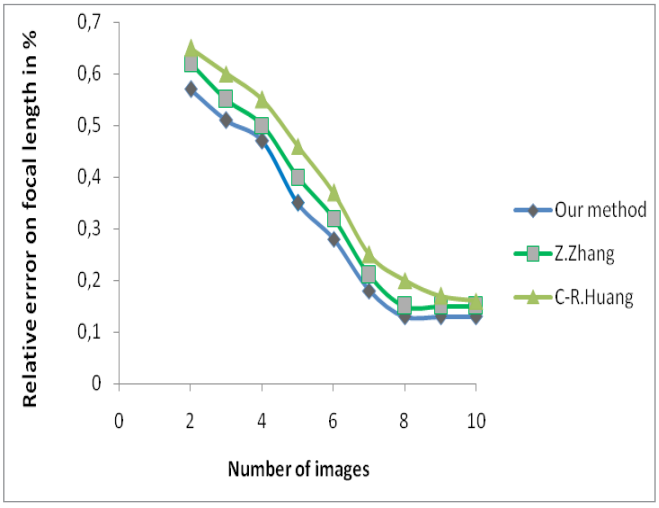


Fig. 2. Relative error on the focal length according to the number of images used.

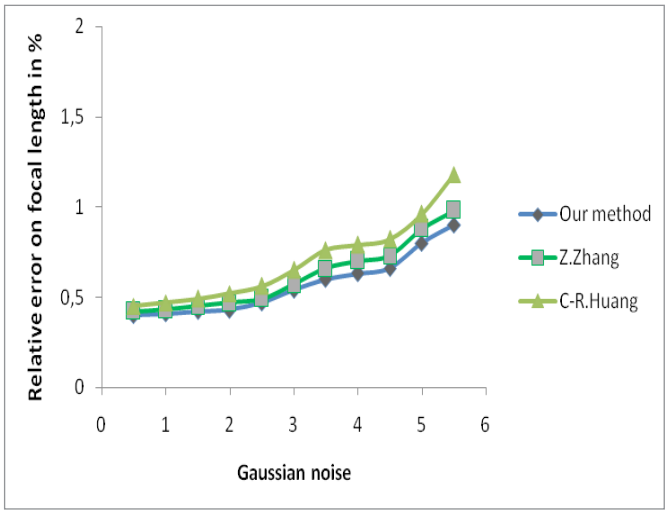


Fig. 3. Relative error on the focal length according to the added noise

The both figures above show that the method developed in this paper and that of Z.Zhang gives the approximate results. By taking into account the robustness of the Z. Zhang's method we can say that our method is effective and efficient. On the other hand, the previous figures prove the preference of our method compared with that presented in [27].

B. Real images

In this second experiment, two 512 × 512 images of a real planar scene are acquired by a CCD camera having constant intrinsic parameters except to the focal length. (See Fig. 4).

Fig. 4. The two real images used

The interest points are determined as well as their matching and calculation of projection matrices.

1. Detection of interest points

The interest points in the two images are detected by using the Harris detector (Fig. 5).

Fig. 5. Detection of interest points

2. Matching of interest points

The interest points previously identified are matched between the two images by using the ZNCC algorithm (Fig. 6).

Fig. 6. Matching of interest points.

3. Estimation of intrinsic parameters

This step consists of estimating the intrinsic parameters of the camera in each view by applying our method. The following table shows the results obtained.

TABLE I
THE RESULTS OBTAINED BY OUR METHOD

		l	η	γ	μ_0	ν_0
Our method	Image 1	1687	0,95	0,09	258	260
	Image 2	1683	0,97	0,05	257	255

The above table presents the results obtained by our method of self-calibration discussed in this article. To verify the effeteness of the above results, we rectified the two used images by using a robust approach of image rectification [28] based on a calibrated camera. The following figure shows the rectified images obtained by applying [28] and using the values of table 1.

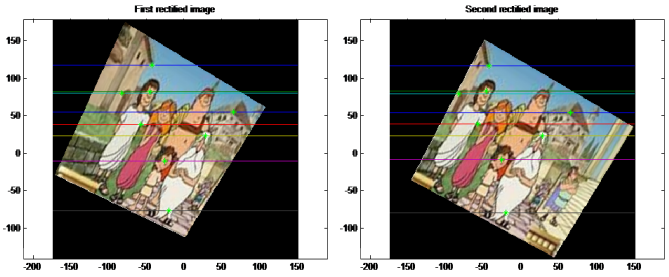


Fig.. 7 The obtained rectified images.

The goal of image rectification is to obtain the epipolar lines horizontal and parallel to the x-axis, according to the Fig. 7 we can deduce that our approach presented in this paper gives satisfactory results and this by taking into account that our method is able to self-calibrate a camera with a variable focal length from two images only based on two points only of an unknown planar scene.

VI. Conclusion

In this work we have proposed a new method of self-calibration of a CCD camera with constant intrinsic parameters except to the focal length which can vary freely from one view to another from an unknown planar scene. We have shown that the determination of the intrinsic parameters is possible by using only two points of the scene. The projections of the two points in the two image planes permit us to determine their projection matrices, these matrices are used to formulate a nonlinear cost function. The minimization of the cost function obtained allowed us to estimate the intrinsic parameters of the camera. The experimental results on the synthetic and real data prove the robustness of our method.

References

[1] B. Triggs. Autocalibration from Planar Scenes. In European Conference on Computer Vision, volume 2, pages 89–105, Friburg, Allemagne, juin 1998.

[2] R. Hartley and A. Zisserman. Multiple View Geometry. Seconde dition. Cambridge University Press, 2003.

[3] M. Merras, N. El Akkad, A. Saaidi and K. Satori. Camera Calibration with Varying Parameters Based On Improved Genetic Algorithm. WSEAS TRANSACTIONS on COMPUTERS. Vol. 13, 2014.

[4] P.Gurdjos, A.Crouzil and R.Payrissat. Another Way of Looking at Plane-Based Calibration:The Centre Circle Constraint.ECCV, 2002.

[5] P.F.Sturm and S.J.Maybank. On Plane-Based Camera Calibration: A General Algorithm, Singularities, Applications. In Proceedings of the CVPR-IEEE, Vol. 1, pp. 432-437, 1999.

[6] Z.Zhang. A Flexible New Technique for camera Calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No.17, pp. 1730-1734, 2000.

[7] M.Wilczkowiak, E.Boyer, P.Sturm. Camera Calibration and 3D Reconstruction from Single Images Using Parallelepipeds.In ICCV, Vancouver, Canada, pp. 162-169, July 2001.

[8] A.Saaidi, A.Halli, H.Tairi and K.Satori. Self–Calibration Using a Particular Motion of Camera. Wseas Transaction on Computer . Issue 4, Vol. 3, April 2008.

[9] P.Gurdjos and P.Sturm. Methods and Geometry for Plane-Based Self-Calibration. CVPR, pp. 491-496, 2003.

[10] B.Triggs. Autocalibration from planar sequences, In Proceedings of 5th EuropeanConference on Computer Vision,Freiburg, Allemagne, Juin 1999.

[11] N. El Akkad, M. Merras, A. Saaidi and K. Satori. Camera self-Calibration with Varying Parameters from Two views. WSEAS TRANSACTIONS on INFORMATION SCIENCE and APPLICATIONS, Issue.11, Vol. 10, 2013.

[12] I.EL Batteoui, A.Saiidi and K.Satori. Efficient planar self-calibration of a stationary camera with variable intrinsic parameters, JATIT, Vol. 50 No.1, April 2013.

[13] A.Hehden and K.Astrom,Euclidean Reconstruction from Images Sequences with Varying and unknown focal length and Principal Point, Proc.IEEE CVPR,pp.438-443,1997.

[14] A. Saaidi, A. Halli, H. Tairi and K. Satori. Self-Calibration Using a Planar Scene and Parallelogram. ICGST-GVIP, ISSN 1697 399X, February 2009 .

[15] L.Agapito,E.Hayman and I.Reid. Self-calibration of rotating and zooming camera. In International Journal of Computer Vision,pp.107-127,2001 .

[16] J. Semple and G. Kneebone. Algebraic Projective Geometry. Oxford Classic Series, Clarendon Press, 1952, réédité, 1998. 2012.

[17] J.More.The levenberg-marquardt Algorithm, implementation and theory. In G.A.Watson, editor, Numerical Analysis, Lecture Notes in Mathematics 630. Springer-Verlag, 1977.

[18] Manolis I.A. Lourakis and R.Deriche. Camera self-calibration using the kruppa equations and the SVD of the fundamental matrix: the case of varying intrinsic parameters. Technical Report 3911, INRIA, 2000.

[19] P.Sturm. Critical motion sequences for the self-calibration of cameras and stereo systems with variable focal length. Image and Vision Computing, 20(5-6) : 415-426, 2002.

[20] N. El Akkad, M. Merras, A.Saaidi and K.Satori, Camera self-calibration with varying intrinsic parameters by an unknown three dimensional scene, Visual Computer, vol.30, No.5, pp.519-530, May 2014.

[21] Y.Zhao and X.D.Lv. A new approach for camera-self-calibration using vanishing-line. Information Technology Journal, 11(2) :276- 282, 2012.

[22] Z.Jiang and S.Liu, The Self-calibration of Varying Internal Camera Parameters Based on Image of Dual Absolute Quadric Transformation, Information and Automation, Communications in Computer and Information Science, Springer, Vol.86, pp. 452-461, 2011.

[23] J.-S.Liu,J.-H, Self-calibration with varying focal length from two images obtained by a camera with small rotation and general translation . Pattern recognition letters- Elsevier 22 (2001) 1393-1404.

[24] Zetao Jiang. Self-calibration of Varying Internal Camera Parameters Algorithm Based on Quasi-affine Reconstruction. Journal of Computers, Vol. 7, No. 3, March 2012.

[25] C.Harris and M.Stephens. A combined Corner and Edge Detector. 4th Alvey vision Conference. Pp. 147-151, 1988.

[26] C.Saravananand, M.Surender. Algorithm for Face Matching Using Normalized Cross-Correlation. IJEAT, Vol.2, Issue 4,April 2013.

[27] C-R. Huang, C-S. Chen, and P-C. Chung, An Improved Algorithm for Two-image Camera Self-Calibration and Euclidean Structure Recovery Using Absolute Quadric, Pattern Recognition , vol. 37, no. 8, pp. 1713-1722, 2004.

[28] Fusiello, A., Trucco, E., & Verri, A. A compact algorithm for rectification of stereo pairs. Machine Vision and Applications, 12(1), 16–22, 2000.

[29] Pacheco, A., H. Bolivar-Baron, R. Gonzalez-Crespo, and J. Pascual-Espada, Reconstruction of High Resolution 3D Objects from Incomplete Images and 3D Information. International Journal of Interactive Multimedia and Artificial Intelligence, 6(2), 7–16, 2014.

Ismail El batteoui received the master's degree from SMBA-Fez University in 2010. He is currently working toward the PhD degree in the LIIAN Laboratory at SMBA Fez University. His current research interests include camera self calibration with variable intrinsic parameters, image rectification and 3D reconstruction.

Abderrahim Saaidi received the PhD degree from SMBA-Fez University in 2010. He is currently a professor of computer science at SMBA-Taza University. He is a member of the LIIAN and LSI Laboratories. His research interests include camera self calibration, 3D reconstruction, genetic algorithms and real-time rendering.

Khalid Satori received the PhD degree from the National Institute for the Applied Sciences INSA at Lyon in 1993. He is currently a professor of computer science at SMBA-Fez University. He is the director of the LIIAN Laboratory. His research interests include real-time rendering, Image-based rendering, virtual reality, biomedical signal, camera self calibration, genetic algorithms and 3D reconstruction.

Multilayer Perceptron: Architecture Optimization and Training

Hassan Ramchoun, Mohammed Amine Janati Idrissi, Youssef Ghanou, Mohamed Ettaouil

Modeling and Scientific Computing Laboratory, Faculty of Science and Technology, University Sidi Mohammed Ben Abdellah, Fez, Morocco

***Abstract* — The multilayer perceptron has a large wide of classification and regression applications in many fields: pattern recognition, voice and classification problems. But the architecture choice has a great impact on the convergence of these networks. In the present paper we introduce a new approach to optimize the network architecture, for solving the obtained model we use the genetic algorithm and we train the network with a back-propagation algorithm. The numerical results assess the effectiveness of the theoretical results shown in this paper, and the advantages of the new modeling compared to the previous model in the literature.**

***Keywords* — Multilayer Perceptron (MLP), Architecture Optimization, Non-Linear Optimization, Genetic Algorithm, Feed-Forward Neural Network Training.**

I. Introduction

IN recent years, neural networks have attracted considerable attention as they proved to be essential in applications such as content-addressable memory, pattern recognition and optimization.

Learning or training of ANN is equivalent to finding the values of all weights such that the desired output is generated to corresponding input, it can be viewed as the minimization of error function computed by the difference between the output of the network and the desired output of a training observations set [1].

Multilayer Perceptron is the most utilized model in neural network applications using the back-propagation training algorithm. The definition of architecture in MLP networks is a very relevant point, as a lack of connections can make the network incapable of solving the problem of insufficient adjustable parameters, while an excess of connections may cause an over-fitting of the training data [3]. Especially, when we use a high number of layer and neurons this is our case in this paper.

Optimizing the number of connection and hidden layer for establishing a multilayer Perceptron to solve the problem remains one of the unsolved tasks in this research area Multilayer Perceptron consists of input layer, output layer and hidden layers between these two layers. The number of these layers is dependent on the problem [8]. In this work, we optimize the number of hidden layers and the number of neurons in each hidden layer and process of to deal with a few connection to increase the speed and efficiency of the neural network. We model this problem of neural architecture in terms of a mixed-integer non-linear problem with non-linear constraints.

The next section present and discuss related works on neural network architecture optimization. Section 3 describes the artificial neural networks. In Section 4, we present the problem of optimization neural architecture and a new modeling is proposed. And before concluding, experimental results are given in the section 4.

II. Related Works

A number of approaches in the literature have taken account the architecture optimization. This section describes only those works that are more or less similar to our work.

Global search may stop the convergence to a non-optimal solution and determine the optimum number of ANN hidden layers. Recently, some studies in the optimization architecture problems have been introduced in order to determine neural networks parameters, but not optimally [3].

Traditional algorithms fix the neural network architecture before learning [4]. Others studies propose constructive learning [5]-[6]. It begins with a minimal structure of hidden layers; these researchers initialized the hidden layers, with a minimal number of hidden layer neurons. The most of researchers treat the construction of neural architecture (structure) without finding the optimal neural architecture [7].

T.B Ludermir et al [14] propose an approach for dealing with a few connections in one hidden layer and training with deferent hybrid optimization algorithms.

In our previous work we take account optimization of hidden layers with introducing one decision variable for layer [1], and in another work we have taken account the hidden node optimization in layers, for training this two models we have used a back-propagation algorithms [2].

III. Units Feed-Forward Neural Networks for Pattern Classification

A data set for pattern classification consists of a number of patterns together with their correct classification. Each pattern consists of a number of measurements (i.e., numerical values).

The goal consists in generating a classifier that takes the measurements of a pattern as input, and provides its correct classification as output. A popular type of classifier is feed-forward NNs [9].

A feed-forward NN consists of an input layer of neurons, an arbitrary number of hidden layers, and an output layer. Feed-forward NNs for pattern classification purposes consist of as many input neurons as the patterns of the data set have measurements, i.e., for each measurement there exists exactly one input neuron. The output layer consists of many neurons as the data set has classes. Given the weights of all the neurons connections, in order to classify a pattern, one provides its measurements as input to the input neurons, propagates the output signals from layer to layer until the output signals of the output neurons are obtained. Each output neuron is identified with one of the possible classes. The output neuron that produces the highest output signal classifies the respective pattern.

A. Multilayer Perceptron

A multilayer Perceptron is a variant of the original Perceptron model proposed by Rosenblatt in the 1950 [10]. It has one or more hidden

DOI: 10.9781/ijimai.2016.415

layers between its input and output layers, the neurons are organized in layers, the connections are always directed from lower layers to upper layers, the neurons in the same layer are not interconnected see Fig. 1.

The neurons number in the input layer equal to the number of measurement for the pattern problem and the neurons number in the output layer equal to the number of class, for the choice of layers number and neurons in each layers and connections called architecture problem, our main objectives is to optimize it for suitable network with sufficient parameters and good generalisation for classification or regression task.

B. Back-propagation and Learning for de MLP

Learning for the MLP is the process to adapt the connections weights in order to obtain a minimal difference between the network output and the desired output, for this raison in the literature some algorithm are used such as Ant colony [11] but the most used called Back- propagation witch based on descent gradient techniques [12].

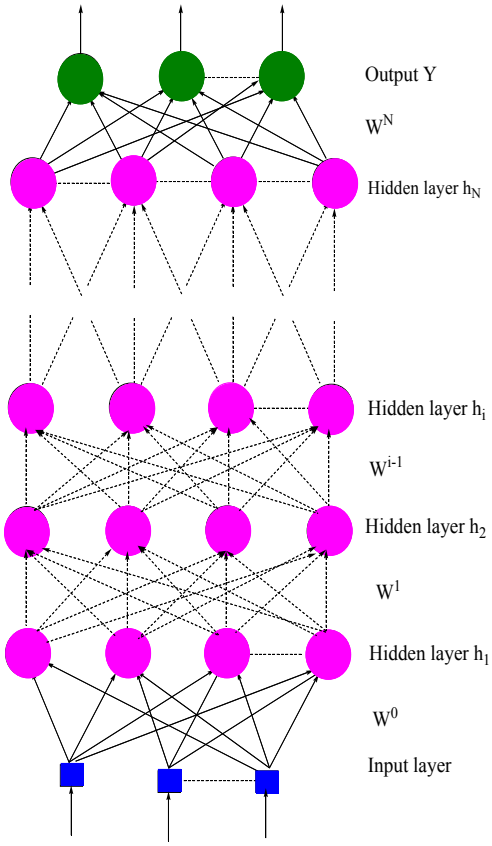


Fig 1. Feed forward neural network structure.

Assuming that we used an input layer with n_0 neurons $X = (x_0, x_1, \dots, x_{n_0})$ and a sigmoid activation function

$$f(x) = \frac{1}{1+e^{-x}} \quad (1)$$

To obtain the network output we need to compute the output of each unit in each layer:

Now consider a set of hidden layers $(h_1, h_2, \dots, h_N)$. Assuming that n_i are the neurons number by each hidden layer h_i.

For the output first hidden layer

$$h_i^j = f\left(\sum_{k=1}^{n_{i-1}} w_{k,j}^0 \, x_k\right) \quad \text{j=1,…,} n_i \quad (2)$$

The outputs h_i^j of neurons in the hidden layers are computing as flows:

$$h_i^j = f\left(\sum_{k=1}^{n_{i-1}} w_{k,j}^{i-1} \, h_{i-1}^k\right) \; i = 2, \dots, N \text{ and } j = 1, \dots, n_i \quad (3)$$

Where $w_{k,j}^i$, is the weight between the neuron k in the hidden layer i and the neuron j in the hidden layer $+1$, n_i is the number of the neurons in the ith hidden layer, The output of the ith can be formulated as following:

$$h_i = {}^{t}(h_i^1, h_i^2, \dots, h_i^{n_i}) \quad (4)$$

The network output are computing by

$$y_i = f\left(\sum_{k=1}^{n_N} w_{k,j}^N \, h_N^k\right)$$
$$Y = (y_1, \dots, y_j, \dots, y_{N+1}) = F(W, X) \quad (5)$$

Where $w_{k,j}^N$ is the weight between the neuron k in the Nth hidden layer and the neuron j in the output layer, $n_N$is the number of the neurons in the Nth hidden layer, Y is the vector of output layer, F is the transfer function and W is the matrix of weights, it's defined as follows:

$$W = [W^0, \dots, W^j, \dots, W^N]$$
$$W^i = \left(w_{j,k}^i\right)_{\substack{0 \le i \le N \\ 1 \le j \le n_{i+1} \\ 1 \le k \le n_i}} \text{ where } w_{j,k}^i \in \mathbb{R}$$

To simplify we can take $n = n_i \;\; \forall i = 1, \dots, N$ for all hidden layers. Where X is the input of neural network and f is the activation function and W^i is the matrix of weights between the i^{th} hidden layer and the $(i+1)^{th}$ hidden layer for $i = 1, \dots, N-1$, W^0 is the matrix of weights between the input layer and the first hidden layer, and W^N is the matrix of weights between the *N*th hidden layer and the output layer.

IV. Proposed Model to Optimize the MLP Weights and Architectures

The MLP architecture definition depends on the choice of the number of layers, the number of hidden nodes in each of these layers and the objective function but another approach which is introduced in this paper allows to control all connections between layers and to delete some of them, when there are no connections between the node n_i and layer $i+1$ and when all neurons are deleted in layers i we delete it.

In this work, we assign to each connection a binary variable which takes the value 1 if the connection exists in the network and 0 otherwise. Also we associate another binary variable for the hidden layers.

Notation:

N: Number of hidden layers.

n_0: Number of neurons in input layer.

n_i : Number of neurons in hidden layer i.

n_{op} : Optimal number of hidden layer.

n_{N+1}: Number of neurons in output layer.

X: Input data of neural network.

Y: Calculated output of neural network.

h_i^j: Output of neuron j in hidden layer i.

f: Activation function.

d: Desired output.

u_i: Binary variable for $i = 1, \dots, N-1$. $U = (u_1, u_2, \dots, u_{N-1})$

$v_{k,j}^i$: Binary variable for $i = 2, \dots, N$, $j = 1, \dots, n_i$ and $k = 1, \dots, n_{i-1}$

We computed the output of neural network by the following formulation:

$$F(U, V, W, X) = Y = \left(y_1, y_2, \dots, y_{n_{N+1}}\right) \tag{6}$$

A. Output of first hidden layer

The neurons of first hidden layer are directly connected to the input layer (data layer) of the neural network.

The output for each neuron in the first hidden layer is calculated by:

$$h_1 = \begin{pmatrix} h_1^1 \\ \vdots \\ h_1^k \\ \vdots \\ h_1^{n_1} \end{pmatrix} = \begin{pmatrix} f\left(\sum_{k=1}^{n_0} w_{k,1}^0 x_k\right) \\ \vdots \\ f\left(\sum_{k=1}^{n_0} w_{k,j}^0 x_k\right) \\ \vdots \\ f\left(\sum_{k=1}^{n_0} w_{k,n_1}^0 x_k\right) \end{pmatrix} \tag{7}$$

B. Output for the hidden layer i = 2,..., N

To calculate the output of each neuron for the hidden layer i, where $i = 2, \dots, N$, we propose the rule:

$$h_i = \begin{pmatrix} h_i^1 \\ \vdots \\ h_i^k \\ \vdots \\ h_i^{n_i} \end{pmatrix} = (1 - u_{i-1}) h_{i-1} +$$

$$u_{i-1} \begin{pmatrix} 1 - \prod_{k=1}^{n_{i-1}} \left(1 - v_{k,1}^{i-1}\right) \\ \vdots \\ 1 - \prod_{k=1}^{n_{i-1}} \left(1 - v_{k,j}^{i-1}\right) \\ \vdots \\ 1 - \prod_{k=1}^{n_{i-1}} \left(1 - v_{k,n_i}^{i-1}\right) \end{pmatrix} \begin{pmatrix} f\left(\sum_{k=1}^{n_{i-1}} v_{k,1}^{i-1} w_{k,1}^{i-1} h_{i-1}^k\right) \\ \vdots \\ f\left(\sum_{k=1}^{n_{i-1}} v_{k,j}^{i-1} w_{k,j}^{i-1} h_{i-1}^k\right) \\ \vdots \\ f\left(\sum_{k=1}^{n_{i-1}} v_{k,n_i}^{i-1} w_{k,n_i}^{i-1} h_{i-1}^k\right) \end{pmatrix} \tag{8}$$

where $k = 1, \dots, n_{i-1}$ and $j = 1, \dots, n_i$

C. Output for the neural network (layer N+1)

The output of the neural network is defined by the following expression:

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_k \\ \vdots \\ y_{n_{N+1}} \end{pmatrix} = \begin{pmatrix} f\left(\sum_{k=1}^{n_N} w_{k,1}^N h_N^k\right) \\ \vdots \\ f\left(\sum_{k=1}^{n_N} w_{k,j}^N h_N^k\right) \\ \vdots \\ f\left(\sum_{k=1}^{n_N} w_{k,N+1}^N h_N^k\right) \end{pmatrix} \tag{9}$$

D. Objective function

The objective function of the proposed model such as in the previous work [2] is the error calculated between the obtained output and desired output:

$$\|F(U, V, X, W) - d\|^2 \tag{10}$$

But we propose to modify this objective function and adding one term for error connections regularized by α in order to control variations of weights in training and optimization phase.

E. Constraints

- The first constraint guarantees the existence of the hidden layers.

$$\sum_{i=1}^{N} u_i \geq 1 \tag{11}$$

- These constraints insured communication between neurons, connections and layers

$$u_{i-1} \times \prod_{j=1}^{n_i} \prod_{k=1}^{n_{i-1}} (1 - v_{k,j}^{i-1}) = 0 \quad \forall i = 2, \dots, N \tag{12}$$

$$(1 - u_{i-1}) \times \sum_{j=1}^{n_i} \sum_{k=1}^{n_{i-1}} v_{k,j}^{i-1} = 0 \quad \forall i = 2, \dots, N \tag{13}$$

- The weights values are the real number.

V. Implementation And Numerical Results

To illustrate the advantages of the proposed approach, we apply our algorithm to a widely used dataset, Iris dataset for classification [13]. It consists of three target classes: Iris Setosa, Iris Virginica and Iris Versicolor. Each species contains 50 data samples. Each sample has four real-valued features: sepal length, sepal width, petal length and petal width. By doing this, all features have the same contribution to Iris classification.

For this task an MLP with sigmoid activation function and a back-propagation training after obtaining an optimal architecture by Genetic algorithm is used, in this section we present a parameters setting, implementations procedures and final results.

A. Parameters setting

We use the genetic algorithm to solve the architecture optimization problems. To this end, we have coded individual by one chromosomes; moreover, the fitness of each individual depends on the value of the objective function obtained by two terms one for network error and the second for the network connections.

The individual of the initial population are randomly generated, and u_i, $v_{k,j}^i$ take the value 0 or 1, and the weights takes random values in space $[-0.5, 0.5]$. After creating the initial population, each individual is evaluated and assigned a fitness value according to the fitness function.

The fitness suggested in our work is the following function:

$$f(i) = M - f(s) \tag{14}$$

We applied crossover and mutation operator, in this step, new individuals called children are created by individuals selected from the population called parents for more exploring the researches spaces of solution.

In Table I we present all parameters used in our experiments

TABLE I
IMPLIMENTATION PARAMETRES

Nmax	16	32	48
Pm	0.200	0.200	0.300
Pc	0.700	0.800	0.700
s	0.520	0.500	0.550
α	0.010	0.300	0.450
β	0.395	0.397	0.340

Nmax: total number of weights, Pm: Probability of mutation, Pc: Probability of crosser, α: Regularization parameters for connection error, β: Training rate, s: Threshold

B. Results for the optimization methodology

After determining the optimal number of hidden layer, in this case three, and the total number of connection. We use an architecture contains four neurons in each layer and we can initialize the neural networks by value of weights obtained with Genetic algorithm. Our algorithm tested on instances for Iris data.

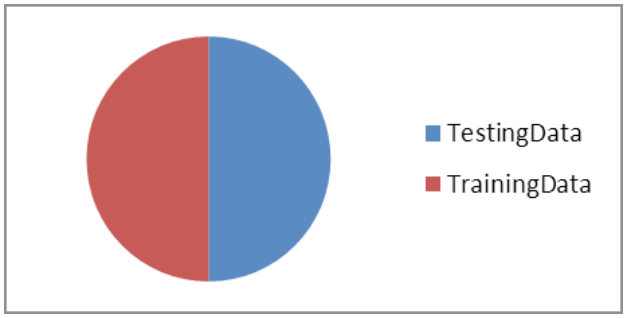


Fig.2 Partition of data base

The Table II presents the obtained clustering results of training and testing data. We remark that the proposed method permits to classify all the training data only one from Versicolor and two from the same type for testing data.

TABLE II
CLASSIFICATION FOR DATA SET (PROPOSED METHOD)

	Connect (%)	MC	Accuracy (%)
Tr.D	50	1	98.7
Tes.D	50	2	97.3

M.C: Misclassified data, Connect (%): percentage of connections weights used in the network between hidden layers, Tr.D: Training Data, Tes.D: Testing Data

We Remark that the obtained clustering results of testing data shows that our method gives the good results, because all the testing data were correctly classified except two. In fact; these elements (misclassified) are from the Versicolor class.

TABLE III
CLASSIFICATION FOR DATA SET (PREVIUOS METHOD)

	Nr. T. D	Connect (%)	MC	Accuracy (%)
Tr.D	75	100	3	96
Tes.D	75	100	2	97.3

From Tables above we can see that the proposed method gets a higher average classification accuracy rate and we have used a few connections than the existing methods. And we can conclude that the proposed approach in this paper gives better results compared to the neural methods optimizing neurons hidden layers.

TABLE IV
COMPARISON FOR IRIS DATA CLASSIFICATION

Method	Connec (%)	It.	M.T.	M.TS	A.T (%)	A.TS (%)
EBP	100	500	3	2	96	97.3
EBP	100	800	2	1	97.3	98.6
RBF	100	85	4	4	94.6	94.6
RBF	100	111	4	2	96	97.3
SVM	_	5000	3	5	94.6	93.3
Previous Method	100	100	3	2	96	97.3
Proposed Method	50	647	1	2	98.7	

It: Number of iterations, M.T: Misclassified for training set, M.TS: Misclassified for testing set, A.T: Accuracy for training set, A.TS: Accuracy for testing set.

The results are shown in the Tables IV, we can see that The comparison of the average classification accuracy rate, convergence iterations and number of connections used of the proposed method with other existing neural networks training algorithms: Error Back-Propagation (EBP), Radial Basis Function (RBF) neural networks and Support Vector Machine (SVM) chow that our model present two quality: few connections and higher average classification accuracy rate.

VI. Conclusion

A Model is developed to optimize the architecture of Artificial Neural Networks. The Genetic Algorithm is especially appropriate to obtain the optimal solution of the nonlinear problem. This method is tested to determine the optimal number of hidden layers and connection weights in the Multilayer Perceptron and the most favorable weights matrix after training. We have proposed a new modeling for the multilayer Perceptron architecture optimization problem as a mixed-integer problem with constraints. Depending on the Iris data, the results obtained demonstrates the good generalization of neural networks architectures. In conclusion, the optimal architecture of artificial neural network can play an important role in the classification problem. We can call the proposed approach to solve our model with other metaheuristics and we are going to tray with many other data base for real problems: Diabetes, Thyroid, Cancer….

References

[1] M. Ettaouil and Y. Ghanou, "Neural architectures optimization and Genetic algorithms", Wseas Transactions On Computer, Issue 3, Volume 8, 2009, pp. 526-537.

[2] M.Ettaouil M.Lazaar and Y.Ghanou "Architecture optimization model for the multilayer perceptron and clustering" Journal of Theoretical and Applied Information Technology 10 January 2013. Vol. 47 No.1.

[3] T.B Ludermir "Hybrid Optimization Algorithm for the Definition of MLP Neural Network Architectures and Weights" Proceedings of the Fifth International Conference on Hybrid Intelligent Systems (HIS'05) 0-7695-2457-5/05 20.00 2005 IEEE.

[4] JOSEPH RAJ V. 'Better Learning of Supervised Neural Networks Based on Functional Graph – An Experimental Approach', WSEAS TRANSACTIONS on COMPUTERS, Issue 8, Volume 7, August 2008.

[5] D. Wang, 'Fast Constructive-Coverting Algorithm for neural networks and its implement in classification', *Applied Soft Computing* 8 (2008) 166-173.

[6] D. Wang, N.S. Chaudhari, 'A constructive unsupervised learning algorithm for Boolean neural networks based on multi-level geometrical expansion', *Neurocomputing* 57C (2004) 455-461.

[7] T. Kohonen, 'Self Organizing Maps', *Springer*, 3eme edition, 2001. *Neural Netw*., vol. 17, no. 6, pp. 1452–1459, Nov. 2006.
[8] E. Egriogglu, C, Hakam Aladag, S. Gunay, 'A new model selection straegy in artificiel neural networks', *Applied Mathematics and Computation* (195) 591-597, 2008.
[9] Bishop CM (2005) Neural networks for pattern recognition. MIT Press, Cambridge.
[10] Rosenblatt, "The Perceptron: A Theory of Statistical Separability in Cognitive Systems", Cornell Aeronautical Laboratory, Report No. VG-1196-G-1, January, 1958.
[11] Krzysztof Socha Christian Blum ''An ant colony optimization algorithm for continuous optimization: application to feed-forward neural network training ''
[12] D.Salamon, "Data compression", Springer, 2004.
[13] S. J. Bolaños, R. G. Crespo, V. H. Medina-García, "Patterns of software development process, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 1, no. 4, pp. 33-40, 2011.
[14] T. B. Ludermir, A. Yamazaki, and C. Zanchettin, "An optimization methodology for neural network weights and architectures," *IEEE Trans*.

Hassan Ramchoun is a PhD student in the Laboratory of Modeling and Scientific computing at the Faculty of Sciences and Technology of Fez, Morocco, he is a member of Operational Research and Computer group. He works on Neural Network, probabilistic modeling, classification problems, and statistical learning methods.

Mohammed Amine Janati Idrissi is a PhD student in Modeling and Scientific Computing laboratory at the Faculty of Sciences and Technology of Fez, Morocco. His research interests include metaheuristics, artificial neural networks, and applications.

Youssef Ghanou is a PhD in Department of Computer Engineering, High School of Technology, Moulay Ismaïl University, B. P. 3103, 50000, Toulal, Meknes, Morocco. His research interests include: Operational Research, metaheuristics, artificial neural network, optimization and applications.

Mohamed Ettaouil is a Doctorate Status in Operational Research and Optimization, FST University Sidi Mohamed Ben Abdellah USMBA, Fez. PhD in Computer Science, University of Paris 13, Galilee Institute, Paris France. He is a professor on the Faculty of Science and technology of Fez FST, and he was responsible of research team in modelization and pattern recognition, operational research and global optimization methods. He was the Director of Unit Formation and Research UFR: Scientific computing and computer science, Engineering Sciences. He is also a responsible of research team in Artificial Neural Networks and Learning, modelization and engineering sciences, FST Fez. He is an expert in the fields of the modelization and optimization, engineering sciences.

Detection of Text Lines of Handwritten Arabic Manuscripts using Markov Decision Processes

Youssef Boulid[1], Abdelghani Souhar[2], Mohamed Youssfi Elkettani[1]

University Ibn Tofail, Faculty of Sciences, [1]Department of Mathematics, [2]Department of Computer Science, Kenitra, Morocco

***Abstract* — In a character recognition systems, the segmentation phase is critical since the accuracy of the recognition depend strongly on it. In this paper we present an approach based on Markov Decision Processes to extract text lines from binary images of Arabic handwritten documents. The proposed approach detects the connected components belonging to the same line by making use of knowledge about features and arrangement of those components. The initial results show that the system is promising for extracting Arabic handwritten lines.**

***Keywords* — Text Line Segmentation, Handwritten Arabic Document, Connected Component Analysis, Markov Decision Processes.**

I. Introduction

MANY documents are stored in their original state (as paper) and have yet to be exploited to produce their electronic versions. There are systems for converting the document image into text version, in order to facilitate the access and the search in the document. These systems often consist of five steps: pre-processing, segmentation, feature extraction, character recognition and post-processing.

Segmentation is the process of partitioning the document into homogeneous entities (lines, words or characters), it is a very important phase in the character recognition process since the recognition rate depends strongly on it. The segmentation of text lines is to assign the same label to units that are spatially aligned (pixels or connected components) [1].

The Arabic script is naturally cursive, which makes the extraction of lines from handwritten Arabic document a real challenge [2]. Among the challenges, we have:

- The writing styles which differ widely from one writer to another.
- The fluctuation of the base line due to the movements of the pen results in different variations within the same text line.
- The cluttered writing style and the ascender and descender introduce ligatures between parts of words which make overlap the adjacent lines.
- The diacritical points, which lie below or above the words, further complicate the task.

To cope with these problems, we propose an approach for the detection and the extraction of Arabic handwritten text lines by exploiting knowledge about the connected components arrangement using Markov Decision Processes. The components that belong to the same line are those satisfying some constraints related to the knowledge.

The rest of this paper is organized as follows: Section II examines some related works. Section III provides insight into Markov Decision Processes. Section IV describes our approach for the segmentation of Arabic handwritten text lines. Section V gives some results and finally Section VI concludes the paper.

II. Related Works

In general there are two approaches for the segmentation of text lines: either through searching for separating locations between lines, or by searching for physical units such as Connected Component (CC) constituting a line.

As mentioned in [3] the extraction techniques can be divided into three classes: top-down, bottom-up and hybrid methods.

In the top-down methods, the document image is partitioned into regions, often recursively based on the global characteristics of the image. These methods are influenced by the large curvatures of the lines and the lines of text that touch each other. For the bottom-up methods, the basic elements such as pixels or CC are grouped to form text line, for that a lot of calculation and heuristic analysis are needed. Hybrid approaches combine the two classes of techniques to give best results. The works in [4,1], present a revision for existing segmentation methods of handwritten text lines. There are also competitions like ICDAR [5] and ICFHR [6] that participate in advancing this field of research.

Table I presents an overview of some related methods.

III. Markov Decision Processes

Markov decision processes (MDP) are defined as controlled stochastic processes satisfying the Markov property, assigning rewards to state transitions [7,8] They are defined by:

- S, is the state space within which the process evolve;
- A, is the space of actions that control the dynamics of the state;
- P() are the transition probabilities between states;
- r() is the reward function on transitions between states.

