# International Journal of Interactive Multimedia and Artificial Intelligence

June 2017, Vol IV, Number 4, ISSN: 1989-1660

*Matching the right kidney to the right patient is one example of an algorithmic artificial intelligence.*
*Hari Sreenivasan*

http://ijimai.unir.net

INTERNATIONAL JOURNAL OF INTERACTIVE MULTIMEDIA AND ARTIFICIAL INTELLIGENCE
ISSN: 1989-1660 –VOL. 4, NUMBER 4

# 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: Radar Clutter, Radar Detectors Performance, Butterfly Optimization Algorithm, Artificial Bee Colony, Evolutionary strategy, Fractal Coding, User Experience, Handwritten Arabic Character Recognition, Feature Extraction, Embedded Hidden Markov Models, Artificial Immune System, Hopfield Neural Network, Browsers, Multimedia, MoCap and Animations.

Machado-Fernandez, J. R. et al. [1] contribute to the improvement of radar detection by suggesting an application of an adaptive scheme which assumes the clutter shape parameter which is known a priori. Offered mathematical expressions are valid for three false alarm probabilities and several windows sizes, covering also a wide range of clutter conditions.

Arora, S. and Singh, S. [2] present a new hybrid optimization algorithm which combines the standard Butterfly Optimization Algorithm (BOA) with Artificial Bee Colony (ABC) algorithm is proposed. The proposed algorithm used the advantages of both algorithms in order to balance the trade-off between exploration and exploitation. Experiments have been conducted on the proposed algorithm using ten benchmark problems having a broad range of dimensions and diverse complexities. The simulation results demonstrate that the convergence speed and accuracy of the proposed algorithm in finding optimal solutions is significantly better than BOA and ABC. for both identification and verification applications.

Habiboghli, A. and Jalali, T. [3]. Biogeography-based Optimization (BBO) is a global optimization algorithm based on population, governed by mathematics of biogeography, and dealing with geographical distribution of biological organisms. The BBO algorithm was used in the present study to provide a solution for the N-queens problem. The performance of the proposed algorithm has been evaluated in terms of the quality of the obtained results, cost function, and execution time. Furthermore, the results of this algorithm were compared against those of genetic and particle swarm algorithms.

Kamble, S.D., et al. [4], present a paper in which the main objective is to develop an approach for video coding using modified three step search (MTSS) block matching algorithm and weighted finite automata (WFA) coding with a specific focus on reducing the encoding time. The MTSS block matching algorithm are used for computing motion vectors between the two frames i.e. displacement of pixels and WFA is used for the coding as it behaves like the Fractal Coding (FC). WFA represents an image (frame or motion compensated prediction error) based on the idea of fractal that the image has self-similarity in itself.

Schrepp, M. et al. [5] talk about questionnaires. These are a cheap and highly efficient tool for achieving a quantitative measure of a product's user experience (UX). However, it is not always easy to decide, if a questionnaire result can show whether a product satisfies this quality aspect. So a benchmark is useful. It allows comparing the results of one product to a large set of other products. In this paper, they describe a benchmark for the User Experience Questionnaire (UEQ), a widely used evaluation tool for interactive products. They also describe how the benchmark can be applied to the quality assurance process for concrete projects.

Boulid, Y. et al. [6] write about handwritten character recognition. A good Arabic handwritten recognition system must consider the characteristics of Arabic letters which can be explicit such as the presence of diacritics or implicit such as the baseline information (a virtual line on which cursive text are aligned and/join). In order to find an adequate method of features extraction, we have taken into consideration the nature of the Arabic characters. The paper investigates two methods based on two different visions: one describes the image in terms of the distribution of pixels, and the other describes it in terms of local patterns. Spatial Distribution of Pixels (SDP) is used according to the first vision; whereas Local Binary Patterns (LBP) are used for the second one. Tested on the Arabic portion of the Isolated Farsi Handwritten Character Database (IFHCDB) and using neural networks as a classifier, SDP achieve a recognition rate around 94% while LBP achieve a recognition rate of about 96%.

Jalal, A. et al. [7] talk about elderly people and they need special care in the form of healthcare monitoring systems. Recent advancements in depth video technologies have made human activity recognition (HAR) realizable for elderly healthcare applications. In this paper, a depth video-based novel method for HAR is presented using robust multi-features and embedded Hidden Markov Models (HMMs) to recognize daily life activities of elderly people living alone in indoor environment such as smart homes. In the proposed HAR framework, initially, depth maps are analyzed by temporal motion identification method to segment human silhouettes from noisy background and compute depth silhouette area for each activity to track human movements in a scene. Several representative features, including invariant, multi-view differentiation and spatiotemporal body joints features were fused together to explore gradient orientation change, intensity differentiation, temporal variation and local motion of specific body parts. Then, these features are processed by the dynamics of their respective class and learned, modeled, trained and recognized with specific embedded HMM having active feature values. Furthermore, we construct a new online human activity dataset by a depth sensor to evaluate the proposed features. Our experiments on three depth datasets demonstrated that the proposed multi-features are efficient and robust over the state of the art features for human action and activity recognition.

Bin Mansor, M. A., et al. [8] in this paper they implement an artificial immune system algorithm incorporated with the Hopfield neural network to solve the restricted MAX-kSAT problem. The proposed paradigm are compared with the traditional method, Brute force search algorithm integrated with Hopfield neural network. The results demonstrate that the artificial immune system integrated with Hopfield network outperforms the conventional Hopfield network in solving restricted MAX-kSAT. All in all, the result has provided a concrete evidence of the effectiveness of our proposed paradigm to be applied in other constraint optimization problem. The work presented here has many profound implications for future studies to counter the variety of satisfiability problem.

Beltrán-Alfonso, R. et al. [9] present a study of the development and evolution of search engines, more specifically, to analyze the relevance of findings based on the number of results displayed in paging systems with Google as a case study. Finally, it is intended to contribute to indexing criteria in search results, based on an approach to Semantic Web as a stage in the evolution of the Web.

Magdin, M., [10] proposes a new easy-to-use system for home usage, through which we are making character animation. In its implementation, they paid attention to the elimination of errors from the previous solutions. In this paper the authors describe the method



DOI: 10.9781/ijimai.2017.440

how motion capture characters on a treadmill and as well as an own Java application that processes the video for its further use in Cinema 4D. This paper describes the implementation of this technology of sensing in a way so that the animated character authentically imitated human movement on a treadmill.

Wall, F.,[11] studies the effects of learning-induced alterations of distributed search systems' organizations. In particular, scenarios where alterations of the search-systems' organizational setup are based on a form of reinforcement learning are compared to scenarios where the organizational setup is kept constant and to scenarios where the setup is changed randomly. The results indicate that learning-induced alterations may lead to high levels of performance combined with high levels of efficiency in terms of reorganization-effort. However, the results also suggest that the complexity of the underlying search problem together with the aspiration level (which drives positive or negative reinforcement) considerably shapes the effects of learning.

Gil, E. and Medinaceli, K. [13] present a paper that focuses on analyzing from the point of view of medical diagnosis the importance of electronic medical records as a unifying element of the information essential for this type of diagnosis, and the use of artificial intelligence techniques in this field. To this end the current situation of electronic medical records is analyzed in a country like Bolivia exhaustively analyzing three of the most important health centers. Is used for this unstructured interview experts on the subject reflect the status of electronic medical records from the point of view of protection of the right to privacy of individuals and will serve as a model for development, not only in Bolivia but also in other Latin American countries.

Dr. Rubén González

## References


[1] Machado-Fernandez, J. R. et al., "CA-CFAR Adjustment Factor Correction with a priori Knowledge of the Clutter Distribution Shape Parameter", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 4, pp. 7-13, 2017.

[2] Arora, S. and Singh, S., "An Effective Hybrid Butter y Optimization Algorithm with Artificial Bee Colony for Numerical Optimization", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 4, pp. 14-21, 2017.

[3] Habiboghli, A. and Jalali, T., "A Solution to the N-Queens Problem Using Biogeography-Based Optimization", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 4, pp. 22-26, 2017.

[4] Kamble, S.D., et al., "Modi ed Three-Step Search Block Matching Motion Estimation and Weighted Finite Automata based Fractal Video Compression", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 4, pp. 27-39, 2017.

[5] Schrepp, M. et al., "Construction of a Benchmark for the User Experience Questionnaire (UEQ)", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 4, pp. 40-44, 2017.

[6] Boulid, Y. et al., "Handwritten Character Recognition Based on the Specificity and the Singularity of the Arabic Language", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 4, pp. 45-53, 2017.

[7] Jalal, A. et al.., "A Depth Video-based Human Detection and Activity Recognition using Multi-features and Embedded Hidden Markov Models for Health Care Monitoring Systems", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 4, pp. 54-62, 2017.

[8] Bin Mansor, M. A., et al., "Robust Artificial Immune System in the Hopfield network for Maximum k-Satisfiability", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 4, pp. 63-71, 2017.

[9] Settouti, N. et al., "Exploring the Relevance of Search Engines: An Overview of Google as a Case Study", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 4, pp. 72-79, 2017.

[10] Magdin, M., "Simple MoCap System for Home Usage", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 4, pp. 80-87, 2017.

[11] Wall, F., "Distributed Search Systems with Self-Adaptive Organizational Setups", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 4, pp. 88-95, 2017.

[12] Gil, E. and Medinaceli, K., "Electronic Health Record in Bolivia and ICT: A Perspective for Latin America", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 4, pp. 96-101, 2017.

TABLE OF CONTENTS



**OPEN ACCESS JOURNAL**

**ISSN: 1989-1660**

# CA-CFAR Adjustment Factor Correction with a priori Knowledge of the Clutter Distribution Shape Parameter

José Raúl Machado Fernández, Shirley Torres Martínez, Jesús de la Concepción Bacallao Vidal

*Instituto Superior Politécnico José Antonio Echeverría, La Habana, Cuba*

***Abstract* — Oceanic and coastal radars operation is affected because the targets information is received mixed with and undesired contribution called sea clutter. Specifically, the popular CA-CFAR processor is incapable of maintaining its design false alarm probability when facing clutter with statistical variations. In opposition to the classic alternative suggesting the use of a fixed adjustment factor, the authors propose a modification of the CA-CFAR scheme where the factor is constantly corrected according on the background signal statistical changes. Mathematically translated as a variation in the shape parameter of the clutter distribution, the background signal changes were simulated through the Weibull, Log-Normal and K distributions, deriving expressions which allow choosing an appropriate factor for each possible statistical state. The investigation contributes to the improvement of radar detection by suggesting the application of an adaptive scheme which assumes the clutter shape parameter is known a priori. The offered mathematical expressions are valid for three false alarm probabilities and several windows sizes, covering also a wide range of clutter conditions.**

***Keywords* — Radar Clutter, CFAR, False Alarm Probability, Radar Detectors Performance.**

## I. Introduction

A radar is a device that emits electromagnetic waves and receives the echo resulting from their interaction with nearby objects [1]. The echoes received from targets of interest are interpreted as useful signal, while those originated from the reflection of the emission on other entities are considered as a distortion and called clutter [2].