The transition probabilities characterize the dynamics of the system state. $P(s'/s,a)$ represents the probability that the system spends in state s' after executing action a in state s. Conventionally requires that $\forall s,a,\sum_{s'}P(s'/s,a)=1$.

As a result of the choice of action a in state s, we receive a reward. Positive values of $r(s,a)$ may be regarded as gains and negative values as costs. This award may depend on the end state s'; $r(s,a,s')$.

A function $\Pi : S \to A$ assigning at each time step an action to every state is called a (stationary) policy.

The value function of a policy $\upsilon_{\Pi} : S \to R$ is defined so that $\upsilon_{\Pi}(s)$ represents the infinite horizon, discounted cumulative reward attached to a policy Π applied to a MDP, with initial state s.

DOI: 10.9781/ijimai.2016.416

TABLE I
AN OVERVIEW OF SOME METHODS OF SEGMENTATION OF HANDWRITTEN TEXT LINES

Ref	Methods	Description	Rate of segmentation
[12]	Projection profile and K-means	The document is split vertically into several strips, and text is detected based on the histogram of the PP. Text blocks are clustered using K-means in three classes; the large blocks are split vertically into a number of average text blocks according to the analysis of the neighborhood and interline spacing.	96%
[13]	Connected Component	The algorithm starts by removing outlier components using a threshold value and then characters belonging to tow lines are detected and divided horizontally at the half distance. For line detection, a rectangular neighborhood is centered on a current component and increases to include those that satisfy certain conditions. The filtered components at the beginning are reallocated to the corresponding lines with respect to distance from their bounding box.	93.35%
[11]	Projection Profile	After removing small components and linking broken characters, Fourier curve fitting within the horizontal PP is used to locate the point of separation. The contour is used to extract the base line of the CC which allows locating the cut point between different adjacent lines. The components are assigned according to the closest line that is approximated by a polynomial curve that fits the pixels in the baselines. Finally, lifted small size components are reassigned to their line according to the closest CCs nearest neighbour in four directions.	93% and 94% for MS 95% and 90% respectively
[14]	Projection profile and morphological dilatation	The horizontal PP is used to estimate the skew line. The slop in each zone is used for smearing it, using dilation with adaptive structuring element that changes the size and the slop according to the zone. In the second stage the big blobs are detected with a recursive function that search for the cut point in order to separate the components and match them with their lines following attraction and repulsion criterion. Touching components are detected and separated using erosion recursively in order to detect the thinness part, which corresponds to the cut point.	96% and 97% for MS 95% and 90% respectively
[15]	Adaptive Local Connectivity map (ALCM)	The grey-level image produced by the ALCM algorithm is binarized using an adaptive thresholding algorithm to detect the location of the text lines (masks). In the second step, CCs are grouped into location masks for each text line. Finally, the text lines are extracted by superimposing the components with text line pattern mask. For touching characters, the contour is divided into pieces of segments and reallocated according to the distance from their center of mass with respect to the nearest text line.	99.5%
[3]	Wigner-Ville distribution and Projection profile	The image is divided into cells and the orientation within each cell is calculated using the Wigner–Ville distribution on the PP histogram. The cells having similar orientation are joined. The method exploits the projection peaks and the orientation within each area to follow the connected components forming text lines. The diacritical components are reassigned to the nearest text line. To separate connected lines, for each touching character and from a starting point, the flow continues beyond the intersection and the strokes of the same component are recovered by analyzing the angular variation corresponding to the curvature of the descending character.	98.6%.
[16]	BreadthFirst Search	The algorithm begins by removing diacritics and then the sparse similarity graph is built based on the local orientation of the component. Text lines are represented by a disjoint set using BreadthFirst Search (BFS). The affinity propagation clustering method is used to assign the blobs to text lines.	95,6%
[17]	Breadth-first-search and affinity-propagation clustering	After removing small components, the direction of text line is detected at each component by locating the region having maximum neighboring components, and then the local orientation is estimated by the least square line passing through the centroid of the components. After that, a graph is constructed where the nodes correspond to components and weights on edges correspond to distance to the estimated orientation line. The shortest path-algorithm is used to complete the graph. Two estimations of text lines are obtained: the Breadth-first-search and the affinity-propagation clustering method. The splitting error are corrected using affinity propagation while merging error are corrected using Expectation Maximization. Touching components are localized by finding the common tangent of the convex-hull of successive components, and then the component is split near the centroid. Finally, the diacritics and accent components are reassigned to their closest component.	98,76% of F1-score for 90% MS
[5]	Divide image into blocks	The Gaussian filter is used and the image is divide into blocks which are then binarized using an adaptive threshold with respect to the skew angle estimated in each block. Then, the blocks are concatenated to get the path of text lines, and finally text lines are extracted by thinning the background of the path image.	94%

This value can be computed as:

$$\upsilon_{\Pi}(s) = \mathrm{E}\left[\sum_{t=0}^{\infty} \gamma^t r(s^t, \Pi(s^t)) \,|\, s^0 = s\right] \quad (1)$$

The expectation is taken over all possible trajectories where $\langle s^0, a^0, s^1, a^1, \ldots, s^t, a^t, \ldots \rangle$ from the initial state s, the policy Π is applied. The factor which ensures that the above infinite sum converges is the discount factor $0 \le \gamma \le 1$.

The problem of finding an optimal policy with respect to the discounted criterion (1) can be written as: $\Pi^* : S \to A$, so that: $\upsilon_{\Pi^*}(s) \ge \upsilon_{\Pi}(s)\ \forall \Pi \in A^S$.

Let's consider the Bellman operator β operating on any value function υ as:

$$h_i^j = f\left(\sum_{k=1}^{n_{i-1}} w_{k,j}^0 \, x_k\right) \qquad \text{j=1,…,} n_i \quad (2)$$

Let υ^* denote the value function attached to an optimal policy $\Pi^*(\upsilon^* = \upsilon_{\Pi})$. It has been shown (see, e.g. [8]) that υ^* can be computed as the unique solution of the system of non-linear equations $\upsilon^* = \beta(\upsilon^*)$:

$$\upsilon^* = \max_{a \in A}\left\{ r(s,a) + \gamma \sum_{s' \in S} P(s'/s,a)\upsilon^*(s') \right\}, \forall s \in S \quad (3)$$

The optimal policy Π^* can then be obtained from υ^*:

$$\Pi^*(s) = \arg\max_{a \in A}\left\{ r(s,a) + \gamma \sum_{s' \in S} P(s'/s,a)\upsilon^*(s') \right\}, \forall s \in S \quad (4)$$

There exist several methods for computing the optimal policy υ^*. In this paper we have used the value iteration algorithm, shown in the figure below:

```
1: Procedure Value_Iteration(S,A,P,R,θ)
2:     Inputs
3:         S is the set of all states
4:         A is the set of all actions
5:         P is state transition function specifying P(s'/s,a)
6:         R is a reward function R(s,a,s')
7:         θ a threshold, θ>0
8:     Output
9:         π[S] approximately optimal policy
10:        V[S] value function
11:    Local
12:        real array V_k[S] is a sequence of value functions
13:        action array π[S]
14:    assign V_0[S] arbitrarily
15:    k ←0
16:    repeat
17:        k ←k+1
18:        for each state s do
19:            V_k[s] = max_a Σ_s' P(s'/s,a) (R(s,a,s')+ γV_{k-1}[s'])
20:    until ∀s |V_k[s]-V_{k-1}[s]| < θ
21:    for each state s do
22:        π[s] = argmax_a Σ_s' P(s'/s,a) (R(s,a,s')+ γV_k[s'])
23:    return π, V_k
```

Fig. 1. Pseudo code of value iteration algorithm.

IV. Proposed Approach

The methods based on CC analysis use geometric information such as shape, orientation, position, and size of CC for grouping them in rows. These methods are more suitable for complex documents compared to methods based on Projection Profile (PP) [9,10], even if they could be sensitive to changes in the size and component structures [11].

Firstly the diacritical components are removed from the document image using an estimated threshold. At the end of the process, they will be reallocated to their corresponding lines. The diagram of the proposed system is shown in Fig.2.

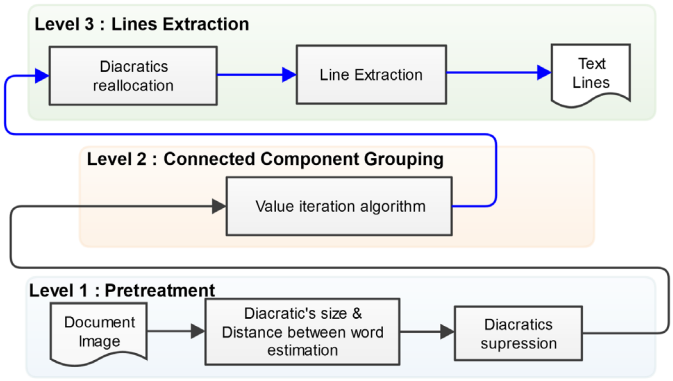


Fig. 2. Block diagram illustrating the sequence of treatments in the proposed approach.

We represent the environment by a set of CCs which are connected to each other, in such way that allows us to link only those which probably belong to the same line.

For this purpose, a circle with an estimated radius is centered on each CC, and from all the components that are located inside the circle, we take only those that are located between 145 ° and -145 ° relative to the current component (Fig. 3).

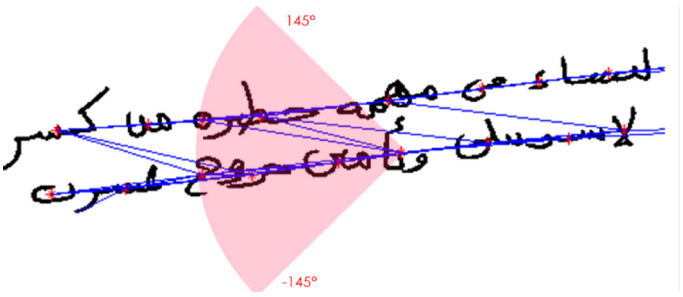


Fig. 3. Example of links (in blue) between components of two lines

The radius of the circle is estimated taking into consideration the average distance between words which is estimated by:

$$D_w = \frac{\sum_{i=1}^{m} W_i}{m} * 2 \quad (5)$$

W_i represents the width of the Ith component and m represents the total number of components in the image after removing diacritics.

We found that there are some components that are connected with others while these latter do not participate in the same line. To solve this problem we introduce in the next paragraph the concept of field of vision.

A. Field of Vision Function

To detect the components that are in the left side of another (which probably belong to the same line), the algorithm loops through pixels in rows and columns from the left side of the current component, and for each line, the index of the first black pixel is added to a list.

Thus, we denote $FV(C)$ the function of field of vision for the component C, such as for the example in Fig. 4, we have: $FV(ح) = \{د, ر, ص\}$

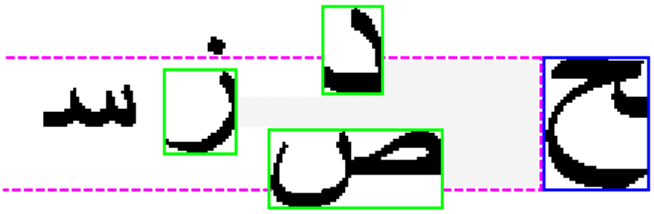


Fig. 4. Example of field vision function.

To improve the previous results we add the criterion of field of vision to keep only the components that are directly visible from the current component. The Fig. 5, illustrates the result of this improvement.

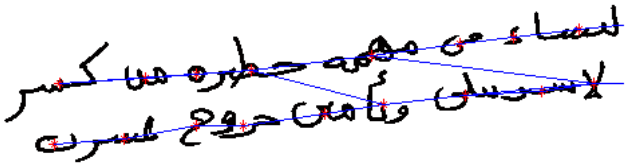


Fig. 5. Example of links (in blue) between components of two lines after using field of vision criterion

The field of vision function improves the environment, which can be used to extract the lines.

The problem now is to choose for each component the one that most likely belong to the same line as the previous component.

In the next paragraph, we will show that this problem could be resolved using MDP, where the environment can be considered as an oriented graph, in which the connected components represent the nodes and the edges represent some measures of alignment between the nodes.

B. Proposed Solution

So, from this point of view, we consider the following:

1. The **States** correspond to connected components:

$$\{S_i\} = \{C_i\}, i = 1...m$$

2. The **Actions** correspond to transitions from a given state source S_i to other possible states S'_{ij} :

$$A(S_i) = \{S'_j\}$$

3. The **Transition Probabilities**; For the example in Fig. 6, the transition probabilities are computed as follows:

$$P(S'_{1,1}, S_1) = 0.8 * \upsilon(S'_{1,1}) + 0.1 * \upsilon(S'_{1,2}) + 0.1 * \upsilon(S'_{1,3})$$

$$P(S'_{1,2}, S_1) = 0.8 * \upsilon(S'_{1,2}) + 0.1 * \upsilon(S'_{1,1}) + 0.1 * \upsilon(S'_{1,3})$$

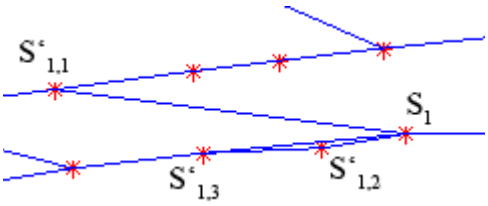


Fig. 6. Example of links (in blue) between components (red dots) of two lines

4. The **Reward Function**, is estimated as a combination of the Euclidean distance between the state S and the next state S', and the percentage of participation of S' related to S.

If we consider that:

- $P_{S'/S}$, The percentage representing the portion of component S' which is included in the field of vision of component S (Fig. 7).
- $P'_{S'/S}$, The percentage representing the portion of component S' which is included in the field of vision of component S, relatively to S' (Fig. 7).

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

Fig. 7. The values of and respectively for a (55%, 0.48%) , for b (0.38%, 0.38%) , for c (0.96%, 0.68%) and for d (0.61%, 100%). S' represents the green rectangle and S represents the blue rectangle, the gray part represents the portion in S' included in S.

So the Reward Function $r(s', s)$, can be estimated by:

$$port = (2 * P_{S'/S} + P'_{S'/S}) / 3 \quad (6)$$

$$dist = \frac{(3 * D_w - D(s, s'))}{3 * D_w} \quad (7)$$

$$r(s', s) = \frac{port + dist}{2} \quad (8)$$

After several experimental tests, we found that the value 0.95 for the **Discount Factor** gives the best results.

At the end of the execution of the value iteration algorithm, we get a list that contains the states and their successors (the next state having maximum utility value). So from this list we recover all the linked components participating to the same line.

Fig. 8, illustrate the result of the value iteration algorithm for a given image.

In order to extract the components with their diacritics, we first search for the polyline that goes through all the pixels of the lower contours of the components. This polyline is then shifted down vertically in order to include diacritical points that are below the words. Secondly we define a polygon which is bounded by the polyline and the two top corners of the image. Finally we extract and suppress the portion of the image represented by this polygon (as shown in Fig. 8.d).

This treatment is executed iteratively for each line, until there are no more rows to be extracted.

(a) (b) (c)

(d)

Fig. 8. Example of result of value iteration algorithm, (a) the initial state of the environment, (b) the result of value iteration algorithm, (c) connected component grouping. (d) Example of components detection, diacritics assignment and line extraction

V. Experimentation

For experimentation, we have used a subset of the publicly available data set "Handwritten Arabic Proximity Datasets", which contains 125 images of Arabic handwritten document [18].

To evaluate the performance of the system, we used the criterion Matching Score (MS) [19] which is calculated as follows:

$$MS(r_i, g_j) = \frac{T(P(r_i) \cap P(g_j))}{T(P(r_i) \cup P(g_j))} \tag{9}$$

With MS(ri,gj) is a real number between 0 and 1 which represents the matching score between the zone ri resulted from the algorithm and the zone gj in the ground truth. P corresponds to pixels that represent the foreground (text) and T is an operator that counts the number of pixels in each zone.

If the MS for a zone is found above a threshold, this is counted as true positive (TP). The resulted zones that do not match any zone in the ground truth are counted as false positive (FP) and the zones in the data set that are left unmatched are considered false negative (FN). Therefore, we count precision, recall and F1-score as follows:

$$precision = \frac{TP}{TP + FP} \tag{10}$$

$$\sum_{i=1}^{N} u_i \geq 1 \tag{11}$$

$$u_{i-1} \times \prod_{j=1}^{n_i} \prod_{k=1}^{n_{i-1}} (1 - v_{k,j}^{i-1}) = 0 \quad \forall i = 2, \ldots, N \tag{12}$$

According to the tests we obtained a F1-Score respectively of 95.8 % and 90.5 % for the MS-threshold values of 90% and 95%. Table II compares the score of the proposed method with other methods tested on the same data set.

TABLE II
RESULTS OBTAINED ON THE ORIGINAL DATA IN [18]

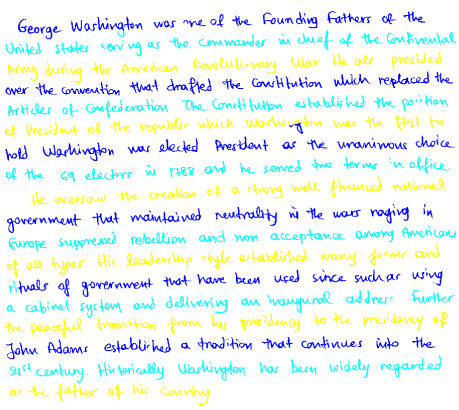
Methods	F1-Mesure	
	MS=90	*MS=95*
The method in [16]	95.6	90.9
The method in [17]	98.8	Not reported
The proposed method	95.8	90.5

Fig. 9 illustrates some visual results of our approach for segmentation of Arabic handwritten documents text line.

As can be seen in Fig.9.e, even though this proposed approach has been adapted here for Arabic script it could be used in the extraction of text lines for Latin script, by just changing the angle parameter and the field of vision direction.

(a) (b)

(c) (d)



(e)

VI. Conclusion

In this paper we presented an approach based on connected components analysis for extracting text lines from the binary Arabic handwritten documents. We formulate the problem in such a way that allows us to use Markov Decision Processes to solve it, which has given promising results.

In the future work we estimate to improve the diacritical assignment in the process of line extraction and to incorporate a preprocessing phase concerning the segmentation of touching and overlapping lines to accurately separate the different strokes belonging to different characters.

References

[1] Likforman-Sulem, L., Zahour, A., & Taconet, B. (2007). Text line segmentation of historical documents: a survey. International Journal of Document Analysis and Recognition (IJDAR), 9(2-4), 123-138.

[2] Li, Y., Zheng, Y., Doermann, D., & Jaeger, S. (2008). Script-independent text line segmentation in freestyle handwritten documents. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 30(8), 1313-1329

[3] Ouwayed, N., & Belaïd, A. (2012). A general approach for multi-oriented text line extraction of handwritten documents. International Journal on Document Analysis and Recognition (IJDAR), 15(4), 297-314

[4] Razak, Z., Zulkiflee, K., Idris, M. Y. I., Tamil, E. M., Noor, M. N. M., Salleh, R., ... & Yaacob, M. (2008). Off-line handwriting text line segmentation: A review. International journal of computer science and network security, 8(7), 12-20.

[5] Stamatopoulos, N., Gatos, B., Louloudis, G., Pal, U., & Alaei, A. (2013, August). Icdar 2013 handwriting segmentation contest. In Document Analysis and Recognition (ICDAR), 2013 12th International Conference on (pp. 1402-1406). IEEE

[6] Gatos, B., Stamatopoulos, N., & Louloudis, G. (2010, November). Icfhr 2010 handwriting segmentation contest. In Frontiers in handwriting recognition (icfhr), 2010 international conference on (pp. 737-742). IEEE.

[7] Bertsekas D. P., Dynamic Programming : Deterministic and Stochastic Models, Prentice-Hall, 1987.

[8] Puterman M., Markov Decision Processes : Discrete Stochastic Dynamic Programming, John Wiley & Sons, Inc., New York, USA, 1994.

[9] Bennasri, A., Zahour, A., & Taconet, B. (1999). Extraction des lignes d'un texte manuscrit arabe. In Vision interface (Vol. 99, pp. 42-48).

[10] Nicolaou, A., & Gatos, B. (2009, July). Handwritten text line segmentation by shredding text into its lines. In Document Analysis and Recognition, 2009. ICDAR'09. 10th International Conference on (pp. 626-630). IEEE

[11] Adiguzel, H.; Sahin, E.; Duygulu, P., "A Hybrid for Line Segmentation in Handwritten Documents," Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on , vol., no., pp.503,508, 18-20 Sept. 2012

[12] Zahour, A., Likforman-Sulem, L., Boussalaa, W., & Taconet, B. (2007, September). Text line segmentation of historical arabic documents. In Document Analysis and Recognition, 2007. ICDAR 2007. Ninth International Conference on (Vol. 1, pp. 138-142). IEEE.

[13] Khandelwal, A., Choudhury, P., Sarkar, R., Basu, S., Nasipuri, M., & Das, N. (2009). Text line segmentation for unconstrained handwritten document images using neighborhood connected component analysis. In Pattern Recognition and Machine Intelligence (pp. 369-374). Springer Berlin Heidelberg

[14] Khayyat, M., Lam, L., Suen, C. Y., Yin, F., & Liu, C. L. (2012, March). Arabic handwritten text line extraction by applying an adaptive mask to morphological dilation. In Document Analysis Systems (DAS), 2012 10th IAPR International Workshop on (pp. 100-104). IEEE

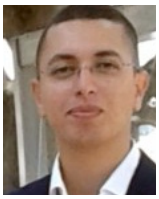
[15] Shi, Z., Setlur, S., & Govindaraju, V. (2009, July). A steerable directional local profile technique for extraction of handwritten arabic text lines. In Document Analysis and Recognition, 2009. ICDAR'09. 10th International Conference on(pp. 176-180). IEEE.

[16] Kumar, J., Abd-Almageed, W., Kang, L., & Doermann, D. (2010, June). Handwritten Arabic text line segmentation using affinity propagation. InProceedings of the 9th IAPR International Workshop on Document Analysis Systems (pp. 135-142). ACM

[17] Kumar, J., Kang, L., Doermann, D., & Abd-Almageed, W. (2011, September). Segmentation of handwritten text lines in presence of touching components. In Document Analysis and Recognition (ICDAR), 2011 International Conference on (pp. 109-113). IEEE

[18] Handwritten Arabic Proximity Datasets. Language and Media Processing Laboratory.http://lampsrv02.umiacs.umd.edu/projdb/project.php?id=65

[19] Phillips, I. T., & Chhabra, A. K. (1999). Empirical performance evaluation of graphics recognition systems. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 21(9), 849-87



Youssef Boulid received his M.S. degree in Decision Support Systems and Project Management in 2012 from univesity Ibn Tofail, Faculty of science, Kenitra- Morocco. Currently he is preparing a PhD in the Computer Science Department at the same faculty. His research interests include image processing, handwritten document analysis, Arabic handwritten recognition and Artificial intelligence.

Abdelghani Souhar received M.S. degree in applied Mathematics in 1992, PhD degree in computer science in 1997 from University Mohammed 5 - faculty of science in Rabat - Morocco. Now he is a Professor at university Ibn Tofail - faculty of science in Kenitra - Morocco. His research interests include Computer Aided Engineering, Computer Aided Design and Artificial Intelligence.

Mohamed Elyoussfi Elkettani received M.S. degree in applied mathematics in 1980 and PhD degree in Statistics in 1984 from Orsay Faculty of Science, University of Paris XI. Now he is a Professor at university Ibn Tofail - faculty of science in Kenitra - Morocco. His research interests include Multivariate statistics and Image recognition algorithms.

MaxHopCount: A New Drop Policy to Optimize Messages Delivery Rate in Delay Tolerant Networks

Youssef Harrati, Abdelmounaim Abdali

Cadi Ayyad University, Faculty of Sciences & Techniques. Laboratoire de Mathématiques Appliquées et Informatique (LAMAI), Marrakech, Morocco

***Abstract* — Communication has become a necessity, not only between every point on the earth, but also on the globe. That includes hard topography, highlands, underwater areas, and also space-crafts on other planets. However, the classic wired internet cannot be implemented in such areas, hence, researchers have invented wireless networks. The big challenge for wireless networking nowadays, is maintaining nodes connected in some difficult conditions, such as intermittent connectivity, power failure, and lot of obstacles for the interplanetary networks. In these challenging circumstances, a new networking model arises; it is Delay Tolerant networking which is based on the Store-Carry-and-Forward mechanism. Thus, a node may keep a message in its buffer for long periods of time; until a delivery or forward chance arises then it transmit it to other nodes. One of the big issues that confront this mechanism is the congestion of nodes buffer due to the big number of messages and the limited buffer size. Here, researchers have proposed buffer management algorithms in order to deal with the buffer overload problem, and they called it Drop Policies. In our present work, we propose a new Drop policy which we have compared to other existing policies in different conditions and with different routing protocols, and it always shows good result in term of number of delivered messages, network overhead and also average of latency.**

***Keywords* — Buffer Management, Delay Tolerant Networks, Drop Policy, Mobility Models, Network Simulator (ONE), Scheduling.**

I. Introduction

Delay Tolerant Networks (DTN) is a new concept of networking, which was proposed by Kevin Fall et .al [1] in 2003. Several recent studies are focused on DTN networks, and consider it to be one of the aspects of mobile network development in the future.

The classic TCP-IP [2] Internet model provides an end-to-end communication, which requires a safe path between source and destination, and a large bandwidth. Nevertheless, when a wireless mobile network suffers from the lack of path between source and destination, intermittent connectivity, as well as long latency and limited bandwidth, the TCP concept can, unfortunately, no longer be applied. Under these particular circumstances, DTN networks were introduced.

Every DTN node has a local storage area where it stores messages until a delivery opportunity arises. However, when the node buffer is full of messages (we call it a congested node) and a new message arrives, we don't know what to do and which message is the best to drop to free space without decreasing the delivery rate. Thus, researchers have proposed a set of drop policies.

In this paper, we compare some existing buffer management policies with existing routing protocols, then after deciding which routing protocol and which buffer policies are optimal for such environment, we will compare these optimal algorithms with our new algorithm "MaxHopCount" which we have developed in our laboratory.

The rest of this paper is organized as follows: Section 2 is about the state of the art where we give brief definition of some existing buffer management policies and other characteristics of DTNs. Section 3 describes our proposed algorithm. Section 4 summarizes performance metrics and shows simulation results and discussion. And finally, Section 6 is reserved for conclusion and future work.

II. State of the Art

A. Existing Buffer management policies (Drop policies)

In order to deal with different problems related to buffer overload, researchers have proposed a set of drop policies which give different results depending on the environment (traffic density, area size, buffer size, TTL …). TABLE 1 shows a brief definition of some existing drop policies:

TABLE I
DESCRIPTION OF SOME BUFFER MANAGEMENT POLICIES

Policy	Description
Drop Front (FIFO)	Using this policy, the message which arrived first in the buffer will be selected first to be dropped. [3]
Drop Tail (LIFO)	The buffer in this strategy is ordered in a LIFO manner, so the message which arrives last to the queue will be the first dropped message
Drop Oldest (SHLI)	Every message has a time to live value (TTL), the message with the lowest TTL value is the oldest one, and is the first selected message to drop.
E-Drop (Equal Drop)	If there is no space to host the new coming message in the buffer, this policy deletes the message having the exact or the nearest size. This strategy minimizes the number of dropped messages. [4]
Drop Youngest	The youngest message is the message with the longest remaining time to live (TTL), and it is the first dropped message when this strategy is used. [5]
Drop Largest	Each message is defined by a size; this buffer management strategy drops the message with the biggest size in the queue to free more space by dropping few messages. [6]
Evict Most Forwarded First (MOFO)	The message that has been forwarded to maximum number of nodes will be dropped first.[7]
Evict Most Favorably Forwarded First (MOPR)	Each message in node has a forwarding predictability FP, initially set to 0. When the message is forwarded, the FP value increases. The message with maximum FP value will be dropped first. [8]

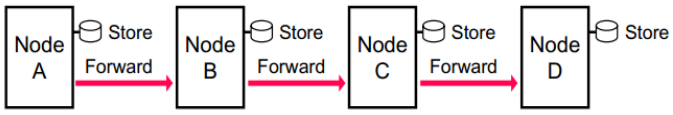
DOI: 10.9781/ijimai.2016.417

B. Routing in DTN

Routing protocols in Delay Tolerant Networks have many characteristics, so they can have many classifications. Some can classify routing algorithms as replication-based algorithms (i.e:the protocol creates message replicas) or forwarding-based (i.e: the protocol don't replicate messages) algorithms. [15]

Some DTN routing protocols can be classified in both categories like "spray and wait" witch benefits from the high delivery ratio of replication-based routing as well as the low resources consumption of forwarding-based routing, and which we will use in our simulation at the bottom of this paper.

NB: All routing protocols in Delay Tolerant Networks are based on Store-carry and Forward mechanism.



1) FirstContact:

A very simple routing algorithm for DTN, and a very quick one, this protocol forwards just a single copy of the message or the fragment of message to the first available contact.

2) Epidemic

From Wikipedia, an epidemic is the rapid spread of infectious disease to a large number of people in a given population within a short period of time. DTN researcher took the same idea to create a routing protocol where every node transmits replicas of its messages to newly discovered node which don't have it already. Theoretically, this algorithm needs to have unlimited buffer size and unlimited energy to give high rate of delivery, but practically this conditions are impossible to implement. [9]

3) Prophet V2

The Prophet routing protocol "probabilistic routing protocol using history of encounters and transitivity." operates on the assumption that human mobility is non-random, and that knowledge of the history of previous encounters is a good indicator of future encounters. In June 2011, Samo Grasic(samo@grasic.net) have proposed a new version of this protocol ProphetV2 white considers also the elapsed time during nodes encounters. [10]

4) Spray & Wait

Spray and Wait combines the speed of epidemic routing with the simplicity of direct transmission. So, for the spray phase: for each message M at the source node, L copies of M are forwarded to L different relays (intermediate nodes). Then in the wait phase: if the destination is not found among the relays, every relay will perform direct transmission to the message destination only, and so on until each message in the source node buffer reaches its destination. [11]

C. Nodes mobility in DTN

Researchers have proposed many mobility models for DTN to track nodes geographic movement in the network by collecting a set of parameters like speed, direction, and acceleration of the node. In DTN there are many types of mobility models:

1) Analytical models

Analytical mobility models are based on mathematic equations for the prediction of the next position of each node

a) Random walk

For this model, the parameters used to determine the next position of the node are the direction and the movement speed. First, the node choose a direction angle between 0 and 2π, then it chooses a random speed between [Vmin, Vmax] and finally after reaching t time it chooses other random values starting from this new position to another.

b) Random waypoint

By the opposite of the precedent model, in this one the node stops for a random pause period when it reaches a destination point before it chooses other speed and direction angle.

2) Traces based models

Traces based model are more realistic mobility models than analytical models, and are obtained from real results. These results are more accurate where the number of mobile nodes is important. The resulted traces are used then to make a new mobility model or to improve existing ones. Some examples are discussed here:

a) ZebaNet

A model made by collected information from some electronic devices placed on zebras at SweetWaters natural reserve in Kenya. [12]

b) DakNet

This model is implemented in developing countries, where the cost of implementing an infrastructure is very high. For this reason the messages are routed by bus or motorcycles traveling between villages. [12]

III. Proposed Policy "MaxHopCount"

A. Description

During its way to destination, every message traverses many nodes and stay a while in every node's buffer, before jumping to another node. Every message has a set of information, such as size, source, destination… etc. among these informations, there is "hop count", it refers to the number of nodes the message has crossed during its path from source to the current node. A high hop count means that the message crossed lot of node, and then there may be lot of copies at the network, so dropping this message from the buffer may not impact its delivery. Otherwise, a low hop count means low number of replicas at the network so dropping this message may lessen the delivery probability.

B. The Flow-sheet:

First, the function checks if the message size is greater than the whole buffer size, if so, the function ends because the message is too big. Otherwise, is initializes m by the message with the highest Hop Count if m is null, so the function cannot delete any more messages and the new message can't be hosted at this node. If m is not null we free the buffer from the message that m is pointing at, and we compare the size of the new coming message with the free buffer size. We repeat those steps until the buffer have enough space to host the new message. Fig. 1 below summarizes this algorithm

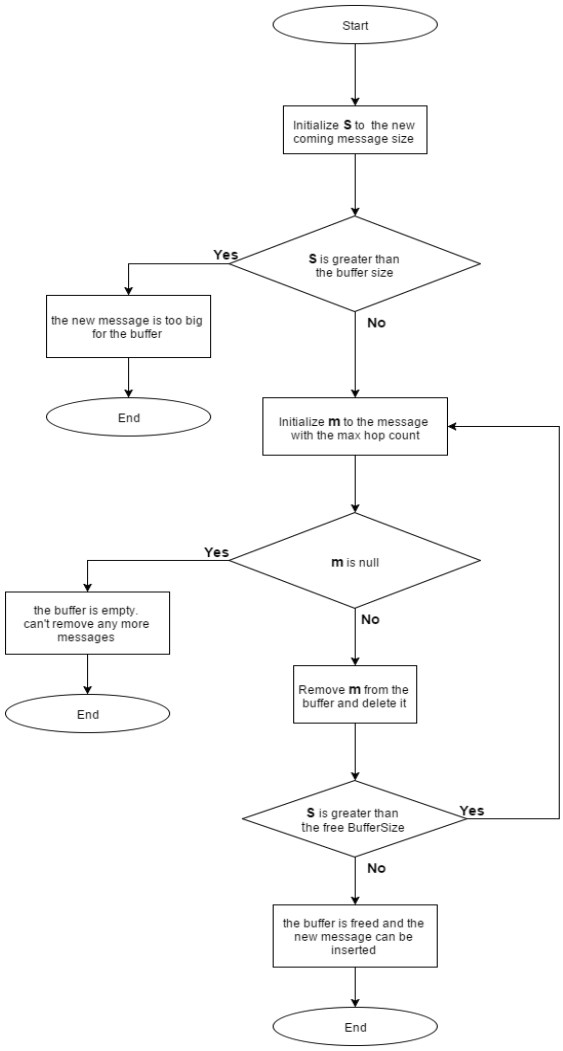


Fig. 1. The flow sheet of the new drop policy MaxHopCount

IV. Results and Discussion

A. The ONE simulator

The ONE Simulator is a discrete event simulator written in Java. The major aspire of simulator is to relate DTN (store-carry-forward) of message for long time, where the probability of disconnections and failures enlarged. [13]

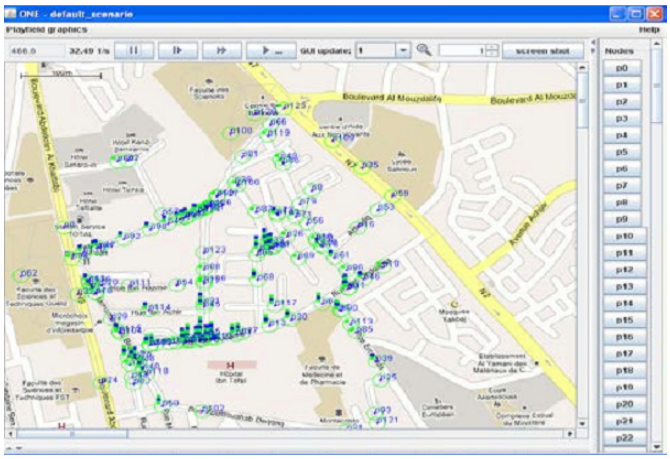


Fig. 2. ONE GUI mode of the ONE Simulator area: Helsinki city.

B. Common metrics for performance evaluation

The following metrics are commonly used when evaluating scenarios related to DTN protocols. [14]

1) Delivery ratio.

Suppose that M be the set of all messages created in the network and Md be the set of all messages delivered. Then, the delivery ratio is computed as:

$$M_d / M$$

2) Average message delivery latency.

Now let the i^{th} delivered message was created at time c_i and delivered at time d_i. Then the average message delivery latency is computed as:

$$\left(\sum_{i=1}^{M_d}(d_i - c_i)\right)\Big/ M_d$$

3) Overhead ratio

Let r_i be the number of replications of any message $m_i \in M$. Then the overhead ratio is determined as:

$$\left(\sum_{i=1}^{M}(r_i - M_d)\right)\Big/ M_d$$

C. Simulation parameters

To simulate our work we switch between different environments and we change many parameters such as routing protocols where we switch between Epidemic and ProphetV2 routers. Table II shows the important parameters of our simulations.

TABLE II
SIMULATION PARAMETERS

Variable	Value
movement Model	RandomWalk
Router	Epidemic – ProphetV2
buffer Size	10M
drop Policy	FIFO – LIFO – DL – DY – SHLI – MaxHopCount - MOFO
Message TTL (in minutes)	60 -120 -180 -240 -300
number of Hosts	200
Message creation interval (in seconds)	25,35
Messages size	500k,1M
World Size (width, height; meters)	500, 500

D. Epedimic

In the first simulation we compare different drop policies with different TTL values, then we analyse the results in term of delivery rate and overhead ratio. The routing protocol of this first simulation is Epidemic.

1) Delivery rate

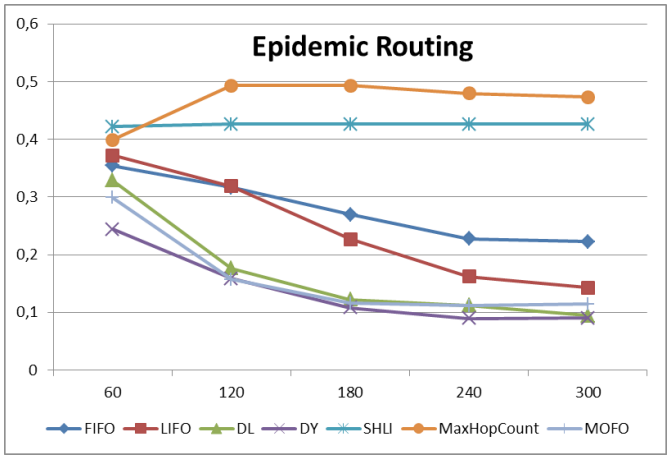


Fig. 3. Delivery rate as a function of TTL for different drop policies. The router is Epidemic.