When operating in coastal and offshore environments, the distortion signal often comes from the sea surface and is thereby called sea clutter. Given the variability of the sea surface, and the influence of others factors, the sea clutter usually introduces an important interference in the detection [3].

The Weibull [4, 5], Log-Normal [6, 7] and K [8-10] distributions are commonly used for sea clutter modeling. They display heavy tails for certain configurations of their respective parameters, according to that observed in radar readings.

CFAR (Constant False Alarm Rate) processors are widely applied for eliminating the effects of the clutter, being the CA-CFAR (Cell Averaging-CFAR) the classical alternative [11]. This scheme constantly moves a reference window across the coverage area for calculating the average of a small region. Then, it decides if the cell at the center of the window is a target by comparing its magnitude with a threshold calculated from the estimated average [12].

When the CA-CFAR mistakenly classifies a cell with clutter as a target, it is said that a false alarm has occurred. Indeed, the false alarm probability ($P_f$) is one of the fundamental parameters of a radar detector [13]. Therefore, the CA-CFAR includes an adjustment factor ($T$) that allows establishing the false alarm probability at a predefined level.

In the traditional operation mode, the adjustment factor is kept constant the entire operation period [14]. This approach provides good results under the assumption that the clutter statistics will remain unaffected. However, when variations occur in the statistics, it becomes necessary to correct the adjustment factor as it was demonstrated in [15]. Otherwise, the operational false alarm probability will deviate from the intended design value.

As a solution to this problem, the authors of the current paper obtained mathematical expressions that allow making the necessary correction in the adjustment factor for a wide range of shape parameters for the Weibull, Log-Normal and K distributions. The results were obtained after processing several millions samples and performing curve fitting procedures. The offered expressions are valid for the false alarm probabilities of $10^{-2}$, $10^{-3}$ and $10^{-4}$ and provide an estimation of the factor regardless of the CA-CFAR sliding window size. The shape parameter of the statistical distributions was assumed to be known in advance, which is acceptable given the accurate methods proposed in [16-18].

The found expressions provide a new alternative for solving a problem ignored by most common CFAR implementations that usually concentrate on modifying the method for estimating the background average [19-22]. This paper focuses on canceling the effect of the clutter slow statistical variations instead of developing techniques for processing non-homogeneities.

The classical solution for this problem was described in [23] and it proposes the use of an empiric formula based on environmental conditions. However, the drawbacks of this approach were pointed out also in [23], where it was explained that the progress is yet reduced because the formulas does not take into account the wind speed and the sea state, two conditions of proven influence on the features of the measurements. The approximation followed in the current paper ignores the complexity introduced by the environmental variables and focuses on the processing of the received samples as a direct approach to the correction of the adjustment factor.

The paper proceeds as follows. The second section, called "Materials and Methods" presents the Probability Density Functions (PDF) of the Weibull, Log-Normal and K models. In the same section, the executed algorithm is described. Subsequently, the results are presented and commented in "Results and Discussion", including the found mathematical expressions from the curve and surface fittings which generalize the obtained outcomes. Finally, in "Conclusions and Future Research" the main contributions of the study are summarized and recommendations are given for ways to continue the investigation.

## II. Materials and Methods

The probability density functions employed in this project are the Weibull (1), the Log-Normal (2) and the K (3). All of them have



DOI: 10.9781/ijimai.2017.441

extensive validation and were categorized as classical distributions in [24].

The K distribution is the more widely accepted model for high resolution sea clutter observed at low grazing angles [23]. The Weibull distribution is a very versatile model that has been applied to ground [25], weather [7] and ice clutter [26]; in addition, in [4] it was selected as the best model fitting sea clutter data. Lastly, the Log-Normal distribution tends to fit data for particular situations such as samples corresponding to HH polarization [27], in the assessment of the clutter spatial distribution [28], and for cells containing mixed target and clutter reflections [4, 7].

$$f_W(x|\alpha,\beta) = \frac{\beta x^{\beta-1}}{\alpha^{\beta}} exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right] \quad (1)$$

$$f_{LN}(x|\mu,\sigma) = \frac{1}{\sigma x\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2\right] \quad (2)$$

$$f_K(x|c,v) = \frac{4c}{\Gamma(v)}(cx)^v K_{v-1}(2cx) \quad (3)$$

In expression (1), $f_W$ is the Weibull PDF whereas $\alpha$ and $\beta$ are scale and shape parameters respectively. Together, in (2) and (3) $f_{LN}$ and $f_K$ are the Log-Normal and K PDFs, whereas $(\mu, c)$ are the scale parameters and $(\sigma, v)$ are the shape parameters. In all cases, the $x$ was utilized as the independent variable. Auxiliary expressions such as the CDF (Cumulative Distribution Function) and the moment generating function can be found in [29-32].

### A. Description of the Experiments

The basic experiment was carried out to simulate the response of a CA-CFAR to one million independent sliding windows filled with clutter samples. The total number of involved samples depended on the window size; for example, when a 64 reference cells window was simulated, there were 65 samples in each essay, which made a total of 65 million samples for the whole experiment.

In the first iteration, a small adjustment factor such as $T = 1{,}1$ was chosen. This value provoked the occurrence of many false alarms, that is, a high number of clutter samples were misclassified as targets. Then, the procedure was repeated increasing the $T$ in each new iteration until the factor that produced a $P_f = 10^{-2}$ was found with less than a $1\%$ deviation. The same process led to the finding of the $T$ factors that guaranteed the occurrence of $P_f = 10^{-3}$ and $P_f = 10^{-4}$.

The three $T$ values extracted from the above algorithm are only valid for the distribution obeyed by the processed samples. Therefore, if the first sequence corresponded to the Weibull model, then the steps had to be repeated for the Log-Normal and K as well.

In addition, three factors are not enough for each distribution because they only represent the CA-CFAR response to a given shape parameter. The authors chose 19 different parameters from the Weibull, Log-Normal and K distributions and re-executed the procedure with each one. The shape parameters were taken from intervals validated in several studies. For the Weibull case, it was used: $1{,}75 < \beta < 6{,}25$ [4, 33, 34]; for the Log-Normal case: $0{,}025 < \sigma < 1{,}25$ [7, 25, 27, 35], and for the K distribution: $0{,}1 < v < 4{,}6$ [34-36]. The reader should note that the scale parameter has no influence in the detection mechanism [15].

Although it includes the handling of a large number of samples, the previously described algorithm is not yet complete. It needs to be repeated for different sliding window dimensions. The authors initially used a size of 64 cells and then recalculated the $T$ s for 32, 16 and 8 cells in the sliding window.

Once completed, the experiments yielded 228 $T$ values per distribution, for a total of 684 figures. These numbers can be placed on a table for searching according to the specifications of a given design. However, to facilitate the implementation of the results, the authors performed a curve fitting procedure that allowed synthesizing the findings into three mathematical expressions per distribution. Each expression belongs to one of the three addressed false alarm probabilities and allows the direct entry of the window size and the shape parameter, returning the $T$ value to be applied.

## III. Results and Discussion

The current section presents the results of the experiments. It begins by describing the influence of each simulation variable on the CA-CFAR adjustment factor. Subsequently, the curve and surface fittings that enabled the generation of the mathematical expressions are discussed. The section concludes by characterizing the deviations introduced by the expressions, and by discussing the application of the paper.

### A. Influence of the variables on the CA-CFAR factor

Fig. 1 plots the adjustment factors estimated for guaranteeing the false alarm probabilities of $P_f = 10^{-2}$, $P_f = 10^{-3}$ and $P_f = 10^{-4}$ for 19 different occurrences of the Weibull shape parameter. The information corresponds to that observed for a 64 cells CA-CFAR.

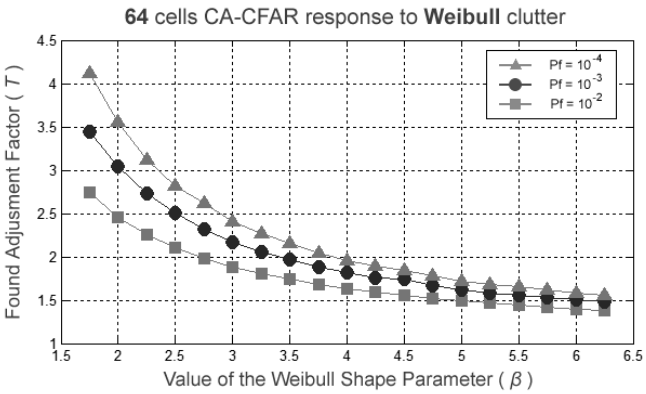



Fig. 1. Adjustment factors found by processing Weibull clutter with a 64 cells CA-CFAR.

As it can be seen, the reduction of the false alarm probability causes the adjustment factor to rise. This represents a logical behavior according on the effect of $T$ on the detection threshold. A high factor causes the subsequent elevation of the threshold which provokes the occurrence of fewer false alarms, since it is less likely that a sample will exceed the defined level.

The effect of the variation of $\beta$ over $T$ is also visible in Fig. 1. The Weibull distribution exhibits heavier tails for the lower figures of the shape parameter, forcing the increase of the adjustment factor. It can be also noted that as the shape parameter increases, it losses influence on the selection of $T$, that is, there is little difference between consecutive $T$ values for $\beta > 5$.

The Weibull distribution shares the above feature with the K model. Actually, as it can be seen in Fig. 2 (left), the saturation of the influence

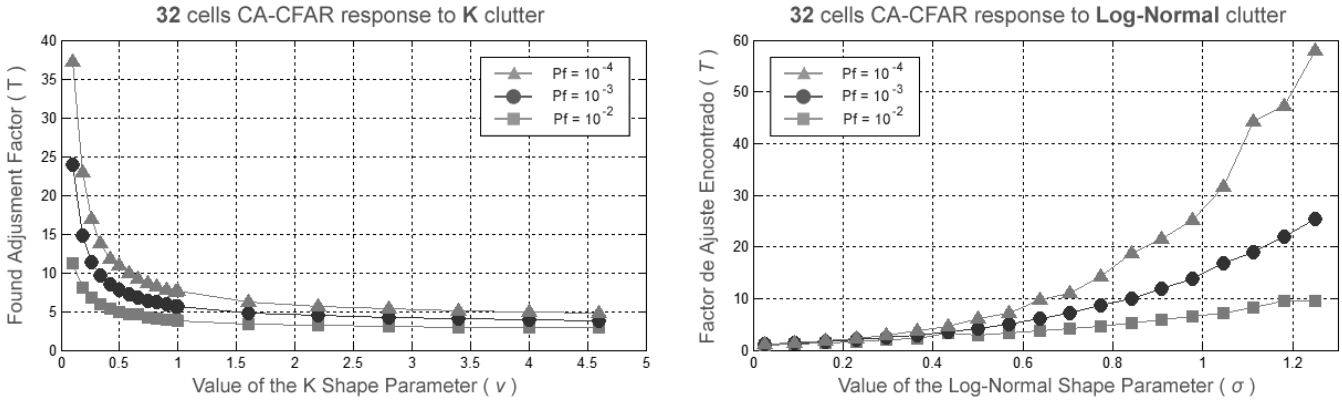



Fig. 2. Adjustment factors extracted from the response of a 32 cells CA-CFAR to K and Log-Normal clutter.

of the shape parameter over $T$ is even higher for the K distribution. The opposite happens in the Log-Normal case (Fig. 2 right) where the increase between consecutive $T$ values is remarkable for the higher $\sigma$s.