As shown in the graph above, MaxHopCount policy, have a good delivery probability when the TTL value is high. the weakness of other policies, apart from SHLI (Shortest TTL), is that the delivery rate keep decreasing while the TTL value is heightening.

2) Overhead Ratio

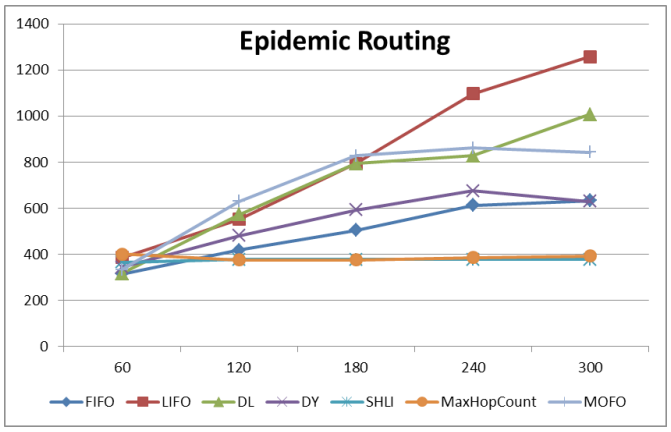


Fig. 4. Overhead ratio as a function of TTL for different drop policies. The router is Epidemic.

As shown above, the Overhead ratio is the number of message replicas in the network, so the main goal of all DTN algorithms is to minimize this ratio.

From Fig. 4, we observe that MaxHopCount policy doesn't overload the networks with message replicas because it gets low overhead ratio, this ratio keeps lessening which means that our new policy is also optimal in term of network overload.

E. ProphetV2

For the second simulation, the chosen routing protocol is ProphetV2. Like the first simulation, we compare here different drop policies with different TTL values, then we analyse the results in term of delivery rate and overhead ratio.

1) Delivery rate

In the opposite of Epidemic, ProphetV2 is an example of a probabilistic routing protocol. In this case, MaxHopCount have the highest delivery rate as well. And the increasing of the TTL value doesn't impact it.

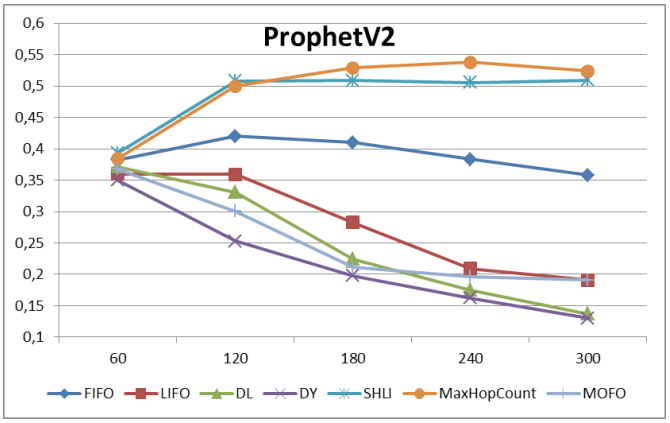


Fig. 5. Delivery rate as a function of TTL for different drop policies. The router here is ProphetV2.

2) Overhead ratio

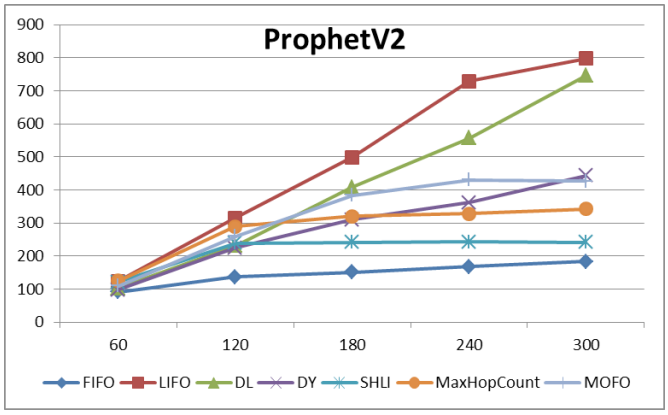


Fig. 6. Overhead ratio as a function of TTL for different drop policies. The router here is ProphetV2.

From Fig. 6 we can notice that the overhead ratio of our policy is not the lowest one, but it has a stable value and it is better than other policies where the high TTL have a very negative impact on this metric

V. Conclusion & Future Works

In this work, we compared some existing DTN drop policies to our new proposed algorithm, and we discussed the simulations results in term of delivery rate, overhead ratio, then we deduct that our new policy MaxHopCount is optimal for high TTL values (greater than 180 minutes or 3 hours), which is more realistic because usually in real scenarios the TTL may reach many days.

As a future work, we are working on this policy to study its efficiency with other parameter such as the area size, the message size... and we have also to choose the policy to be applied in ambiguous situations like equals numbers of hop.

References

[1] Fall, K. (2003). A Delay Tolerant Networking Architecture for Challenged Internets. SIGCOMM.

[2] RFC 1155

[3] Sulma Rashid, Q. A., M. Soperi Mohd Zahid, A.Hanan. Abdullah (2011). "E-DROP: An Effective Drop Buffer Management Policy for DTN Routing Protocols." International Journal of Computer Applications 13(7): 8-13.

[4] Yun Li, L. Z., Zhanjun Liu, Qilie Liu (2009). "N-Drop: congestion control strategy under epidemic routing in DTN." IWCMC '09 Proceedings of the

2009 International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly: 457-460

[5] Thrasyvoulos Spyropoulos, K. P., Cauligi S. Raghavendra (2005). Spray and Wait: An Efficient Routing Scheme for Intermittently Connected Mobile Networks. SIGCOMM'05 Workshops. Philadelphia, PA, USA.

[6] Sulma Rashid, Q. A., M. Soperi Mohd Zahid, A.Hanan. Abdullah (2011). "Impact of Mobility Models on DLA (Drop Largest) Optimized DTN Epidemic routing protocol." International Journal of Computer Applications 18: 35-39.

[7] Krifa, A. B., C. (2008). Optimal Buffer Management Policies for Delay Tolerant Networks. 5th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks. San Francisco, CA: 260 - 268.

[8] Samina Mansuri, H. S., Yogeshwar Kosta "Performance Analysis of Drop Policies for Different Mobility Models in DTN." nternational Journal of Computer Applications 59 (14).

[9] Amin Vahdat, D. B. "Epidemic Routing for Partially-Connected Ad Hoc Networks."

[10] Ahmed El Ouadrhiri, I. R., Mohamed El-Kamili, Ismail Berrada (2014). Controlling messages for probabilistic routing protocols in Delay-Tolerant Networks. ISCC: 1-6.

[11] Thrasyvoulos Spyropoulos, K. P., Cauligi S. Raghavendra (2005). Spray and Wait: An Efficient Routing Scheme for Intermittently Connected Mobile Networks. SIGCOMM'05 Workshops. Philadelphia, PA, USA.

[12] (2012). Une solution tolérante aux délais pour des applications de localisation et de traçabilité a posteriori en milieux confinés, UNIVERSITÉ BLAISE PASCAL.

[13] Ari Keränen, J. O., Teemu Kärkkäinen (2009). The ONE Simulator for DTN Protocol Evaluation. 2nd International Conference on Simulation Tools and Techniques. Rome, Italy.

[14] Saha, B. K. (2014). "Commonly Used Metrics for Performance Evaluation." Retrieved 20 Sep 2014, 2014, from http://delay-tolerant-networks.blogspot.com/2014/03/commonly-used-metrics.html.

[15] Torres-Franco, E., et al. "A Quantitative Justification to Dynamic Partial Replication of Web Contents through an Agent Architecture", International Journal of Interactive Multimedia and Artificial Intelligence, 3(3), pp. 82-88, 2015.

Youssef Harrati IT Engineer graduated in 2013 from the Cadi Ayyad University of Marrakech / Morocco. PhD student at the Laboratory of Applied Mathematics and Computer Science (LAMAI) of the doctoral studies center of Cadi ayyad University Marrakech / Morocco. His research interests are computer science, Delay Tolerant Networks, sensor networks and interplanetary networks.

Abdelmounaïm Abdali PhD in Solid Mechanics and Structures in University of Amiens in 1996, France. He is a Professor in computer science at the University Cadi Ayyad, Faculty of Sciences and Technics, Marrakech, Morocco. Member at the Laboratory of Applied Mathematics and Computer Science (LAMAI) Marrakech,Morocco. His research interests are computer science, DTN Network, business intelligence, Semantic web, Big Data , ubiquitous computing, biomechanics, bone remodeling, numerical simulation in solid mechanics and structures.

A Distributed Intelligent System for Emergency Convoy

M. Benalla, B. Achchab, H. Hrimech

Univ. Hassan 1, Higher School of Technology, Laboratory of Analysis and Modeling Systems for Decision Support, Berrechid, Morocco

***Abstract* — The general problem that guides this research is the ability to design a distributed intelligent system for guiding the emergency convoys; a solution that will be based on a group of agents and on the analysis of traffic in order to generate collective functional response. It fits into the broader issue of Distributed Artificial System (DAI), which is to operate a cooperatively computer agent into multi-agents system (MAS). This article describes conceptually two fundamental questions of emergency convoys. The first question is dedicated to find a response to the traffic situation (i.e. fluid way), while the second is devoted to the convoy orientation; while putting the point on the distributed and cooperative resolution for the general problem.**

***Keyword* — Traffic Flow, Multi Agent System, Shortest Path, Iot, Graph Theory.**

I. Introduction

The Distributed artificial Intelligence had its inauguration in the late 1970s [1]. The definition, in its most common acceptance of research of artificial intelligence is the study and design of artificial agents organization for intelligent systems [2]. DAI has emerged specifically to develop and solve problems often encountered in everyday life and which require to lend them a collective intelligence. Existing works consist essentially on predictive simulations (e.g. traffic flow) and on research in metaheuristic optimization (e.g. swarm intelligence) and so on.

The emergency convoy problem is a large issue, an issue centered optimization, so the arrival of the convoy to its destination in a reasonable time, is one of the main factors that reduce damage. The convoy guidance from the source to the target in the shortest time and in a safety environment is challenge that takes in its implication the following interrelated aspects: traffic density, road signs, and roads infrastructure quality.

For this purpose, the taking into account of these three aspects constitutes a basic foundation of emergency convoy problem. Traffic density which reflects the state of the environmental circulation leads a movement resistance of the convoy, more that this density is high, the friction to move increases in parallel direction. Sometimes, traffic situation reaches a case where vehicles remain motionless during seconds; thus, representing the worst case of a convoy and certainly not a wishing state. Road signs are also one of the formative elements of traffic density level, by means of their predetermined rules process the vehicles as the same level without considered priority cases of traffic (e.g. convoy of prisoners). While the quality of infrastructure plays a vital role for credible and fluid passage of emergency convoys. For this introduction, the next section makes clear the problem of emergency convoy.

II. The Emergency Convoy Problem

Nowadays, the emergency convoy confronts difficulties to reach the destination because of the complex configuration of road traffic network. Various road users continuously adjust their actions to avoid collisions in order to get into the desired environment. This problem becomes much serious when heterogeneous goals of road users meet; the thing that creates traffic jams and from time to time accidents, frequently trick to surpass by an emergency convoy. The distributed artificial intelligence is the driving force behind the resolution of this type of phenomenon which is inherently distributed. It provides a distributed resolution mechanisms through the focus on the communication and cooperation principle. For [3] the faster response, increased flexibility, robustness, resource sharing, graceful degradation and better adaptability are all advantages taking place into concept (i.e. DAI). In addition, with the emergence of the internet of things (IoT) paradigm, the development of autonomous entities becomes a promoter tool since the (IoT) exploits the internet network for information exchange.

This problem can be seen from two perspectives; a traffic manager's point of view, and an emergency conveyor's point of view: for the emergency conveyor's view are only problems to arrive safety and in a reasonable way to the destination; for the traffic manager's view it is to ensure control of the traffic flow.

The urban traffic system is composed of a network, operating rules, management and control system, and entities taking the network. The network is a set of communications and intersection axes [4]. This latter is in perpetually dynamic, which raises the complexity level. The paper will consider the problem of guiding emergency convoy from its departure point to the target, while using the shortest path with an easy passage (i.e. fluid traffic).

In the classical model of emergency convoy routing, conveyors exploit their knowledge about traffic situations, constantly watching the environment state and necessarily react to achieve the goal. So the anticipation field is limited to human feelings, particularly on the look.

III. Related Works

Existing studies do not address the emergency convoy problem itself, but the problem separately, There are those who are interested in the scope of road convoy problems (i.e. travel in platoon) as [5] [6] [7], while others put the points on the optimization problem of shortest path as [8] [9][15].

IV. An Agent- Based Model for Traffic Fluidization for Emergency Convoy

A. The notion of agent

There is a large number of works on the agents, the definitions

DOI: 10.9781/ijimai.2016.418

of agents and multi-agent systems. According to [10] the author, he considers the agent as a physical or abstract entity capable of acting on itself and on its environment, having a partial representation thereof that can communicate with other agents and whose behavior is the result of his observations, his knowledge and interactions with other agents.

Nowadays, a number of more and more software systems are designed in terms of agents and multi-agent systems in which software components rather act as independent individual entities, notably agents, instead of being only the components of the system. In problems of road traffic, this can be a vehicle, traffic lights, crossing, and so on.

B. Agent and traffic fluidization

The guiding complexity of emergency convoy is subjected to two levels, the first level focuses on finding the shortest paths, and the second level aims to improve traffic flow. The agent model for traffic fluidization in this paper puts back on the hierarchical architecture of information systems architecture that consists of three levels: a high level of decision, an intermediate level responsible for communication and coordination, and a low level for execution.

The hierarchical form has a set of privileges in an agent-based configuration. It allows in one hand to reduce redundancy and the amount of information exchanged between levels. On the other hand, to limit the number of communications between agents, and makes direct scheduling instructions.

The principal agents that coexist in the proposed model are the first and the last vehicle of emergency convoy, traffic lights, crossroads, coordinators, and the shortest path. All agents endow computational capabilities and communications models. A general modeling of model is given in Fig. 1.

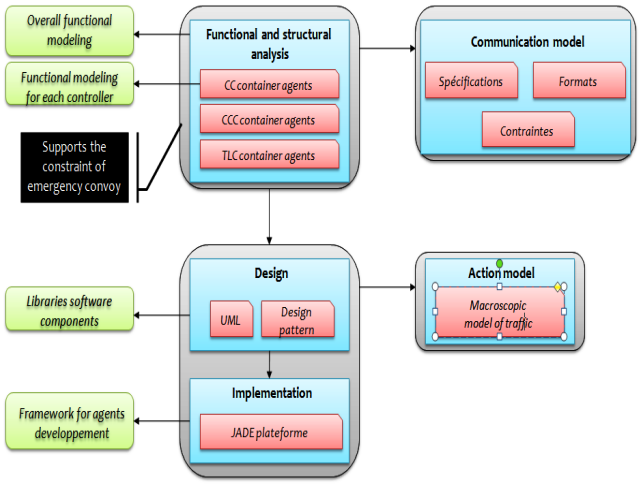


Fig. 1. A general modeling of the model

The communication model is a key factor in the effectiveness of the system, as we speak of a society of agents. It describes the communication forms between the model structures (i.e. CC, TLC, and CCC). These structures also containers of the above agents will undergo a definition in section V. The communication model was evoked to ensure system integrality and to ensure the autonomy of structures as well.

In the model, the traffic lights in the intersection and in the singleton are the decision point to solve the traffic fluidization problem. Generally, traffic lights are in two modes, a mode for vehicles (i.e. light red, orange and green) and a mode for pedestrians (i.e. red and green lights). Before proceeding to the description of the traffic fluidization components model of emergency convoy which is based on a group of agents in cooperation, it is necessary to highlight the strategies used to control traffic lights in the intersections since it is the decision point for the model as described above.

In the literature, there are three types of strategies [11], each strategy has its own rules to follow. The control variables of traffic lights are: the length of the cycle; the length of "active time"; the point of time when the cycle starts.

a) *Strategy 1: the semiactuated regime*

This strategy is used when the traffic flow in a route gets to an intersection which is considered important in term of vehicle flow. Then the traffic light remains green until the other road (i.e. not busy) marks the approach of vehicles. Note that the busy road (i.e. green light), the light cycle is predetermined to a maximum time even if the unoccupied road does not mark the presence of vehicles.

b) *Strategy 2: the platooning regime*

It is a technique exploited in moderate traffic flow. The vehicles travel in platoon separated by distances (i.e. time interval) facilitates the task to the control system of urban traffic. The traffic light is green when vehicles are approaching and red when a distance appears. It is a recommended strategy when it is possible to apply it.

c) *Strategy 3: the regime to control individual characteristics separately*

This mode of operation is based on three sets of rules which are ranked in the order of usage priority: modify cycle length; change cycle splits; change cycle start time. Each control has different variables

According to these strategies and as this is the issue of emergency convoys, an emergency strategy must take place, dealing with the emergency phenomenon in its entirety. From this point of view, the emergency strategy will help in emergency cases (e.g. ambulances, prisoner convoy, firefighter convoy, gear trains and so on) to put the traffic light in green mode as soon as the emergence of the convoy and in red mode every traffic light of the supervised environment by the Central Controller (CC) (i.e. crossroads lights, pedestrians lights).

V. A Distributed Intelligent System for Emergency Convoy (SDICU)

Through the multi-agent definition of our hierarchical system and a general modeling in Fig. 1, a distributed system for emergency convoys is designed. The general scheme consists of three hierarchical levels, Fig. 2: the Central Controller (decision-maker); the Communication and Coordination Controller (information provider); the Traffic Lights Controller (executor decision).

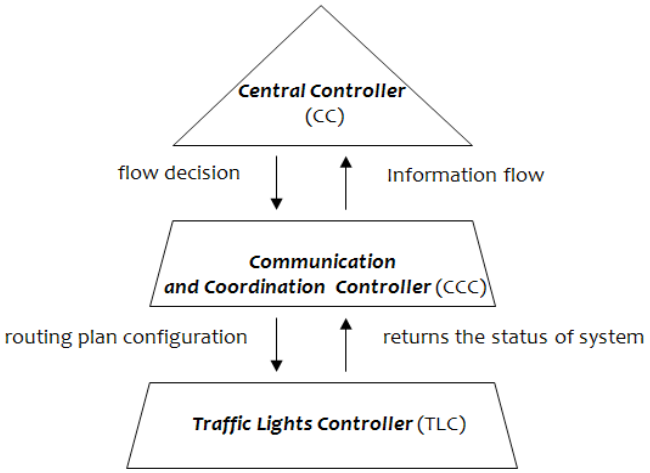


Fig. 2. levels of the multi-agent environment

As shown in Fig. 2, the (SDICU) consists of the functional components, Fig. 3. The central controller as the name indicates, establishes a policy decision on road map taking into account the information required by the subject (i.e. the communication and coordination controller); because the color of the traffic lights depends on its control strategy (i.e. emergency strategy). While the communication and coordination controller (CCC) reflects all the operations that will be applied on the traffic lights controller, based on the decisions flow provided by the (CC). In addition, the (CCC) is responsible for collecting information about the environment in which it is located. The (TLC) remains an application agent of orders received by the (CCC).

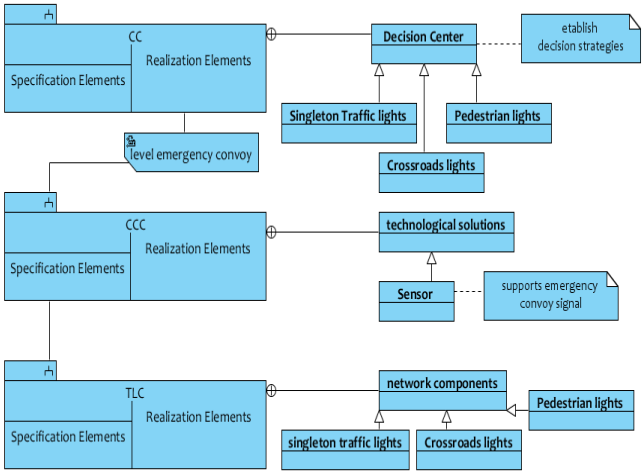


Fig. 3. Some functional components of environment

A. Finding the shortest paths

The orientation of emergency convoys is a complex challenge, regarding the multitude of ways to be followed. The emergency convoy first tries to reach the destination by travelling a short distance. It is mentioned that among agents that coexist in the system, an agent is in charge of calculating the shortest path.

The itinerary choice by an emergency convoy is governed by a prejudiced law on its capability in real time or by a stochastic law determining the probability of a point that will be visited by an exploration process. From a graph theory viewpoint, the problem can be formulated as follows: a graph G exposing all itineraries which emergency convoy can take; a set of end points (i.e. intersection) representing one or more issues on an edge (i.e. road segment). An edge is a combination linking each two points. In other words, let $G = (V, E)$ be a complete graph with a collection of nodes $V = (v_0, v_1, v_2, ..., v_n)$ and a collection of edges $E = (v_i, v_j) : v_i, v_j \in V, i \neq j$

In the literature there is a large number of calculation algorithms for the shortest path [12]. The Dijkstra's method, Fig 4, is a well known calculation problem of shortest path. It consists of a weighted directed graph for which all evaluations are positive to find the shortest between a node v and all other nodes. For each node v, we designate the distance between u and v by $D[u]$. By convention if there is no path between u and v we have $d(v,u) = \infty$.

Input : A simple undirected weighted graph G with nonnegative edge weights, and a distinguished vertex v of G
Output : A label $D[u]$, for each vertex u of G, such that $D[u]$ is the distance from v to u in G

$D[v] \leftarrow 0$
for each vertex $u \neq v$ of $\vec{G}$ **do**
  $D[u] \leftarrow +\infty$
Let a priority queue Q contain all the vertices of G using the D labels as keys
while Q is not empty **do**
  {pull a new vertex u into the cloud}
  $u \leftarrow Q.removeMin()$
  for each vertex z adjacent to u such that z is in Q **do**
    {perform the relaxation procedure on edge (u,z)}
    if $D[u] + w((u,z)) < D[z]$ **then**
      $D[z] \leftarrow D[u] + w((u,z))$
      Change to $D[z]$ the key of vertex z in Q
return the label $D[u]$ of each vertex u

Fig. 4. Dijkstra's algorithm for a graph G, starting from a vertex v.

B. Functioning

Initially, the study in this level (research the shortest distance) includes a step process.

1) Determination of the map of the city pointing out all possible routes and available.

2) On the map, the application of a points identification algorithm which will assign different symbols under the following rule:

- *For each traffic intersection, assign the symbol "Inter",*
- *For each traffic intersection with traffic lights, affect the symbol "InterPF",*
- *For each roundabout, assign the symbol "RPT",*

3) Calculation distance in meters between each two points connected,

4) Assignment of random characters for each point,

5) Build a matrix labeled with these characters. The combinations of values are distances in meters,

6) Running the Dijkstra algorithm for finding the shortest path.

After this, each emergency convoy will be integrated in a telecommunications network through the Internet of Things concept. At this moment, the shortest path agent communicates with each itinerary point where there are agents controllers responsible for communication and coordination. Each communication and coordination controller agent calculates the estimated time of arrival of the convoy based on their speed which is also provided by the shortest path agent.

At the time of approach of emergency convoy, the agent (CCC) determines the time for the active traffic light to turn green in the convoy sense based on velocity convoy (i.e. advanced preparation of environmental pass) to avoid road collisions. Subsequently, the same agent sends this information to the (CC) agent who charges to apply the aforementioned emergency strategy by returning instructions to the (CCC) agent that communicates them with the controller agent of any traffic light. After the convoy pass (i.e. pass timer), the (CCC) from information collected from its environment (i.e. sensor) requests by forwarding this information to (CC) the traffic strategy to be in place for the supervised area, Fig. 5.

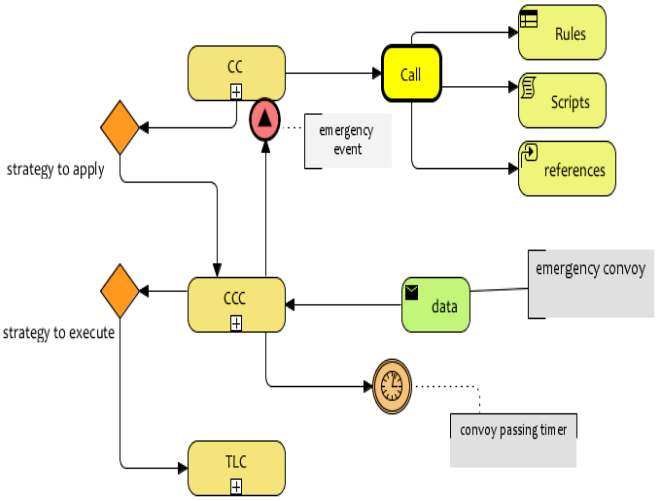


Fig. 5. operating process of the solution

VI. Conclusion

The work presented in this paper is an opening on a number of possibilities. The key of this opening passes undoubtedly by experimental work. The complex configuration of the road network makes it difficult to present a complete solution. In some emergency cases, convoys are face to face where a scheduling problem.

This paper only provides ideas for solving this kind problem of road traffic. We have seen how aspects: distributed artificial intelligence, multi-agent systems are essential to achieve a whole imitating artificial intelligent behavior. We wish to continue for some time our research on the experimental basis. The main features are multi-agent simulation and distributed problem solving, algorithms traffic, implementation of hardware technology; other issues will be raised by prediction but it is not impossible to envisage forthwith transposing a number of our results into reality, which would provide some opportunity to the solution that we defend.

Acknowledgment

This work is supported by the grant of National Center for Scientific and Technical Research (CNRST- Morocco): No. K005/006 for the first author.

References

[1] G. Weiss, "A modern Approach to Distributed Artificial Intelligence", page 1-2, MIT Press, 1999.

[2] J. Erceau, J. Ferber, "l'Intelligence Artificielle Distribuée, vol. 233

[3] N.V. Findler and G.D. Elder, "Multiagent Coordination and Cooperation in a Distributed Dynamic Environement with Limited Resources", Artificial Intelligence in Engineering, 9(1995), pp 229-238.

[4] M. Chabrol, D. Sarramia, « Modélisation orientée objet du système d'information des Systèmes de Trafic Urbain : une approche multi-agents »

[5] D. Goldstein, T. Shehab, J Casse, Hsiu-Chin Lin, "On the formulation and solution of the convoy routing problem", Transportation Research Part E, 2009.

[6] A. Rasheed, H. Zia, F. Hashmi, U. Hadi, W. Naim, S. Ajmal, "Fleet & Convoy Management Using VANET", Journal of Computer Networks, 2013, Vol. 1, No 1, 1-9.

[7] V. Baines, J. Padget, "Communication and metrics in agent convoy organization", Dept. of computer science, university of Bath.

[8] Y. Huang, G. Zhang, J. Wang, "An Optimization Dijkstra Algorithm Based on Two-function Limitation Strategy, Image and signal Processing, 2009, 1-4.

[9] S. Kadry, A. Abdallah, J. Chibli, "On The Optimization of Dijkstra's Algorithm", Electrical Engineering, 2012, Vol. 2, pp 393-397.

[10] J. Ferber, « Les systèmes multi-agents : vers une intelligence collective », InterEditions, ISBN : 2-72-96-0572-X, 1995.

[11] N.V. Findler, "Contibutions to a Computer-Based Theory of Strategies", Springer-Verlag, 1990.

[12] M.T. Goodrich, R. Tamassia, "Algorithm Design: Foundations, Analysis, and Internet Examples", pp 341-345

[13] M. Li, J. Hallam, L. Pryor, S. Chan, K. Chong, "A Cooperative Intelligent System for Urban Traffic Problems", Proceedings of the 1996 IEEE International Symposium on Intelligent Control Dearbor, 1996.

[14] J.M. Fernandes, E. Oliveira, "TraMas: Traffic Control through Behavior-based Multi-Agent System", ResearchGate, 2015.

[15] Guillermo Cueva-Fernandez et al., "Fuzzy decision method to improving the information exchange in a vehicle sensor tracking system.", Applied Soft Computing, 35, pp. 708-716, 2015.

Mohammed Benalla received his B.Sc. in Distributed Information System from Hassan II University. He is a Ph.D. student of the Higher School of Technology at Hassan 1st University. His research interests include distributed artificial intelligence and parallel and distributed computing. His research topic is about adaptation of flow traffic in urban growth during the Ph.D. program.

Boujemâa Achchab is a Full Professor in the Computer Science and Mathematics Department of the Higher School of Technology at Hassan 1st University and the director of the Laboratory of Analysis and Modeling Systems for Decision Support. He has a Ph.D. in applied mathematics from the University Lyon 1, France. His research interests include numerical analysis, data mining and econometric modeling.

Hamid Hrimech is an assistant Professor in the Computer Science and Mathematics Department of the Higher School of Technology at Hassan 1st University. He has a Ph.D. from Arts et Métiers ParisTech (France) in computer science. His research interests include collaborative interactions in collaborative virtual environment, AI, and driving simulation.

Statistical Comparisons of the Top 10 Algorithms in Data Mining for Classification Task

Nesma Settouti, Mohammed El Amine Bechar and Mohammed Amine Chikh

Biomedical Engineering Department, University of Tlemcen, Algeria

***Abstract* — This work is builds on the study of the 10 top data mining algorithms identified by the IEEE International Conference on Data Mining (ICDM) community in December 2006. We address the same study, but with the application of statistical tests to establish, a more appropriate and justified ranking classifier for classification tasks. Current studies and practices on theoretical and empirical comparison of several methods, approaches, advocated tests that are more appropriate. Thereby, recent studies recommend a set of simple and robust non-parametric tests for statistical comparisons classifiers. In this paper, we propose to perform non-parametric statistical tests by the Friedman test with post-hoc tests corresponding to the comparison of several classifiers on multiple data sets. The tests provide a better judge for the relevance of these algorithms.**

***Keywords* — The Top 10 Data Mining Algorithms, Classification, Statistical Comparisons Of Classifiers, Non-Parametric Test, Friedman Test, Post-Hoc Procedures.**

I. Introduction

Today, in the field of pattern recognition exists a large number of classifiers and feature selection methods. It is clear that no single model exists for all pattern recognition problems and no single technique is applicable to all problems. Rather, what we have is a bag of tools and a bag of problems [1]. Despite the numerous work in the field, that did not allow to highlight the indisputable superiority of one method of classification to another or a feature selection on another.

The identification of the top 10 algorithms by Wu et al. [2], [3] inspire us to explore these 10 algorithms, including their impact and new research issues. These 10 algorithms cover classification, clustering, statistical learning, association analysis, and link mining, which are all among the most important topics in data mining research and development, as well as for curriculum design for related data mining, machine learning, and artificial intelligence courses.

The initiative of identifying the top 10 data mining algorithms comes from a nomination and voting process. As the first step in the identification process, in September 2006 Wu et al. invited ACM KDD[1] Innovation Award and IEEE ICDM[2] Research Contributions Award winners to each nominate up to 10 best-known algorithms in data mining.

After the nominations, the step 2, concerns the verification of each nomination form its citations, and removing those nominations that did not have at least 50 citations. All remaining (18) nominations were then organized in 10 topics: association analysis, classification, clustering, statistical learning, bagging and boosting, sequential patterns, integrated mining, rough sets, link mining, and graph mining.

In the third step of the identification process, a wider involvement of the research community was necessary. Divers Program Committee members (KDD, ICDM, SDM[3], and ACM KDD) were invited to vote for the 10 well-known algorithms for the 18-algorithms candidate list. The voting results of this step were presented at the ICDM06 panel on Top 10 Algorithms in Data Mining.

In classification task, the choice of classifier in this long list of methods is primordial for a better recognition and this especially on the medical field. Modern medicine needs computer assistance for detection, diagnosis and classification of certain diseases in a very short time hence the need for a classification system. The use of intelligent methods to perform this classification is becoming more frequent. Although the decision of the doctor is the most critical factor in the diagnosis, a diagnostic to medical aid has developed and gained popularity, these systems are even considered as essential in many medical disciplines. In practice, there are already many applications using automatic learning that allows assisting clinicians in their diagnostic procedures. These approaches can provide a more accurate diagnosis and reduce the maximum errors due to fatigue and doubts of the doctor.

Therefore, in order to have an effective classification / regression systems from a set of representative examples of a population dataset. We must make the best choice of classifier. In this condition gives rise to a series of questions:

- How to know which is the most suitable classifier for a specific dataset?
- Are there cases to identify a classifier as a “logic” choice?
- What are the principles of selecting a classifier?

A recommended approach is to test several different classifiers as well as different parameter sets within each algorithm, and then to select the most effective using the non-parametric statistical tests. We talk about non-parametric tests when we make no assumptions about the distribution of variables. Also known as, free distribution tests, i.e. the quality of the results do not depend, a priori, on the underlying data distribution. An extensive study is performed by the Friedman test with post-hoc tests corresponding to the comparison of several classifiers on multiple data sets, in order to validate the most appropriate classifier structure, in terms of the correct classification rate and the generalization ability.

In this paper, we focus on the problem of determining the most suitable classifier to solve a given problem of classification. The choice of the classifier is already guided by operational constraints, but beyond these constraints, and after that the classifier is configured through a learning basis, the rate of generalization of the classifier (or Accuracy) which is the criterion characterizing its performance. This rate, usually unknown, is estimated using a generalized basis. This estimate, therefore, depends on the classification problem studied, the classifier uses the learning base and widespread basis. These dependencies are studied theoretically and experimentally over a dozen different classifiers. The problem of the validity of the comparison of two or more classifiers by estimates of their generalization rate is also studied by using non-

1. Association for Computing Machinery Knowledge Discovery and Data Mining
2. IEEE International Conference on Data Mining
3 SIAM International Conference on Data Mining

DOI: 10.9781/ijimai.2016.419

parametric tests. A ranking of classifiers goal provides in this work by testing this top 10 algorithms on different databases.

This paper is organized as follows: first, in section 2, the definitions of non-parametric tests are exposed. We present briefly in section 3, the 10 classifiers candidates in the standings. In Section 4, the step of experimentations and results, we discuss and analyze the results performed on 10 medical databases of different distributions and sizes. We conclude with a synthesis of this approach and a ranking of the 10 best methods for classification. (Section 5).

II. The Non-Parametric Tests

A non-parametric test is a test where the model does not specify the conditions that must fulfill the parameters of the population, which the sample was extracted. However, certain conditions of application should be checked. The samples must be considered random (when all people have the same probability to be a part of the sample) and single (all individuals who should form the sample has taken independently of each other) [4], and possibly independent from each other (use of random number tables). The random variables considered are generally assumed continuous.

Instead of entering in a debate "for or against" the nonparametric tests by opposing their parametric counterparts based on the normal data distribution. We try to characterize the situations where it is more (or less) advantageous to use them.

a) Advantages:

1. Their use is justified when the conditions for application of other methods are not satisfied even after possible variable transformations.
2. The probability of the results for most non-parametric tests are exact probabilities regardless of the distribution and the population shape with the sample is drawn.
3. For samples of very small size to N = 6, the only possibility is using a non-parametric test, unless the exact nature of the distribution of the population is precisely known. This allows a reduction in the cost or time needed to collect the information.
4. There are non-parametric tests for processing composite samples based on observations from different populations. Such data may only be processed by the parametric tests without making unrealistic assumptions.
5. Only non-parametric tests exist that allow the treatment of qualitative data either in rows or expressed more or less (ordinal scale) or nominal.
6. Non-parametric tests are easier to learn and apply than the parametric tests. Their relative simplicity often results from the replacement of the values observed either by alternative variables indicating membership in one or the other class observation or by the rows, i.e. the number order of observed values arranged in ascending order. Thus, the median is generally preferred to the mean, as seating position.

b) Disadvantages:

1. The parametric tests, when their conditions are satisfied, are more potent than the non-parametric tests.
2. A second disadvantage is the difficulty in finding the description of the tests and their tables of significant values. Fortunately, the standard statistical software gives the significance levels.

We selected the appropriate tests depending on the type of measurement, the shape of the frequency distribution and the number of samples that are available (see diagram Fig.1). Therefore, the Friedman test with the post-hoc approach applies to the static comparison studied in this work.

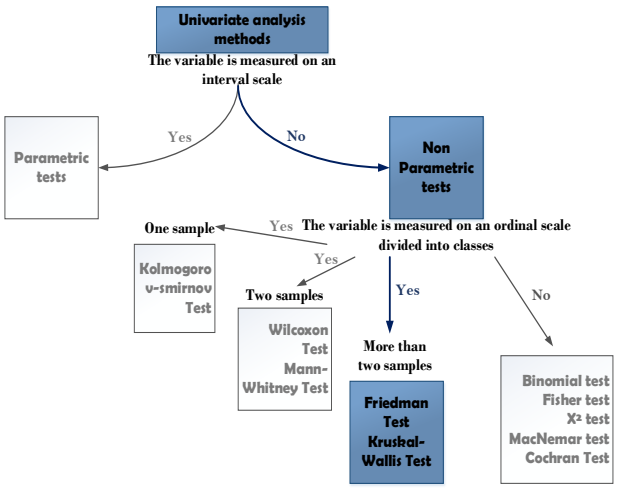


Fig 1. Univariate analysis methods diagram.

A. Friedman Test for Multiple Comparisons

In order to better assess the results obtained for each algorithm, we adopt in this study the post-hoc Friedman test methodology proposed by Demsar [5] for the comparison of several algorithms over multiple datasets.