Fig. 2 plots the performed estimates for a 32 cells processor and variations of the K and Log-Normal shape parameters. Note that what was observed in Fig. 1 regarding the relationship between the false alarm probability and $T$ also applies to these charts.

Moreover, Fig. 3 displays the effect of modifying the size of the sliding window on the adjustment factor selection. The reader may notice that the graph placed in the upper right corner is a zoomed view of the behavior in the $4 < \beta < 6,25$ region.

After processing Weibull samples with CA-CFAR schemes whose windows sizes were changed, the need to increase $T$ with the decrease of the number of cells in the window became obvious (see Fig. 3). A system with fewer cells for computing the background mean will produce less accurate threshold estimations. Accordingly, the processor will require a higher adjustment factor to maintain its design $P_f$.

Here ends the description of the influence of the different simulation variables on the outcomes. The behavior described for the presented examples generalizes that observed in all trials.

### *B. Curve and surface fittings*

In order to concentrate the information gathered in all the experiments, curve and surface fittings were conducted. Fig. 4 shows some of the curve fits made for the $T$ factors obtained from Log-Normal distributed samples.

The following fits were tested for each distribution: polynomial from first to eighth order; rational from first to fifth order both in the numerator and in the denominator; and power fits. It was found that the rational fit provided the best results, accurately matching the shape parameters relation with $T$ increase. Weibull and K distributions were fitted with a second degree polynomial in the numerator and a

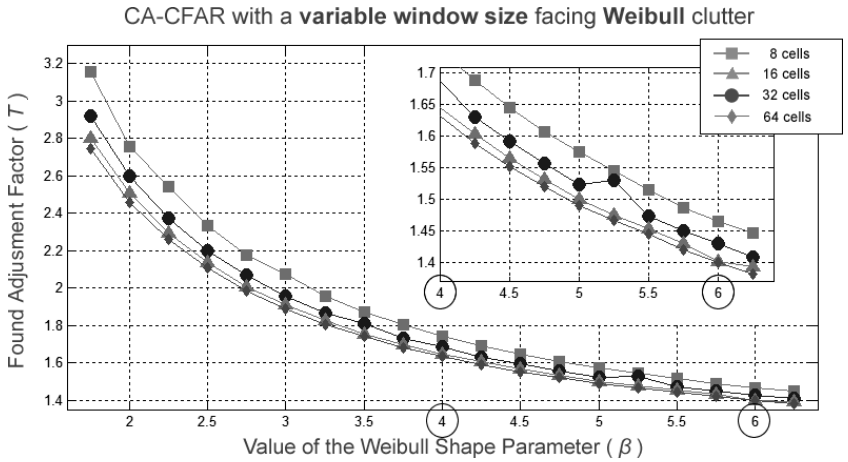



Fig. 3. Adjustment factors found by processing Weibull clutter samples with different window sizes.

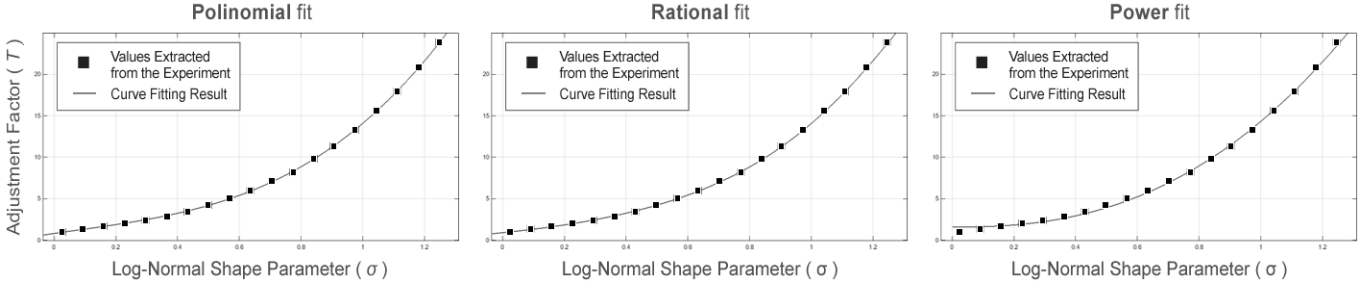



Fig. 4. Curve fittings for a 64 cells CA-CFAR operating for $P_f=10^{-3}$ and facing Log-Normal clutter.

first degree in the denominator. On the other hand, the Log-Normal distribution was fitted with a first degree polynomial in the numerator and a second degree in the denominator.

After completing the curve fittings, the authors proceeded to execute surface fittings which included the influence of the window size on the selection of the $T$ value. A representation of the problem at hand is offered in Fig. 5 where a three-dimensional surface relating $T$ with the shape parameter of the K distribution and the number of cells is presented. Expressions obtained from the surface fitting are given in Table 1 that includes nine formulas for different false alarm probabilities and distributions, together with their respective values of RMSE (Root Mean Squared Error). As it can be seen, an exponential term that takes into account the contribution of the number of cells in the window was added in all cases.

### C. Evaluating the quality of the fit

The quality of the fit was measured by generating new independent sample sets with a similar structure to those employed in the initial simulations. Then, the sets were processed using the $T$ values extracted from the formulae provided in Table I. The measured $P_f$ values that were extracted from these experiments proved that the expressions guarantee an average deviation of a 32,3% for the Weibull distribution, of a 30,4% for the Log-Normal case and of a 37,66% for the K model.

The same experiment was reproduced using expressions obtained from the curve fittings instead of those from the surface fittings. The average deviation reduced to a 12,1% for the Weibull distribution, a 17,4% for the Log-Normal and a 15,85% for the K distribution when applying formulae from Tables II, III and IV. This indicated that the surface fit was unable to correctly adapt to the data at hand without losing precision. The drawback of using the curve fit is that it produces 36 expressions instead of the 9 displayed in Table 1.

Then, a selection have to be made according on the characteristics of the implementation at hand. If a fixed window size is used in the receiver, then a curve fit expression can be used for an improved precision. On the other hand, if the window size is not known in advance, or if it may vary given certain situations, a surface fitting expression will provide a good overall performance. Besides, expressions from Table I also allow estimating the $T$ values for window sizes different from 64, 32, 16 and 8.

TABLE I. EXPRESSIONS OBTAINED FROM THE CURVE FITTINGS FOR ESTIMATING $T$ USING THE SHAPE PARAMETER FROM THE WEIBULL, LOG-NORMAL AND K DISTRIBUTIONS AND THE WINDOW SIZE $(w)$.

| $P_f$ | Expressions found through the curve fitting procedure | RMSE |
|---|---|---|
| K Distribution | | |
| - | $T = \frac{A * v^2 + B * v + C}{v + D} + e^{-(E * w^2)}$ | - |
| $10^{-2}$ | $A = -0{,}1656 \quad B = 2{,}98 \quad C = 0{,}8969$ $D = -0{,}01431 \quad E = 0{,}04643$ | 1,36 |
| $10^{-3}$ | $A = -0{,}1076 \quad B = 3{,}439 \quad C = 2{,}189$ $D = -0{,}0367 \quad E = 0{,}04659$ | 6,641 |
| $10^{-4}$ | $A = 0{,}3314 \quad B = 2{,}98 \quad C = 4{,}105$ $D = -0{,}05309 \quad E = 0{,}04716$ | 12,67 |
| Log-Normal Distribution | | |
| - | $T = \frac{A * \sigma + B}{\sigma^2 + C * \sigma + D} + e^{-(E * w)}$ | - |
| $10^{-2}$ | $A = 9{,}595 \quad B = 2{,}664 \quad C = -3{,}544$ $D = 4{,}306 \quad E = 0{,}0517$ | 0,4308 |
| $10^{-3}$ | $A = 9{,}365 \quad B = 1{,}548 \quad C = -3{,}185$ $D = 2{,}892 \quad E = 0{,}04424$ | 1,564 |
| $10^{-4}$ | $A = 24{,}95 \quad B = -0{,}6269 \quad C = -3{,}626$ $D = 3{,}42 \quad E = 0{,}07412$ | 3,435 |
| Weibull Distribution | | |
| - | $T = \frac{A * \beta^2 + B * \beta + C}{\beta + D} + e^{-(E * w^2)}$ | - |
| $10^{-2}$ | $A = 0{,}003044 \quad B = 0{,}9155 \quad C = 2{,}158$ $D = 0{,}4476 \quad E = 0{,}2679$ | 0,04264 |
| $10^{-3}$ | $A = 0{,}1517 \quad B = -0{,}8014 \quad C = -8{,}978$ $D = -0{,}5195 \quad E = 0{,}1586$ | 0,1898 |
| $10^{-4}$ | $A = 0{,}03097 \quad B = -0{,}1254 \quad C = 4{,}483$ $D = -0{,}6789 \quad E = 0{,}01009$ | 0,1581 |

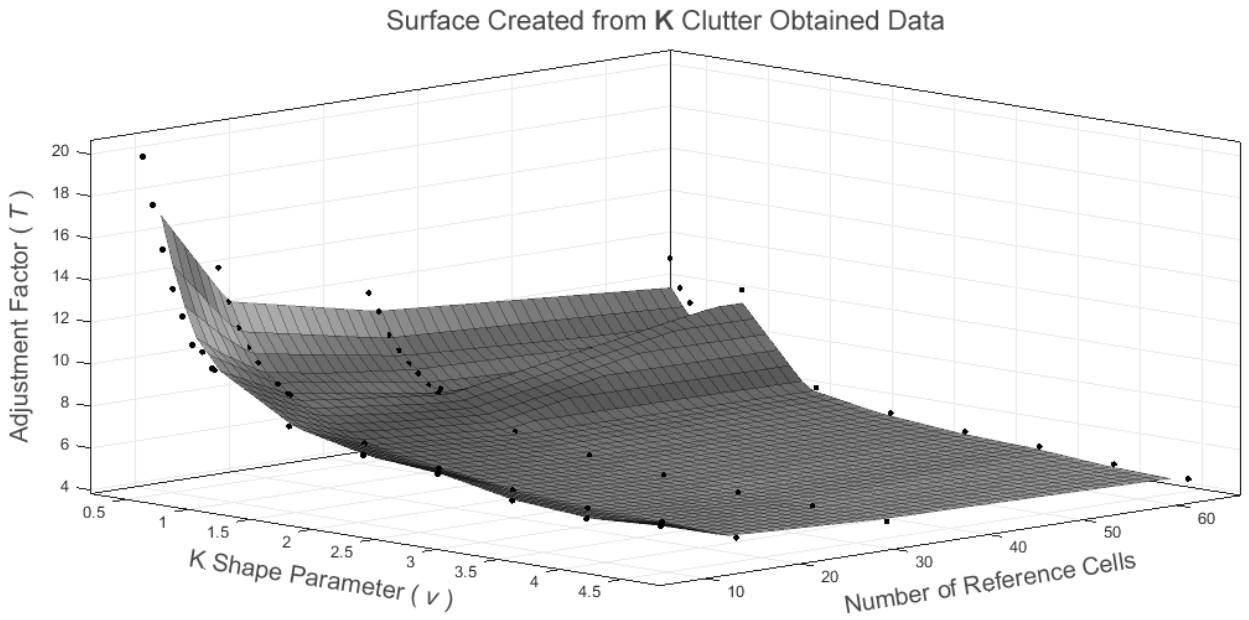



Fig. 5. Adjustment factors extracted from the response of a 32 cells CA-CFAR to K and Log-Normal Clutter.

TABLE II. EXPRESSIONS OBTAINED FROM THE CURVE FITTINGS FOR ESTIMATING $T$ USING THE SHAPE PARAMETER FROM THE K DISTRIBUTION AND THE WINDOW SIZE $(w)$.

| $P_f$ | Expressions found through the curve fitting procedure (A; B; C; D) |
|---|---|
| $T = \frac{A * v^2 + B * v + C}{v + D}$ | |
| $w = 8$ | |
| $10^{-2}$ | ( −0,5805 ; 5,161 ; 0,06531 ; −0,07565 ) |
| $10^{-3}$ | ( 0,3245 ; 2,406 ; 3,794 ; −0,05263 ) |
| $10^{-4}$ | ( 1,649 ; −2,001 ; 8,859 ; −0,06449 ) |
| $w = 16$ | |
| $10^{-2}$ | ( 0,01706 ; 2,612 ; 1,528 ; 0,02671 ) |
| $10^{-3}$ | ( −0,03579 ; 3,685 ; 2,327 ; −0,01915 ) |
| $10^{-4}$ | ( 0,2441 ; 3,306 ; 4,331 ; −0,02897 ) |
| $w = 32$ | |
| $10^{-2}$ | ( −0,04266 ; 2,79 ; 1,295 ; 0,03994 ) |
| $10^{-3}$ | ( −0,5496 ; 4,68 ; 1,596 ; −0,01426 ) |
| $10^{-4}$ | ( 0,01038 ; 4,059 ; 3,391 ; 0,001708) |
| $w = 64$ | |
| $10^{-2}$ | ( 0,1018 ; 2,029 ; 2,272 ; 0,1796 ) |
| $10^{-3}$ | ( −0,01864 ; −3,37 ; 2,281 ; 0,03967 ) |
| $10^{-4}$ | ( 0,06958 ; 3,647 ; 3,874 ; −0,0428 ) |

TABLE III. EXPRESSIONS OBTAINED FROM THE CURVE FITTINGS FOR ESTIMATING $T$ USING THE SHAPE PARAMETER FROM THE LOG-NORMAL DISTRIBUTION AND THE WINDOW SIZE $(w)$.

| $P_f$ | Expressions found through the curve fitting procedure (A; B; C; D) |
|---|---|
| $T = \frac{A * \sigma + B}{\sigma^2 + C * \sigma + D}$ | |
| $w = 8$ | |
| $10^{-2}$ | ( 10,22 ; 4,442 ; −3,953 ; 4,774 ) |
| $10^{-3}$ | ( 6,645 ; 2,34 ; −2,991 ; 2,475 ) |
| $10^{-4}$ | ( 11,71 ; 1,117 ; −2,848 ; 2,196 ) |
| $w = 16$ | |
| $10^{-2}$ | ( 5,704 ; 3,577 ; −3,303 ; 3,582 ) |
| $10^{-3}$ | ( 10,34 ; 2,686 ; −3,36 ; 3,185 ) |
| $10^{-4}$ | ( 12,04 ; 1,473 ; −2,955 ; 2,396 ) |
| $w = 32$ | |
| $10^{-2}$ | ( 8,004 ; 4,658 ; −3,692 ; 4,528 ) |
| $10^{-3}$ | ( 7,839 ; 2,6 ; −3,129 ; 2,836 ) |
| $10^{-4}$ | ( 8,235 ; 1,879 ; −2,808 ; 2,157) |
| $w = 64$ | |
| $10^{-2}$ | ( 17,88 ; 5,59 ; −3,997 ; 6,555 ) |
| $10^{-3}$ | ( 8,548 ; 2,733 ; −3,209 ; 3,011 ) |
| $10^{-4}$ | ( 3,468 ; 0,7**375** ; 2,874 ; −1,5428 ) |

TABLE IV. EXPRESSIONS OBTAINED FROM THE CURVE FITTINGS FOR ESTIMATING $T$ USING THE SHAPE PARAMETER FROM THE WEIBULL DISTRIBUTION AND THE WINDOW SIZE $(w)$.

| $P_f$ | Expressions found through the curve fitting procedure (A; B; C; D) |
|---|---|
| $T = \frac{A * \beta^2 + B * \beta + C}{\beta + D}$ | |
| $w = 8$ | |
| $10^{-2}$ | ( 0,005302 ; 0,9185 ; 2,397 ; −0,4731 ) |
| $10^{-3}$ | ( 0,03274 ; 0,6309 ; 3,795 ; −0,6078 ) |
| $10^{-4}$ | ( 0,06138 ; 0,319 ; 5,247 ; −0,6935 ) |
| $w = 16$ | |
| $10^{-2}$ | ( −0,005079 ; 1,041 ; 1,762 ; −0,5287 ) |
| $10^{-3}$ | ( 0,02276 ;0,7381; 3,26 ; −0,541 ) |
| $10^{-4}$ | ( 0,02564 ;0,6743 ; 3,853 ; −0,6671 ) |
| $w = 32$ | |
| $10^{-2}$ | ( 0,0002637 ; 0,978 ; 1,942 ; −0,4444 ) |
| $10^{-3}$ | ( 0,009078 ; 0,9082 ; 2,447 ; −0,625 ) |
| $10^{-4}$ | ( 0,01197 ; 0,8352 ; 3,092 ; −0,6961) |
| $w = 64$ | |
| $10^{-2}$ | ( 0,002336 ; 0,9605 ; 1,976 ; −0,4131 ) |
| $10^{-3}$ | ( 0,1037 ; 0,2026 ; 7,177 ; 0,4165 ) |
| $10^{-4}$ | ( 0,02288 ; 0,7079 ; 3,534 ; −0,5758 ) |

### *D. Application and assessment of the results*

The current research led to the creation of nine mathematical expressions that allow selecting optimal CA-CFAR adjustment factors for maintaining the false alarm probabilities of $\mathbf{10^{-2}}$, $\mathbf{10^{-3}}$ and $\mathbf{10^{-4}}$ for any window size between 8 and 64. The offered formulae take into account the contribution of the shape parameter of the Weibull, Log-Normal and K distributions on the selection of the CA-CFAR factor, as well as the influence of the window size of the processor itself.

The research assumed the shape parameter was known a priori for all distributions. Its main application is in the improvement of the NATE-CFAR detector presented in [17, 37]. This detector uses an enhanced neuronal parameter estimation technique to obtain a statistical characterization of clutter related distributions such as the Weibull [16], K [18] and Pareto [38]. The current contribution allows the expansion of the NATE method to a CA-CFAR with different numbers of cells in the sliding window; the previous NATE implementation was only suitable for 64 cells. Moreover, this research is particularly useful in the adaptation to statistical clutter changes of CFAR mechanisms that modify the size of the window such as the CI-CFAR [39] or TL-CFAR [40].

Regarding the accuracy achieved by the expressions, it can be noted that previous researches had displayed deviation percentages below 20% for the false alarm probability [41-44]. However, the authors believe that the 33,45% achieved in the current project constitutes a positive outcome and justify the augmentation in the percentage of mistakes on the inclusion of several window sizes in the essays. Indeed, the papers [41-44] were concentrated on a single window size. In addition, in order to improve the performance of the presented solution, the authors also offered 36 particular expressions that can be used for fixed window sizes, reducing the deviation to a 15,11%.

## IV. Conclusions and Future Research

After processing several million computer-generated samples corresponding to Weibull, Log-Normal and K distributions, the authors were able to create mathematical expressions relating the shape parameter of each distribution with the CA-CFAR adjustment factor for several sliding window sizes. By using the values extracted from the expressions for correcting the adjustment factor, a cell averaging processor can guarantee its false alarm probability will only deviate a 33,45% from the value conceived in the design even when the clutter exhibits statistical variations and the size of the sliding window varies from 8 to 64 cells. The system requires a priori knowledge of the shape parameter of the clutter distribution; therefore the formulae are to be used together with other previously presented mechanisms that include the estimation of the parameters.

The authors will focus next on searching a method for generalizing the results for false alarm probabilities lower than $10^{-4}$. The reproduction of the study for the popular OS-CFAR processor is also a future goal.

## References


[1] W. L. Melvin and J. A. Scheer, *Principles of Modern Radar, Vol III Radar Applications*: Scitech Publishing, 2014.

[2] H. Meikle, *Modern Radar Systems*, 2nd Edition: Artech House 2008.

[3] T. Oluwale Oyedokun, "Sea Clutter Simulation", Master of Science in Engineering, Department of Electrical Engineering, University of Cape Town, 2012.

[4] S. Ishii, S. Sayama, and K. Mizutani, "Effect of Changes in Sea-Surface State on Statistical Characteristics of Sea Clutter with X-band Radar", *Wireless Engineering and Technology,* 2(3), pp. 175-183, 2011.

[5] Y. Dong, "Distribution of X-Band High Resolution and High Grazing Angle Sea Clutter", Technical Report DSTO-RR-0316, Electronic Warfare and Radar Division, Defence Science and Technology Organization, Edinburgh, South Australia, 2006.

[6] J. Z. Yim, C.-R. Chou, and W.-K. Wong, "A Study of the Statistics of Sea Clutter in the Northern Coast of Taiwan", in *Proceedings of the Seventeenth International Offshore and Polar Engineering Conference*, Lisbon, Portugal, pp. 1-6, 2007.

[7] S. Sayama and S. Ishii, "Suppression of Log-Normal Distributed Weather Clutter Observed by an S-Band Radar", *Wireless Engineering and Technology,* 4(3), pp. 125-133, 2013.

[8] Z. Zhijian, X. Ruilai, H. Yong, and G. Jian, "New Nonparametric Detectors under K-Distributed Sea Clutter in Radar Applications", in *IEEE 2011 CIE International Conference on Radar*, pp. 1752-1755, 2011.

[9] X. Meng, G. Feng, H. Xue, and Z. He, "Wideband Radar Target Detection Theory in Coherent K Distributed Clutter", *Research Journal of Applied Sciences, Engineering and Technology,* 5(5), pp. 1528-1532, 2013.