The Friedman test [6] is a non-parametric test (free distribution) used to compare observations repeated on the same subjects. This is also called a non-parametric randomized black analysis of variance. The test statistic for the Friedman's test is a Chi-square with a-1 degrees of freedom, where a is the number of repeated measures. When the p−value for this test is small (usually < 0.05) you have evidence to reject the null hypothesis. The goal of this test is to determine whether there are significant differences among the algorithms considered over given sets of data. The test determines the ranks of the algorithms for each individual data set.

Garcia et al. [7] and Derrac et al. [8] considered non-parametric tests for multiple comparison as well as post-hoc procedures for NxN comparisons, for classification tasks. The studies illustrate that first the Friedman test should be conducted in order to detect whether statistically significant differences occur among the examined algorithms. Moreover, these tests rank the algorithms from the best performing one to the poorest one. If statistical significance is revealed, then the researcher may proceed to accomplish post-hoc procedures to point out which pair of algorithms differ significantly.

B. Post-hoc procedures for NxN comparisons

The Friedman can only detect significant differences over the whole multiple comparison, although they are not in a position to establish interrelations between the algorithms under consideration. If the null hypothesis of equivalence of rankings is rejected by these tests, the researcher may proceed with post-hoc procedures. Fig. 2 present the different procedures for multiple comparison NxN.

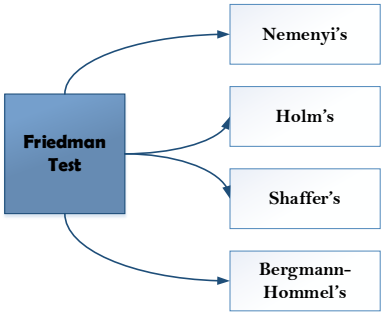


Fig 2. Non-parametric tests and post-hoc procedures for NxN comparisons.

TABLE I.
POST-HOC PROCEDURES FORNXNCOMPARISONS

PROCEDURE	DESCRIPTION	APV FORMULA
Nemenyi	Calculates the adjusted value of α in a single step by dividing it by the number of comparisons accomplished, i.e., k(k 1)/2.	min {v; 1}, where v =k(k- 1)pi/2.
Holm	Step-down method, it rejects H1 to Hi-1 if i is the smallest integer such that pi >α (k(k-1)/2-i+1).	min {v; 1}, where v = Max{(k(k-1)/2-j+ 1)pj : 1≤j≤i}
Shaffer	Following Holms step-down method, at stage j, instead of discarding Hi if pi ≤α(k(k-1)/2-i+ 1), discards Hi if pi≤α/ti, where ti is the maximum number of hypotheses which can be true given that any (i,…,1)hypotheses are false.	min {v; 1}, where v = max{tj pj : 1≤j≤i}
Bergmann and Hommel	Rejects all Hj with j ∉ A, where the acceptance set A, given as A= ∪ { I : I exhaustive, min{Pi: i ∈ g} > α/‖I‖}, is the index set of null hypotheses which are retained.	Min{v; 1}, where v = Max{‖I‖ min{pj, j ∈ I} : I exhaustive;i ∈ I}

In Table I a set of post-hoc procedures is presented for NxN comparisons. Trawinski et al. [9] summarized for each procedure a brief outline of its scheme and the formula for computation of the Adjusted P-Value (APV). The notation used in Table I is as follows:

- Indexes *i* and *j* apply to a given comparison or hypothesis in the family of hypotheses. Index i always concerns the hypothesis whose APV is being determined and index j refers to another hypothesis in the family;
- p_j is the p-value calculated for the j-th hypothesis;
- k is the number of predictors being compared.

III. The 10 Classifiers Candidates

There have been a large number of data mining algorithms rooted in these fields to perform different data analysis tasks. The 10 algorithms identified by the IEEE International Conference on Data Mining (ICDM) and presented in this article are among the most influential algorithms for classification, clustering, statistical learning and association analysis. We focus only on the Classification task of evaluating experiments with the listed algorithms in Table II over a set of selected medical databases.

TABLE II
DESCRIPTION OF THE 10 CLASSIFIERS CANDIDATES

METHODS	DESCRIPTION	REF
AdaBoost	Adaptive Boosting Negative Correlation Learning Extension with C4.5 Decision Tree as Base Classifier.	[10]
Apriori	Association rule mining using the Apriori algorithm.	[11]
Bagging	Multi-classifier learning approach with C4.5 as baseline algorithm.	[12], [13]
C4.5	Generate classifier expressed as decision trees	[14]
CART	Classification and Regression Tree.	[15]
EM	Expectation-Maximization algorithm	[16]
K-means	K means Classifier.	[17]
KNN	K-Nearest Neighbors Classifier.	[18]
NB	Nave-Bayes.	[19], [20]
SVM	Support vector networks.	[21]

IV. Experimentations and Results

A. Tools for the experimentations

All experiments were conducted using KEEL (Knowledge Extraction based on Evolutionary Learning) [22], [23], an open source Java software tool that can be used for a large number of different knowledge data discovery tasks. It contains a wide variety of algorithms for creating, learning, optimizing and evaluating various models ranging from soft computing ones to support vector machines, decision trees for regression, and linear regression. KEEL algorithms are employed to carry out.

The experiments listed in Table II, where references to source papers are shown. Details of the algorithms can also be found on the KEEL web site http://www.keel.es/.

B. Benchmark data sets

Twelve Medical and Biological datasets mainly selected from the UCI Machine Learning Repository [24], and ASU feature selection Repository [25]. These are used to evaluate the performance of the top 10 algorithms; their characteristics are described in Table III.

TABLE III
DESCRIPTION OF EXPERIMENTAL MEDICAL DATABASES

DATA SETS	# INSTANCES	# FEATURES	# LABELS
Appendicitis	106	9	2
Breast cancer	699	9	2
Dermatology	358	34	6
Diabetes	672	8	2
Heart	270	13	2
Heberman	306	3	2
Hepatitis	155	20	2
Liver disorder	345	7	2
Lymphoma	96	19	4
Mammographic	830	5	2
New-thyroid	215	5	3
Post-operative	87	8	3

C. Results

The top 10 machine learning algorithms were run in KEEL individually for 12 data sets using 10-fold cross validation (10cv) and the prediction error was measured with the Root Mean Square Error (RMSE) in Table IV.

Firstly, we used the non-parametric Friedman test to evaluate the rejection of the hypothesis that all the classifiers perform equally well for a given level. It ranks the algorithms for each dataset separately, the best performing algorithm getting the higher rank, for example, in the case with 4 classifiers, the best rank its equals to 1, the second best rank 2 etc.

TABLE IV
THE RMSE OF THE TOP 10 COMPARED ALGORITHMS OVER TWELVE MEDICAL AND BIOLOGICAL DATA SETS

DATA SETS	ADABOOST	APRIORI	BAGGING	C4.5	CART	EM	K-MEANS	KNN	NB	SVM
Appendicitis	0.197	0.182	0.153	0.153	0.137	0.181	0.176	0.164	0.198	0.127
Breast cancer	0.275	0.311	0.236	0.236	0.291	0.322	0.315	0.319	0.292	0.250
Dermatology	0.178	0.503	0.038	0.260	0.299	0.300	0.153	0.304	0.489	0.365
Diabetes	0.178	0.503	0.038	0.260	0.299	0.300	0.153	0.304	0.489	0.365
Heart	0.310	0.340	0.236	0.236	0.269	0.244	0.531	0.242	0.444	0.370
Heberman	0.121	0.218	0.434	0.270	0.300	0.330	0.510	0.335	0.435	0.294
Hepatitis	0.125	0.422	0.182	0.182	0.125	0.187	0.449	0.193	0.138	0.182
Liver disorder	0.181	0.245	0.139	0.330	0.333	0.389	0.256	0.389	0.255	0.368
Lymphoma	0.207	0.499	0.263	0.239	0.228	0.250	0.543	0.240	0.485	0.165
Mammographic	0.162	0.146	0.054	0.163	0.175	0.260	0.412	0.256	0.302	0.160
New-thyroid	0.452	0.319	0.260	0.059	0.067	0.027	0.506	0.025	0.348	0.060
Post operative	0.377	0.260	0.301	0.301	0.342	0.479	0.458	0.496	0.284	0.302

TABLE VI
ADJUSTED P-VALUES FOR N X N COMPARISONS OF THE TOP 10 ALGORITHMS OVER 12 DATA SETS

i	Hypothesis	Unadjusted p	pNeme	pHolm	pShaf	pBerg
1	C4.5 vs .K-means	***0.000209***	***0.009397***	***0.009397***	***0.009397***	***0.009397***
2	Bagging vs .K-means	***0.000455***	0.020483	0.020027	0.016386	0.012845
3	C4.5 vs .NB	***0.003012***	0.135554	0.129529	0.108443	0.098299
4	AdaBoost vs .K-means	***0.00336***	0.151179	0.1411	0.120943	0.012596
5	CART vs .K-means	***0.00336***	0.151179	0.1411	0.120943	0.012596
6	AdaBoost vs .SVM	***0.003743***	0.168428	0.149714	0.134742	0.129529
7	Bagging vs .NB	***0.005706***	0.256766	0.222531	0.205413	0.188684
8	C4.5 vs .EM	0.02391	1.075958	0.908587	0.860766	0.739714
9	AdaBoost vs .NB	0.028441	1.279844	1.052316	1.023875	0.962429
10	CART vs .NB	0.028441	1.279844	1.052316	1.023875	0.962429
11	NB vs .SVM	0.030971	1.393714	1.084	1.023875	1.01455
12	Bagging vs .EM	0.039753	1.788871	1.351591	1.152828	1.123586
13	Apriori vs .C4.5	0.043114	1.94015	1.422777	1.250319	1.205413
14	C4.5 vs .KNN	0.046713	2.102103	1.494828	1.354688	1.333108
15	Apriori vs .Bagging	0.068707	3.091828	2.129926	1.992512	1.71346
16	Bagging vs .KNN	0.073997	3.329882	2.219921	2.145924	2.091892
17	K-means vs .KNN	0.085576	3.850929	2.48171	2.48171	2.397533
18	Apriori vs .K-means	0.091892	4.135149	2.572982	2.572982	2.572982
19	AdaBoost vs .EM	0.138011	6.210483	3.72629	3.312258	3.238016
20	CART vs .EM	0.138011	6.210483	3.72629	3.312258	3.238016
21	EM vs .SVM	0.14719	6.623537	3.72629	3.532553	3.277384
22	EM vs .K-means	0.14719	6.623537	3.72629	3.532553	3.277456
23	AdaBoost vs .Apriori	0.212299	9.553437	4.882868	4.670569	4.312559
24	Apriori vs .CART	0.212299	9.553437	4.882868	4.670569	4.312559
25	AdaBoost vs .KNN	0.224916	10.121215	4.882868	4.723234	4.599059
26	Apriori vs .SVM	0.224916	10.121215	4.882868	4.723234	4.599059
27	CART vs .KNN	0.224916	10.121215	4.882868	4.723234	4.599059
28	KNN vs .SVM	0.23806	10.712699	4.882868	4.723234	4.599059
29	KNN vs .NB	0.328277	14.772476	5.580713	5.580713	5.580713
30	Apriori vs .NB	0.345231	15.535398	5.580713	5.580713	5.580713
31	C4.5 vs .SVM	0.418492	18.832151	6.277384	6.277384	6.277384
32	C4.5 vs .CART	0.438145	19.716515	6.277384	6.277384	6.277384
33	AdaBoost vs .C4.5	0.438145	19.716515	6.277384	6.277384	6.277384
34	K-means vs .NB	0.458318	20.624296	6.277384	6.277384	6.277384
35	EM vs .NB	0.479001	21.555056	6.277384	6.277384	6.277384
36	Bagging vs .SVM	0.543997	24.479865	6.277384	6.277384	6.277384
37	Bagging vs .CART	0.566597	25.496882	6.277384	6.277384	6.277384
38	AdaBoost vs .Bagging	0.566597	25.496882	6.277384	6.277384	6.277384
39	EM vs .KNN	0.787406	35.433292	6.277384	6.277384	6.277384
40	Apriori vs .EM	0.813456	36.605519	6.277384	6.277384	6.277384
41	Bagging vs .C4.5	0.839714	37.787108	6.277384	6.277384	6.277384
42	AdaBoost vs .SVM	0.973108	43.789878	6.277384	6.277384	6.277384
43	Apriori vs .KNN	0.973108	43.789878	6.277384	6.277384	6.277384
44	CART vs .SVM	0.973108	43.789878	6.277384	6.277384	6.277384
45	AdaBoost vs .CART	1	45	6.277384	6.277384	6.277384

Then, the Friedman test compares the average ranks of the algorithms and calculates the Friedman statistic. If a statistically significant difference in the performance is detected, which means that some of the hypotheses in the experimentation have different distribution from one another, therefore, our next step will be to try to find out which pairs of our algorithms are significantly different then each other. We proceed with a post-hoc test.

We use the Nemenyi, Holm, Shaffer, Bergmann, and Hommel tests, find out which of the tested methods are distinctive among an NxN comparison. The post-hoc procedure is based on a specific value on the significance level α. Additionally, the obtained pvalue should be examined in order to check how different given two algorithms are. We fix the significance level α= 0.10 for all comparisons. Average rankings of the 10 algorithms over 12 medical and biological data sets for produced by the Friedman test are shown in Table V.

TABLE V
AVERAGE RANKINGS OF THE ALGORITHMS

ALGORITHM	RANKING
AdaBoost	4.5417
Apriori	6.0833
Bagging	3.8333
C4.5	3.5833
CART	4.5417
EM	6.375
K-means	8.1667
KNN	6.0417
NB	7.25
SVM	4.5833

D. Discussion

The results achieved in post-hoc comparisons for α = 0.10 are depicted in Table VI. The unadjusted values and adjusted p-values for Nemenyi, Holm, Shaffer, and Bergmann-Hommel tests for NxN comparisons for all possible 45 pairs of algorithms are placed in Table VI. The pvalues below 0.10 indicate that respective algorithms differ significantly in prediction errors; they were marked with an italic font.

Among the NxN procedures, the Bergmann-Hommels procedure is the most powerful one, but it requires intensive computation in comparisons comprising a bigger number of predictors. Thus, the Shaffers static routine or the Holms step down method is recommended. It should be noted that with 45 hypotheses Holm, Nemenyis, Shaffer and Bergmann-Hommel ones discard only four methods. C4.5 and Bagging revealed significantly better performance than most of the 10 algorithm, thus propelling it to the top spot of classification algorithms.

However, for multiple comparisons the more data sets used in tests the bigger the number of null-hypotheses rejected. Our investigation proved the usefulness and strength of multiple comparison, statistical procedures to analyses and select machine learning algorithms. The ranking reveals that C4.5, Bagging are the most influential for classification tasks. In the third place, we have CART and Adaboost with an equal score, after that SVM, Apriori, KNN, NB and finally the unsupervised classifier K-means.

V. Conclusion

In contemporary machine learning, one cannot say that a given algorithm is superior over another one, without the use of statistical analysis. Experimental results must be accompanied by a thorough statistical analysis, to prove that the reported differences between analyzed models are significant.

In this paper, we studied the application of non-parametric statistical tests and post-hoc procedures devised to perform multiple comparisons of classification algorithms over medical and biological benchmark data sets. We conducted experiments on statistical procedures designed especially for multiple NxN comparisons with the top 10 algorithms in data mining. The tests provide a ranking of the top 10 algorithms, revealing the C4.5, Bagging, Adaboost, CART and SVM for the five most relevant classification algorithms.

References

[1] L. Kanal, "Patterns in pattern recognition: 1968-1974," Information Theory, IEEE Transactions on, vol. 20, no. 6, pp. 697–722, Nov 1974.

[2] Xindong Wu, Vipin Kumar, J. Ross Quinlan, Joydeep Ghosh, Qiang Yang, Hiroshi Motoda, Geoffrey J. McLachlan, Angus Ng, Bing Liu,

[3] Philip S. Yu, Zhi-Hua Zhou, Michael Steinbach, David J. Hand, and Dan Steinberg, "Top 10 algorithms in data mining," Knowl. Inf. Syst., vol. 14, no. 1, pp. 1–37, Dec. 2007.

[4] Xindong Wu and Vipin Kumar, The Top Ten Algorithms in Data Mining, Chapman & Hall/CRC, 1st edition, 2009.

[5] R. A. Fisher, Statistical Methods for Research Workers, Cosmo study guides. Cosmo Publications, 1925.

[6] Janez Demsar, "Statistical comparisons of classifiers over multiple data˘ sets," J. Mach. Learn. Res., vol. 7, pp. 1–30, Dec. 2006.

[7] M. Friedman, "The use of ranks to avoid the assumption of normality implicit in the analysis of variance," Journal of the American Statistical Association, vol. 32, no. 200, pp. 675–701, 1937.

[8] Salvador Garcıa, Alberto Fernandez, Julian Luengo, and Francisco Herrera, "Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power," Inf. Sci., vol. 180, no. 10, pp. 2044–2064, May 2010.

[9] Joaqun Derrac, Salvador Garca, Daniel Molina, and Francisco Herrera, "A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms," Swarm and Evolutionary Computation, vol. 1, no. 1, pp. 3 – 18, 2011.

[10] Bogdan Trawinski, Magdalena Smetek, Zbigniew Telec, and Tadeusz Lasota, "Nonparametric statistical analysis for multiple comparison of machine learning regression algorithms.," Applied Mathematics and Computer Science, vol. 22, no. 4, pp. 867–881, 2012.

[11] S. Wang and X. Yao, "Multiclass imbalance problems: Analysis and potential solutions," IEEE Transactions on Systems and Man and Cybernetics Part B, vol. 42, no. 4, pp. 1119–1130, 2012.

[12] R. Srikant and R. Agrawal, "Mining quantitative association rules in large relational tables," in ACM SIGMOD International Conference on Management of Data, 1996, pp. 1–12.

[13] Leo Breiman, "Bagging predictors," Mach. Learn., vol. 24, no. 2, pp. 123–140, Aug. 1996.

[14] Mikel Galar, Alberto Fernandez, Edurne Barrenechea,´ Humberto Bustince, and Francisco Herrera, "An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes," Pattern Recogn., vol. 44, no. 8, pp. 1761–1776, Aug. 2011.

[15] J.R. Quinlan, C4.5: Programs for Machine Learning, Morgan Kauffman, 1993.

[16] L. Breiman, J.H. Friedman, R.A. Olshen, and C.J. Stone, Classification and Regression Trees, Chapman and Hall (Wadsworth and Inc.), 1984.

[17] Tapio Schneider, "Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values," Journal of Climate, vol. 14, no. 5, Mar. 2001.

[18] J.B. MacQueen, "Some methods for classification and analysis of multivariate observations," in 5th Berkeley Symposium on Mathematical Statistics and Probability, 1967, pp. 281–297.

[19] T. Cover and P. Hart, "Nearest neighbor pattern classification," IEEE Trans. Inf. Theor., vol. 13, no. 1, pp. 21–27, Sept. 2006.

[20] M.E. Maron, "Automatic indexing: An experimental inquiry," Journal of the ACM (JACM), vol. 8, no. 3, pp. 404–417, 1961.
[21] P. Domingos and M. Pazzani, "On the optimality of the simple bayesian classifier under zero-one loss," Machine Learning, vol. 29, pp. 103–137, 1997.
[22] C. Cortes and V. Vapnik, "Support vector networks," Machine Learning, vol. 20, pp. 273–297, 1995.
[23] J. Alcala Fdez, L. Sunchez, S. Garcia, M. J. del Jesus, S. Ventura, J. M. Garrell, J. Otero, C. Romero, J. Bacardit, V. M. Rivas, J. C. Fernández, and F. Herrera, "Keel: A software tool to assess evolutionary algorithms for data mining problems," Soft Comput., vol. 13, no. 3, pp. 307–318, Oct. 2008.
[24] Jesus Alcala-Fdez, Alberto Fernandez, Julian Luengo, Joaquın Derrac, and Salvador Garcıa, "KEEL data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework," Multiple-Valued Logic and Soft Computing, vol. 17, no. 2-3, pp. 255– 287, 2011.
[25] D.J. Newman, S. Hettich, C.L. Blake, and C.J. Merz, "Uci repository of machine learning databases," 1998.
[26] Reza Zafarani and Huan Liu, "Asu repository of social computing databases," 1998.

Nesma Settouti is an Assistant professor at Tlemcen University, Algeria. She received her Engineer degree in Electrical Biomedical from the Tlemcen University in 2009. In 2011, she obtains a Magisterial degree in the same option. She is currently pursuing her PhD Thesis in the Biomedical Engineering Laboratory and LIMOS of Aubière (France), her research interests are in computer assisted medical decision support systems, ensemble methods, neural networks, clustering methods, optimization, classification and artificial intelligence. She had published a great deal of research studies published at national and international journals, conference proceedings as well as chapter book in computer assisted medical decision support systems.

Mohammed El Amine Bechar got his master degree in Biomedical Engineering with an emphasis in medical images processing in 2013, from the University of Tlemcen, Algeria. Currently, he is working towards his Ph.D. In Biomedical engineering laboratory at the University of Tlemcen, Algeria. His main interests are in the areas of computer assisted medical decision support systems, classification, artificial intelligence and image processing.

Mohammed Amine Chikh is a Professor at the Tlemcen University. He is graduated from The Electrical Engineering Institut (INELEC) of Boumerdes –Algeria in 1985 with engineering degree in Computer science and in 1992 with a Magister of Electronic from Tlemcen University. He also received a Ph.D in electrical engineering from the University of Tlemcen (Algeria) and INSA of Rennes (France) in 2005. Actually, he is currently Professor at Tlemcen University, Algeria, and head of CREDOM research team at Biomedical Engineering Laboratory. He conducted post-doctoral teaching and research at the University of Tlemcen. Pr Chikh has published over 90 journal and conference papers to date and is involved in a variety of funded research projects related to biomedical engineering. His is a member of several scientific conferences. His research interests have been in artificial intelligence, machine learning, and medical data.

PI Stabilization for Congestion Control of AQM Routers with Tuning Parameter Optimization

S.Chebli[1], A.Elakkary[1], N.Sefiani[1], N.Elalami[2]

[1]LASTIMI, Graduate School of Technology of Salé, Crown Prince Street, B.P : 227 Salé medina –Morocco
[2]Mohamdia School of Engineers, P.O. Box. 765, Ibn sina Street, Agdal Rabat Morocco

***Abstract* — In this paper, we consider the problem of stabilizing network using a new proportional- integral (PI) based congestion controller in active queue management (AQM) router; with appropriate model approximation in the first order delay systems, we seek a stability region of the controller by using the Hermite-Biehler theorem, which isapplicable to quasipolynomials. A Genetic Algorithm technique is employed to derive optimal or near optimal PI controller parameters.**

***Keywords* — Congestion Control, Activequeue Management (AQM), Stabilization, PI Controller, First Order Delay System,Genetic Algorithm (GA).**

I. Introduction

Recently and due to the explosive growth of computer networks, several non linear phenomena of the internet have been discovered. For instance, it has been shown that network traffic congestion results in long time delays for data transmission and often makes the queue length in the buffer of the intermediate router overflow, and can even lead to network collapse. To face the congestion problems, TCP chiefly uses a network congestion avoidance which comprehend different aspects of an additive-increase multiplicative-decrease (AIMD) scheme, with other schemes such as slow-start[1][2] and [3].

It was shown that the problem of congestion control and development of AQM could be considered as a problem of regulating. Thus, due to the limitations of packet-dropping probability and the effects of propagation delays in TCP networks, the TCP AQM network was modeled as a time-delayed system based on a fluid-based model of the dynamics of the TCP and RED, developed by the stochastic theory [4][5][6] and [7].

From the control theory, the AQMs types PI [8], PID were built to control the congestion phenomenon. In this paper, we extend our previous work [9] by implementing Genetic Algorithm (GA) in determining PI Controller parameters to compensate the delay in the system. The Genetic Algorithm has been considered as a useful optimization techniques employing the principles of natural genetic systems [10] to search a global solution of the optimization problem.

II. TCP/AQM Model

Our study will focus on the sharing of a communication link between multiple transmitters at remote locations (Fig. 1).

In [6], the authors modeled TCP process with stochastic differential equations without taking slow start and timeoutmechanisms into account. In this model, the congestion window w(t) increase linearly if no packet loss is detected;otherwise it halves. Based on some reasonable assumptions, the following relations were gotten:

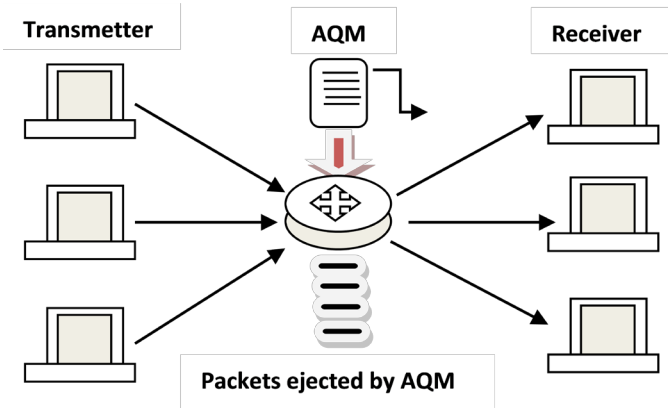


Fig. 1. Topology studied

$$\begin{cases} \dot{w}(t) = \frac{1}{R(t)} - \frac{w(t).w\left(t-R(t)\right)}{2R\left(t-R(t)\right)} p\left(t-R(t)\right) & (1) \\ \dot{q}(t) = \frac{w(t)}{R(t)} N(t) - C & (2) \end{cases}$$

With $R(t) = \frac{q(t)}{C} + Tp$, where Tp is the propagation delay, $q(t)$ is the queue length at the router, C is the router's transmission capacity, thus $\frac{q(t)}{C}$ is the queuing delay and $R(t)$ is the round trip time delay, $N(t)$ denotes the TCP load factor.

$\dot{w}(t)$ and $\dot{q}(t)$ denote the time-derivative of $W(t)$ and $q(t)$, respectively $p(t)$ is the probability of packet mark due to the AQM mechanism at the router.

The linearization of (1) and (2) about the operating point is carried out in (1) and the perturbed variables about the operating point satisfy

$$\begin{cases} \delta\dot{w}(t) = -\frac{N}{R_0^2 C}\left(\delta w(t) + \delta w(t-R_0)\right) & (3) \\ -\frac{1}{R_0^2 C}\left(\delta q(t) - \delta q(t-R_0)\right) - \frac{R_0 C^2}{2N^2}\delta p(t-R_0) \\ \delta\dot{q}(t) = \frac{N}{R_0}\delta w(t) - \frac{1}{R_0}\delta q(t) \end{cases}$$

From (3) and (4) we derive the transfer function from $\delta p$to δq:

$$\frac{\delta q(s)}{\delta p(s)} = -\frac{R^3C^3}{2N^2}\frac{A(s)e^{-Rs}}{1 + A(s)Rse^{-Rs}} \quad (5)$$

Where $A(s) = \frac{1}{\frac{R^3C}{N}s^2 + \left(R + \frac{R^2C}{N}\right)s + 2}$ (6)

Considering a negative feedback control system with the AQM being the controller, the system to be controlled is given by[9]:

DOI: 10.9781/ijimai.2016.4110

$$G(s) = \frac{R^3C^3}{2N^2}\frac{A(s)e^{-Rs}}{1 + A(s)Rse^{-Rs}} \quad (7)$$

$$= \frac{R^3C^3}{2N^2}\frac{e^{-Rs}}{\frac{R^3C}{N}s^2 + \left(R + \frac{R^2C}{N}\right)s + 2 + Rse^{-Rs}} \quad (8)$$

III. AQM for Congestion Control

A. Approximation of the control model of a first order delay system

In the process control industry, most systems can be roughly modeled as a first order system with delay:

$$\hat{G}(s) = \frac{Ke^{-Ls}}{Ts+1} \quad (9)$$

Where K, T and L represent respectively the state gain, the constant time and the time delay of the plant. These three parameters are supposed to be positive.

We already carry out this approximation calculation in [9], we found:

$$\begin{cases} K = \frac{R^3C^3}{4N^2} \\ T = R\left(\frac{R^4C^2}{4N^2} + \frac{2RC}{N} - \frac{2R^2C}{N} + 2\right)^{\frac{1}{2}} \\ L = \frac{R^2C}{2N} + 2R - T \end{cases} \quad (10)$$

B. Preliminary results

Large number problems in process control engineering are related to the presence of delays. These delays are involved in the dynamic models whose equations characteristics are of the following form [12, 13]:

$$\delta(s) = d(s) + e^{-L_1 s}n_1(s) + e^{-L_2 s}n_2(s) + \cdots + e^{-L_m s}n_m(s) \quad (11)$$

Where: d(s) and n(s) are polynomials with real coefficients and Li represent time delays. These characteristic equations are recognized as quasi-polynomials. We consider the case which meets the following assumptions:

$$A1: 0 < L_1 < L_2 < \cdots < L_m \quad (12)$$

$$A2: \deg(d(s)) = n \text{ and } \deg(n_i(s)) < n \quad (13)$$

$$i = 1,2,3,\ldots,m$$

One can consider the quasi-polynomials δ*(s) described by:

$$\delta^*(s) = e^{L_m s}\delta(s)$$

$$\delta^*(s) = e^{L_m s}d(s) + e^{(L_m - L_1)s}n_1(s) + e^{(L_m - L_2)s}n_2(s) + \cdots + n_m(s) \quad (14)$$

The zeros of δ(s) are identical to those of δ*(s) since $e^{L_m s}$ does not have any finite zeros in the complex plan. However, the quasi-polynomial δ*(s) has a principal term since the coefficient of the term containing the highest powers of s and e^s is nonzero. If δ*(s) does not have a principal term, then it has an infinity roots with positive real parts [15].

The stability of the system with characteristic equation (13) is equivalent to the condition that all the zeros of δ*(s) must be in the open left half of the complex plan. We said that δ*(s) is Hurwitz or is stable. The following theorem gives a necessary and sufficient condition for the stability of δ*(s).

Theorem 1[11][12][13][14][15]

Let $\delta^*(s)$ be given by (3) and write:

$$\delta^*(j\omega) = \delta_r(\omega) + j\delta_i \quad (15)$$

Where $\delta_r(\omega)$ and $\delta_i(\omega)$ represent respectively the real and imaginary parts of $\delta^*(j\omega)$. Under conditions (A1) and (A2), $\delta^*(s)$ is stable if and only if:

1: $\delta_r(\omega)$ and $\delta_r(\omega)$ have only simple, real roots and these interlace,

2: $\delta_i'(\omega_0)\delta_r(\omega_0) - \delta_i(\omega_0)\delta_r'(\omega_0) > 0$ for some ω_0 in $[-\infty, +\infty]$ Where $\delta_r'(\omega)$ and $\delta_i'(\omega)$ denote the first derivative with respect to ω of $\delta_r(\omega)$ and $\delta_i(\omega)$, respectively.

A crucial stage in the application of the precedent theorem is to verify that and have only real roots. Such a property can be checked while using the following theorem.

Theorem 2 [8][12][13][14][15]

Let M and N designate the highest powers of s and e^s which appear in $\delta^*(s)$. Let η be an appropriate constant such that the coefficient of terms of highest degree in $\delta_r(\omega)$ and $\delta_i(\omega)$ do not vanish at $\omega = \eta$. Then a necessary and sufficient condition that $\delta_r(\omega)$ and $\delta_i(\omega)$ have only real roots is that in each of the intervals $-2l\pi + \eta < \omega < 2l\pi + \eta, l = l_0, l_0 + 1, l_0 + 2 \ldots$

$\delta_r(\omega)$ or $\delta_i(\omega)$ have exactly $4lN + M$ real roots for a sufficiently large l_0.

C. Stabilization using PI Controller

We consider now the feedback control system shownin Fig. 2, where r is the command signal, y is the output of the plant, $\hat{G}(s)$ is the plant to be controlled, and C(s) is the controller.

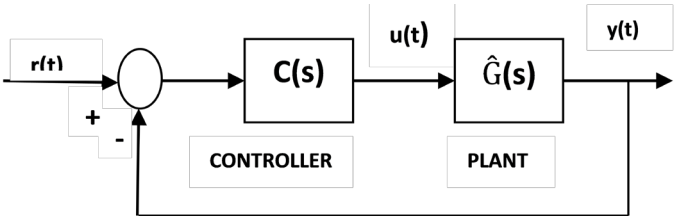


Fig. 2. Feedback control system.

We will focus on the case when the controller is the PI controller described by the following transferfunction:

$$C(s) = K_p + \frac{K_i}{s} \quad (16)$$

The objective is to determine the set of controller parameters (Kp, Ki) for which the closed-loop system is stable.

Theorem 3 [11][12][13][14][15]

The range of K_p value, for which a solution to PI stabilization problem for a given stable open-loop plant exists, is given by:

$$-\frac{1}{K} < K_p < \frac{T}{KL}\sqrt{\alpha_1^2 + \frac{L^2}{T^2}} \quad (17)$$

Where α_1 *the solution of the equation* $tan(\alpha) = -\frac{T}{L}\alpha$ *in the interval* $[\frac{\pi}{2}, \pi]$

The proof of this theorem is detailed in[9].

To determine the stabilizing PI parameters for a first order delayed system, analgorithm has been proposed deduced from the theorems above [9].

IV. Genetic Algorithm Optimization

GeneticAlgorithms (GAs) are adaptive heuristic search algorithm based on the evolutionary ideas of natural selection and genetics. As such they represent an intelligent exploitation of a random search used to solve optimization problems. The main idea in this technique is derived from natural evolution so there are some biological operators such as Crossover, Mutation and Selection [16]. The evolution from one generation to the following is based on the use of the three operators; selection, crossover and mutation which are applied to all the elements of the population. There are three part of this process which are randomly generated, namely the Initial Population, the Crossover and the Mutation. In the first step of the process, the Initial Population in the Genetic Algorithm generates some random solutions. In the second step, the random value of Crossover aid to make new offspring and in the third step, with random value of Mutation let to change a few of gens [17][22][23].

The principle of regulator parameters optimization by the Genetic Algorithms is shown by Fig. 3. It is about the search of parameters k_p and k_i in the area of stability.

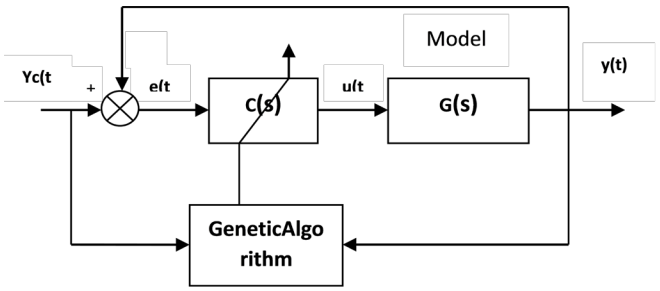


Fig. 3. The optimization principle by genetic algorithm.

A. The objective functions fitness values

The most crucial step in applying GA is to choose the objective functions that are used to evaluate fitness of each chromosome. Some works, like [18] and [19] use performance indices as the objective functions.The both authors use Integral of Time multiplied by Absolute Error (ITAE), Integral of Absolute Magnitude of the Error (IAE), and Integral of the Squared Error (ISE). Here we use all three performance indices stated above and Integral of Time multiplied by the Squared Error (ITSE) to minimize the error signal e(s) and compare them to find the most suitable one. The performance indices are defined as follow:

$$ISE = \sum_0^{t_{max}} e(t)^2, \quad IAE = \sum_0^{t_{max}} |e(t)|, \tag{18}$$

$$ITAE = \sum_0^{t_{max}} t|e(t)|, \qquad ITSE = \sum_0^{t_{max}} te(t)^2$$

Where e(t) is the error signal in time domain.

If we want to minimize the tuning energy, the ITAE criteria and the IAE are considered. In the case where we privilege the rise time, we take the ITSE criterion. In order to guarantee the tuning energetic cost, we choose the ISE criterion.

The calculation steps of the control law are summarized by the following algorithm [20]:

1. Introduction of the following parameters:
   - max_{pop} individuals number by population.
   - initial population.
   - gen_{max} generation number.
2. Initialization of the generation counter ($gen = 1$).
3. Initialization of the individual counter ($j = 1$).
4. For $t = 1s$ to $t = t_{max}$ efficiency evaluation of j^{th} population individual.
5. Individual counter incrementing $(j = j + 1)$.
   - If $j < max_{pop}$, going back to step 4.
   - If not: application of the genetic operators (selection, crossover, mutation) for founding a newpopulation.
6. Generation counter incrementing $gen = gen + 1$,
7. If $gen < gen_{max}$; going back to step 3.
8. Taking $K_p opt$ and $K_i opt$ which correspond to the best individual in the last population (individual making the best fitness).

On the following, the genetic algorithm is characterized by generation number equal to 100, Pc = 0.8, Pm =0.08 and individual number by population equivalent to 50.

We now present an example that illustrates the application of the results presented in this section.

V. Simulation Results

The performance of the closed-loop system with the new PI controller is tested by simulation. The number of TCP flows is taken to be N = 60, the link capacity C = 3750 packets/sec, round trip time R = 400s, the desired queue size q_0= 150 packets, and the buffer size q_{max} = 200 packets [9].