[10] G. Tanriverdi, "Arma Model Based Clutter Estimation and its Effect on Clutter Supression Algorithms", Master of Science in Electrical and Electronics Engineering, The graduate School of Natural and Applied Sciences, 2012.

[11] J. C. Bacallao Vidal, "Un modelo Teórico de la Técnica DRACEC. Metodología del Proceso de Adaptación al Fondo", Doctor en Ciencias Técnicas, Instituto Técnico Militar "José Martí", La Habana, Cuba, 2003.

[12] J. R. Machado Fernández and R. C. Sánchez Rams, "Implementación de un Detector de Promediación de Clutter (CA-CFAR) usando VHDL (accepted)", *Telem@tica,* 2016.

[13] M. A. Richards, J. A. Scheer, and W. A. Holm, *Principles of Modern Radar Vol I Basic Principles*: Scitech Publishing, 2010.

[14] R. C. Sánchez Rams, "Implementación del detector CA-CFAR en VHDL para crear un PC-Radar Cubano usando FPGA", Ing. en Telecomunicaciones y Electrónica, Departamento de Telecomunicaciones y Telemática, Facultad de Ingeniería Eléctrica, Instituto Superior Politécnico José Antonio Echeverría (ISPJAE), La Habana, Cuba, 2014.

[15] J. R. Machado Fernández and J. C. Bacallao Vidal, "MATE-CFAR: Ambiente de Pruebas para Detectores CFAR en MATLAB", *Telem@tica,* 13(3), pp. 86-98, 2014.

[16] J. R. Machado Fernández, J. C. Bacallao Vidal, and N. Chávez Ferry, "A Neural Network Approach to Weibull Distributed Sea Clutter Parameter's Estimation", *Inteligencia Artificial,* 18(56), pp. 3-13, 2015.

[17] J. R. Machado Fernández and J. C. Bacallao Vidal, "Log-Normal clutter-Method of Moments - Cell Averaging - Constant False Alarm Rate (LN-MoM-CA-CFAR) Detector (bajo revisión)", *Ingeniería e Investigación, Colombia,* 2016.

[18] J. R. Machado Fernández and J. C. Bacallao Vidal, "Improved Shape Parameter Estimation in K Clutter with Neural Networks and Deep Learning", *International Journal of Interactive Multimedia and Artificial Intelligence,* 3(7), pp. 96-103, 2016.

[19] D. Ivkovic, M. Andric, and B. zrnic, "False Alarm Analysis of the CATM-CFAR in Presence of Clutter Edge", *International Journal of Engineering and Innovative Technology (IJEIT),* 23(1), pp. 66-72, 2014.

[20] D. S. Ranjan and H. K. Moorthy, "Development of Adaptive Algorithm for CFAR in non-homogenous environment", *International Journal of Engineering and Innovative Technology (IJEIT),* 3(1), 2013.

[21] W. K. Abd-Ali and N. Abd-Ullah, "Evaluation of AND-CFAR and OR-CFAR Processors under Different Clutter Models", *Eng. & Tech. Journal,* 31(5), pp. 964-975, 2013.

[22] J.-W. Shin, Y.-K. Seo, D.-W. Do, S.-M. Choi, and H.-N. Kim, "Modified Variability-Index CFAR Detection Robust to Heterogeneous Environment", presented at the International Conference on Systems and Electronic Engineering, Phuket (Thailand), 2012.

[23] K. Ward, R. Tough, and S. Watts, *Sea Clutter Scattering, the K Distribution and Radar Performance*, 2nd Edition, London, United Kingdom: The Institution of Engineering and Technology, 2013.

[24] J. R. Machado Fernández and A. Bueno González, "Clasificación del Clutter Marino utilizando Redes Neuronales Artificiales", Grupo de Radares, Departamento de Telecomunicaciones y Telemática, Instituto Superior Politécnico José Antonio Echeverría (ISPJAE), Habana, Cuba, 2012.

[25] S. Sayama and M. Sekine, "Weibull, Log-Weibull and K-Distributed Ground Clutter Modelling Analyzed by AIC", *IEEE Transactions on Aerospace and Electronic Systems,* 37(3), 2001.

[26] R. Vicen Bueno, M. Rosa Zurera, M. P. Jarabo Amores, and D. de la Mata Moya, "Coherent Detection of Swerling 0 Targets in Sea-Ice Weibull-Distributed Clutter Using Neural Networks", *IEEE Transactions on Instrumentation and Measurement,* 59(12), 2010.

[27] A. Farina, F. Gini, M. V. Greco, and L. Verrazzani, "High Resolution Sea Clutter Data: Statistical Analysis of Recorded Live Data", *IEE Proceedings on Radar, Sonar and Navigation,* 144(3), pp. 121-130, 1997.

[28] Y. Dong, "Clutter Spatial Distribution and New Approaches of Parameter Estimation for Weibull and K-Distributions", DSTO-RR-0274, DSTO Systems Sciences Labpratory, Edingburgh South Asutralia, 2004.

[29] A. N. O'Connor, *Probability Distributions Used in Reliability Engineering*: University of Maryland, 2011.

[30] M. P. McLaughlin, *Compendium of Common Probability Distributions*, 2014.

[31] J. R. Machado Fernández and J. C. Bacallao Vidal, "Modelación de la Distribución K en MATLAB para Aplicaciones de Radar (aceptado para publicación)", *Revista de Ingeniería Electrónica, Automática y Comunicaciones (RIELAC),* 2016.

[32] J. R. Machado Fernández, "Modelación de las Distribuciones Weibull y Log-Normal para Aplicaciones de Radar (aceptado para publicación)", *Ciencias Holguín,* 2016.

[33] S. Sayama and S. Ishii, "Amplitude Statistics of Sea Clutter by MDL Principle", *IEEJ Transactions Fundamentals and Materials,* 132(10), pp. 886-892, 2011.

[34] I. Antipov, "Statistical Analysis of Northern Australian Coastline Sea Clutter Data", Surveillance Systems Division, Electronics and Surveillance Research Laboratory, Edinburgh, South Australia, 2001.

[35] M. Greco, F. Bordoni, and F. Gini, "X-Band Sea-Clutter nonstationarity: Influence of Long Waves," *IEEE Journal of Oceanic Engineering,* 29(2), 2004.

[36] S. Sayama, S. Ishii, and M. Sekine, "Amplitude Statistics of Sea Clutter Observed by L-band Radar", *IEEE Trans-actions on Fundamentals and Material,* 126(6), pp. 438-442, 2006.

[37] J. R. Machado Fernández and J. C. Bacallao Vidal, "Procesador de Promediación con Corrección Adaptativa del FActor de Ajuste mediante

Redes Neuronales Artificiales", presented at the VIII Congreso Internacional de Telemática y Telecomunicaciones, Convención Científica de Ingeniería y Arquitectura '16, Palacio de las Convenciones, La Habana, Cuba, 2016.

[38] J. R. Machado Fernández and J. C. Bacallao Vidal, "Improved Shape Parameter Estimation in Pareto Distributed Clutter with Neural Networks (accepted)", *International Journal of Artificial Intelligence and Interactive Multimedia,* 3(9), 2016.

[39] J. H. Kim and M. R. Bell, "A computationally efficient CFAR algorithm based on a goodness-of-fit test for piecewise homogeneous environments", *IEEE Transactions on Aerospace and Electronic Systems,* 49(3), pp. 1519-1535, 2013.

[40] P. P. Gandhi and S. A. Kassam, "Two level CFAR detector (TL-CFAR) Analysis of CFAR processors in Homogeneous Background", *IEEE Transactions on Aerospace and Electronic Systems,* 24(4), pp. 427-445, 1988.

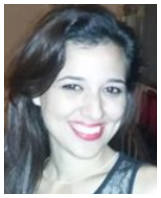
[41] J. R. Machado Fernández, "Estimation of the Relation between Weibull Distributed Sea clutter and the CA-CFAR Scale Factor", *Journal of Tropical Engineering,* 25(2), pp. 19-28, 2015.

[42] J. R. Machado Fernández, "CA-CFAR Scale Factor Optimal Selection for Several Probabilities of False Alarm under Log-Normal Slowly Variable Clutter (under revision)", *Ingeniería y Desarrollo, Colombia,* 2015.

[43] J. R. Machado Fernández and J. C. Bacallao Vidal, "Optimal Selection of the CA-CFAR Adjustment Factor for K Distributed Amplitude Samples with a Fluctuating Shape Parameter (Accepted)", *Nova Scientia,* 2016.

[44] J. R. Machado Fernández and J. C. Bacallao Vidal, "Estimation of the Optimal CA-CFAR Threshold Multiplier in Pareto Clutter with Known Parameters (under revision)", *Revista Entramado,* 2016.

**José Raúl Machado Fernández** received his Telecommunications and Electronics Engineering Degree from the Instituto Superior José Antonio Echeverría (ISPJAE-CUJAE) in 2012. He is currently a Ph.D. student at the same institution. His research topics include teledetection, digital signal processing, sea clutter modeling and the application of artificial intelligence for solving diverse engineering problems.



**Shirley Torres Martínez** received his Telecommunications and Electronics Engineering Degree from the Instituto Superior José Antonio Echeverría (ISPJAE-CUJAE) in 2016. Her research interests include the performance of CFAR processors and the creation of enhanced digital detection proposals.

**Jesús de la Concepción Bacallao Vidal** received his Electrical Engineering Degree and Master Degree from the Instituto Superior Politécnico José Antonio Echeverría (ISPJAE-CUJAE), and the Ph.D. Degree from the Instituto Técnico Militar José Martí in 2003. Since 2013, he has been co-directing the CUJAE radar research team. His research topics include teledetection, the evaluation of CFAR detectors performance and the creation of alternative detection schemes based on signal processing in the moment domain (DRACEC).

# An Effective Hybrid Butterfly Optimization Algorithm with Artificial Bee Colony for Numerical Optimization

Sankalap Arora[1], Satvir Singh[2]

[1]Research Scholar, I.K. Gujral Punjab Technical University, Jalandhar, Punjab, India
[2]Shaheed Bhagat Singh State Technical Campus, Ferozpur, Punjab,India

***Abstract* — In this paper, a new hybrid optimization algorithm which combines the standard Butterfly Optimization Algorithm (BOA) with Artificial Bee Colony (ABC) algorithm is proposed. The proposed algorithm used the advantages of both the algorithms in order to balance the trade-off between exploration and exploitation. Experiments have been conducted on the proposed algorithm using ten benchmark problems having a broad range of dimensions and diverse complexities. The simulation results demonstrate that the convergence speed and accuracy of the proposed algorithm in finding optimal solutions is significantly better than BOA and ABC.**

***Keywords* — Butterfly Optimization Algorithm, Artificial Bee Colony, Hybrid algorithm, Optimization.**

## I. Introduction

Nature-inspired metaheuristic algorithms have received much attention by researchers in the past as they have the ability to solve real world complex problems. These problems require optimal solution in less computational time [1]. Their potential has recognized themselves as numerical optimization techniques in various real world complex problems [2]. These algorithms find their source of inspiration in nature. Various algorithms have been proposed in the past like Particle Swarm Optimization (PSO) [3], Firefly Algorithm (FA) [4], Cuckoo Search (CS) [5], Ant Colony Optimization (ACO) [6] and many more [7,8].