We obtain the stability region in Kp, Ki plane, presented in Fig. 4:

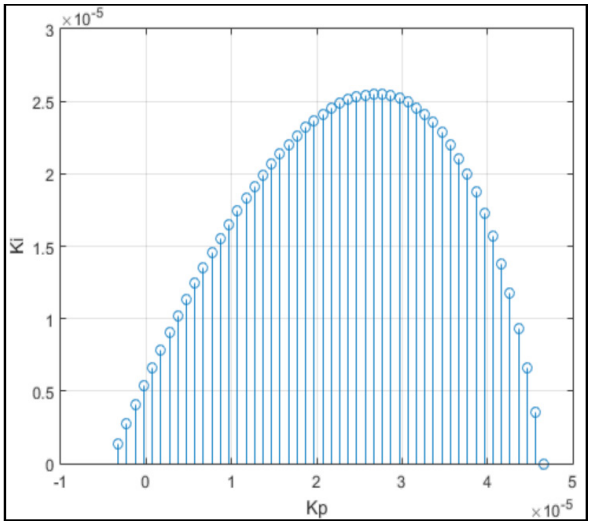


Fig. 4. Stability region of the system

The closed loop step responses of PI conroller optimum parameters supplied by Genetic Algorithm are illustrated in the following Fig. 5:

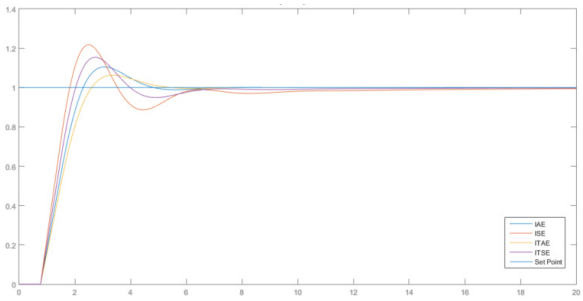


Fig. 5. Step response for the different optimal PI controllers.

The performances are rise time, settling time and overshoot. The aim of the controllers is to minimize the transient state response performances. Numerical values of these performances of the closed loops are given in Table 1 for comparison:

TABLE 1:
NUMERICAL VALUES OF STANDARD PERFORMANCE MEASURES.

Criterion	ISE	IAE	ITAE	ITSE
$\overline{K_p opt(10^{-4})}$	0.22	0.1642	0.1489	0.0296
$K_i opt(10^{-4})$	0.0293	0.0325	0.0296	0.0321
Rise Time (s)	0.8183	1.1453	1.3229	0.9630
Settling Time (s)	9.7619	4.4026	4.6105	6.1405
Overshot (%)	21.7792	10.5320	6.3206	15.4224

A. Standard performance measures:

It can be seen from Table 1 that GA optimized by the four criterions provides a very small rise time which is beneficial for the congestion control of the AQM routers, chiefly the one obtained by the ISE criterion. The best value of the settling time is provided by IAE and ITAE and that of the overshoot is given by GA optimized by ITAE.

VI. Conclusion

Congestion is one of the most important problems faced in communications networks. In this paper we proposed to solve this problem by applying a new approach to compute the stability region for first order delay system controlled by PI controllers. This result is based on an extension of the Hermite-Biehler Theorem to quasipolynomials. Then, we were interested in search of optimal PI for performance criteria, inside the stability region by resorting to the Genetic Algorithm techniques as a method of optimization.

Finally, it is pointed out that the effectiveness of the proposed approach has been verified via simulation.

References

[1] V. Jacobson, "Congestion avoidance and control," in Proceedings of ACM SIGCOMM, Stanford, CA, August 1988.

[2] D. Chiu and R. Jain, "Analysis of the increase/decrease algorithms for congestion avoidance in computer networks," Journal of Computer Networks and ISDN, vol. 17, no. 1, pp. 1–14, June 1989.

[3] Ohsaki H., Sugiyama K. and Imase M., Congestion Propagation among Routers with TCP Flows, International Journal of Computer Networks & Communications, vol. 1, no. 2,pp. 112-127. 2009

[4] V. Misra, W. Gong, and D. Towsley. Stochastic differential equation modeling and analysis of TCP windowsize behavior. Technical Report ECE-TR-CCS-99-10-01, University of Massachusetts, October 1999.

[5] H. I. Alshimaa, A. EL-Sayed & Z. Elsaghir, I. Z. Morsi, "Enhanced Random Early Detection (ENRED)" International Journal of Computer Applications (0975 – 8887) Vol. 92 – No.9, April 2014.

[6] V. Misra, W. Gong, and D Towsley. Fluid-based analysis of a network of AQM routers supporting TCP flows with an application to RED. In ACM SIGCOMM, pages 151–160, August 2000.

[7] H. S. Low, F. Paganini, and J.C. Doyle. Internet congestion control.IEEE Control Systems Magazine, 22:28–43, Feb 2002.

[8] G. J. Silva, A. Datta and S. P. Bhattacharyya, "PI stabilization of first-orde systems with time-delay", Automatica, 37(12) :2025-2031, 2001.

[9] S. Chebli, A. Elakkary, N. Sefiani, N. Elalami, "New PI stabilization of first-order congestion control of active queue management routers", in Proceeding of Electrical and Information Technologies (ICEIT) Conference, pp.12 – 217, Marrakech, Morocco, March, 2015.

[10] Goldberg, D. E. Genetic algorithm in search, optimization, and machine learning, reading. MA: Addison Wesley, 1989.

[11] G. J. Silva, A. Datta and S. P. Bhattacharyya, "PID controllers for time delay systems", Birkhäuser, MA, Boston, 2001.

[12] G. J. Silva, A. Datta and S. P. Bhattacharyya, Stabilization of Time Delay Systems, Proceedings of the American Control Conference, pp.963-970, 2000.

[13] G. J. Silva, A. Datta and S. P. Bhattacharyya, Stabilization of First-order Systems with Time Delay using the PID controller, Proceedings of the American Control Conference, pp. 4650-4655, 2001.

[14] G. J. Silva, A. Datta and S. P. Bhattacharyya, New synthesis of PID controller, IEEE transactions on automatic control, vol.47,No.2, 2002.

[15] G.J. Silva, A. Datta, S.P. Bhattacharyya, PID controllers for time delay systems, Springer, London, 2005.

[16] M. Martínez, S.García-Nieto, J.Sanchis, X.BlascoFerragud: Genetic algorithms optimization for normalized normal constraint method under Pareto construction. Advances in Engineering Software (AES) 40(4):260-267 (2009).

[17] G. Ye, Z. Tao,"Improved genetic algorithms based on chaotic mutation operation and its application". Multimedia Technology (ICMT), 2010 International Conference on 1-3 (2010).

[18] Ian Griffin, On-line PID Controller Tuning using Genetic Algorithms, Dublin City University, 2003.

[19] T. O'Mahony& CJ. Downing (Cork Institute of Technology, Ireland), KlaudiuszFatla (Wroclaw University of Technology, Poland), Genetic Algorithms for PID Parameter Optimization, Minimizing Error Criteria.

[20] R. Farkh, K. Laabidi, M. Ksouri, "PI control of second order delay system with tuning parameter optimization", International Journal of Electrical and Electronics Engineering, vol.3, pp. 1-7, 2009.

[21] Zaldaña, H. R., and E. Castañeda, "The Use of Genetic Algorithms in UV Disinfection of Drinking Water", International Journal of Interactive Multimedia and Artificial Intelligence, 3(3), pp. 43-48, 2015

[22] Arranz, Á., and M. Alvar, "GPGPU Implementation of a Genetic Algorithm for Stereo Refinement", International Journal of Interactive Multimedia and Artificial Intelligence, 2(3), pp. 69-76, 2015

S. Chebli was born in Ouezzane, Morocco, in 1990. She is a PhD Student at Graduate School of Technology Salé/ Mohammadia Engineers School Rabat, in Automatic Engineering . She obtained the Engineer degree in Electromechanical Engineering from the National School of Mineral Industry of Rabat in July, 2013. Her research interests are in control systems, optimization methods, and network theories.

Ahmed El Akkary was born in Settat in Morocco, assistant professor at the Graduate School of Technology in Sale Morocco, winner of ENSET Rabat in 1995, in 2003 he obtained his DESA specialty automatic and manufacturing, then, in 2009, he obtained his doctorat from the Mohammadia School of Engineers of Rabat, specialty Automatic, researcher at the laboratory LASTIMI of EST Sale Morocco. His research interests are in control systems.

Mr Noureddine El Alami received his Engineering in Automatic and Industrial Computing from Mohammadia School of Engineering, Morocco, in 1980. He holds a degree in Applied Mathematics from the Faculty of Sciences, Rabat, Morocco. In 1986, he received his Doctor of Philosohy in Automatic from the university of Perpignann, France. He was a professor at Mohammadia School of Engineering, Morocco, in Applied Mathematics, Optimal Control, Non – linear Systems and Signal Theory. His rechearch interests have revolved around the control of non-linear systems and their applications in engineering. He is the author or co-author of some papers in international conferences and journals.

Solving the Weighted Constraint Satisfaction Problems Via the Neural Network Approach

Khalid Haddouch[1], Karim Elmoutaoukil[1], Mohamed Ettaouil[2]

[1]National School of Applied Sciences of Al Hoceima, University Mohammed First, Box 03, Al Hoceima, MOROCCO
[2]Modeling and Scientific Computing Laboratory, Faculty of Science and Technology of Fez, University Sidi Mohammed ben Abdellah Box 2202, Fez, MOROCCO

***Abstract* — A wide variety of real world optimization problems can be modelled as Weighted Constraint Satisfaction Problems (WCSPs). In this paper, we model this problem in terms of in original 0-1 quadratic programming subject to leaner constraints. View it performance, we use the continuous Hopfield network to solve the obtained model basing on original energy function. To validate our model, we solve several instance of benchmarking WCSP. In this regard, our approach recognizes the optimal solution of the said instances.**

***Keywords* — Weighted Constraint Satisfaction Problems, Quadratic 0-1 Programming, Continuous Hopfield Network, Energy Function.**

I. Introduction

Constraint programming is a successful technology for solving combinatorial problems modelled as constraint satisfaction problems (CSPs). In the last few years, the CSP framework has been extended to the soft constraints permits to express preferences among solutions [2][3]. Soft constraint frameworks associate costs to tuples and the goal is to find a complete assignment with minimum aggregated cost. Costs from different constraints are aggregated with a domain dependent operator. This case is known as the Weighted Constraint satisfaction problems (WCSPs)[15][16][21]. This latter is an extension of the CSP problem. Each constraint of the problem has an associated weight. If a solution violates a constraint, then the weight associated to the corresponding solution is incurred. Every solution has a cost as the sum of all weight incurred by the solution. The objective of this problem is to find the solution with the minimum cost.

In this paper, we focus on the weighted constraint satisfaction problem (WCSP), a well-known non-idempotent soft-constraint framework with several applications in several domains such as resource allocation, scheduling, bioinformatics, CP networks and probabilistic reasoning [16]-[17]. In the literature, a number of different approaches have been developed to solve this problem [7]-[9]-[10]-[14]-[15]. In this work we propose a new model of WCSP problem consists in minimizing the quadratic objective function subject to linear constraints (QP). To solve the QP problem, many different methods are tried and tested such as interior point, semi definite relaxations and lagrangian relaxations [20]. In our case, we introduce the continuous Hopfield network for solving The QP problem in order to validate our model.

Hopfield neural network was introduced by Hopfield and Tank [10] [11][13][23]. It was first applied to solve combinatorial optimization problems. It has been extensively studied, developed and has found many applications in many areas, such as pattern recognition, model identification, and optimization. CHN also demonstrated capability of finding solutions to difficult optimization problems [5].

In this paper, our main objective is to propose a new approach for solving the weighted constraint satisfaction problems using continuous Hopfield network. This paper is organized as follows: In section 2, we define weighted constraint satisfaction problems. In the last section, we propose and describe the new model of the binary WCSP problem. This problem is formulated as a quadratic assignment problem with linear constraints. A new theorem, which consists to define the relation between the WCSP problem and the quadratic programming, is demonstrated. In section 4, an introduction of CHN is presented, the generalized energy function associated to WCSP problem is defined and a direct parameter setting procedure is computed. Finally, the implementation details of the proposed approach and experimental results are presented in the last section.

II. New Model of Weighted Constraints Satisfaction Problems

A constraint satisfaction problem refers to the problem of finding values to a set of variables, subject to constraints on the acceptable combination of values. Solving this problem requires finding values for problem variables from each domain, which satisfies all members of the set of constraints. In some cases a privilege tuples relative to others. This case is known as a weighted constraint satisfaction problem (WCSP). The solution of this later is similar then the classical CSP with the cost according to this solution is optimal.

A. Weighted constraint satisfaction problem

Each constraint of the problem has an associated weight (or error). If a solution violates a constraint, then the weight associated to this solution is incurred. Every solution has a cost which consists of the sum of all weight incurred by the solution. For a WCSP problem, a relevant question, both theoretically and practically, is to determine an assignment of values to variables with the minimum cost. This problem is calcified as NP-hard problems [8].

In general, a WCSP problem forms a class of models representing problems that have as common properties, a set of variables and a set of constraints [2]–[3]. The variables should be instantiated from a discrete domain. The study of WCSP problem has become focused on binary forms. More precisely, a weighted constraint satisfaction problem is defined by a quadruplet sets (*Y, D, C, S(k)*) where:

- $Y = \{y_1, ..., y_n\}$ is the set of n variables,
- $D = \{D(y_1), ..., D(y_n)\}$ where each $D(y_i)$ is the set of d_i possible values for y_i,
- $C = \{C_1, ..., C_m\}$ is the set of m constraints which restricts the values that the variables can simultaneously take.
- *S(k)* is the valuation structure, where $K \in \mathbb{R}^+$ denotes the maximum cost.
- The valuation structure is defined as *S(k)* = ({*0, 1, ..., k*},⊕,>),

DOI: 10.9781/ijimai.2016.4111

where:

- $\{0, 1, ..., k\}$ is the set of costs, which are natural numbers bounded by k.
- $\oplus$ is the sum of costs $\forall a, b \in \{0, 1, ..., k\}$, $a \oplus b = min\{k, a + b\}$
- $>$ is the standard order among naturals.

For each constraint C_{ij} and each tuple t_p represented by two values (v_r, v_s) from the domains associated with the variables involved in C_{ij}, a cost $c_{rs} \in \{0,1, \ldots, k\}$ is assigned to t_p. When a constraint C_{ij} assigns the cost k to a tuple t_k, it means that C_i forbids t_p. Otherwise, t_p is permitted by C_i with the corresponding cost. The cost of an instantiation of variables is the sum (using operator $\oplus$) over all constraints involving variables instantiated.

B. Modelling the WCSP problems

The solution of weighted constraint satisfaction problem is based on assigning each variable, a value from its domain with the minimum cost. In this context, we propose a new model of the WCSP problem as 0-1 quadratic programming, which consists in minimizing a quadratic function subject to linear constraints. During this phase of modelling, we use different mathematical notations motioned bellow.

In this case, we want to propose a formulation of the WCSP problem. This formulation is based on the selection variables. Then, in the first time, for each variable y_i of the WCSP problem, we introduce d_i binary variables x_{ir} such that:

$$x_{ir} = \begin{cases} 1 & if\ y_i = v_r \\ 0 & Otherwise \end{cases} \qquad v_r \in D(y_i) \tag{1}$$

The vector of variables is represented by:

$$x \equiv \left(x_{11} \quad \cdots \quad x_{1d_1} \quad \cdots \quad \cdots \quad x_{n1} \quad \cdots \quad x_{nd_n} \right)^T \text{ with}$$

$$N = \sum_{i=1}^{n} d_i \text{ and } d_i = |\, D(y_i)\,|$$

Based on this binary variable we have for each couple (v_r, v_s) :

$$x_{ir}x_{js} = \begin{cases} 0 & if\ (v_r, v_s) \notin R_{ij} \\ 1 & if\ (v_r, v_s) \in R_{ij} \end{cases} \tag{2}$$

The mean property of the solution to the WCSP is that each variable y_i must take an unique value v_r from its domain $D(y_i)$. Then the linear constraints of WCSP problem are defined bellow:

$$\sum_{r=1}^{d_i} x_{ir} = 1 \tag{3}$$

In the second time, it is necessarily to treat two cases. The first one is the case of the unary constraints, but the second one is case of the binary constraints.

Each unary constraint C_i is defined by its relation R_i, specifying the privileged values using the notion of cost. Recall that, for each constraint C_i and each value v_r, a cost $c_{ir}\{0,1, \ldots, k\}$ is assigned to v_r. Then, for each value v_r we generate a constant:

$$q_{ir} = \begin{cases} k & if\ \ v_r \notin R_i \\ \mathrm{c}_{ir} & if\ \ v_r \in R_i \end{cases} \tag{4}$$

Based on these propositions (equations (1) and (4)), each unary constraint C_i can be characterized by the following expression:

$$\sum_{r=1}^{d_i} q_{ir} x_{ir} \tag{5}$$

Finally, we can generalize this expression for all constraints of the WCSP problem by the following global expression:

$$\sum_{i=1}^{n} \sum_{r=1}^{d_i} q_{ir} x_{ir} \tag{6}$$

In the same, each binary constraint C_{ij} between variables y_i and y_j is defined by its relation R_{ij} specifying the compatible values between y_i and y_j with certain cost. Recall that, for each constraint C_{ij} and each tuple tl represented by two values (vr,vs) from the domains associated with the variables involved in C_{ij}, a cost crs∈{0,1, . . . , k} is assigned to tl. Then, for each couple (v_r, v_s) we generate a constant:

$$q_{irjs} = \begin{cases} k & if\ (v_r, v_s) \notin R_{ij} \\ \mathrm{c}_{rs} & if\ (v_r, v_s) \in R_{ij} \end{cases} \tag{7}$$

Each constraint C_{ij} can be characterized by the following expression:

$$\sum_{r=1}^{d_i} \sum_{s=1}^{d_j} q_{irjs} x_{ir} x_{js} \tag{8}$$

Finally, we can generalize this expression for all constraints of the WCSP problem by the following global expression:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{r=1}^{d_i} \sum_{s=1}^{d_j} q_{irjs} x_{ir} x_{js} \tag{9}$$

Based on these two expressions (equations (6) and (9)), the objective function $f(x)$ can be formulating in the following form:

$$f(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{r=1}^{d_i} \sum_{s=1}^{d_j} q_{irjs} x_{ir} x_{js} + \sum_{i=1}^{n} \sum_{r=1}^{d_i} q_{ir} x_{ir} \tag{10}$$

Then, the objective function can be written in the following matrix form:

$$f(x) = \frac{1}{2} x^T Q x + q^T x \tag{11}$$

Finally, the binary WCSP problem is modelled as a 0-1 quadratic programming with a quadratic function subject to linear constraints:

$$(QP) \quad \begin{cases} Min & f(x) = \frac{1}{2} x^T Q x + q^T x \\ Subject\ to & \\ & Ax = b \\ & x \in \{0,1\}^N \end{cases}$$

Where Q is an $N \times N$ symmetric matrix, A is an $N \times n$ matrix, q is an N vector and b is an N vector.

The following theorem determines the relation between a binary WCSP problem and optimization model QP. In order to validate this new model of WCSP, we use the continuous Hopfield network for solving the resulting model.

III. The Proposed Model Solved by Continuous Hopfield Networks

In this section, we present a general approach to solve this problem using the continuous Hopfield networks. The neural network approaches are the efficient approaches for solving different problems in different areas [2]–[4]–[9]–[1]–[19]. Moreover, Hopfield and Tank [10]–[11] presented the energy function approach in order to solve several optimization problems [2]-[6]. Their results encouraged a number of researchers to apply this network to different problems. The continuous Hopfield neural network is a generalization of the discrete case. The common output functions used in the networks are hyperbolic tangent functions. Afterwards, many researchers implemented CHN to solve the optimization problem, especially in mathematical programming problems.

The CHN is a fully connected neural network, which means that every neuron is connected to all other neurons. The connection weights between the neuron *i* and neuron *j* is represented by W_{ij} and each neuron *i* has an offset bias i_i^b [12].

For solving any combinatorial problems, it is necessarily to map it's in the form of the energy function associated to the continuous Hopfield network. The expression of this energy function is the following:

$$E(x) = -\frac{1}{2} x^T W x - (i^b)^T x \tag{12}$$

In this work, our main objective is to solve the *QP* problem, i. e., solving weighted constraint satisfaction problem using the continuous Hopfield network. According to the proposed model, which consists modeling the *WCSP* problem into a quadratic programming *QP*, this step of representation becomes easy and more general. Then, the continuous Hopfield network can be used to solve the weighted constraint satisfaction problem[1]-[18].

In order to represent this latter, the energy function must be defined by two expressions $E^O(x)$ and $E^C(x)$. The first one is directly proportional to the objective function of the *QP* problem and the second one is a quadratic function that penalizes the violated constraints of the *QP* problem. Therefore the energy function associated to the *CHN* is:

$$E(x) = E^O(x) + E^C(x) \qquad \forall\ x \in H \tag{13}$$

Where *H* is set of the Hamming hypercube:

$$H \equiv \{x \in [0,1]^N\}$$

The algebraic form of the generalized energy function is:

$$E(x) = \frac{\alpha}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{r=1}^{d_i}\sum_{s=1}^{d_j} q_{irjs} x_{ir} x_{js} + \alpha\sum_{i=1}^{n}\sum_{r=1}^{d_i} q_{ir} x_{ir} + \frac{1}{2}\phi\sum_{i=1}^{n}\sum_{r=1}^{d_i}\sum_{s=1}^{d_i} x_{ir} x_{is} + \beta\sum_{i=1}^{n}\sum_{r=1}^{d_i} x_{ir} + \gamma\sum_{i=1}^{n}\sum_{r=1}^{d_i} x_{ir}(1 - x_{ir}) \tag{14}$$

The weights and thresholds of the connections between N neurons are:

$$\begin{cases} W_{irjs} = -\alpha(1-\delta_{ij}) q_{irjs} - \delta_{ij}\phi + 2\delta_{ij}\delta_{rs}\gamma \\ i_{ir}^b = -\alpha q_{ir} - \beta - \gamma \end{cases} \tag{15}$$

Where δ_{ij} is the Kroenecker delta.

In this way, the quadratic programming has been presented as an energy function of continuous Hopfield network. To solve an instance of the *QP* problem, the parameter setting procedure is used. This procedure assigns the particular values for all parameters of the network, so that any equilibrium points are associated with a valid affectation of all variables when all the constraints in *QP* problem are satisfied. We observe that, the weights and thresholds of the continuous Hopfield network depend on the parameters α, ϕ, β and γ. To solve the *QP* problem via the *CHN*, an appropriate setting of these parameters is needed[1]-[19]. The parameter-setting procedure is based on the partial derivatives of the generalized energy function:

$$\frac{\partial E(x)}{\partial x_{ir}} = \alpha\sum_{j=1}^{n}\sum_{s=1}^{d_j} q_{irjs} x_{js} + \phi\sum_{s=1}^{d_i} x_{is} + \beta + \gamma(1 - 2x_{ir}) \tag{16}$$

Based on this hyperplane method and the associated half-spaces, the complementary corners set of the feasible solutions for the *QP* problem is partitioned and a set of analytical equations of the *CHN* parameter is proposed[19]. When we apply this method, the obtained analytical equations system is:

$$\begin{cases} \alpha > 0 \quad , \phi \geq 0 & \\ -\phi + 2\gamma \geq 0 & (17.a) \\ \alpha d_{\min} + 2\phi + \beta - \gamma = \varepsilon & (17.b) \\ \alpha d_{\max} + \beta + \gamma = -\varepsilon & (17.c) \end{cases}$$

Where $d_{\min} = Q_{\min} + q_{\min}$, $d_{\max} = Q_{\max} + q_{\max}$

With $Q_{\min} = Min\{q_{irjs}\}$ and $q_{\min} = Min\{q_{ir}\}$

$Q_{\max} = Max\{q_{irjs}\}$ and $q_{\max} = Max\{q_{ir}\}$

The inequation (17.a) guaranteed the satisfaction of the integrity constraints ($x_{ir} \in \{0,1\}$), but the equations (17.b) and (17.c) guaranteed the satisfaction of the linear constraints.

Finally, the weights and thresholds of *CHN* (system 15) can be calculated using these parameters setting. Finally, we obtain an equilibrium point for the *CHN* using the algorithm depicted in[17], so compute the solution of constraint satisfaction problem.

IV. Computational Experiments

In order to validate the proposed approach, some experiments are effectuated to solve some typical problems of WCSP problem[16]. These experiments are effectuated in personal computer with a 2.79 GHz processor and 512 MB RAM. This approach is implemented by java language. The performance has been measured in terms the minimum obtained cost.

Recall that *n* is the number of variables. Based on a series of experiments, α and ε are determined by the following values:

$$\alpha = \frac{1}{n}\ ,\ \varepsilon = 10^{-4}$$

In comparison with wbo 1.72WCSP solvers, the optimum cost obtained by CHN is very interesting, it equal to optimum cost obtained by this solver[22]. Morever, these results are obtained in the minimum time(See TABLE I). Noted that, this exeperement stady is integrated just for validate our proposed model. Generally, our model is very successful, it happens to reperent fidelly the weithted constraint satisfaction problems. Finally, we can concluded that the best results are obtained by this approach.

TABLE I: COMPUTATIONAL RESULTS OF THE TYPICAL WCSP INSTANCES

INSTANCES NAME	VARIABLE NUMBERS	CONSTRAINT NUMBERS	MAXIMAL COST	SOLVER WBO 1.72 OPTIMUM COST	CHN	
					OPTIMUM COST	EXECUTION TIME (S)
4wqueens	4	10	5	0	0	0.002
spot5-8	8	15	13	2	2	0.003
langford-2-4	8	32	1	0	0	0.007
8queens	7	28	1	0	0	0.008
8wqueens	8	36	9	2	2	0.013
zebre-ext	23	62	1	0	3	0.014
geom30a-3_wcsp	30	81	82	11	--	--
geom30a-5_wcsp	30	81	82	1	1	0.008
geom30a-6_wcsp	30	81	82	0	0	0.008
geom40-2_wcsp	40	78	79	22	--	--
geom40-3_wcsp	40	78	79	7	7	0.003
bwt3cc_wcsp	45	685	1073	--	1073	0.016
mprime03bc_wcsp	49	625	4941	--	2843	0.062
myciel5g-3_wcsp	47	236	237	--	--	0.010
myciel5g-4_wcsp	47	236	237	--	--	0.015
spot5-29_wcsp	82	462	20092	--	5038	0.072

V. Conclusions

In this paper, we have proposed a new approach for solving binary weighted constraint satisfaction problems. The interesting steps of this approach are: proposing the new model of weighted constraint satisfaction problem as a 0-1 quadratic program subject to linear constraints and using the continuous Hopfield network to solve this problem. The most interesting propriety of this approach is used to give the solution of the binary WCSP. The experimental results show that our method can find a good optimal solution in short time. The future directions of this research are reducing the architecture of Hopfield neural network and applying the proposed approach to give a good solution of real world problems [4]-[5].

References

[1] Ettaouil, M.; Haddouch, K.; Hami, Y.; Chakir, L., *Neural networks approach for solving the Maximal Constraint SatisfactionProblems*, Intelligent Systems: Theories and Applications (SITA), 2013 8th International Conference, 2013, pp: 1 – 6.

[2] M. Ettaouill, and C. Loqman, *A New Optimization Model for Solving the Constraint Satisfaction Problem*, Journal of Advanced Research in Computer Science, 1(1), 2009, 13--31.

[3] M. Ettaouil, C. Loqman, K. Haddouch. *Job Shop Scheduling Problem solved by the continuous Hopfield networks*, Journal of Advanced Research in Computer Science (JARCS), Volume 2, Issue 1, pp. 31 - 47, 2010.

[4] El Moutaouakil K., Haddouch K. and Ettaouil M., *A reduced continuous hopfield networks For the traveling salesman problem*, Theory and Application of Logistics and Transport TALT'15) », 25-26 May, Tetuan.

[5] M. Ettaouil, K. Elmoutaouakil, Y.Ghanou. *The continuous Hopfield networks (CHN) for the placement of the electronic circuits problem*, WSEAS Transactions on Computer, 8 (12), December 2009.

[6] D. J. Evansi, and M. N. Sulaiman, *Solving optimization problems using neucomp-a neuralnetwork compiler*, International Journal of Computer Mathematics, 62(1), 1996, pp. 1--21.

[7] Hoong C. L., A *new approach for Weighted Constraint Satisfaction: Theoretical and computational results*, Principles and Practice of Constraint Programming — CP96, Lecture Notes in Computer Science Volume 1118, 1996, pp 323-337.

[8] M.R Garey, and D.S Johnson, *A Guide to the Theory of NP-Completeness*, Computers and Intractability, New York: W.H. Freeman and Company, 1979.

[9] Huguet M.-J., Lopez P., Karoui W. *Weight-based Heuristics for Constraint Satisfaction and Combinatorial Optimization Problems* June 2012, Volume 11, Issue 2, pp 193-215.

[10] Levasseur, N.; Boizumault, P.; Loudni, S., *A Value Ordering Heuristic for Weighted CSP*,Tools with Artificial Intelligence, 2007. ICTAI 2007. 19th IEEE International Conference on Year: 2007, Volume: 1 pp: 259 – 262.

[11] J. J. Hopfield, *Neural networks and physical systems with emergent collective computational abilities*, Proceedings of the National Academy of Sciences of the United States of America, 79, 1982, pp. 2554—2558.

[12] J.J. Hopfield, *Neurons with graded response have collective computational properties like those of two-states neurons*, proceedings of the National academy of sciences of the USA 81, pp. 3088-3092, 1984.

[13] J.J. Hopfield, D.W. Tank, *Neural computation of decisions in optimization problems*. Biological Cybernetics 52, , 1985, pp. 1-25.

[14] V. Kumar, *Algorithms for Constraint Satisfaction Problems*, T A Survey, AI Magazine, 13(1), 1992, pp. 32--44.

[15] Lecoutre, C.; Paris, N.; Roussel, O.; Tabary, S. *Solving WCSP by Extraction of Minimal Unsatisfiable Cores,* Tools with Artificial Intelligence (ICTAI), 2013 IEEE 25th International Conference on Year: 2013, pp: 915 – 922.

[16] Lecoutre, C., instances aviable at : http://www.cril.univ-artois.fr/~lecoutre/benchmarks.html#.

[17] Lecoutre, C. *Constraint Networks: Targeting Simplicity for Techniques and Algorithms* ISBN: 978-1-84821-106-3, 320 pages July 2009, Wiley-ISTE.

[18] P.M. Talavàn and J. Yànez, *A continuous Hpfield network equilibrium points algorithm*. Computers and operations research 32, pp. 2179-2196., 2005.

[19] Talavàn and J. Yànez, *The generalized quadratic knapsack problem. A neuronal network approach*, Neural Networks 19, pp. 416-428, 2006.

[20] Thiongan, and al., *An Adapted Step Size Algorithm for a 0-1 BiknapsackLagrangean Dual*, Annals of Operations Research, 139(1), 2005, pp. 353–373.

[21] E. Tsang, *A Foundations of Constraint Satisfaction*, Academic Press, 1993.

[22] Result page for solver wbo *1.72,* available results at: http://www.cril.univartois.fr/PB12/results/solver.php?idev=68&idsolver=2315.

[23] Raveane, W., and M. A. G. Arrieta, *Neural Networks through Shared Maps in Mobile Devices,* International Journal of Interactive Multimedia and Artificial Intelligence, 1(3), pp. 28-35, 2014

K. Haddouch Has a Ph.D. in Artificial Intelligence by Faculty of Science and Technology of Fez with a thesis titled "Constraint programming and neural network approach applied to reels problems". His main research topics are neural networks and constraint programming. He has published an active part in different congress and jornals. He has participated in directed and coordinated conferences, courses and projects of education of the computer sciences. He is a member of Scientific Committee for different international congress in Morocco. Since 2014 he is a Professor at National School of Applied Sciences of Al Hoceima, University Mohammed First, Box 03, Al Hoceima, MOROCCO.

K. Elmoutaoukil: Has a Ph.D. in Artificial Intelligence by Faculty of Science and Technology of Fez with a thesis titled "Contribution to the optimization of architectures in artificial neural networks and nonlinear programming, application to calculate optimal topology problem of SDH telecommunication networks". His main research topics are neural networks, optimization, clustering and machine learning. He has published an big part in different congress and jornals. He has participated in directed and coordinated conferences, courses and projects of education of the computer sciences operational research. He is a member of Scientific Committee for different international congress. He is a Professor at National School of Applied Sciences of Al Hoceima, University Mohammed First, Box 03, Al Hoceima, MOROCCO.

M. Ettaouil: is a Professor at Faculty of Science and Technology of Fez, University Sidi Mohammed ben Abdellah Box 2202, Fez, MOROCCO, is a member of Modeling and Scientific Computing Laboratory. His main research topics are neural networks, optimization, modelling and machine learning. He has published an big part in different congress and papers. He is a responsible of teams "Digital and Computer Engineering, Artificial Neural Networks and Learning (N2I-RNA)" at FST Fez. He is a member of Scientific Committee for different international congress. He has published a big part in different congress and papers. He has participated in directed and coordinated conferences, courses and projects of education of the computer sciences.

Evaluating the Emotional State of a User Using a Webcam

M. Magdin[1], M. Turčáni[1], and Lukáš Hudec[2]

[1]*Department of Computer Science, Faculty of Natural Science, Constantine the Philosopher University in Nitra*

[2]*Java developer in EmbedIT*

***Abstract* — In online learning is more difficult for teachers identify to see how individual students behave. Student's emotions like self-esteem, motivation, commitment, and others that are believed to be determinant in student's performance can not be ignored, as they are known (affective states and also learning styles) to greatly influence student's learning. The ability of the computer to evaluate the emotional state of the user is getting bigger attention. By evaluating the emotional state, there is an attempt to overcome the barrier between man and non-emotional machine. Recognition of a real time emotion in e-learning by using webcams is research area in the last decade. Improving learning through webcams and microphones offers relevant feedback based upon learner's facial expressions and verbalizations. The majority of current software does not work in real time – scans face and progressively evaluates its features. The designed software works by the use neural networks in real time which enable to apply the software into various fields of our lives and thus actively influence its quality. Validation of face emotion recognition software was annotated by using various experts. These expert findings were contrasted with the software results. An overall accuracy of our software based on the requested emotions and the recognized emotions is 78%. Online evaluation of emotions is an appropriate technology for enhancing the quality and efficacy of e-learning by including the learner´s emotional states.**

***Keywords* — Emotion Recognition, Face Recognition, Facial Expressions, Emotion Classification.**

I. Introduction

Today, ICT are fundamental for our society. Their task is to make information accessible from every place in the world, quickly and without effort [47]. ICT are the means, which from the global point of view can contribute to the development of knowledge and skills [37],[22]. The ICT with its basic character enables us to increase the quality of educational processes [26]. Nowadays, we can't even imagine education in our information society without its electronic form [25]. E-learning is multimedia support of a learning process, combined with modern information and communication technologies [24]. During the last decade, several new technologies have been adopted by e-learning specialists for enhancing the effectiveness, efficiency and attractiveness of e-learning. Development of new adaptive techniques, as well as modernization of old-fashioned technologies, causes also fundamental changes in the development of our society. Development of adaptive techniques for the sphere of e-learning systems, which allow for personalization of the student, has been known for long [18]. However, during the last 5 years it has come to a real progress in opportunities for use of these technologies [8], [45], [27]. Wide spectrum of opportunities and open space for experimenting allows us applying our creativity on the topmost level [12], [31]. Such system of learning by means of ICT, for example in cooperation with the opportunities provided by e-learning systems and technologies Web 2.0 has been actively used mainly in the last two decades [20],[33].

However, recent developments of ICT, specifically input devices (such as webcams) for interacting with such environments are still underexploited. Such devices firstly offer opportunities for more natural interactions with the e-learning applications [1]. Secondly, they offer better ways for gathering affective user data, as they do not interfere with the learning like questionnaires often do. This is because of their unobtrusive and continuously nature of data gathering.

Education belongs to areas where extensive data exploration is needed [7]. There are various methods of gathering data by the use of which it is possible to adapt the learning process to the learner. These methods can be divided into direct and indirect. Direct methods are those which are used by the user directly during the learning process, they are known and can be partly influenced (e.g. a questionnaire). Indirect methods represent a away how to individualize the learning style of the learner without participating in entering input data into the adaptive process. The learner does not fill a questionnaire but his activity is evaluated during the learning process and the learning style of the learner is determined according to association rules. As an example of data mining of indirect methods is the use of interactive animations through which an overview can be gained about cognitive and intellectual skills of the learner (module IES – Interactive Element Stat, [30]).

Existing methods for gathering affective user data, like psychological sensors and questionnaires, are either obtrusive or discontinuous. They can hamper learning as well as issues in its suitability for elearning [16], [41].

Previous software primarily dealt with offline emotion recognition that cause post-processing of the learner's data [1]. They have a couple of limitations that mainly restrict their application context and might impede their accuracy. The application context is restricted by the fact that such software can only manage a small set of expressions from frontal view faces without facial hair, glasses provided that there is constant illumination. Furthermore, the software requires postprocessing steps for analysing videos and images and cannot analyse extracted facial information on different time scales [39]. In addition, their accuracy might also be impeded as this software used no databases for authentic emotions. In our research we will investigate the opportunities of a webcam for continuous online and unobtrusive gathering of affective user data in an e-learning context.

Emotions are a critical component of effective learning and problem solving, especially when it comes to interacting with computer-based learning environments (CBLEs; multi-agent systems, intelligent tutoring systems, serious games [17].

II. Related Work

Major component of human communication are facial expressions which constitute around 55 percent of total communicated message

DOI: 10.9781/ijimai.2016.4112

[32]. We use facial expressions not only to express our emotions but also to provide important communicative cues during social interaction, such as our level of interest, our desire to take a speaking turn and continuous feedback signaling understanding of the information conveyed.

The facial expression research is an actual topic. The basic essays about expressions which are forming the current ones can be found in the 17th century. A detailed description of various expressions and facial muscle movements was provided by John Bulwer in 1649 in his book "Pathomyotomia".

Another important work dealing with facial expression analysis was by Charles Darwin. In his work he described and assorted groups of expressions into categories according to similarities. He described deformations as well by which each expression is formed [6].

Though facial expressions obviously does not necessarily convey emotions, in the computer vision community, the term ―facial expression recognition often refers to the classification of facial features into one of the six so called basic emotions: happiness, sadness, fear, disgust, surprise and anger, as introduced by Ekman. The advantage of this categorization is its universality among races and various cultures. An alternative of this category is a categorization designed by Baron-Cohen which includes more complex emotions. The group contains 412 various emotions divided into 24 groups. Ther are emotions such as boredom, interest, frustration [5]. The problem of the classification is a smattering knowledge about the universality towards various cultures and races [4]. Researches have shown that, in fact, various facial expressions are hard to record because while following, people are changing them subconsciously. Thus differences between real spontaneous expression and played expression are created [14].

There are two approaches by the use of which emotional state evaluation can occur. The first approach is to have native coders see the images or videotapes, and then make holistic judgments concerning the degree to which they see emotions on target faces in those images. While relatively simple and quick to perform, this technique is limited in that the coders may miss subtle facial movements, and in that the coding may be based by idiosyncratic morphological features of various faces. Furthermore, this technique does not allow for isolating exactly which features in the face are responsible for driving particular emotional expressions. The second approach is to use componential coding schemes in which trained coders use a highly regulated procedural technique to detect facial actions. For example, the Facial Action Coding System [15] is a comprehensive measurement system that uses frame-by-frame ratings of anatomically based facial features (''action units'') [2]. While lot of work on FACS has been done and also FACS is an efficient, objective method to describe facial expressions, but coding a subject's video is a time- and labor-intensive process that must be performed frame by frame. A trained, certified FACS coder takes on average 2 hours to code 2 minutes of video. In situations where real-time feedback is desired and necessary, manual FACS coding is not a viable option [42].

Automatic facial expression recognition and emotion recognition have been researched extensively [2]. In the last decade, both mentioned approaches have been used in facial expression and emotion recognition. In their cases, we cannot speak about automatic recognition of facial expressions and emotions. Gained images have been manually compared and progressively evaluated. The work of [39] is considered to be the first automatic comparison system. They developed a system for automatic recognition of facial action units and analyzed those units using temporal models from profile-view face image sequences. For quicker way of comparison, neural networks were starting to be used. Progressively, they way of facial expression recognition has shifted to the other approach – recognition in real time. They have developed a general computational model for facial affect inference and have implemented it as a real-time system. This approach used dynamic Bayesian networks for recognizing six classes of complex emotions. Their experimental results demonstrated that it is more efficient to assess a human's emotion by looking at the person's face historically over a two second window instead of just the current frame. Their system was designed to classify discrete emotional classes as opposed to the intensity of each emotion [2]. Others, such as Corzilus and Smids related to the work but their designed system had no response [10].

The problem of real time smile detection related is facial expression recognition. Sensing component company Omron [38] in 2009 has released smile measurement software. It can automatically detect and identify faces of one or more people and assign each smile a factor from 0% to 100%. Omron uses 3D face mapping technology and claim its detection rate is more than 90%.

III. Automatic System Of Emotion Recognition (Theoretical Analysis)

The system for automatic facial expression recognition has to deal with these problems: facial detection and location in a chaotic environment, extraction of facial features and correct classification of the expression on face [53].

Drawing upon various published works [20], [16], [48], [1], [43], main system elements from various similar system have been identified. The main elements of each system to evaluate facial expression are:

1. facial detection,
2. extraction of features,
3. expression evaluation.

Fig. 1. Database operation – testing application (own creation, used image from Jaffe database)

Other common features can be found in all works as well. The general proposition for systems with automatic emotional state evaluation of the user can be divided into six parts [21]:

1. image acquisition,
2. facial detection,

3. preprocessing,
4. feature acquisition,
5. classification,
6. final processing and correction.

The proposition can be extended by access to emotions. According to them it would be decided which kinds of features are appropriate to use. Each solution is different only in approaches they have to each part of the proposition. In the proposition, the model has been enriched by other elements in order to acquire fully generalized proposition.

The aim was to create an application with a simple GUI which on the bases of the webcam inputs would evaluate the actual emotional state of the user. From the results it would provide output in the form of the given situation result.

Software which was designed for uses webcams allows to interpret the emotional state of students during their interactions with an e-learning environment. This can trigger timely feedback based upon learner's facial expressions and verbalizations. The following emotions were observed: sadness, anger, disgust, fear, happiness, surprise, and neutral.

The primary objective when realizing the system was a progressive problem analysis, feature detection, state recognition and state evaluation. The original six-level model has been extended, edited and parts were added concerning the data preparation and basis. The resulting model has the form:

1. choice of monitored emotions,
2. choice of monitored features,
3. database acquisition,
4. image acquisition,
5. facial detection,
6. preprocessing,
7. feature approximation,
8. feature acquisition,
9. classification and correction.

IV. Test Sample Database

One of the most important aspects of creating a new system to detect or recognize, is a database choice which would be used for new system testing. If only one database was used for each research, then new system testing, comparing with other top-class systems and efficiency testing would be rather trivial tasks [6].

The standard database for testing the systems is currently the FERET Face database [6], [46]. The next used database if Jaffe (Japanese Female Facial Expression). There are images of women in gray scales with checked background and with seven different expression based on Ekman´s classification. The advantage of the database is that the expressions are already classified and described.

The main system testing is carried out on a custom sample of data. The advantage of the sample is that it is directly aimed to reveal and test weaknesses of the system. Images of various quality, size, with barriers on the face (hairstyle or glasses) were raised. The images were taken from webcams with various settings and in contrast to the database Jaffe, they do not have checked background. They are characterized by wider snapshot of face rotation and tilt. Thus the sample reflected the conditions the most realistically in which the system would work (bad light conditions, averted face, glasses, etc.). According to results gained by the system, other optimization would take place and more exact conditions would be evaluated for the image parameters.

V. Detailed Analysis Of The Six-level Module

This section lists six basic steps that make up the analysis of emotions:

1. Image acquisition,
2. Detection and facial rejection,
3. Preprocessing,
4. Feature Acquisition,
5. Emotion classification,
6. Final processing and correction.

A. Image Acquisition

When evaluating the emotional state:

1. static images,
2. sequences

can be used as input.

The static images are enough to acquire the expression. Although, the result contains less information according to which it would be possible to evaluate the expression. A common problem is for example the found out neural emotion when transforming from one expression to another. The sequences carry more information and provide better possibilities to optimize, to evaluate emotions and facial detection [21]. From the programming point of view, the use of sequences is demanding because a new problem is needed to be solved – facial observation. By combining the techniques, a compromise can be done – from sequence images faces can be acquired and then assigned to faces found in other images from the sequence [6].

B. Detection and facial rejection

For facial detection, it is possible to use various methods, and according to evaluation access they can be divided into four groups:

1. knowledge-based methods
2. feature invariant approaches
3. template matching methods
4. appearance-based methods.

The first approach, knowledge-based method, is based on a man´s knowledge about a typical facial appearance. Usually, it represents the resolution and the relation among the facial features.

The second approach, feature invariant approache, is based on determining the signs and the rules which define the face when changing poses, or at bad light conditions. The face is being searched for according to these signs.

The third approach, template matching method, compares parts of image with the templates and facial patterns or each features.

The fourth approach, appearance-based method, the templates/ models are created from groups of images which correctly represent tha variability of the face. Thus learned models are used for detection [52].

Those methods are considered to be the best ones which have a high percentage of successful recognition even in unsuitable conditions, unchecked environment (uniform background, various levels of light conditions). They also fulfil a requirement of an image evaluation in real time (Fig. 2).

The majority of software's work with a successfulness of about 70%. It was possible to reach higher scores by removing the common imperfections of the software (see the implemented research).

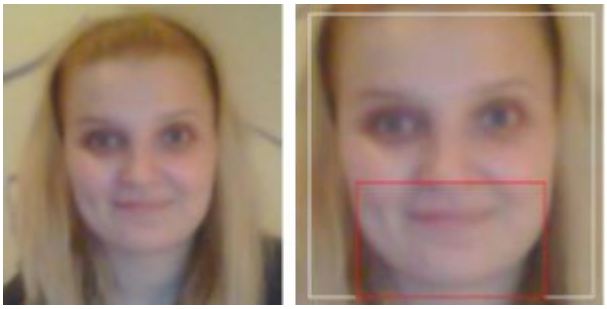


Fig. 2. Samples of custom database (on the left – an image taken by a webcam, on the right is facial and mouth detection – emotion evaluation)

The most used method is Viola-Jones, based on Haar cascade [50]. The method was used in the solution. Another method searched for connected groups of pixels which have the skin colour. According to Dadgostar and Sarrafzadeh [11] the skin colour is found in HSI model between values 0-33 for H and 15_250 for value I. The group of the pixels is processed and facial geometry is applied on them. The pixel group which fulfil the requirements would be evaluated as a human face. For facial detection, neural networks might be used, eigenfaces methods [49], support vector machine and others. For more complex view, we draw upon the work Yang, Kriegman and Ahuja which deals with description of various methods [52].

An important requirement is face rejection. All faces would be rejected where the wanted features or features needed for further evaluation are impossible to find. Also, all faces might be rejected which do not fulfil the requirement of size. Haar cascade is able to find features which have minimal size and it can be defined and calculated according to face size. That is why it is possible to exclude faces which do not fulfil the size requirement and thus there would be no wanted features. When finding more faces, the most dominant face would be left and that would mean, it has the biggest size. Next rejections would follow after feature approximation. In Fig.3 there is an image before and after rejection of incorrect faces.

Fig. 3. Before and after face rejection (own creation)

After finding the face, it is possible to use normalization methods and more accurate face designation. An example can be detection, cropped hair, removing background, brightness compensation and other.

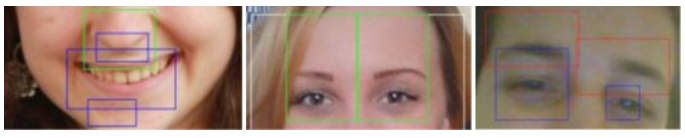


Fig. 4. Examples of mouth, eye and eyebrow detection (own creation)

C. Preprocessing

Image preprocessing often precedes face detection. The main way of preprocessing is removing noise and other processes done in order to improve quality and image appropriateness [21]. Other ways of preprocessing can be size changes, colour converting of the image into grey colour spectrum, intensity increase or applying other colour transformations. Thus, information about blushing can be lost which is lowering the accuracy of calculation [21].

D. Feature acquisition

Feature acquisition on the face is similar to facial detection on an image. Procedures and methods of facial detection can be easily modified and used for feature detection. To each feature should be approached distinctively to get the most precise feature detection and choose a method to acquire it [36], [34], [40], [3]. Precision is also important by which the feature is acquired. Optimized haar cascade adds approximated feature position but is not enough for expression evaluation and is more appropriate as approximator and a support for other methods.

When searching for lips position, it is possible to use colour transformation designed in ??? [9]. To acquire eyebrow position it is better to use method based on another principle [40].

Features can be processed according to two approaches [51]:

1. holistic,
2. local.

Holistic approach studies the face as a whole. The local approach is oriented on each features or face characteristics which can be the subjects of change [51].

The paper deals with the local approach mainly because of the easy implementation by the use of haar cascades.

According to feature detection it can be divided into:

1. approach based on geometry,
2. method based on template,
3. methods based on appearance,
4. methods based on colour and texture [36], [34], [40], [3].

E. Emotion classification

In classification, we draw on Ekman´s definition of six emotions. We dare to complete it as below:

1. joy
2. sadness
3. surprise
4. fear
5. anger
6. disgust
7. neutral expression.

An important factor when creating the system to recognize expressions is a set of expressions which are needed to be recognized. In analysis a few approaches to emotion categorization have been mentioned. The observed features depend on the categorization. The system serves mainly as support in e-learning applications so the final list of expressions can be edited. Extreme expressions can be omitted such as scare, wrath, big shock and other the final categories can be divided into three groups:

1. positive
2. negative
3. neutral.

Drawing upon the hypothesis it can be possible to reduce wanted features. The nose and eyes play an important role in escalated expression and thus are not that important, in comparison to mouth and eyebrows which would be the important features.

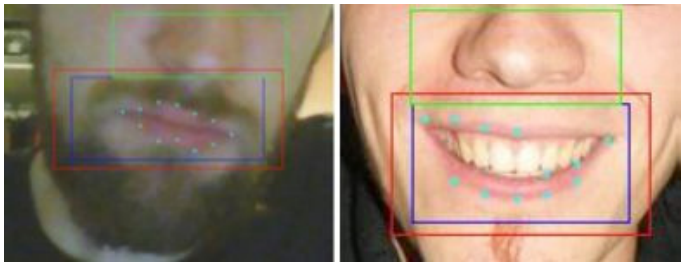


Fig. 5. Mouth extraction (own creation)

Although, both are needed to be approached differently due to anatomic differences.

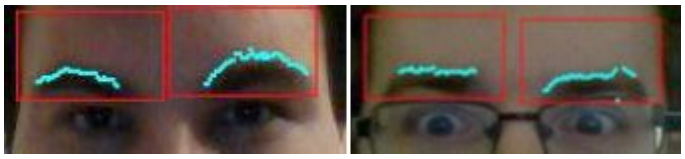


Fig. 6. Eyebrow extraction (own creation)

F. Final processing and correction

The last part of the evaluation of the emotional state of the user's is facial expression evaluation. As Sumathi, Santhanam and Mahadevi indicate to evaluate the expression, two approaches are used [44]:

1. frame based expression recognition,
2. sequence based recognition.

Frame based expression recognition uses a single image as an input. In the case of evaluation of other images, it approaches distinctively to each one and does not save information about their connection and progress. The main methods of frame based recognition are neural networks, SVM (Support Vector Machine), rule based classifiers and linear discriminant analysis.

Sequence based recognition uses image sequence as an input. Besides of images, it saves information about their progress ans uses face observation. The main methods of sequence approach are recurrent neural networks, hidden Markov models, rule based classifiers [44].

According to Chibelushi another step can be result correction based on for example knowledge about common mistakes and incorrect classifications [21].

VI. Percentage Software Successfulness – Comparison Of Results Acquired From The Software With Real Observed Emotions

When verifying the software reliability, we draw on a similarly orientated research [1]. The following hypothesis has been made.

Hypothesis: There is a reliable use of data acquired by webcam and the designed software for user´s emotion recognition.

Images with facial expressions were not used, but observations have been made about what emotions are evoked by the images. Each image represented a certain emotion and was provided to assistants at the department of psychology. This way has been chosen not for the users to not realize their emotions but to reach the truest result. Samples of same size have been chosen – 10 students. These numbers are taken from all 1000 emotions (10 test persons displaying 100 emotions each) including the cases that one or more of the rates judged that the test person was unable to mimic the requested emotion correctly. Each requested emotion is separated in two rows that intersect with the recognized emotions by the software.

In the following table (Table 1) there are students´ emotions in contrast to software results determined for recognition. To be able to compare the students´ emotions with the acquired results, students were stealthily shot by the camera (with their permission). The video underwent an analysis. It was done by an assistant from the department of psychology and he provided detailed explanation to each analysed emotion (how did it contrast with the result achieved by the designed software).

Software that uses Bahreini, Nadolski and Westera [1] has the highest recognition rate for the neutral expression (77,2%) and the lowest recognition rate for the fear expression (50%). Note that the obtained differences between software and requested emotions are not necessarily software faults but could also indicate that participants were sometimes unable to mimic the requested emotions. The software had in particular problems to distinguish surprise from neutral. Error rates are typically between 1% and 14%. The software confused 11,3% of the neutral emotions as surprise and confused 12,5% of surprise as neutral.

The achieved results considerably differentiate from the results of Bahreini, Nadolski and Westera [1]. It is caused because of different recognition technology of each facial parts which contributes to detailed analysis and to more relevant results. The reliability of our software is 78%.

According to Table 1, the software has the highest rate of recognition for neutral expression (98,7%) and the lowest rate for sad expression (54,7%) (Table 1). Similarly, in our case there were relatively large deviations. We are inclined to the view of Bahreini, Nadolski and Westera [1] that the acquired differences between the software and the required emotions are not necessarily software mistakes. Mistake rate is in interval of 0 to 15,3%, specifically in the case of differentiating anger from disgust. What is interesting, that the software shows very different results in this cases: confused 12,5% of the neutral emotions as surprise and confused 7,2% of surprise as neutral.

TABLE I
REQUESTED EMOTIONS AND RECOGNIZED EMOTIONS BY THE SOFTWARE – THESE NUMBERS ARE TAKEN FROM ALL 1000 EMOTIONS INCLUDING 'UNABLE TO MIMIC' BY THE PARTICIPANTS (10 PARTICIPANTS DISPLAYING 100 EMOTIONS EACH).

		Recognized Emotion by the Software							
		Happy	Sad	Surprise	Fear	Disgust	Angry	Neutral	Total
Requested Emotions	Happy	94	8	7	12	3	1	5	130
		78,3%	5,4%	2,9%	8,0%	1,3%	0,7%	3,3%	100%
	Sad	0	52	7	5	12	7	12	95
		0,0%	54,7%	7,4%	5,3%	12,6%	7,4%	12,6%	100%
	Surprise	5	2	73	1	2	5	8	96
		4,8%	2,1%	76,8%	1,1%	2,1%	5,1%	8,0%	100%
	Fear	0	5	9	45	13	4	15	91
		0,0%	5,2%	10,2%	60,0%	10,8%	2,6%	11,1%	100%
	Disgust	8	2	1	3	68	4	3	89
		8,4%	1,7%	1,2%	3,2%	80,0%	3,4%	2,1%	100%
	Anger	0	4	2	2	13	63	1	85
		0,0%	4,7%	2,4%	2,4%	15,3%	74,1%	1,2%	100%
	Neutral	8	17	42	12	18	3	314	414
		0,0%	0,2%	0,2%	0,4%	0,4%	0,0%	98,7%	100%
Total		113	87	145	74	130	89	362	1000

Table 1 is designed in a way it would be possible to differentiate all the seven basic emotions and easily identify the software results. Disgust has the second biggest value (80%)which is contrast to Bahreini, Nadolski and Westera. Apart from neutral, the emotion that shows best discrimination from other emotions is disgust, as disgust has a high score of 80% and is not confused with happy, sad, and

angry. The most difficult emotion is sad (54,7%) and fear (60%) and is easily confused with anger .74,1%. The difference is not a software mistake. In our opinion, the difference might have happened because the students observe the image differently and simultaneously achieve more emotions (anger and disgust). The affirmation is in accordance with Murthy [35] and Zhang [53] - the most difficult emotion to mimic accurately is fear and this emotion is processed differently from other basic facial emotions. According to various researches [35] the three emotions sad, disgust, and angry are difficult to distinguish from each other and are therefore often wrongly classified. This is confirmed by our acquired and processed results.

In following table it is visible that the analysis results of Agreements and disagreements about 1000 as students perceive various images - evaluation emotions from professor assistant at the department of psychology.

TABLE II
AGREEMENTS AND DISAGREEMENTS ABOUT 1000 AS STUDENTS PERCEIVE VARIOUS IMAGES - EVALUATION EMOTIONS FROM PROFESSOR ASSISTANT AT THE DEPARTMENT OF PSYCHOLOGY.

	Happiness	Sadness	Surprise	Fear	Disgust	Anger	Neutral	Total
Raters agree:	125	19	65	13	46	54	398	720
Able to perceive	91%	21,3%	67,8	15,4%	51,9%	61,7	89,3%	72%
Raters disagree:	21	24	25	16	21	15	61	183
Able/unable to	14,5%	25,6%	28,7%	23,9%	29,3	24,6	18,2	18,3%
Raters agree:	5	29	8	33	8	4	10	97
Unable to perceive	2,1%	35,6%	12,4%	51,3%	12,4%	5,8%	3,6%	9,7%
								100%

According to the assistant of psychology, Table 2 specifies that the students were able to perceive various images (requested emotion in 72% of the occurrences). In 183 occurrences (18,3%) there was disagreement between rater and software evaluation In 9,7% of the cases the rater stated that students were unable (on the basis of each other) to perceive requested emotions (97 times). It is interesting that students are best at perceiving neutral (89,3%) and worst at fear (15,4%). It should be noted that students are not apart. Therefore, an imitation of emotions there is possible.

For correct percentage software successfulness the results of each emotion separately and in total were re-calculated. In Table 3 the requested emotions of participants are shown (these numbers are taken by the raters from 720 emotions of the participants that were able to perceive the requested emotions) contrasted with software recognition results. Difference between Table 1 and Table 3 is that we removed both the 'unable to perceive' various images and the records from assistant psychology disagreed from the dataset.

TABLE III
REQUESTED EMOTIONS AND RECOGNIZED EMOTIONS BY THE SOFTWARE

		Recognized Emotion by the Software							
		Happiness	Sadness	Surprise	Fear	Disgust	Anger	Neutral	Total
Requested Emotions	Happiness	93	5	10	5	3	2	7	125
		84,5%	5,3%	6,3%	1,9%	0,6%	0,3%	1,2%	100%
	Sadness	0	12	4	1	1	0	1	19
		0,0%	75,0%	6,7%	5,2%	10,4%	0,0%	2,8%	100%
	Surprise	1	2	52	1	2	3	4	65
		2,9%	5,0%	63,4%	2,9%	5,7%	8,6%	11,4%	100%
	Fear	0	0	3	8	0	1	1	13
		0,0%	0,0%	6,1%	72,7%	0,0%	11,1%	10,1%	100%
	Disgust	6	0	1	1	33	2	3	46
		7,1%	0,0%	1,1%	1,1%	82,5%	5,1%	3,1%	100%
	Anger	0	1	0	0	3	49	1	54
		0,0%	2,0%	1,1%	1,4%	10,0%	84,7%	0,8%	100%
	Neutral	7	10	42	9	15	1	314	398
		3,0%	3,4%	8,1%	1,1%	1,5%	0,1%	82,6%	100%
Total		107	30	112	25	57	58	331	720

In Table 3, there are results that have been achieved by removing incorrect records and repeatedly re-calculate the results achieved from evaluation when recognizing each expression of the students. The software´s success rate has been set as follows: all percentage result values have been added for each expression on the diagonal in Table 3 and divided by the number of evaluated expressions (7). Thus was the success rate of 78% achieved.

VII. Discussion

This study contrasted the requested emotions of participants with our designed software for the face emotion recognition. Results from assistant of psychology were used for evaluation. The best recognized emotion is anger 84,7% followed by happiness 84,5%, neutral 82,6%, disgust 82.5%, sadness 75%, fear 72.7%, and surprise 63.4%. These results are in stark contrast with the results of Bahrein, Nadolski and Westera [1]. Also, it has not been confirmed that the most intensive emotions are ranked higher than the less intensive emotions except the neutral emotion.

Anger and disgust have relatively lot of common facial signs [15], [1] and that is why high scores have been reached in Table 3 (82,5% and 84,7%).

Software precision can be verified in various ways. In the previous studies [28], [19], [1] the software has been verified on the basis of miming emotions. The respondents (note: they were not only university students, there were respondents of various age levels as well) were exposed to images with relevant expressions and their task was to mime them. The software scanned the faces and evaluated them. Consequently, results from the software were compared to required expressions.

The problem of this kind of determining emotion is the fact that the respondents might be aware about the testing and it might influence the results disconcertingly. In our case, it was not the student´s task to mime the expressions from the images, but to express the emotion that the image evokes in them. Thus evoked emotions were evaluated by the designed software and, consequently, software results were confronted with the statements of an expert assistant from the departments of psychology. The comparison results are in Table 2 while in Table 3 there are results achieved by removing incorrect records and repeated calculations of results achieved from evaluation while observing each students´ expressions. The software´s success rate has been set as follows: all percentage result values have been added for each expression on the diagonal in Table 3 and divided by the number of evaluated expressions. Thus the success rate is on the limit of 78%. Thus the stated hypothesis can be accepted.

For this method of determining the level of success we have chosen because of that the youngsters and older adults are not equally good in miming different basic emotions (e.g., older adults are less good in miming sadness and happiness than youngsters, but older adults mimic disgust better than youngsters), it is acknowledged that the sample of test persons might influence the findings of the software accuracy [19].

VIII. Conclusion

In this paper we have proposed a system that automatically detects human emotions on the basis of facial expressions. We have implemented this system in LMS Moodle and is used to test students. Our interest is to determine what emotions have students in testing and help them to overcome for example stress, anger or disgust. The system is still in the testing phase.

The system works well for faces with different shapes, complexions as well as skin tones and senses basic seven emotional expressions. The success rate is on the limit of 78%. Facial expression recognition is a challenging problem in the field of image analysis and computer vision. Inclusion of emotions in human computer interface is an emerging field mainly in fields of acquiring data to support education. The solution provides us with many new opportunities. It is an assumption that the development of computationally of effective and robust solutions will

lead to increased importance of user in the process and set stage for revolutionary interactivity.

References

[1] K. Bahreini, R. Nadolski, & W. Westera, Towards multimodal emotion recognition in e-learning environments. *Interactive Learning Environments,* no. Ahead-of-print, 2014, p. 1-16.

[2] J. N. Bailenson, E. D. Pontikakis, I. B. Mauss, J. J. Gross, M. E. Jabon, C. A. C. Hutcherson, . . . O. John, Real-time classification of evoked emotions using facial feature tracking and physiological responses. *International Journal of Human Computer Studies, 66*(5), 2008. 303-317.

[3] U. Bakshi, & R. Singhal, A Survey of face detection methods and feature extraction techniques of face recognition. *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, 3(3), 2014, 233-237.

[4] S. Baron-Cohen, Reading the mind in the face: A cross-cultural and developmental study. *Visual Cognition*, 3(1), 1996, 39-60.

[5] S. Baron-Cohen, O. Golan, S. Wheelwright, & J. Hill, Mind reading: the interactive guide to emotions. 2004.

[6] V. Bettadapura, Face Expression Recognition and Analysis: The State of the Art. arXiv: Tech Report, (4), 2012, 1-27.

[7] I. Birzniece, P. Rudzajs, D. Kalibatiene, O. Vasilecas, & E. Rencis, Application of interactive classification system in university study course comparison. *Informatics in Education, 14*(1), 2015, 15-36.

[8] P. Brusilovsky, I.H. Hsiao, & Y, Folajimi, QuizMap: open social student modeling and adaptive navigation support with TreeMaps. *Proceedings of the 6th European conference on Technology enhanced learning: towards ubiquitous learning* (EC-TEL'11), Carlos Delgado Kloos, Denis Gillet, Raquel M. Crespo García, Fridolin Wild, and Martin Wolpers (Eds.). Springer-Verlag, Berlin, Heidelberg, 2011, 71-82.

[9] U. Canzlerm & T. Dziurzyk, T. Extraction of Non Manual Features for Video based Sign Language Recognition. *Proceedings of IAPR Workshop*, 2002, 318-321.

[10] M. Corzilus, & M. Smids, *A Real-time Emotion Mirror*, 2006, Available: http://www.inf.unideb.hu/~sajolevente/papers2/emotion/emotion_mirror.pdf

[11] F. Dadgostar, & A. Sarrafzadeh, An adaptive real-time skin detector based on hue thresholding: A comparison on two motion tracking methods. *Pattern Recognition Letters, 27*(12), 2006, 1342-1352.

[12] K.E. DeLeeuw, R.E. Mayer, & B. Giesbrecht, How does text affect the processing of diagrams in multimedia learning. *Proceedings of the 6th international conference on Diagrammatic representation and inference* (Diagrams'10), Ashok K. Goel, Mateja Jamnik, and N. Hari Narayanan (Eds.). Springer-Verlag, Berlin, Heidelberg, 2010, 304-306.

[13] M. Drlík, & J. Skalka, Virtual Faculty Development Using Top-down Implementation Strategy and Adapted EES Model. *Proceedings of the World Conference on Educational Technology Research*. East Univ, Nicosia, Cyprus, Paper presented at the Procedia - Social and Behavioral Sciences, 28, 2011, 616-621.

[14] P. Ekman & E. L. Rosenberg, *What the face reveals: basic and applied studies of spontaneous expression using the facial action coding system (FACS)*, Illustrated Edition, Oxford University Press, 1996, 3-559.

[15] P. Ekman, & W. V. Friesen, Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*(2), 1971, 124-129.

[16] M. Feidakis, T. Daradoumis, & S. Caballe, Emotion Measurement in Intelligent Tutoring Systems: What, When and How to Measure. *Third International Conference on Intelligent Networking and Collaborative Systems*, 2011, 807-812.

[17] J. M. Harley, F. Bouchet, M. S. Hussain, R. Azevedo, & R. Calvo, A multi-componential analysis of emotions during complex learning with an intelligent multi-agent system. *Computers in Human Behavior, 48*, 2015, 615-625.

[18] I. -. Hsiao, P. Brusilovsky, M. Yudelson & A.Ortigosa, The value of adaptive link annotation in e-learning: A study of a portal-based approach. Paper presented at the *HT'10 - Proceedings of the 21st ACM Conference on Hypertext and Hypermedia,* 2010, 223-227.

[19] I. Huhnel, M. Fölster, K. Werheid & U. Hess, Empathic reactions of younger and older adults: No age related decline in affective responding. *Journal of Experimental Social Psychology*, 50, 2014, 136-143.

[20] Y. Chen, R. Hwang, & C. Wang, Development and evaluation of a web 2.0 annotation system as a learning tool in an e-learning environment, *Computers and Education*, 58(4), 2011, p. 1094-1105.

[21] CC. Chibelushi, & F. Bourel, Facial expression recognition: a brief tutorial overview, CVonline: On-Line Compendium of Computer Vision, *Proceedings of the 13th Annual ACM International Conference on Multimedia*, 5, 2002, 669-676.

[22] K. Jakobs, (2009). ICT standardisation in china, the EU, and the US. Paper presented at the *International Telecommunication Union - Proceedings of the 2009 ITU-T Kaleidoscope Academic Conference: Innovations for Digital Inclusion, K-IDI 2009, 2009,* 177-182.

[23] Š. Koprda, P. Brečka, & M. Maroš, Project and realization of WIFI net. *Proceedings of the 2009 conference* Trends in education: Information Technologies and technical education, Olomouc, Votobia, 2009, 301-304.

[24] Š. Koprda, M. Munk, D. Klocoková, & J. Záhorec, Základy práce s PC a v prostredí operačného systému MS Windows. In *Elearning 2006*. Hradec Králové: Gaudeamus, 2006, 239-249.

[25] K. Kostolányová, E-learning Form of Adaptive Instruction. Proceedings *of the 9th International Scientific Conference on Distance Learning in Applied Informatics (DIVAI 2012)*. Nitra: Faculty of Natural Sciences Constantine the Philosopher University in Nitra, 2012, 193-201.

[26] K. Kostolányová, O. Takács, & J. Šarmanová, Automated Assessment of Students' Knowledge in Adaptive Instruction. *Information and Communication Technology in Education (ICTE 2012)*. Univ Ostrave, Pedag Fac, Roznov pod Radhostem, Czech Republic, 2012, 119-130.

[27] K. Kostolányová, O. Takács, & J. Šarmanová, Adaptive Education Process Modeling. *Proceedings of the 10th International Conference on Efficiency and Responsibility in Education (ERIE 2013)*, Prague 6th - 7th June 2013. Prague : Czech University of Life sciences, 2013, 300-308.

[28] E. Krahmer, & M., Swerts, Audiovisual Expression of Emotions in Communication. *Philips Research Book* Series. Springer Netherlands, 12, 2011, 85-106.

[29] H.-T. Lin, Ch.-H. Wang,, CH.-F. Lin, & S.-M. Yuan, Annotating Learning Materials on Moodle LMS. *Proceedings of the 2009 International Conference on Computer Technology and Development, Volume 02* (ICCTD '09), IEEE Computer Society, Washington, DC, USA, 2009, 455-459.

[30] M. Magdin, M. Cápay, New Module for Statistical Evaluation in LMS Moodle. *Proceedings of the 11th European Conference on e-Learning*, ECEL 2012 : 26-27 October, Groningen, The Netherlands. - Groningen : Academic Publishing International Limited, 2012. 304-310.

[31] R. E. Mayer, Applying the science of learning to medical education. *Medical Education*, 44(6), 2010, 543-549.

[32] A. Mehrabian, Communication without words. *Psychology Today*, 2(4), 1968, 53–56.

[33] P. Miller, Web 2.0: Building the New Library. *Ariadne,* Issue 45, 2005, Available: http://www.ariadne.ac.uk/issue45/miller/

[34] J. L. Moreira, A.Braun, & S. R. Musse, Eyes and eyebrows detection for performance driven animation. Paper presented at the *Proceedings - 23rd SIBGRAPI Conference on Graphics, Patterns and Images, SIBGRAPI,* 2010*,* 17-24.

[35] G. R. S. Murthy, & R. S. Jadon, Effectiveness of Eigenspaces for facial expression recognition. *International Journal of Computer Theory and Engineering*, 1 (5), 2009, 638-642.

[36] M. Nixon, & A. S. Aguado, *Feature Extraction and Image Processing*, 2nd ed. Academic Press, 2007.

[37] A. Ollo-López, & M. E. Aramendía-Muneta, ICT impact on competitiveness, innovation and environment. *Telematics and Informatics*, 29 (2), 2012, 204-210.

[38] Omron. *OKAO Vision*, 2009, Available: http://www.omron.com/r_d/coretech/vision/okao.html.

[39] M. Pantic, N. Sebe, J. F. Cohn, & T. Huang, .Affective Multimodal Human computer interaction, 2005

[40] U. Saeed, & J. C. Dugelay, Combining edge detection and region segmentation for lip contour extraction. *Proceedings of the 6th international conference on Articulated motion and deformable objects* (AMDO'10), Francisco J. Perales and Robert B. Fisher (Eds.). Springer-Verlag, Berlin, Heidelberg, 2010, 11-20.

[41] A. Sarrafzadeh, S. Alexander, F. Dadgostar, C. Fan, & A. Bigdeli, How

do you know that I don't understand? A look at the future of intelligent tutoring systems. *Computers in Human Behavior*, 24(4), 2008, 1342-1363.

[42] S. Srivastava, Real Time Facial Expression Recognition Using A Novel Method. International Journal of Multimedia & Its Applications (IJMA), 4(2), 2012, 49-57.

[43] M. Suk, & B. Prabhakaran, Real-time facial expression recognition on smartphones. Paper presented at the Proceedings - 2015 IEEE Winter Conference on Applications of Computer Vision, WACV 2015, 2015, 1054-1059.

[44] C.P. Sumathi, T. Santhanam, & M. Mahadevi, Automatic Facial Expression Analysis A survey. International Journal of Computer Science & Engineering Survey (IJCSES), 3(6), 2012, 47-59.

[45] J. Sun, Y. M. Chai, Guo, & H. F. Li, On-line adaptive principal component extraction algorithms using iteration approach. Signal Processing, 92(4), 2012, 1044-1068.

[46] The FERET Database [online]. Available: http://www.itl.nist.gov/iad/humanid/feret/

[47] V. T. Tinio, ICT in Education. Presented By UNDP for the Benefit of Participants To The World Summit On The Information Society. UNDP's Regional Project, the Asia-Pacific Development Information Program (APDIP), In Association with the Secretariat of ASEAN, 2009, Available: http://www.apdip.net/publications/iespprimers/eprimer-edu.pdf

[48] M. Trojahn, F. Arndt, M. Weinmann, & F. Ortmeier, Emotion recognition through keystroke dynamics on touchscreen keyboards. Paper presented at the ICEIS 2013 - Proceedings of the 15th International Conference on Enterprise Information Systems, 3, 2013, 31-37.

[49] M. Turk, & A. Pentland, Eigenfaces for recognition. Journal of Cognitive Neuroscience, 3(1), 1991, 71–86.

[50] P. Viola, & M. J. Jones, Robust real-time object detection. International Journal of Computer Vision, 57(2), 2004, 137–154.

[51] J. Whitehill, G. Littlewort, I. Fasel, M. Bartlett, & J. Movellan, Toward Practical Smile Detection. Pattern Analysis and Machine Intelligence, 31(11), 2009.

[52] M. Yang, D. J. Kriegman, & N. Ahuja, Detecting faces in images: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(1), 2002, 34-58.

[53] Z. Zhang, Feature-Based Facial Expression Recognition: Sensitivity Analysis and Experiment with a Multi-Layer Perceptron. International Journal of Pattern Recognition Artificial Intelligence, 13 (6), 1999, 893-911.

Martin Magdin works as a professor assistant at the Department of Computer Science. He deals with the theory of teaching informatics subjects, mainly implementation interactivity elements in e-learning courses, face detection and emotion recognition using a webcam. He participates in the projects aimed at the usage of new competencies in teaching and also in the projects dealing with learning in virtual environment using e-learning courses.

Milan Turčáni is head of Department of Computer Science and works as a professor at the Department of Computer Science. He deals with the theory of teaching informatics subjects, mainly implementation e-learning in learning process. He participates in the projects aimed at the usage of new competencies in teaching and also in the projects dealing with learning in virtual environment using e-learning courses. He is supervisor of important projects, with a focus to the area of adaptive hypermedia systems.

Lukáš Hudec was a student at the Department of Computer Science. In present is Java developer in company EmbedIT. As a student deals with facial expression recognition. At present, it continues as a consultant for the development and improvement of software for recognition of emotions students through webcams.

Comparison between Famous Game Engines and Eminent Games

Prerna Mishra and Urmila Shrawankar

Department of Computer Science & Engineering, RTMNU, Nagpur (MS), India

***Abstract* — Nowadays game engines are imperative for building 3D applications and games. This is for the reason that the engines appreciably reduce resources for employing obligatory but intricate utilities. This paper elucidates about a game engine, popular games developed by these engines and its foremost elements. It portrays a number of special kinds of contemporary game developed by engines in the way of their aspects, procedure and deliberates their stipulations with comparison.**

***Keywords* — Game Engine, Non-Cloud Gaming Platforms, Cloud Gaming Platforms, Popular Games, GPGPU**

I. Introduction

IN general, the notion of game engine is very easy to comprehend. It is a platform for performing game related tasks like interpretation, physics related reckoning, and to facilitate developers for focusing on the niceties that make the game inimitable. Engines are in reality an assemblage of reusable modules that can be manipulated in order to carry a game towards realism. Indeed, there are certain disparities between a game and a game engine [1]. Graphics, animation, audio, physics, UI and AI are the major different constituents of an engine. Conversely the subject matter of a game, its definite characters and background, real world avatar and its behaviours etc. are the components that create the real game. Game engines are middleware's.