Another example is Butterfly Optimization Algorithm (BOA) which is inspired by the food foraging behavior of the butterflies [9]. The underlying mechanism of BOA is to mimic the food searching abilities of biological butterflies. It has demonstrated better results over other population based algorithm [9]. Recently, chaos is introduced in BOA so as to increase the global search mobility for global optimization problems [10]. BOA is a powerful algorithm in exploitation (i.e., local search) but at times it may trap into some local optima so that it cannot perform global search well.

On the other hand, Artificial Bee Colony (ABC) which was based on the intelligent behavior of the honey bee swarm was proposed by karaboga in 2007 [11]. In the past, it has been applied to various real world problems which demonstrate its superiority over many other algorithms [12].

The main strength of ABC lies in its strategy which allows the solutions to move towards those solutions which have better fitness probability. The aim of the paper is to propose a hybrid algorithm, namely BOA/ABC which have strengths of both the algorithms viz. BOA and ABC. The proposed BOA/ABC algorithm will have advantages of both the algorithms which will enable the algorithm to demonstrate fast convergence and avoid local optima trap problem. The remainder of this paper is organized as follows. Previous work is described in Section II. The conventional BOA and ABC are reviewed in Sections III and IV, respectively. Section V describes the proposed hybrid algorithm BOA/ABC. Experimental results demonstrating the performance of BOA/ABC in comparison with the conventional BOA and ABC over a subset of ten numerical optimization problems are presented in Section VI and conclusions are drawn in Section VII.

## II. Related Work

In the past, various researchers have successfully hybridized optimization algorithms in order to increase performance of algorithms in terms of solution quality. One of the finest example is Differential Evolution (DE) hybridized with Biogeography Based Optimization (BBO) algorithm [13]. DE/BBO utilizes the exploration capability of DE along with the exploitation capability of BBO effectively and efficiently in order to generate the promising solutions. In [14] a hybrid algorithm Harmony Search (HS) and FA is proposed which uses the exploration of HS and the exploitation of FA, efficiently, so that HS/FA shows faster convergence speed than individual algorithms i.e. HS and FA. In [15] PSO is hybridized with Genetic Algorithm (GA) in order to utilize the unique advantages of both the algorithms. The hybridized algorithm efficiently uses the operations of PSO and GA such as single or multiple crossover, mutation, and the PSO formula. Further the selection of these operators is based on fuzzy probability. In [16] a hybrid optimization method named hybrid evolutionary FA is proposed. The algorithm combines the classical FA with the evolutionary operations of DE algorithm aiming to improve the searching accurateness and information sharing among the solutions in the search space.

In [17] a hybrid algorithm combining two swarm intelligence algorithms i.e., ABC and PSO is presented. The proposed method is component based technique in which PSO is augmented with ABC in order to improve the overall efficiency of the algorithm. Another hybridized algorithm is proposed integrating ABC and DE [18]. In this hybridized algorithm the basic drawback of DE i.e. it requires a relatively large population size to avoid premature convergence is overcome by the use of ABC which have proven to demonstrate excellent abilities of global searching. Furthermore, there are various algorithms which have been proposed in the past in order to accelerate the performance of algorithms [19-23].

## III. Butterfly Optimization Algorithm

Butterfly Optimization Algorithm (BOA) is a recently developed nature inspired optimization algorithm by Arora [9, 24]. It finds its source of inspiration in the food foraging behavior of butterflies. Biologically, the butterflies are very efficient in finding their food and the same food searching mechanism is used in the algorithm. In BOA, the butterflies are used as search agents in order to perform optimization. Nature has equipped the butterflies with sense receptors


DOI: 10.9781/ijimai.2017.442

which allow them to sense the fragrance of food and consequently they move towards the particular direction. These sense receptors, also called chemoreceptors, are scattered over the butterfly's body parts like legs, palps, antennae etc. [25]. In BOA, it is assumed that each butterfly is able to generate fragrance with some intensity. This fragrance is further correlated with fitness of the butterfly. It means that whenever a butterfly moves from one position to other particular position in the search space, its fitness will vary accordingly. Further, the fragrance which is being generated by the butterflies is propagated over distance to all the other butterflies in that region. The propagated fragrance is sensed by the other butterflies and a collective social knowledge network is formed. Whenever a butterfly is able to sense fragrance from the best butterfly in the region, it moves towards the best butterfly and this phase is termed as global search phase of BOA. In the second scenario, when a butterfly is not able to sense fragrance of any other butterfly in the search space, it will move randomly in the region and this phase is termed as local search phase in BOA.

The underlying strength of BOA lies in the mechanism of modulating the fragrance in the whole searching process. In order to understand the modulation of fragrance, first, it should be discussed that how any sense like sound, smell, heat, light etc. is processed by a stimulus of a living organism. The basic concept of sensing is dependent on three vital parameters i.e., sensory modality (*c*), stimulus intensity (*I*) and power exponent (*a*). Sensory modality defines the method by which the form of energy is measured and processed by the stimulus. Different modalities/senses can be smell, sound, light, temperature or pressure etc. and in BOA, it is fragrance. *I* represents the magnitude of the physical/actual stimulus and in BOA, it is correlated with the fitness of the butterfly/solution i.e. a butterfly with higher fragrance or greater fitness value attracts other butterflies in the search space. The parameter *a* allows response compression i.e. as the stimulus gets stronger; insects become increasingly less sensitive to the stimulus changes [26, 27].

```
1: Objective function f(x), x_i(i = 1, 2, . . . , n)
2: Generate initial population of butterflies
3: Find the best solution  in the initial population
4: Define switch probability p
5: while stopping criteria not met do
6:     for each butterfly in population do
7:       Draw rand from a uniform distribution in
8:       Calculate fragrance of the butterfly using Eq. 1
9:       if rand < p then
10:            Global search using Eq. 2
11:      else
12:            Do Local search using Eq. 3
13:      end if
14:      Evaluate new solutions
15:      Update Better Solutions.
16:    end for
17:    Find the current best solution
18: end while
19: Output the best solution found.
```

Algorithm1. Pseudocode of Butterfly Optimization Algorithm.

Considering the facts of biological butterflies, the searching phenomenon is based on two important issues: (1) variation of *I*, (2) formulation of *f*. For simplicity, *I* of a butterfly is associated with encoded objective function in BOA. However, *f* is relative i.e. it should be sensed by other butterflies in the search space. Therefore, bearing in mind these concepts of biological butterflies, the fragrance is formulated as a function of the physical intensity of stimulus in BOA [26] as follows:

$$f_i = cI^a \tag{1}$$

where is the perceived magnitude of fragrance, i.e., how stronger the fragrance is perceived by *i*-th butterfly, *c* is the sensory modality, *I* is the stimulus intensity and *a* is the power exponent dependent on modality, which accounts degree of absorption. There are two important phases in the BOA, they are; global search phase and local search phase. In global search phase, the butterfly takes a step towards the fittest butterfly/solution which can be represented as:

$$x_i^{t+1} = x_i^t + ( l\grave{e}vy(\lambda) \times g^* - x_i^t) \times f_i \tag{2}$$

where is the solution vector for *i*-th butterfly in iteration *t*. Here represents the best solution found among all the solutions in current generation. The fragrance of *i*-th butterfly is represented by while step size is represented as . Local search phase can be represented as:

$$x_i^{t+1} = x_i^t + ( l\grave{e}vy(\lambda) \times x_k^t - x_j^t) \times f_i \tag{3}$$

where and are *j*-th and *k*-th butterflies chosen randomly from the solution space. If and belongs to the same swarm and is the step size, then Eq. 3 becomes a local random walk. The food searching process can occur at local as well as global level, so considering this; a switch probability *p* is used in BOA to control the common global search and intensive local search. The above mentioned steps frame the complete algorithm of Butterfly Optimization Algorithm and its pseudocode is presented in Algorithm 1.

## IV. Artificial Bee Colony Algorithm

The artificial bee colony algorithm draws its inspiration from the intelligent behavior of real honey bees [11]. In ABC, The honey bees are divided into three categories: employed bees, onlooker bees, and scout bees. The colony is divided into two halves; first half consists of employed bees whereas the second half consists of onlookers bees. Each solution in the search space consists of a set of optimization parameters which represent a food source population [28]. The number of food sources is equal to the number of employed bees, i.e., one employed bee is there for every food source. Employed bees share the information regarding food sources with onlooker bees, which wait in the hive. Based on the shared information, a food source is selected to be exploited. A few employed bees whose food source has been exhausted are translated to scout bees. After the initialization of population, the whole iterative process of ABC is divided into three phases; (1) employed bee search phase, (2) onlooker bee selection phase and (3) scout bee phase. In the first phase, a candidate food position is produced using the old one. It is achieved by the following equation:

$$v_i^j = x_i^j + \varphi_i^j (x_i^j - x_k^j) \tag{4}$$

where and are the indexes, chosen randomly whereas *k* is determined randomly, different from *i*. represents a random number in the range [-1, 1]. A food source $V_i$ within the neighborhood of every food source site represented by $x_i$, is calculated by modifying one parameter of $x_i$. After the $V_i$ is calculated, it will be evaluated and compared to $x_i$. A greedy selection method is used to select the better one between $x_i$ and $V_i$, depending on fitness values representing the

nectar amount of the food sources at $x_i$ and $V_i$respectively. $V_i$will replace $x_i$ if the fitness of $V_i$ is equal to or better than that of $x_i$, and become a new member of the populations; otherwise $x_i$ is retained. In the second phase, one food source is selected by each onlooker bee on the basis of fitness value obtained from the employed bees. Now, any fitness based probability election strategy can be used like rank based, roulette wheel, tournament selection etc. In the basic version of ABC, roulette wheel selection strategy is used, which can be mathematically defined as:

$$P_i = \frac{fit(x_i)}{\sum_{m=1}^{n} fit(x_m)} \quad (5)$$

In Eq. 5 the fitness of the solution *i* is represented as fit($x_i$). Obviously, higher fitness value means more probability of getting selected. After the selection of food source, onlooker bees will move towards the selected food source and a new candidate position is produced in the neighborhood of the selected food source by using Eq. 4.

1: Generate the initial population $x_i (i = 1, 2, \ldots, n)$
2: Evaluate the fitness of the population
3: **while** stopping criteria not met **do**
4: **for each** employed bee **do**
5: Produce new solution $V_i$ by using Eq. 4
6: Calculate its fitness value fit($V_i$)
7: Apply greedy selection process
8: Calculate $P_i$ for the solution ($x_i$) by Eq. 5
9: **end for**
10: **for each** onlooker bee **do**
11: Select a solution $x_i$ depending on $P_i$
12: Produce new solution $V_j$
13: Calculate its fitness value fit($V_j$)
14: Apply greedy selection process
**15: end for**
16: If there is an abandoned solution for the scout, then replace it with a new solution using Eq. 6
17: Memorize better solutions
18: **end while**
19: Output the best solution found.