Game engines produce the replication of actual world in the digital world by controlling the elementary physics. Games developed by these engines make user, casual or die-hard player. In general, smartphone operators, as mobile game players, are alike casual gamers rather than die-hard gamers. In contradiction of die-hard gamers, casual gamers are outlined as less dedicated, less spirited, and more tranquil users. Casual gamers incline to be not as much of ardent and less fascinated to classy or multifarious games (comparative to gameplay, achieve target, environs, graphics, chaps, etc.).The market for video games is growing, with sales in 2015 of $91.5 billion marking an 11.84% increase over 2014, at this price global revenues are expected to reach $107 billion in 2017. However, growth is not only in sales but also in the miscellany of matter offered, vacillating from scholastic games to first-person shooters. In addition, a captivating conjunction of mass media is proceeding with video games, having motion picture eminence cut-scenes and voiceover.

In a cloud environment, the task of service provider is separated into two: the infrastructure providers managing cloud platforms and rent assets conferring to a usage, and service providers leasing assets from one or more infrastructure providers to assist the users. Lately, a novel type of cloud service has been familiarized, which have the utmost severe exigencies on network Quality of Service (QoS) to date known as cloud gaming. However, in cloud gaming the complete user experience is provided through the network. This creates dissimilarity between cloud gaming and conventional online Gaming in stipulations of network quality of experience (QoE). While in conformist Online Gaming the user experience is spawned at the client side so the network does not have any impact on the performance, affecting the worth of Cloud Gaming [2, 4]

With the constituent of video gaming disappearing, people merely favour to finance time in real-time mobile games that are compatible with an extensive array of platforms and operating systems. These engines takes the gaming experience to an entirely new-fangled echelon, avoiding poor graphics and quality experiences with the similar joysticks from the past to play around.

This paper portrays the comparative study of few diverse cloud and non-cloud platforms that are currently associated with gaming. As games are evolving progressively delivering profounder and added a biding experience for players, their prospective for psychosomatic impression is growing in proportion. Some renowned companies of games are Microsoft Game Studios, Electronic Arts, etc [3]. Console producers are a company yielding and disseminating video game consoles. Some of the utmost familiar console producers are Atari, Microsoft Corporation, Nintendo Company, Sega, and Sony Computer Entertainment Inc. However, currently there are three major popular platforms Microsoft Xbox 360, Sony PlayStation3 and Nintendo apart from mobile gaming apps as shown in figure 1.

II. Comparative Study

A. Non-Cloud Platform Stage

Analysis amid numerous game engines is a daunting errand for the reason that of their innumerable field, kinds, hypermedia sustenance, middleware support, language and platform enslavements and many other deputized characteristics. It has to be acknowledged that the non-cloud game platforms have dissimilar features, modules, benefits and restrictions relating some eminent game engines. Six popular game engines are compared with each other namely Cry Engine 3, Hero Engine, Source 2 Engine, Unity 4 Engine, Unreal Engine 4 and Vision Engine 8.Table 1 expresses the efficacy in terms of platform reliance, interface and language, intrinsic physics and AI engine supported with forward and backward compatibilities.

An analysis was completed with the help of 15 students (beginner coders) and 20developers from different game industries who gave their views on a number of game engines currently accessible in the marketplace in India. Figure 1 depicts the popular gaming platforms in India.

B. Cloud Gaming Platforms

A cloud gaming [5-7] structure assembles the participant's activities, transfers it to the cloud server, concocts the act, extract the consequences, and conceals the resultant modifications into the gaming simulating environment and rivulets the gaming sequences in return to the player.

DOI: 10.9781/ijimai.2016.4113

TABLE 1
ASSESSMENT OF ENGINES ON PERFORMANCE ASPECTS

Game Engine	Platforms	Language Support	AI Engine	Physics Engine	Forward compatibility	Backward compatibility
CryEngine 3	Win, X360, PS3, Wii U	C++, Visual Script, Lua	Lua-driven AI	Soft-body	No	Yes
Hero Engine	Win	Hero Script	AIseek	PhysX	Partial	Yes
Source 2 Engine	Win, Mac, Xbox 360, Wii, Linux, Android	C++	AI Director	Ipion	No	Partial
Unity 4	BlackBerry, Win Phone, Win, OS X, Android, iOS, Apple TV, PS3/4, PS Vita, Xbox 360, Xbox One, Wii U, Wii.	C#, JavaScript, Boo	RAIN	PhysX	Partial	Yes
Unreal 4 Engine	Windows, OS X Linux, Xbox 360/ One, PS3/4, Wii U, Android, iOS, WinRT, PS Vita	C++, C#, GLSL, CG, HLSL	Kynapse	PhysX	Partial	Yes
Vision Engine 8	Windows, Xbox 360, PS3, Wii, Wii U, iOS, Android, Win Phone, PS Vita	C++	Kynapse	Bullet, ODE, PhysX	No	Partial

For safeguarding inter communication, all the sequential actions tends to happen in a period of milliseconds. Subliminally, the total time, which can be said as communication interruption or delay, essentially are kept negligibly conceivable for delivering an amusing involvement to cloud game players. One of the modest methodologies for supporting cloud gaming is to use the services of general desktop streaming thin clients, such as Ubitus, GamingAnywhere [6], and VirtualGL.

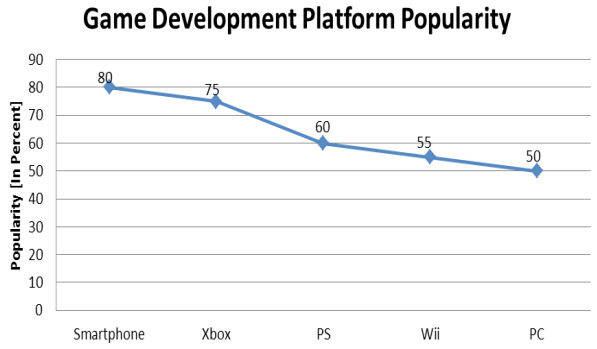


Fig 1. Popularity of gaming platforms in India

Cloud gaming not only necessitates elevated continual downlink bandwidth but low latency too [5-7]. Cloud gaming by now has engendered a prodigious contract of curiosity amongst business persons, venture investors, and the end users. Some prominent cloud platforms are discussed below.

1) *Gaikai*

In GaiKai, games are uploaded to the datacenters situated largely around the globe. From these centers, they are cascaded using high-end servers to the devices connected to each other by medium of internet, analogous to the manner videos are streamed to the user's computer. It endows the capability of streaming the graphically rich and real games and additional related data rapidly to more or lessons any devices from anywhere in the world. Gaikai has developed the utmost excellence; quickest communicating cloud-streaming platform on the globe, aiding a platform to developer's proficient enough for providing games and other communicating matter instantaneously to the end users via the Internet. Nevertheless Gaikai do not sustain packages on devices like digital TVs and tablets

2) *StreamMyGame*

StreamMyGame [6] is a wide-ranging software elucidation that empowers games and applications to be played remotely. One accesses and plays games tenuously via their local/home network. This gaming platform gains access and engages in gameplay tenuously by means of their broadband network, note downs gameplay to High definition Video archives, one can upload recorded HD video files to any online video sites and publicize games, so anybody on their native system can perceive it and participants gets access to amenities of meeting other players by means of creating the groups with forums, chats, recognize new and former participants presence.

3) *OnLive*

OnLive [5] conveys on request real-time communal experiences with opulent content via the Internet. OnLive uses cloud to deliver the potency and astuteness needed to immediately convey comprehensive applications having animations, graphics, AI, physics, etc.

OnLive make use of virtual machines on customized servers with graphics processing units (GPUs) and commercial compression procedures, which has to work-out two issues for game respectively. Live stream is augmented for gameplay depending on physical-world Internet circumstances and Media stream is a server-side full HD stream managing viewers or players for recording and reviewing progressions of their games. In spite of the benefits, Onlive has quite a few restrictions. Actually OnLive is not reasonably noble enough for accomplishing1080p resolution. That is to say for the reason that it requires broadband speediness of up to 10 MB per second.

4) *GamingAnyWhere*

GamingAnyWhere is the paramount open-source game development platform [6]. It concedes scholars to experiment notions and concepts on a cloud gaming proving-ground, gaming bringers to cultivate services on it and player scan fabricate their personal gaming clouds from their computers [11]. Moreover, of its candidness, it is designed for high affability, maneuverability, and re-assessability. Platforms supported by GamingAnyWhere are Windows, Linux, and OS X, iOS and Android. GamingAnyWhere is multi-platform, proficient in footings of time and space complexities and offers accessible unrestricted platform.

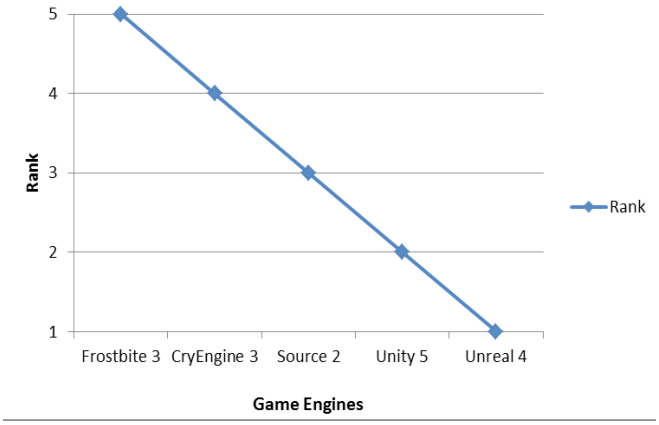


Fig 2. Best Game Engines for Developers as per usability

III. Testing

Impact and efficiency of game engines were tested against the most famous games developed using these game engines. Every game is unique on its own and has set the benchmark for other similar games. The games taken for testing are:

- Cry Engine: Crysis 3
- Source Engine: Counter Strike Global Offensive
- Hero Engine: Star Wars The Old Republic
- Unity Engine: Assassin's Creed
- Unreal Engine: Batman Arkham City
- Vision Engine: Max Payne 3

Most popular and best six games stated above were compared to each other on theirGPU usage and its efficiency to operate high end graphics, their CPU utilization and memory usage. These three testing drive are explained in detail.

For testing, four standard budget GPU's i.e. GE Force GTX 680, Radeon HD 7870, GE Force GTX 670 and Radeon HD 7970 are used. These GPU's were tested on two different dual core CPU's namely, Intel Core i7-3770k and AMD FX-8150. The games were tested at two common desktop display resolutions: 1680x1050, 1920x1200, using normal and very high quality settings respectively.

A. GPU Performance

For Max Payne, attaining the target frame rate of 80fps was arduous at 1680x1050 than expectation. This called for the utilization of either the GeForce GTX 670 Ti or Radeon HD 7870 which averaged 77fps and 70fps respectively. Radeon HD 7870 declined the desired target framerate with 77fps. Graphics cards such as the GE Force GTX 680, GE Force GTX 670 and Radeon HD 7970 managed to achieve the average by 108 fps, 101 fps, and 87 fps respectively. At 1920x1200, it takes the Radeon HD 7870 to surpass an average of 80fps. The Radeon HD 7970 averaged 100fps being consistent with the GeForce GTX 670 which exceeded by 22fps. In the meantime, the GeForce GTX 680 and Radeon HD 7970 both scored 134fps and 87fps as shown in figure 3 and figure 4.

Comparing Star Wars: The Old Republic (SWTOR) to all GPUs shows that, Star Wars: The Old Republic requires a graphics card that is capable of DX 9 or OpenGL 4.1. It is seen that star wars executes best compared to other games. All of the cards in these graphs deliver more than a 90 FPS minimum at 1680x1050 and 80fps at 1920x1080, a playable result across the panel The Radeon cards impede behind similarly-priced GeForce panels.

Testing Crysis 3 at 1680x1050 on medium quality depicted that GTX 680 toped the graph with 79fps, while the GTX 670 trailed closely with 75fps and the HD 7970 GHz edition delivered 70fps, 3fps than the usual7970.At 1920x1200, Radeon HD 7970 or GTX 670 is needed for an average of 60fps, while the GTX 680 is compulsory for an attempt to break the 60fps hurdle. On the contrary HD 7870 failed to reach the average fps.

Starting by the top of the graph, it can be seen that GTX 680 and Radeon 7970 for Batman Arkham City, averages a 36fps. While the GTX 670 R9 290 performs unpersuasively, delivering an average of 30fps. Ardently GTX 680 and RD 7970 tops the chart being the suitable high value performance GPU at 1680x1050.Moving further, in the graph it is observed, the RD 7970 delivered playable performance along with the GTX 680 4.0 GHz Edition. The old GTX 670 also performed quite well, despite dipping down to 45fps at times.

The Assassin's Creeds thrives to match up the intensity, pushing the GTX 680 to its edge at 1680x1050. It is observed that GTX 680's framerate is fully playable in ACU. Performance declined slightly for Radeon 7970 and GTX 670 but the smoothness and efficiency of motion is appreciative even for the scarcely perceptible fps drops for GTX users. At 1920x1200, everything until the GTX 680 perseveres at playable framerate. Certain graphics optimization adjusts on the GTX 670 enables an amalgam high-medium setup with playable framerates.

The limited hardware requirement is one of the robust benefits of Counter Strike: Global Offensive (CS: GO). Even if with Intel's former HD Graphics 3000, multiplayer mode is feasible depending on the resolution at lower or normal settings. Midrange GPU's such as the GeForce GTX 680 deals with higher details and 2 x anti-aliasing qualities with no tribulations for 1680x1050 avoiding the need a high-end GPU at all. These settings are promising with the GeForce GTX 670. For the newest fervent levels of GPUs like the Radeon HD 7970M and the GeForce GTX 670M, Counter Strike does not appear to pretense any confront.

B. CPU Performance

On testing the games on quad and octa-core processors, it is seen, high CPU utilization across quad processors, while those with eight cores spreads the load equally well and evenly depicted on figure Y.

Crysis 3, Core i7-3770K at its default clock frequency of 4.0GHz, impelled the GTX 680 to a mellifluously flowing to 69fps, but on overclocking to 4.5GHz profits more frames and gets a 37% growth of the total of 95fps for GTX panel. Similar fashions is seen with the AMD FX-8150 but since the chip already comes clocked at 4.0GHz, it does not leave much margin. Effortlessly it managed to increase the average frame rate from 80fps to 85fps thus giving better play experience as shown in figure 5.

For beginners, it is much desirable to use a quad-core processor, by preference a latest Core i7 or the FX-8150. Additionally, more likely there is a need of dual GTX 680s or HD 7970 GHz versions in order to play with high settings at 1920x1200 or more. For the classic regular game player, the preeminent graphics card for playing Crysis 3 is the GTX 670, as only one card delivers an average of 70fps even at 1920x1200.

Although Batman: Arkham City is extremely demanding on Virtual RAM. If sufficient VRAM is accessible, the game uses about 4.2GB's at 1080p, 5GB at 1440p and a slightly above 6GB at 4K. This might be the reasons triggering several performance disputes for those who are executing on lower-end hardware. Heavy CPU utilization is witnessed across quad-threaded processors; the Core i7 presented a real astonishment by delivering a consonant performance giving 89 fps which is reasonable. However FX 8150 executed rather poorly on comparison with core i7. It provided highest of 79fps when

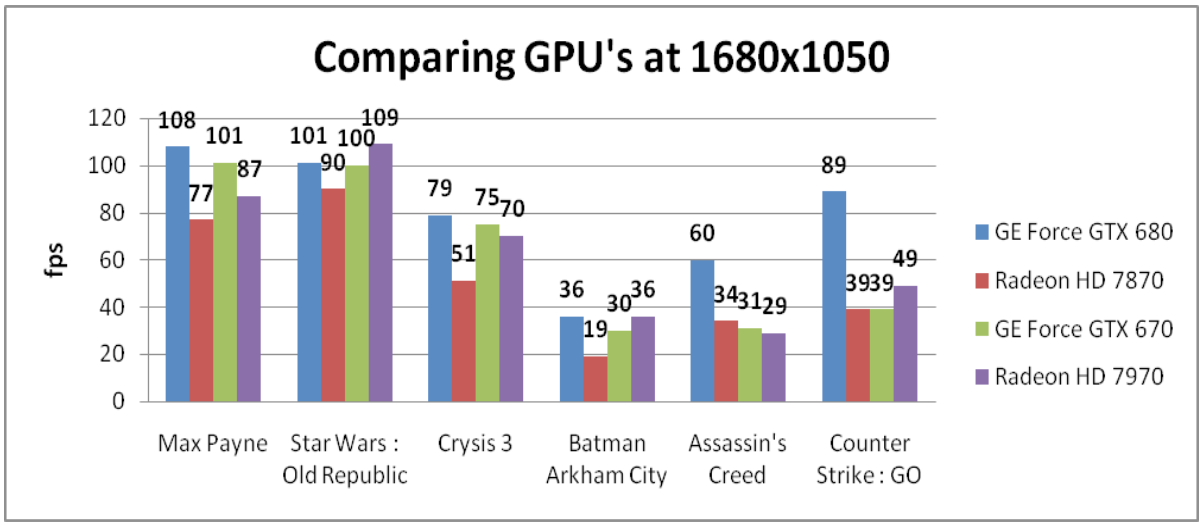


Fig 3. GPU performance at resolution 1680x 1050

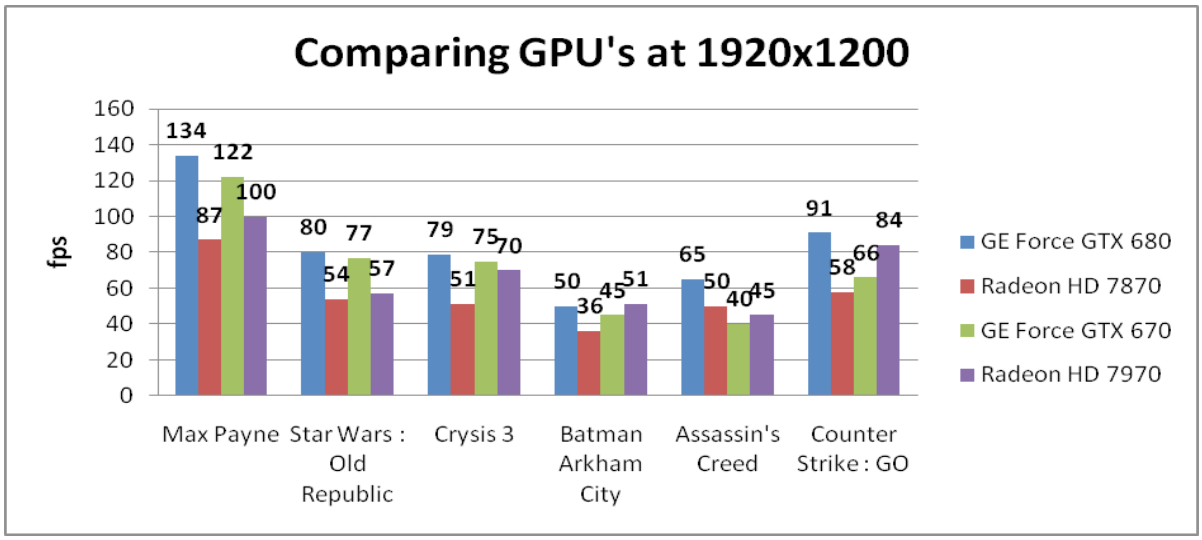


Fig 4. GPU performance at resolution 1920x 1200

overclocked, touching the average of core i7.

Although Batman: Arkham City is extremely demanding on Virtual RAM. If sufficient VRAM is accessible, the game uses about 4.2GB's at 1080p, 5GB at 1440p and a slightly above 6GB at 4K. This might be the reasons triggering several performance disputes for those who are executing on lower-end hardware. Heavy CPU utilization is witnessed across quad-threaded processors; the Core i7 presented a real astonishment by delivering a consonant performance giving 89 fps which is reasonable. However FX 8150 executed rather poorly on comparison with core i7. It provided highest of 79fps when overclocked, touching the average of core i7.

For core i7 in Max Payne, there was a minimal variation in the fps at default clock frequency giving peak of 70fps but this variation rose in the case of FX-8150. Radeon panel gave 75fps providing better play to user. But GTX performed poorly giving only 61 fps. Overall it can be concluded that, the averages appear playable.

Counter Strike: Global Offensive is CPU heavy game, utilizing much of CPU it can. It can reach high fps over 90 and by no means dropping below 70fps on most areas on both the processors. Core i7 gave consistent performance though FX-8150 had consonant but improved performance on the GPU's.

Assassin's Creed on core i7 at default clock frequency, gives 90 fps thus making the game intensely playable. But FX 8150 performed poorly when overclocked, manages to give average of 60fps.

Unfortunately, Star Wars Old Republic foiled to even reach 60 fps, with averages of 58 fps and 59 fps at 1680x1050 and 1920x1200, respectively. The only system that hit the 60 fps average was Intel Core i7with the GTX GeForce 680 builds, which hit 61 fps as lowest. At 4.5GHz, the FX series is nearly capable to meet the Intel Core i7 processors. Core i7 is able to reach the peak on GTX 680 at just 2.5GHz, while the FX-8150 involves its full 5 GHz express ascent clock frequency to attain the equal point.

In the situation of quad-core chips, to enable fast and liquefied multitasking, gleaming 2D/3D gameplay, and super-rapid camera presentation, among other things, each core can be put to work concurrently on a particular job. Current octa-core chips merely have two sets of quad-core processors, which split several tasks between them accordingly. Many a times, the set of cores which are low-powered will be assigned tasks. With the need of advanced tasks, eventually, the set of faster four cores will take effect.

It is always influenced by the power-per-core and its architecture. For example, Intel quad cores are about the same as AMD octa-cores, because of the architecture and the separate cores have more influence.

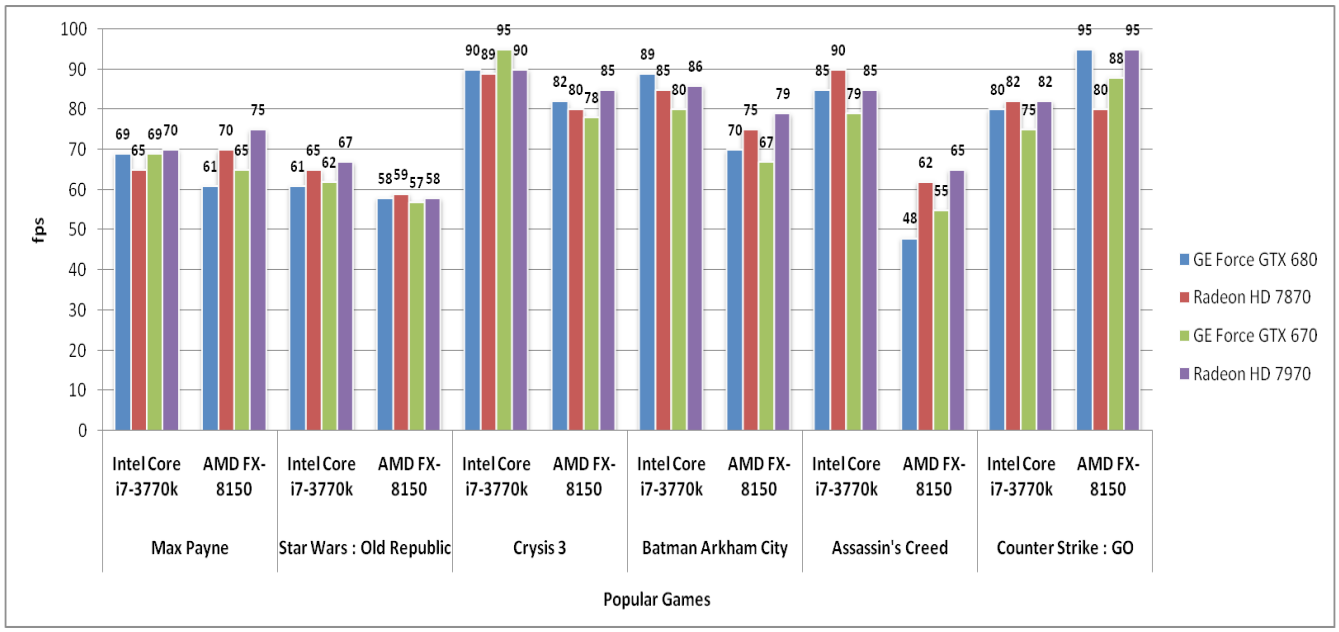


Fig 5. CPU performance of games for four selected GPU's

C. Memory Usage

It was equally essential to check the memory required communally for these games to run smoothly and effectively. On both aspects, minimum required memory and recommended memory availability should be taken into account while playing the games, for the reason that lack of available memory might cause the system to run leisurely while actively playing these games. Table 2 shows the assessment between memory usage by these games.

Overall it can be seen that minimum requirements of all games are satisfactory but Max Payne proves to be memory effective in case of HDD. Along with Max Payne, batman and crysis3 uses 2 GB RAM but both these games are reasonable for HHD required. On the other hand, Assassin's creed proves to be very affordable.

TABLE 2
MEMORY REQUIREMENT BY GAMES

Games	Minimum Requirement	Recommended Memory
CS: GO	1 GB RAM 7.6 GB HDD	2 GB RAM 7.6 GB HDD
Max Payne 3	2 GB RAM 35 GB HDD	3 GB RAM 35 GB HDD
Assassin's Creed	1 GB RAM 8 GB HDD	1 GB RAM 8 GB HDD
SWTOR	1 GB RAM 15 GB HDD	4 GB RAM 15 GB HDD
Batman	2 GB RAM 18 GB HDD	4 GB RAM 18 GB HDD
Crysis 3	2 GB RAM 20 GB HDD	8 GB RAM 20 GB HDD

IV. Empirical Study As Per Views Given By Gamers

Various game devotees and intensive players were asked to give their comments and judgment on these games. A Google form was created and players were asked to submit their opinion for the game based on certain criteria stated below. Table 3 shows the overall comparative analysis of games on the basis of comments sent by them.

It can be seen that, people admires playing Batman Arkham City, Assassin's Creed and Crysis 3. These games have tremendous followers around the world as they offer intensive graphics, high image quality and enmeshing music

V. Role Of GPGPU In Games

GPUs are potent processors dedicated to graphics computation, much faster than CPU on considering all the parallel processors. With the advancements of programmable GPUs has facilitated new potential for general purpose GPU computation (GPGPU) which is used to embellish the altitude of realism in virtual gaming simulations.

Mostly GPGPU accords only with graphics steps of the game loop, while the CPU processes primarily utmost of the game logic. This work diverges from the conventional approach; GPU offers practical implementation of the entire game loop. This is indeed a leap for gaming world, as the CPU's are advancing towards multi-core processors thus imminent games needs related parallelism as in the case of GPU programs.

Contemporary GPUs are fully programmable many-core chips build up roundabouts an array of lateral processors. GPU has scalable array of multithreaded Streaming Multiprocessors (SM), each of which is competent enough to support thousands of inmate parallel hardware threads. Each SM here is a collection of processors. A multiprocessor executes the same instruction on a these assembly of threads at every clock cycle, thus called as warp. To handle such a bulky number of threads, it utilizes an exclusive scheme called SIMT (Single-Instruction, Multiple-Thread). The thread management consisting of creation, scheduling, and synchronization is performed entirely within hardware by the SM with basically zilch cost.

Existing GPUs consist of a huge number of splinter processors with high memory bandwidth. By several form, the current

TABLE 3:
COMPARATIVE EVALUATION OF GAMES

	CRYSIS 3	STAR WARS:OR	Max Payne	Assassin's Creed	Batman Arkham City	Counter Strike: GO
Graphic quality	Stunning and crisp textures but more visually pleasing in 1080P	Blurry at 720p but enjoyable at 1080p	Impressive graphics	Good	Astonishing graphics	Normal graphics
Art Style	Visually rich	Blend of realistic and more animated styles	Gothic but incredible art style	Striking art style with smooth and satisfying movement	Modern Hyperrealism style	Art style is mixed with the real world and replicates real-world finishing well
Story	*Crysis 3*'s revolves around Prophet, a Nanosuit holder who is on a quest to take revenge on the CELL Corporation, and the Alpha Ceph, the leader of the Ceph alien race.	Story builds on an ensemble cast, part and parcel to Star Wars' movie charm	Max investigates on a case related to his family murder and finds himself framed in a murder	Two linked story. One is of an Assassin who want to find the "Apple of Eden", and other is of a person in the present who's experiencing the assassin's memories	The main storyline revolves around Batman's imprisonment in Arkham City later stopping the chaos created by Joker	Each player joins either the Terrorist or Counter-Terrorist team and attempts to complete objectives or eliminate the enemy team.
Pacing of story	Pacing is very slow	Well-paced story, likeable original characters	Game pace keeps moving forward speedily	Well paced story	Insanely tight pacing	Fairly consistent pace
Character emotional development	Convincing enough to have emotional attachment to the character and face decisions	Emotional attachment to your character and face decisions in the game which will guide the development of your character and determine the direction of the character's storyline	Max is relentless and even makes you feel his pain. Good facial expressions and emotions.	Minimal cold emotion	Subtle emotional responses on the characters faces	Minimal emotional involvement
Customization	Weapon and equipment's, multiplayer character customization	UI, Character and companion customization	Character and multiplayer avatars customization	Character, equipment and outfit customization	Customize playing modes	Customization of crosshair, in-game text, view model, the mini map
Ability to modify game with ease	On unlocking suits, players has an advantage over his enemies by using the suit's enhanced strength, stealth capabilities to modify gaming actions	Ability to switch factions with modified base game storyline	Modifications can perform several functions, such as the ability to add new weapons, skills, perspectives, surroundings, and characters	Chest piece reduces the ability of guards to see the hero to a specific percentage. One can add options to equip the assassin's ability	Ability to change predator challenge	Modification can be done on completion of tasks
Gameplay	The gameplay of Crysis 3 is almost exclusively about killing	The social gameplay and challenging, puzzle combat. The project's key focus is to differentiate amid the player's faction and morality	Max Payne 2 is a third-person shooter, in which the player assumes the role of Max Payne	Jacob and Evie Frye, assassin twins who come to London in 1868 to dismantle the Templar conspiracy	Batman is devoted to protect city from crime	Modern-military first-person shooter game
Smooth transition between story and gameplay sequences	Seamlessly transition between story and action sequences by the input controller	Nice flow and has the ability to queue up multiple actions in a sequences	Story sequences bleed directly into gameplay wonderfully telling story between actions	Smooth transition between story and gameplay, splitting the story between both characters	A gap between story and action	Seamless transition
Good Ending/Few loose ends	Prophet, the hero is alive and thinks to live peaceful life by hiding his identification/ glaring gameplay bugs, the bow is very powerful making too easy to play	Story ends in a different place than shown in the movie. The ending is dumb	It has a sad ending, max is left all alone/ Has glitches and other technical problems	The competitive multiplayer component acts as a loose end	Realistic story/ some missions need to be downloaded by internet	Counter-Strike is an entirely multiplayer experience. One can play offline against AI bots
Difficulty of game	Once selected a difficulty level, one can jump to lower difficulty level only The five difficulty levels in order of easiest to hardest are: Recruit, Soldier, Veteran, Super soldier, and Post-Human Warrior	They come in Normal, Difficult, Elite and Boss difficulties. Each level increases the difficulty in space mission	The Max Payne 3 Story can be played at five difficulty settings that can be adjusted on the fly and provide varying levels of challenge: easy, medium, hard, hardcore and old school	Mission difficulty is indicated by diamonds	The ability to change the difficulty settings mid-game.	Hardest version of CS series, require the most skill out of all the games
Voice Acting quality	Amazing and A high-pitched keening sound	Ambient sound effects	Incredible visuals and sound to go along with actions	Inconsistent voice acting	Acceptable quality	Finest quality
Music soundtrack	Light melodic music playing constantly in the background is great	Incredibly successful soundtrack	Bold and majestic	Best music, extremely enjoyable	Dark and powerful instrumental enhances the action and emotion of the gaming experience	Energetic and cinematic
Background noises	Background noises enhances the story	Constant buzzing noise	The music fits nicely with the game but not overwhelming	Background is louder than the cut scene dialogues	Ambient sounds	Satisfying but heavy noises
Image Quality	The very high quality image is stunning, the textures are impressive, lighting is great	Image quality improves by running the game at high resolutions	HD graphics and high resolution textures	Better but could have been best	High resolution images	Realistic and highly crisp graphical images

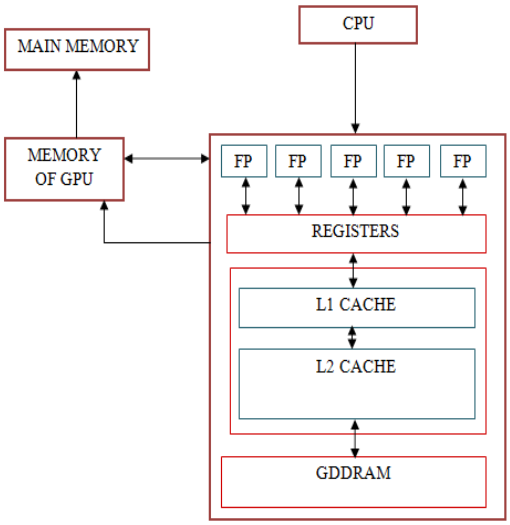


Fig. 6. Working flow of GPGPU

GPU architecture is alike to a many-core processor, which accomplishes higher performance on parallel code which is on the contrary with multi-core CPUs, containing finest single-thread performing cores. GPUs are chiefly upgraded for 2D arrays. Caches inherited on the GPU are shared in a great extent by fragment processors (FP).

GPU kernel code is also loaded on CPU. For superior performance, data movement from main memory to the GPU should be averted. On completing the computation, GPU copies the data back to the CPU memory. All SM is coupled with a personal L1 Data Cache and stable caches along with a shared scratchpad memory as shown in figure 6.

A. Improving Game Engine Performance in Terms of Task and Data Level Parallelism

A contemporary game engine is usually encompasses user input, game logic, artificial intelligence (AI), physics, audio and graphics. A renderer is needed for 2D/ 3D graphics.

Lately popularized CUDA/GPGPU computing possesses the capability to amplify the acceleration persistently [15]. Though the GPU was produced to quickly maneuver and modify memory to expedite the formation of images in a frame buffer aimed for output display, GPGPU computing is now bestowing excellent performance by unloading exhaustive computation component of the application, whereas the remaining code executes on the CPU.

Games are innately serialized making multithreaded application intricate to run. The initial attempt made in the game engine was to simultaneously run the client and the server on the individual cores respectively by means of coarse grained threading mechanism. Thus, GPGPU has now become a vital component of today's conventional computing systems. In [11], four GPGPU triumphs in game physics and computational AI succeed above the optimized CPU applications. Since games have become progressively restricted by CPU performance, divesting intricate CPU jobs to the GPGPU concedes improved overall performance.

Task parallelism is employed in a game engine by executing each module task in its individual thread [12-14]. Graphics renderer and physics model are good for laterally analogous because of their processor-severe tasks. In a game engine, data parallelism allows similar category of data in a module to run parallel in several threads. This scheme balances fine for a huge figure of processors because of the reason that, the magnitude of the data running on each thread is partitioned uniformly. Communication overhead can be reduced by grouping the application objects that might interact within the same thread.

B. Improving Game Engine Performance in Terms of Physics and AI

GPU currently offer hardware acceleration for physics transformation, sustaining amassed general computation. This conception is recognized as General Purpose processing on Graphics Processing Unit. Latest AMD and NVIDIA graphics cards enables base for rigid body dynamics computations. When such style of hardware acceleration is unavailable, Physics computations are implemented in software which causes inferior CPU-confined performance. Development of PhysX engine articulates on the GPGPU competence of modern GPUs. GPU's are the best room to work out physics calculations for the reason that, physics execution is propelled by thousands of concurrent calculations and execution. These days, NVIDIA's GPUs, is having as many as 480 cores, complementarily profits PhysX engine. NVIDIA is dedicated to create the gaming experience and gameplay exhilarating, lively, and vibrant [27].

The subsequent huge obsession for GPGPU processing is AI [30], to be confronted in games. In 2005, an Israeli company called AIseek proclaimed a devoted processor called the Intia, which was created to speed up various AI characteristics. This makes AI to be an element of a customary PC, thus making AI a suitable preference in view of GPGPU technology. nearly every player has a graphics card compatible with Nvidia's CUDA or AMD's Stream technology, it may possibly take even a little of the total load from the CPU while processing recurring AI tasks and features. Even a minute segment of a GPU's resources enhances AI processing power for several jobs. Possibly OpenCL or DirectX 11's Compute Shader, might present a method to speed up AI on extensive GPU array. Moreover, it seems as AI possibly can be a chief progression in GPGPU processing.

Concerning the cutting edge cloud computing, most of the computations are eventually carried out by multi-core CPUs and many-core GPUs, despite of their allocation. GPGPU cloud computing [16-21, 24] is more broadly applicable, offering general purpose computing ability in the manner of on-demand virtual resources. In GPGPU cloud computing, the processing strength of physical resources is divided by virtual resources and summoned obliquely. The whole GPGPU cloud system is separated into three layers with this analysis articulating on the task characteristics, hardware features, scheduling mechanism, and execution mechanism of each layer [cloud].

Disparity in the architectures between CPU and GPU specifies that in order to attain superior performance, code for the GPU must be optimized in a different way. GPU is dedicated for extremely parallel computation and thus premeditated to include more dedicated transistors for data processing more willingly than caching and flow control. More explicitly, the GPU is particularly compatible to deal with tribulations that are articulated as data-parallel computations with intense arithmetic computation. At this point, GPGPU is yet to be an established technology. In the next few years, graphic processors will turn out to be a superior technology, best suited to sustain the cascaded processed applications.