Algorithm 2. Pseudocode of Artificial Bee Colony Algorithm.

In the last phase, after the completion of searches by employed and onlooker bees, the algorithm checks whether the source which is exhausted, needs to be abandoned. This means that if a better position cannot be attained in predetermined number of chances i.e. cycles, then that particular food source is assumed as abandoned and a new food source is calculated using:

$$x_i = x_{min} + r(x_{max} - x_{min}) \quad (6)$$

In Eq. 6, *r* is a random number in the range [0, 1] whereas$x_{min}$ and $x_{max}$ represents the respective lower and upper bounds of variable $x_i$. The last phase helps the algorithm to avoid suboptimal solutions. Detailed pseudocode of the ABC algorithm is given in Algorithm 2.

## V. The Proposed Hybrid BOA/ABC Algorithm

In the past, various hybrid optimization algorithms have been developed which demonstrate efficient results [13-23]. So, based on the description of BOA and ABC in the previous sections, the two approaches are combined and a hybrid BOA/ABC algorithm is proposed. Using the strengths of both the approaches, BOA/ABC is able to update the poor solutions which accelerate its convergence speed. BOA and ABC are very efficient in exploring the search space and exploiting the solutions. In the proposed hybrid algorithm, the optimization process of ABC and BOA is used effectively. Therefore, the lack of exploitation is overcome in BOA. In the current study, a modified version of ABC is hybridized with BOA to solve numerical optimization problems [12, 29]. Another modification is that in BOA, lèvy flights are used but in BOA/ABC pseudorandom numbers are used [9]. By incorporating ABC into BOA, the BOA/ABC algorithm is developed as shown in Algorithm 3.

1: Objective function $f(\mathbf{x})$, $x_i (i = 1, 2, \ldots, n)$
2: Generate initial population of individuals
3: Find the best solution in the initial population
4: Define switch probability *p*
5: **while** stopping criteria not met **do**
6: **for each** butterfly in population **do**
7: Draw *r1*and *r2* from a uniform distribution in
8:**if** *r1<0.5***then**
9: Calculate *fragrance* of the butterfly using Eq. 1
10: **if** *r2< p* **then**
11: Global search using Eq. 2
12:**else**
13: Do Local search using Eq. 3
14: **end if**
15: Evaluate new solutions
16: Update Better Solutions.
17:**end if**
18: **else**
19: **for each** employed bee **do**
20: Produce new solution $V_i$ by using Eq. 4
21: Calculate its fitness value fit($V_i$)
22: Apply greedy selection process
23: Calculate $P_i$ for the solution ($x_i$) by Eq. 5
24:**end for**
25: **for each** onlooker bee **do**
26: Select a solution $x_i$ depending on $P_i$
27: Produce new solution $V_j$
28: Calculate its fitness value fit($V_j$)
29: Apply greedy selection process
30:**end for**
31: If there is an abandoned solution for the scout, then replace it with a new solution using Eq. 6
32: Memorize better solutions
33: **end else**
34: **end while**
35: Output the best solution found

Algorithm 3. Pseudocode of the proposed BOA/ABC Algorithm.

In comparison to the original BOA, the proposed approach needs a very small amount of computational cost additionally. The incorporation of ABC in BOA enables the proposed algorithm to avoid

the local optima trap problem and increase its convergence speed. The proposed BOA/ABC algorithm is aimed at hybridizing components from both BOA and ABC in order to have an algorithm that easily solve separable problems as BOA while having a rotationally invariant behavior as ABC, at the same time. Detailed pseudocode of the BOA/ABC algorithm is given in Algorithm 3.

## VI. Benchmark Problems and Experiment Settings

Every novel optimization algorithm must be subject to a testbed of benchmark functions in order to validate the algorithm. Ideally, the selected test functions should have characteristics alike to those of a real-world problem in order to better assess an algorithm [30]. However, no standard benchmark function testbed available. Still there are many benchmark functions which are well known and recommended by various researchers in the past [31]. So in this study, ten benchmark functions are used to validate the proposed algorithm. All the functions used in this study are minimization problems. The benchmark functions which are selected in this study are chosen in such a way that the proposed algorithm is tested on almost all types of problems. Considering this viewpoint a diverse subset of benchmark functions is chosen in this study. This subset can be classified into four major categories. In the first category, function can either have single optima or multiple optima. In second category, the number of dimensions can be low or high. High dimensional problems are very difficult to solve that's why most of the benchmark functions used in this work are high dimensional.

Another category is that some functions are separable and some are non-separable. In the last category, functions with noisy data are used. These functions are alike real world problems which contains noisy data which makes them difficult to solve. New algorithms must be tested on all these kinds of test functions in order to properly validate and demonstrate the efficiency of the algorithm. These benchmark functions are described in Table 1 along with their dimensions and range.

Rigorous nonparametric statistical framework is used to compare the performance of the proposed algorithm with BOA and ABC algorithm. All initial solutions of the population are randomly generated for each run of the algorithm. The population size is fixed to 30 for all the algorithms. In order to avoid discrepancy due to

TABLE I.

Benchmark functions used in the present computational analysis.

| S. No. | Benchmark functions | Formula | Dimensions | Range | Optima |
|---|---|---|---|---|---|
| $f_1$ | Beale | $P_i = \frac{fit(x_i)}{\sum_{m=1}^{n} fit(x_m)}$ | 2 | (-4.5,4.5) | 0 |
| $f_2$ | Cigar | $f(x) = x_1^2 + 10^6 \sum_{i=2}^{n} x_i^2$ | 30 | (-10,10) | 0 |
| $f_3$ | Easom | $f(x) = -\cos(x_1)\cos(x_2)\exp(-(x_1-\pi)^2-(x_2-\pi)^2)$ | 2 | (-100,100) | -1 |
| $f_4$ | Griewank | $f(x) = 1 + \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}}\right)$ | 30 | (-600, +600) | 0 |
| $f_5$ | Quartic function with noise | $f(x) = \sum_{i=1}^{n} i{x_i}^4 + rand(0,1)$ | 30 | (-1.28 1.28) | 0 |
| $f_6$ | Rastrigin | $f(x) = \sum_{i=1}^{n} (x_i^2 - 10\cos(2\pi x_i) + 10)$ | 30 | (-5.12, 5.12) | 0 |
| $f_7$ | Shubert | $f(x) = \left(\sum_{i=1}^{5} icos((i+1)x_1 + i)\right)\left(\sum_{i=1}^{5} icos((i+1)x_2 + i)\right)$ | 2 | (-10,10) | -186.73 |
| $f_8$ | Step | $f(x) = \sum_{i=1}^{n} (\lfloor x_i \rfloor + 0.5)^2$ | 30 | (-100,100) | 0 |
| $f_9$ | Levy | $f(x) = sin^2(\pi w_1) + \sum_{i=1}^{d-1} (w_i - 1)^2 [1 + 10\sin^2(\pi w_i + 1)] + (w_d - 1)^2[1 + \sin^2(2\pi w_d)]$ | 30 | (-10,10) | -21.5023 |
| $f_{10}$ | Power Sum | $f(x) = \sum_{i=1}^{d} \left(\left[\sum_{j=1}^{d} x_j^i\right] - b_i\right)^2 \;;\; b = [8\ 18\ 44\ 144]$ | 4 | (0,d) | 0 |

the stochastic properties of the algorithms, 30 independent runs for each optimization algorithm having 30 different initial population. The proposed algorithm is implemented in C++ and compiled using Qt Creator 2.4.1 (MinGW) under Microsoft Windows 8 operating system. All simulations are carried out on a computer with an Intel(R) Core(TM) i5-3210@2.50Ghz CPU.

## VII. Results and Discussions

Our proposed BOA/ABC approach is compared with classical BOA and ABC algorithm in order to demonstrate its superiority. The values of mean and standard deviation of all the algorithms are presented in Table 2. The best values are highlighted in bold. According to the simulation results, it can be analyzed that BOA/ABC is significantly better than BOA on all the ten benchmark functions used in the study. It can be observed from the Table 2 that the BOA/ABC has ability to converge faster and escape from local optima. The underlying reason behind the better performance of BOA/ABC can be explained by the fact that exploration of BOA is good but it lacks the exploitation. Compared to the ABC algorithm, the proposed BOA/ABC can converge faster, if proper terminating conditions are set.

For example, after a number of iterations the best fitness value is not improved whereas the computation time can be reduced significantly. BOA has a disadvantage that exploitation to found good solutions is very bad whereas the exploration of solution of BOA algorithm is good. On the other hand the BOA/ABC can overcome this shortcoming by incorporating the strengths of ABC which allows the BOA/ABC to avoid the local optima trap problem and on the same side it maintains the overall good solution quality during the optimization process. The proposed BOA/ABC requires less population size and needs less computational time to reach global optima.

The proposed BOA/ABC demonstrates its better search ability than the individual algorithms i.e. BOA and ABC algorithm. Integrating the BOA and ABC's local and global search abilities, BOA/ABC hybrid algorithm demonstrate better global and local search ability than the original BOA. If we consider the same population size of BOA/ABC and ABC, then BOA/ABC obtains optimized results at least as good as the results obtain by ABC algorithm. When the population size of BOA/ABC is less than ABC, then the BOA/ABC performs better whereas when ABC has large population size than BOA/ABC, ABC shows better performance.

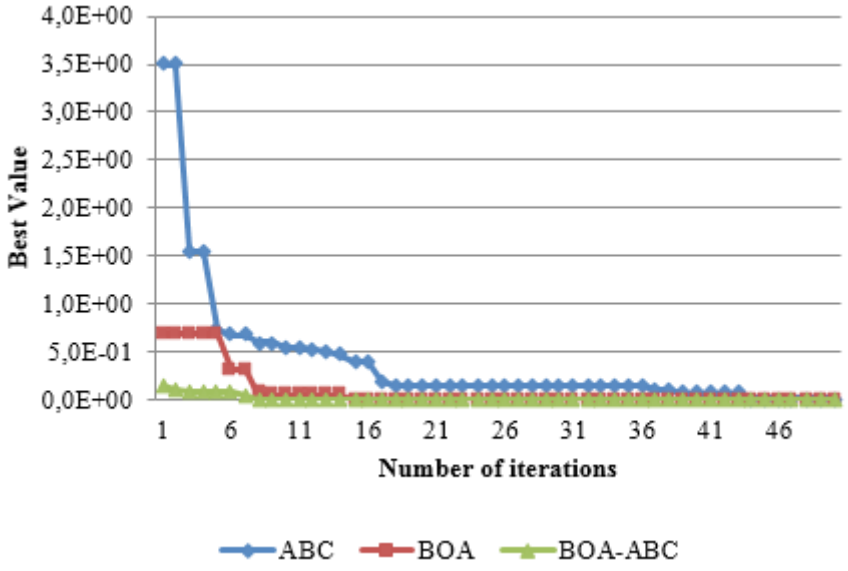



Fig.1. Convergence curves of $f_1$ function.