VI. From Advancements To The Future Of Gaming Technology

Gaming technology has been booming since its exhumed roots of tic-tac-toe, brick-games, snakes and ladders, and chess. In leaps and bounds, digital games have gone from novelties that performed simple tasks, to intricate code that facilitates immersive gameplay experiences. On exploring achievements in gaming world, AI, portability, ubiquitous, open web, 3D, voice and motion controls, one

can envision how lifelike and experiential video games will get the breakthrough, as the technology progresses.

With enormous investment in designing the next big gaming innovation, the potential of what the futurity of games will bring are infinite.

Game developers have always been adept at combining cutting edge and existing technologies. For instance, movement control gaming achieved a mainstream audience when it amalgamated with the mobile world. And open-world games will be played without any skills to involve players collectively from all over the world by global network. A best mechanism for predicting the future of gaming is to review what is already present and by visualizing the how it could evolve effectively. Just by taking 3D image or augmented reality a step further and then can experience and get pleasure from hologram gaming, an edge that does not give the impression of having any covenance and regulations.

Video games has emanated from a long way ever since they overlapped into the conventional in the 1980s, but some incredible improvements in the technologies made the future of gaming even dazzling. 3D scanning and facial recognition technology has permitted developers to actually develop their likeness in the game or to ingeniously transfer their expressions to additional digital creations. For example, Intel's RealSense 3D camera allows gaming developers to produce games adapting the emotions of the gamer by scanning different locations of a player's face.

Voice controlled gaming has remained all over the place for a while, but the prospective of using this technology in gaming systems has finally captivated to realism, now even computers are capable to effortlessly diagnose voice commands from the player. Player can now turn the game console on and off, supervise gameplay, interact on social game group, all by commanding to your gaming system. Using a 3D camera that trails distinct area of your body, gesture control consents players to connect to their gaming experience by means of the usual body movements. State of the art advancements deliver to players a rich and high quality experience in form of fully rendered worlds with photo realistic textures.

Whether it's smart watches or goggles, wearable games make gaming mobile without being too intrusive. Wearable games are extensions of the body in addition to extensions of the consoles used while playing. With the dawn of smartphones, the entire gaming experience is now in the hand of the players. Instead of developing gaming structures necessitating more influential hardware, developers have lessened the load with the use of cloud computing. By usage of clouds, game streaming is now a reality, streaming similar to movie streaming.

However many Virtual Reality (VR) gaming consoles are not released commercially yet, but this developing VR technology is dignified to concede gamers a wholly immersive gaming experience. VR are sometimes stated as immersive multimedia or computer-simulated reality duplicates environs by simulating a corporeal existence at places in the physical-world or a fantasy world, letting the players to communicate with that world within the simulated environment. Virtual Reality is imminent. It's no overstatement to state that VR is one of the gigantic advancements in gaming world and eradicating the blockade in the midst of player and game world.

Augmented reality gaming starts to take hold through smartphones and glasses technology from the likes of Google. Till 2025, Virtual reality simulations embarks on to turn out to be feasible in the home market, and unlike the augmented reality techniques that have been in favor for the last ten years these will begin to feature fully immersive systems. These experiences will still experience to some extent disconnected from realism though, as control mechanisms will require various forms of static activity. Graphics and physics technology by this stage will be able to perfectly mimic real-world resolutions and depth, with added emergent behavioral algorithms allowing sandbox gaming and interaction with characters and environments that promote personally crafted narrative experiences.

VII. Conclusion

From the commencement of game engines to the latest 3D jazzed-up game engines, the objective of development was to endure the equivalent i.e. giving game coders a rostrum for creating their unique games into reality. These engines offer the rudimentary central design with codes and the manifestation as a middleware. The progression of gaming engines is at the present are proceeding en route for supplementary realistic and technically sound games in innumerable grounds like physics, sounds, AI, graphics, and animations etc. from the analysis it can be stated that, Max Payne 3 is rich to deliver high and crisp graphics but is costly in terms of memory whereas Crysis 3 bestows better CPU utilization. Max Payne 3 has high level of emotional intelligence compared to other games, engaging the player's in the gameplay.

Future machines will be progressively more assorted. Individual processor chips will probably enclose processing component with diverse functioning features, memory hierarchy, and levels of physical concurrency. Till 2018, it can anticipated that GPU's will no longer be an external driving force of CPU; rather, both CPU's and GPU's will be incorporated by the same dye through a unified memory architecture..

Acknowledgment

We wish to thank all the students, game player's and game developers around the country who actively contributed in this comparative analysis.

References

[1] Julie Wilmore, "Dissecting the Video Game Engine and a Brief History", Retrieved From: juliewilmore.files.wordpress.com/2010/07/tc339finalpaper.pdf.

[2] Kuan-Ta Chen et al, "On the Quality of Service of Cloud Gaming Systems" Published in IEEE Transactions on Multimedia, Volume: 16, Issue: 2, 19 November 2013, 480 – 495.

[3] Kowit Rapeepisarn et al, "A Comparative Study of Digital Game Platforms for Educational Purposes", Published in Ninth Postgraduate Electrical Engineering and Computing Symposium, 4th November 2008, Perth, W.A. pp. 37-40 Chen, *Linear Networks and Systems* (Book style). Belmont, CA: Wadsworth, 1993, pp. 123–135.

[4] Michael Jarschel, et al., "Gaming in the clouds: QoE and the users perspective", Published in Mathematical and Computer Modelling, Volume 57, Issues 11–12, June 2013, Pages 2883–2894.

[5] Rahma Gharsallaoui et al, "A comparative Study on Cloud Gaming Platforms", Published in 7th International Conference on Control and Automation, 20-23 Dec. 2014, pp 28 – 32.

[6] Chun-Ying Huang et al, "GamingAnywhere: An Open Cloud Gaming System", Published in MMSys Proceedings of the 4th ACM Multimedia Systems Conference, Pages 36-47, pp 36-41

[7] Ryan Shea et al, "Cloud Gaming: Architecture and Performance" Published in IEEE Network, Volume 27, Issue 4, July-August 2013, 16-21.

[8] Georgios N. Yannakakis, John Hallam, Henrik Hautop Lund, "Comparative Fun Analysis in the Innovative Playware Game Platform", Published in Proceedings of the 1st World Conference for Fun 'n Games 2006, pp. 64-70.

[9] M. Joselli and E. Clua, "Gpu wars: Design and implementation of a GPGPU game," Published in Proceedings of the 8th Brazilian Symposium on Games and Digital Entertainment, Brazil, October, 8th-10th 2009.

[10] Richard M. Ryan • C. Scott Rigby • Andrew Przybylski, "The Motivational Pull of Video Games: A Self-Determination Theory Approach", Published in Springer Journal of Motivation and Emotion, December 2006, Volume

30, Issue 4, pp 344-360.

[11] Owens, J.D., Houston, M., Luebke, D., Green, S., Stone, J.E. & Phillips, J.C. "GPU Computing: Graphics Processing Units-Powerful, Programmable and Highly Parallel-Are Increasingly Targeting General -Purpose Computing Applications", Published in Proceedings of the IEEE, pp. 879-899, 2008.

[12] Ooste, J.V., 3D Game Engine Programming: Helping You Build Your Dream Game Engine, 2011, Published online at http://3dgep.com/?p=1821,1 August 2013.

[13] Guevara, M., Gregg, C., Hazelwood, K., & Skadron, K., "Enabling Task Parallelism in the CUDA Scheduler", Published in IEEE PEMA. 2009.

[14] Baumstark, L.Jr. & Wills, L., Exposing Data - Level Parallelism in Sequential Image Processing Algorithms, Published in Proceedings of the Ninth Working Conference on Reverse Engineering (WCRE'02), IEEE, 2002.

[15] Designing the Framework of a Parallel Game Engine (PGE), Intel, Published online: http://www.intel.com (1 August 2013).

[16] Liang Hu, Xilong Che and Zhenzhen Xie, "GPGPU Cloud: A Paradigm for General Purpose Computing", Published in Journal of Tsinghua Science And Technology, ISSN:1007-0214l l03/12, pp22-33, Volume 18, Number 1, February 2013.

[17] J. Nickolls and W. J. Dally, "The GPU computing era", Published in IEEE Micro., vol. 30, no. 2, pp. 56-69, March-April 2010.

[18] L. Shi, H. Chen, J. Sun, and K. Li, "CUDA: GPU-accelerated high-performance computing in virtual machines", Published in IEEE Transactions on Computers, vol. 61, no. 6, pp. 804-816, June 2012.

[19] V. Gupta, A. Gavrilovska, K. Schwan, H. Kharche, N. Tolia, V. Talwar, and P. Ranganathan, "GViM: GPU-accelerated virtual machines", Published in Proceedings of the 3rd ACM Workshop on System-Level Virtualization for High Performance Computing (HPCVirt 09), Nuremberg, Germany, March 31, 2009, pp. 17-24.

[20] C. Yang, C. Huang, C. Lin, and T. Chang, "Hybrid parallel programming on GPU clusters", Published in Proceedings of the IEEE International Symposium on Parallel and Distributed Processing with Applications (ISPA 2010), Taipei, China, September 6-9, 2010, pp. 142-147.

[21] V. V. Kindratenko, J. J. Enos, G. Shi, M. T. Showerman, G. W. Arnold, J. E. Stone, J. C. Phillips, and W. Hwu, "GPU clusters for high-performance computing", published in Proceedings of the 2009 IEEE International Conference on Cluster Computing (CLUSTER 2009), New Orleans, Louisiana, USA, August 31-September 4, 2009, pp. 1-8.

[22] John D. Owens et al., "GPU Computing", Published in Proceedings of the IEEE, Volume: 96, No. 5, May 2008.

[23] William Blewitt, Gary Ushaw, and Graham Morgan, "Applicability of GPGPU Computing to Real-Time AI Solutions in Games", Published in IEEE Transactions on Computational Intelligence and AI in Games, April 11, 2013.

[24] R. Shea, Simon Fraser, "On GPU pass-through performance for cloud gaming: Experiments and analysis", Published in IEEE 12th Annual Workshop on Network and Systems Support for Games (NetGames), pp. 1 - 6 , Denver, CO, USA, 9-10 Dec. 2013.

[25] Mark Joselli et al., "A new physics engine with automatic process distribution between CPU-GPU", Published in Proceeding Sandbox '08 ACM SIGGRAPH symposium on Video games, pp.149-156.

[26] M. Mittring and B. Dudash, "The technology behind the directx 11 unreal engine "samaritan" demo," Published in Proceedings of Game Developers Conference 2011.

[27] AIseek, "AIseek - intelligence for new worlds," White Paper, 2006

[28] E. Lindholm, J. Nickolls, S. Oberman, and J. Montrym, "Nvidia tesla: A unified graphics and computing architecture," Published in IEEE Micro, vol. 28, no. 2, pp. 39–55, 2008.

[29] NVIDIA, "Nvidia's next generation cuda compute architecture: Kepler gk110," White Paper, 2012

[30] C. J. Darken, E. R. Pursel, and J. S. Correia, "AI on the GPU," Published in ACM Workshop on General Purpose Computing on Graphics Processors, 2004.

[31] J. D. Owens, D. Luebke, N. Govindaraju, M. Harris, J. Kr¨uger, A. E. Lefohn, and T. J. Purcell, "A survey of general-purpose computation on graphics hardware," Published in Journal of Computer graphics forum, Blackwell Publishing Ltd, pp. 80-113, 2007.

[32] D. Kirk, "Nvidia CUDA software and GPU parallel computing architecture," Published in Proceedings of the ACM 6th International Symposium on Memory Management, Montreal, Canada — October 21 - 22, 2007.

[33] G. Gunasekaran, "Study of performance for the CPU and GPU architecture," Published in International Journal of Research in IT, Management and Engineering, vol. 2, no. 2, pp. 33–44, February 2012.

[34] X. Zhang, B. Wang, and C. Geng, "Gpu-based background generation method," Published in Proceedings of the IET International Communication Conference on Wireless Mobile and Computing (CCWMC 2011), 2011.

[35]] C. Gregg and K. Hazelwood, "Where is the data? Why you cannot debate CPU vs. GPU performance without the answer," Published n Proceedings of the IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS '11), Washington, DC, USA, pp. 134–144, 2011.

[36] Raúl Lara-Cabrera, Mariela Nogueira-Collazo, Carlos Cotta, Antonio J. Fernández-Leiva, "Procedural Content Generation for Real-Time Strategy Games", Published in Special Issue on Digital Economy, International journal of Interactive Multimedia and Artificial Intelligence, Vol. 3, No 2, 2014.

[37] Arranz, Á., and M. Alvar, "GPGPU Implementation of a Genetic Algorithm for Stereo Refinement", Published in Special Issue on Digital Economy in International Journal of Interactive Multimedia and Artificial Intelligence, ISSN 1989-1660, pp. 69-76, March 2015.

[38] Holman Bolívar Baron, Sandra Castillo Salinas ; Rubén Gonzalez Crespo, "An approach to assessment of video game-based learning using structural equation model", Published in IEEE 9th Iberian Conference on Information Systems and Technologies (CISTI), Barcelona, 18-21 June 2014.

[39] Lara-Cabrera, R., M. Nogueira-Collazo, C. Cotta, and A. J. Fernández-Leiva, "Procedural Content Generation for Real-Time Strategy Games", Published in Special Issue on Digital Economy in International Journal of Interactive Multimedia and Artificial Intelligence, ISSN 1989-1660, pp 40-48, March 2015.

Ms. Prerna Mishra has completed her M. Tech degree in Computer Science & Engg from G.H. Raisoni College of Engineering, Nagpur. She is presently working as Assistant Prof in CSE Dept, J. D. College of Engineering, Nagpur, India. Her area of interest includes Assistive Technology and Human Computer Interaction. She is member of various professional bodies like IEEE, IAENG, UACEE, IACSIT, CSI, ICST, ISTE.

Dr. Urmila Shrawankar completed her Ph.D. degree in Computer Science and Engg from SGB Amravati University and M. Tech. degree in Computer Science & Engg from RTM Nagpur University. She is presently working as Associates Prof. in CSE Dept, G H Raisoni College of Engineering, Nagpur, India. Her area of interest includes Advanced Operating Systems, High performance Computing, Distributed and Parallel computing, Cloud Computing, Algorithms, Assistive Technology etc. She has many research publications in refereed international journal and conferences in her account with good number of citations. She is a receiptant of DST Travel Grant from Govt. of India, Ministry of Science & Technology for attending Springer - International Conference at Manchester, UK, Year 2010 and UGC Minor Project Grant. She is a reviewer of many refereed Inter-National journals and has worked as a core organizing committee member, Technical program committee member and Session Chair at many international conferences. She is Member of various Professional bodies like IEEE, ACM, CSI, ISTE, IE. Under her guidance 22 B.E (CSE) project groups, 60 M. Tech (CSE) project scholars and 03 PhD Research scholars are working and completed their projects.

Correlation Between Coupling Metrics Values and Number of Classes in Multimedia Java Projects: A Case Study

Mr. V. S. Bidve[1], Dr. P. Sarasu[2]

[1]Ph.D. Scholar, [2]Director R & D, Veltech Dr. RR & Dr. SR Technical University, Avadi, Chennai, India

***Abstract* — Coupling is an interdependence relationship between the modules of object-oriented software. It is a property with the most influence on quality attributes of the object-oriented software. Coupling with high values results in complex software design hence software professionals try to keep the coupling as low as possible. The values of coupling metrics are dependent on the type of input source code. Reusability is the main feature of object-oriented languages, so coupling occurs due to reuse of code modules. This paper investigates a correlation between the values of coupling metrics and the number of classes in the multimedia Java code. Here, a case study of a banking multimedia Java project with its forty different versions is conducted to comments on this correlation. The analysis of the results shows that, if the input source code is with a large number of classes then it results in high coupling values.**

***Keywords* — Coupling, Object-oriented, Metrics, Classes, Software.**

I. Introduction

Software coupling is an important metrics from the quality point of view. Software professionals need to evaluate continuously and study relations and correlations of all attributes and factors that can affect their developed software products [1]. Software metrics values are the indicators of one or more software quality attributes. Coupling metrics values are also the indicators different quality attributes like reliability, efficiency, integrity, maintainability, flexibility, portability, reusability, and interoperability, etc. [2], [3], [4].

In this paper, we focused on studying the correlation between coupling metrics values and the total number of classes in multimedia Java projects. The focus of this study is to monitor the behavior of coupling values with raise in numbers of classes of multimedia Java projects. The study considers banking software named *Cyclos* with its forty different versions downloaded from the open source repository sourceforge.net. The versions of *Cyclos* taken into account for this study possess the different number of classes and coupling values.

The rest of the paper is organized as follows. In section 2 data collection for the case study is given. Section 3 provides detailed of results and analysis. The correlation analysis between coupling metrics values and number of classes in input multimedia Java code is provided in section 4. Section 5 concludes the paper.

II. Data Collection for the Case Study

This section gives details of input multimedia Java projects, selection of coupling metrics, coupling data collection procedure used in the work.

A. Input multimedia Java projects

This study considers the projects developed in multimedia Java programming language. The requirement of the input multimedia Java projects is a folder containing .java files. The multimedia Java project folders with .exe or .jar files are not useful for this study. This study referred the open-source repository sourceforge.net to download the multimedia Java projects. The sourceforge.net repository has ten different category projects. This study selected the projects only from the *Banking* category. The *Banking* project *Cyclos* with its **forty different versions** is considered for this case study. *Cyclos* offers a complete on-line banking system with additional modules such as e-commerce and communication tools. *Cyclos* is used for Microfinance institutions, local banks (in developing countries) and complementary currency systems like LETS, Barter networks, and Timebanks. The detailed listing of forty versions of *Cyclos* is given in Table II.

TABLE I
COUPLING METRICS

Sr. No.	Coupling Metrics	Type of interactions	Coupling Mechanisms	Source
1	Parameter coupling	Method-Method,	The method of one class invokes method/passes parameter/passes message to methods of another class or to make a call to the constructor of another class.	[5]
2	Inheritance coupling	Class-Class	One class is a superclass of another class (Inheritance).	[5, 6, 7, 8]
3	Global coupling	Method-Attribute/Class-Attribute (Friend)	The method of one class can directly access parts of the internal structure, of another class method (friend). Also to access common, shared, non-local variables of another class.	[5]
4	Data Abstraction Coupling	Class-Method/Class-Attribute	One class is used in the implementation of methods of another class. One class is the domain of the instance variable, the local variable of another class.	[8]
5	Import Coupling	ALL	All type of coupling due to any import mechanism.	[9]
6	Export Coupling	ALL	All type of coupling due to an export mechanism.	[9]
7	External coupling	Sharing of global devices.	Sharing an external device like the printer, HDD, external file by the two classes.	[9]

DOI: 10.9781/ijimai.2016.4114

TABLE II
COUPLING VALUES OBTAINED FROM JCMT

Project Name	Number of Classes in Project	Parameter coupling	Inheritance Coupling	Global Coupling	Data Abstraction Coupling	Import Coupling	Export Coupling	External Coupling
cyclos_3.0	1095	6884	1902	10	4004	4684	4680	22
cyclos_3.0.1	1096	6888	1904	10	4007	4688	4684	22
cyclos_3.0.2	1095	6862	1904	10	3995	4665	4661	22
cyclos_3.0.3	1097	6916	1907	10	4011	4684	4680	22
cyclos_3.0.4	1098	6937	1907	10	4018	4706	4702	22
cyclos_3.0.5	1098	6937	1907	10	4018	4706	4702	22
cyclos_3.0.6	1098	6950	1907	10	4022	4716	4712	22
cyclos_3.0.7	1098	7004	1907	10	4024	4732	4728	22
cyclos_3.0.8	1098	7008	1909	10	4029	4741	4737	22
cyclos_3.0.9	1099	7026	1909	10	4032	4758	4754	22
cyclos_3.0.10	1098	7027	1909	10	4031	4759	4755	22
cyclos_3.0.11	1098	7029	1909	10	4032	4760	4756	22
cyclos_3.0.B1	990	5624	1716	10	3526	3853	3849	21
cyclos_3.0.B2	1040	6254	1799	10	3748	4285	4281	21
cyclos_3.5	1809	14438	3232	77	7442	9946	9953	28
cyclos_3.5.1	1809	14449	3232	77	7449	9959	9966	28
cyclos_3.5.2	1812	14485	3238	79	7475	9985	9992	28
cyclos_3.5.3	1812	14546	3238	80	7486	10019	10026	28
cyclos_3.5.4	1819	14608	3253	81	7524	10079	10086	28
cyclos_3.5.5	1820	14631	3255	81	7531	10091	10098	28
cyclos_3.5.6	1821	14633	3256	81	7531	10093	10100	28
cyclos_3.5.beta1	1698	12975	3010	62	6886	8914	8925	29
cyclos_3.5.beta2	1699	13065	3011	62	6906	8992	9003	29
cyclos_3.5.beta3	1750	13531	3113	63	7105	9368	9379	29
cyclos_3.5.beta4	1756	13701	3127	65	7154	9461	9472	29
cyclos_3.5.RC1	1755	13726	3126	65	7162	9474	9485	29
cyclos_3.5.RC1a	1754	13725	3126	65	7161	9473	9484	29
cyclos_3.5.RC2	1768	14111	3154	67	7303	9669	9678	29
cyclos_3.6	2357	16615	4459	109	8937	11544	11573	30
cyclos_3.6.1	2367	16711	4466	111	9026	11637	11666	29
cyclos_3.6.beta1	1936	15714	3487	106	8107	10889	10897	28
cyclos_3.6.beta2	1935	15718	3485	106	8110	10911	10919	28
cyclos_3.6.beta3	1936	15617	3491	107	8069	10840	10848	28
cyclos_3.6.RC1	2035	16978	3659	110	8558	11873	11899	29
cyclos_3.6.RC2	2357	16615	4460	109	8944	11561	11590	30
cyclos_3.7	2435	17275	4602	123	9552	12018	12062	30
cyclos_3.7.1	2438	17320	4606	125	9569	12057	12101	30
cyclos_3.7.2	2443	17405	4615	125	9600	12113	12157	30
cyclos_3.7.3	2443	17403	4617	125	9602	12109	12153	30
cyclos_3.7_RC1	2431	17253	4599	123	9537	11994	12038	30

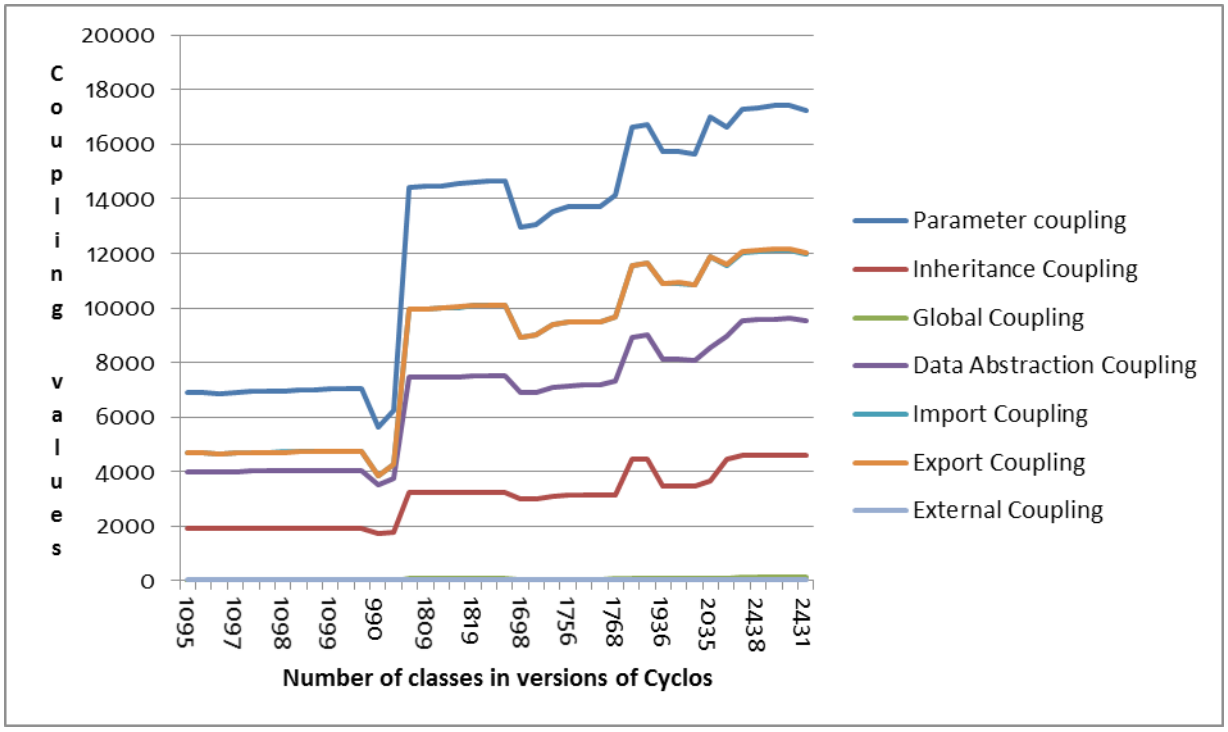


Fig. 1 Behavior of coupling values of forty versions of Cyclos

B. *Selection of coupling metrics*

So far, there are many coupling metrics described in the literature by various authors. Every author has a diverse set of coupling metrics focusing on different aspects of coupling measurement. In this study, we considered seven coupling metrics as given in Table I. The metrics selected in this study are comprehensive and considers all aspects of coupling measurement. Every coupling metric chosen for the study have the entirely different type of coupling interactions and mechanisms. The coupling metrics considered is a part of our previous study.

C. *Coupling data collection*

The authors have developed a specific tool named Java Coupling Measurement Tool (JCMT) to compute coupling between all pairs of classes of the input multimedia Java source code. The multimedia Java project folder possessing .java files is an input required to the JCMT tool. The JCMT calculates the seven types of couplings described in Table I.

III. Results and Analysis

In this study, forty different versions of a Banking project Cyclos are evaluated using JCMT to compute values of coupling metrics described in Table I. The sum of the coupling values of all classes belonging to a project is calculated for each coupling metrics by JCMT. The coupling values of forty versions of banking software Cyclos are computed using JCMT and specified in Table II.

The behavior of the coupling values of all forty versions of Cyclos project is plotted using a graph shown in figure 1.

In figure 1, X-axis represents the number of classes and Y-axis represents the coupling values of the different versions of Cyclos project. The curves of import and export couplings are overlapped in figure 1 because of the minor difference in their coupling values. Similarly, the curves of global and external couplings are also overlapped and are near to X axis because of their smaller coupling values. Primarily from figure 1, it is observed that every coupling metrics has a different range of values but a geometrically similar behavior. Secondly, it seems that the behavior of coupling values is changing as per the number of classes of a project. If the numbers of classes in project increases then coupling metrics values are also increases.

IV. Correlation Analysis

The correlation analysis aims to determine the relationship between the number of classes and coupling metrics values of a project. The goal of this correlation analysis is to answer the following questions.

Q1. Is there a correlation between the number of classes and coupling metrics values of a project?

Q2. Which correlation exists between the number of classes and coupling metrics values of a project?

To answer the above questions, data of Table II is sorted as per the ascending order of the number of classes in the projects.

The graph is plotted for the values of Table III and presented in figure 2. In figure 2 X-axis represents the number of classes in the project and y-axis represents coupling metrics values of the Cyclos project.

From the figure 2 following observations are made,

1. **As the number of classes in the project increases the coupling values also increases.** For e.g. in the version cyclos_3.0.B1, the numbers of classes are 990 means the lowest number of classes, so the project has lowest coupling values. Where, in subsequent projects numbers of classes are increased, so coupling values are also increased. There are some exceptions to this rule, in cyclos_3.0.9 the total classes are increased by one than its earlier version cyclos_3.0.11 in Table II, but the coupling values of cyclos_3.0.9 are smaller or same as compared to its earlier version cyclos_3.0.11. These exceptions happen when there is very small rise in the total number of classes i.e. rise of one to five classes or no rise. If

TABLE III
DATA OF TABLE 2 SORTED AS PER ASCENDING ORDER OF NUMBER OF CLASSES

Project Name	Number of Classes in project	Parameter coupling	Inheritance Coupling	Global Coupling	Data Abstraction Coupling	Import Coupling	Export Coupling	External Coupling
cyclos_3.0.B1	990	5624	1716	10	3526	3853	3849	21
cyclos_3.0.B2	1040	6254	1799	10	3748	4285	4281	21
cyclos_3.0	1095	6884	1902	10	4004	4684	4680	22
cyclos_3.0.2	1095	6862	1904	10	3995	4665	4661	22
cyclos_3.0.1	1096	6888	1904	10	4007	4688	4684	22
cyclos_3.0.3	1097	6916	1907	10	4011	4684	4680	22
cyclos_3.0.4	1098	6937	1907	10	4018	4706	4702	22
cyclos_3.0.5	1098	6937	1907	10	4018	4706	4702	22
cyclos_3.0.6	1098	6950	1907	10	4022	4716	4712	22
cyclos_3.0.7	1098	7004	1907	10	4024	4732	4728	22
cyclos_3.0.8	1098	7008	1909	10	4029	4741	4737	22
cyclos_3.0.10	1098	7027	1909	10	4031	4759	4755	22
cyclos_3.0.11	1098	7029	1909	10	4032	4760	4756	22
cyclos_3.0.9	1099	7026	1909	10	4032	4758	4754	22
cyclos_3.5.beta1	1698	12975	3010	62	6886	8914	8925	29
cyclos_3.5.beta2	1699	13065	3011	62	6906	8992	9003	29
cyclos_3.5.beta3	1750	13531	3113	63	7105	9368	9379	29
cyclos_3.5.RC1a	1754	13725	3126	65	7161	9473	9484	29
cyclos_3.5.RC1	1755	13726	3126	65	7162	9474	9485	29
cyclos_3.5.beta4	1756	13701	3127	65	7154	9461	9472	29
cyclos_3.5.RC2	1768	14111	3154	67	7303	9669	9678	29
cyclos_3.5	1809	14438	3232	77	7442	9946	9953	28
cyclos_3.5.1	1809	14449	3232	77	7449	9959	9966	28
cyclos_3.5.2	1812	14485	3238	79	7475	9985	9992	28
cyclos_3.5.3	1812	14546	3238	80	7486	10019	10026	28
cyclos_3.5.4	1819	14608	3253	81	7524	10079	10086	28
cyclos_3.5.5	1820	14631	3255	81	7531	10091	10098	28
cyclos_3.5.6	1821	14633	3256	81	7531	10093	10100	28
cyclos_3.6.beta2	1935	15718	3485	106	8110	10911	10919	28
cyclos_3.6.beta1	1936	15714	3487	106	8107	10889	10897	28
cyclos_3.6.beta3	1936	15617	3491	107	8069	10840	10848	28
cyclos_3.6.RC1	2035	16978	3659	110	8558	11873	11899	29
cyclos_3.6	2357	16615	4459	109	8937	11544	11573	30
cyclos_3.6.RC2	2357	16615	4460	109	8944	11561	11590	30
cyclos_3.6.1	2367	16711	4466	111	9026	11637	11666	29
cyclos_3.7_RC1	2431	17253	4599	123	9537	11994	12038	30
cyclos_3.7	2435	17275	4602	123	9552	12018	12062	30
cyclos_3.7.1	2438	17320	4606	125	9569	12057	12101	30
cyclos_3.7.2	2443	17405	4615	125	9600	12113	12157	30
cyclos_3.7.3	2443	17403	4617	125	9602	12109	12153	30

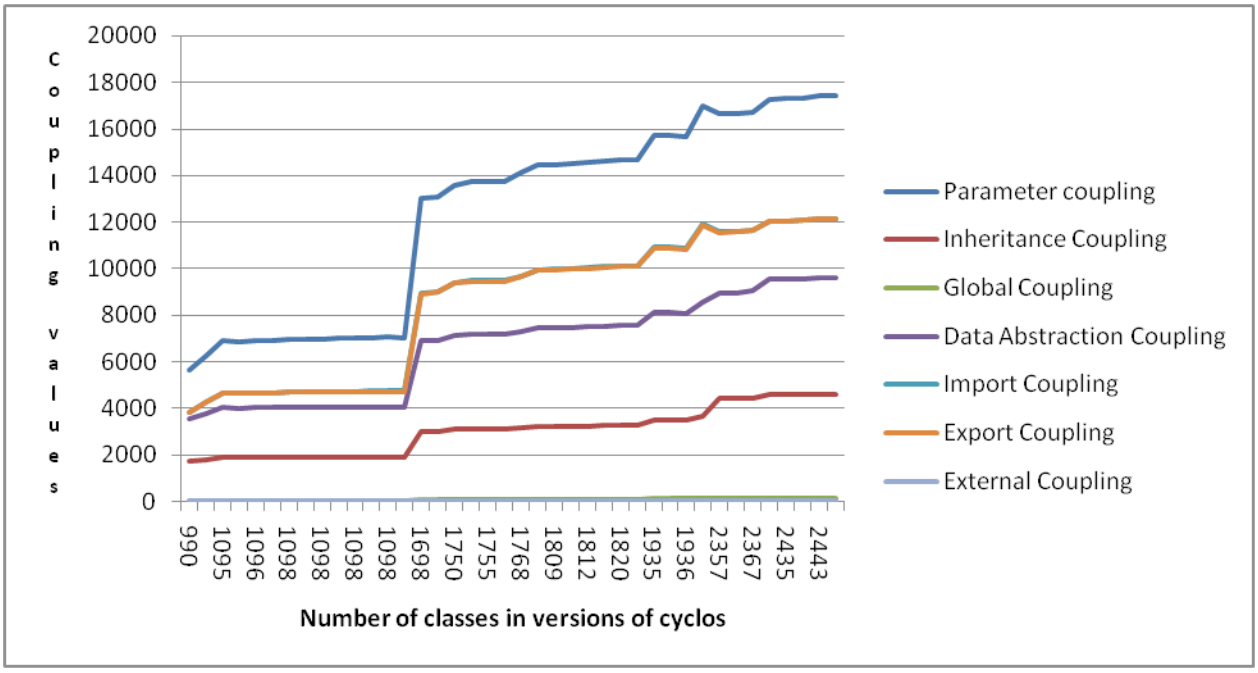


Fig. 2 Behavior of coupling values of forty versions of Cyclos after sorting

there is a major rise (i.e. rise of more than five) in the number of classes of the project, then the coupling values of the projects are increased. It means if there is a major rise in the number of classes of the project the coupling values also increases.

2. **Even if the numbers of classes are stable, the coupling values are not stable.** From the versions cyclos_3.0.4 to cyclos_3.0.11 it is observed that even if the number of classes is same for multiple versions, the coupling values can differ. This is due to changes in the software as per requirements.

Answer 1. With the above observations, the questions can be answered as below.

Answer 2. The correlation exists between the number of classes and coupling values of the projects.

(1) The second question is answered in two parts.

(2) Numbers of classes in the project are directly proportional to the coupling values of the project in case of the significant rise in the number of classes.

If the numbers of classes are stable for many versions then coupling values of the versions may differ.

V. Conclusion

In this paper, a case study is used to find the correlation between the numbers of classes and coupling values of the multimedia Java projects. Results showed that the numbers of classes in the project are directly proportional to the coupling values of the project in case of the major (more than five) rise in the number of classes. Also, if the numbers of classes are stable for many versions, then coupling values of the versions may differ.

References

[1] Arwa Abuasad, Izzat M. Alsmadi, "Evaluating the Correlation Between Software Defect and Design Coupling Metrics", 2012 International Conference on Computer, Information and Telecommunication Systems, CITS 2012, Amman, Jordan. 978-1-4673-1550.

[2] Lucas Batista Leite de Souza, Marcelo de Almeida Maia, "Do Software Categories Impact Coupling Metrics?", Proceedings of the Working Conference on Mining Software Repositories, pp. 217–220. IEEE 2013.

[3] Marcio F. S. Oliveira, Ricardo Miotto Redin, Luigi Carro, Luis da Cunha Lamb, Flavio Rech Wagner, "Software quality metrics and their impact on embedded software", In 5th International Workshop on Model-based Methodologies for Pervasive and Embedded Software (MOMPES), pages 68–77, 2008.

[4] Yeresime Suresh, Jayadeep Pati, Santanu Ku Rath, "Effectiveness of software metrics for the object-oriented system." Procedia Technology 6 (2012): 420-427.

[5] Jeff Offutt, Aynur Abdurazik, and Steve Schach, "Quantitatively Measuring Object-Oriented Couplings", Springer's Software Quality Journal, 6(4):489-517, December 2008.

[6] J. Eder, G. Kappel, M. Schrefl, "Coupling and Cohesion in Object-Oriented Systems", Technical Report, University of Klagenfurt, 1994.

[7] Huan Li, "A Novel Coupling Metric for Object –Oriented Software Systems", IEEE International Symposium on International Journal of Computer Applications (0975 – 8887) Volume 27– No.10, August 2011Knowledge Acquisition and Modeling Workshop, pp. 609-612, 2008.

[8] Husein Sukainah, Oxley Alan, "A Coupling and Cohesion Metrics Suite for Object-Oriented Software", International Conference on Computer Technology and Development, vol.1, no., pp.421-425, 13-15 Nov. 2009.

[9] Briand, Lionel C., John W. Daly, and Jurgen K. Wust, "A unified framework for coupling measurement in object-oriented systems." Software Engineering, IEEE Transactions on 25.1 (1999):91-121.

Mr. V. S. Bidve Ph.D. scholar in #Veltech Dr. RR & Dr. SR Technical University, Avadi, Chennai, India. Works as a Assistance professor in SKNCOE, Pune. Obtained B.E. in Computer Science and Engineering, and M.Tech. in Information Technology. Has 14+ years teaching experience. Working in the area software engineering, web technology, Advanced databases etc.

Dr. P. Sarasu. Working as a Research Dean in #Veltech Dr. RR & Dr. SR Technical University, Avadi, Chennai, India. Also working as a Research supervisor for Ph.D. scholars. Having 20+ years experience in teaching. Working in the area software engineering, mobile computing, wireless network, Advanced Databases etc. Guiding eight research scholars.