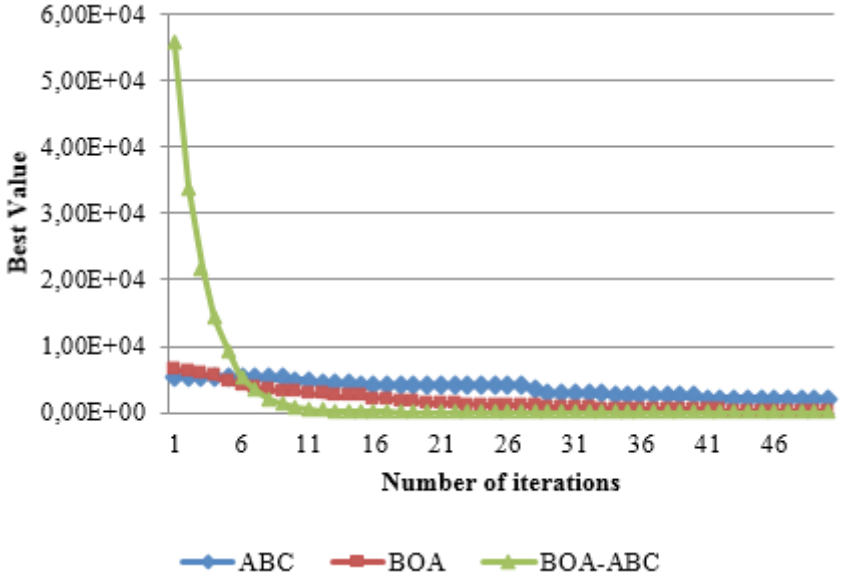



Fig. 2. Convergence curves of $f_8$ function.

TABLE II.

STATISTICAL RESULTS OF DIFFERENT METHODS.

| | | ABC | BOA | BOA/ABC |
|---|---|---|---|---|
| $f_1$ | Mean | 3.636E-15 | 9.322E-08 | **1.537E-15** |
| | Std. Dev | 1.001E-14 | 1.626E-07 | **2.687E-15** |
| $f_2$ | Mean | 3.679E-16 | 1.830E-07 | **1.162E-16** |
| | Std. Dev | 7.568E-17 | 3.010E-08 | **1.525E-17** |
| $f_3$ | Mean | -1.000E+00 | -1.850E-04 | **-1.00E+00** |
| | Std. Dev | 1.826E-07 | 1.013E-03 | **0.000E+00** |
| $f_4$ | Mean | 3.896E-16 | 1.539E-05 | **1.153E-16** |
| | Std. Dev | 6.874E-17 | 2.131E-06 | **6.220E-18** |
| $f_5$ | Mean | 1.983E-01 | 1.333E-01 | **2.622E-03** |
| | Std. Dev | 5.231E-02 | 1.142E-02 | **8.868E-04** |
| $f_6$ | Mean | 2.632E-08 | 2.689E+02 | **2.309E-15** |
| | Std. Dev | 1.440E-07 | 2.111E+01 | **1.054E-15** |
| $f_7$ | Mean | -8.481E+01 | -4.741E+01 | **-8.481E+01** |
| | Std. Dev | 1.445E-14 | 2.145E+01 | **2.340E-03** |
| $f_8$ | Mean | 0.000E+00 | 5.573E+01 | **0.000E+00** |
| | Std. Dev | **0.000E+00** | 9.036E+00 | **0.000E+00** |
| $f_9$ | Mean | 4.860E-16 | 8.838E-01 | **4.116E-16** |
| | Std. Dev | 1.151E-16 | 1.964E-01 | **1.134E-16** |
| $f_{10}$ | Mean | 2.513E-02 | 2.542E-01 | **6.028E-03** |
| | Std. Dev | 2.081E-02 | 2.405E-10 | **5.369E-03** |

The strategy of BOA/ABC is designed in such a way that it balances the local and global search processes of BOA and further it should enhance the performance of BOA in order to use limited computation resources. The faster convergence of BOA/ABC is shown in the Fig.1 and Fig.2. In addition, the BOA/ABC allows the BOA to enhance the exploration and exploitation capabilities. It is worth mentioning that using ABC in BOA enables the algorithm to improve the poor solutions and meanwhile, it is able to avert the good solution from being destroyed during the optimization process. Hence, the better performance of BOA/ABC.

Due to limitations of space, few representative convergence graphs on benchmark functions are shown in Fig. 1-2. By carefully analyzing the Fig. 1-2, it can be observed that the solution quality of BOA/ABC improves continuously during the optimization process. Moreover, in the whole iteration process, BOA/ABC keeps an edge over the BOA and ABC algorithm. The reason might be that BOA/ABC better manages the global search and local search process in the optimization process. This indicates that BOA/ABC has the ability to escape from poor local optima and locate a good near-global optimum. Overall, the performance of BOA/ABC is highly competitive with ABC and BOA, particularly for the high-dimensional and noisy problems. Besides, BOA/ABC is much better than BOA for all the problems.

In order to better demonstrate the performance of the proposed algorithm, pair-wise comparison of the simulation results is done using the Wilcoxon signed rank test. In this method, the algorithm having a lower rank is considered better on the particular benchmark function. If Algo-1 and Algo-2 perform better than Algo-3, then it can be concluded that Algo-1 and

Algo-2 outperform Algo-3 on the specified benchmark function. However, if Algo-1outperforms Algo-2 but both Algo-1 and Algo-2 perform similarly with Algo-3, all of Algo-1, Algo-2, and Algo-3 are positioned in the same rank. The order of algorithms within same rank is alphabetical and has no implication on performance. In Table 3, Algo-1 < Algo-2 stands for "Algo-1 shows better performance than Algo-2", and Algo-1 ≈ Algo-2 stands for "Algo-1 and Algo-2 performs similarly" on the particular benchmark function. On the basis of the pair-wise findings, partial orderings of the algorithms is constructed as shown in Table 3 [33].

TABLE III.

PAIR-WISE WILCOXON SIGNED RANK TEST RESULTS

| Function | Wilcoxon Signed Rank Test Order |
|---|---|
| $f_1$ | BOA/ABC< ABC < BOA |
| $f_2$ | BOA/ABC< ABC < BOA |
| $f_3$ | ABC≈BOA/ABC< BOA |
| $f_4$ | BOA/ABC< ABC < BOA |
| $f_5$ | BOA/ABC< BOA < ABC |
| $f_6$ | BOA/ABC< ABC < BOA |
| $f_7$ | ABC≈BOA/ABC< BOA |
| $f_8$ | ABC≈BOA/ABC< BOA |
| $f_9$ | BOA/ABC< ABC < BOA |
| $f_{10}$ | BOA/ABC< ABC < BOA |

According to the results shown in the Table 3, BOA/ABC indicates its outstanding capability in fast converging to the global optimum while avoiding premature convergence. It can be analyzed that BOA/ABC demonstrates better statistical results as well as higher robustness. This statement is supported by the results in Table 2 and Table 3 which shows the better performance of BOA/ABC over other algorithms.

With the intention to better demonstrate the performance of proposed BOA/ABC, the results in the Wilcoxon signed rank test are used to further rank the algorithms according to their performance [32]. For every benchmark function, the best/first algorithm is assigned value 1; the second best is assigned value 2 and so on. In case of ties,

an average rank is assigned to the algorithms involved in the tie. For example, in case of function $f_3$, ABC and BOA/ABC shares same rank, and they are first and second rank algorithm for the particular function. As a result, an average rank value of 1.5 is assigned to these two algorithms. Further, the rank values of each algorithm is summed together in order to have overall assessment of the algorithm in solving general optimization problems. Similar evaluation methods have been adopted in the past by various researchers for testing of different metaheuristic algorithms [34].The test results are presented in Table 4.The results obtained by the statistical assessment analysis supports our previous observations. In all the benchmark functions used in this study, BOA/ABC possesses a superior position on all the functions. It can be observed from Table 3 and Table 4 that BOA/ABC has the most stable position on the benchmark functions. The performance of BOA and ABC is reasonable but neither of these two catches up with BOA/ABC. So it can be concluded that in general BOA/ABC has better optimization performance in terms of efficient results and stability.

TABLE IV.
Rank summary of statistical assessment results

| Function | ABC | BOA | BOA/ABC |
|---|---|---|---|
| $f_1$ | 2 | 3 | 1 |
| $f_2$ | 2 | 3 | 1 |
| $f_3$ | 1.5 | 3 | 1.5 |
| $f_4$ | 2 | 3 | 1 |
| $f_5$ | 3 | 2 | 1 |
| $f_6$ | 2 | 3 | 1 |
| $f_7$ | 1.5 | 3 | 1.5 |
| $f_8$ | 1.5 | 3 | 1.5 |
| $f_9$ | 2 | 3 | 1 |
| $f_{10}$ | 2 | 3 | 1 |
| Sum | **19.5** | **29** | **11.5** |

Furthermore, in order to demonstrate the efficiency of the proposed BOA/ABC algorithm in terms of finding the optimum value, an experiment in conducted on the particular algorithms. In this experiment, all the algorithms were initialized in regions that include the global optimum for a fair evaluation. The algorithms were run for 50 times to catch their stochastic properties. The goal of this experiment is not to find the global optimum values but to find out the potential of the algorithms. The success rate of algorithm is defined in Eq. (7) which has been used for comparison in this study.

$$S = 100\frac{NT_{successful}}{NT_{all}}\Big|_{Qlevel} \quad (7)$$

Where  is the number of trials, which found the solution on thein the allowable maximum iteration which is set to 500 in this case study. is the number of alltrials. The results of this experiment are shown in the Table 5. It can be analyzed from these simulation results that BOA/ABC has the ability to reach the global optima in almost every time it is executed. The simulation results of ABC and BOA are also satisfactory but not as good as of BOA/ABC. The underlying reason behind the better performance of BOA/ABC is that incorporating the ABC in BOA has improved the reliability of the global optimality and it has also enhanced the quality of the results.

TABLE V.
Success rates of different methods

| Function | ABC | BOA | BOA/ABC |
|---|---|---|---|
| $f_1$ | 82 | 78 | 100 |
| $f_2$ | 80 | 75 | 100 |
| $f_3$ | 95 | 88 | 95 |
| $f_4$ | 80 | 75 | 93 |
| $f_5$ | 80 | 87 | 98 |
| $f_6$ | 78 | 70 | 90 |
| $f_7$ | 92 | 90 | 94 |
| $f_8$ | 90 | 86 | 92 |
| $f_9$ | 89 | 86 | 94 |
| $f_{10}$ | 90 | 87 | 95 |

These simulation results do not indicate that BOA/ABC is "better" than BOA and ABC. As this common statement would be an unjustified remark, particularly in view of the no free lunch theorem [35]. However, these results illustrate that BOA/ABC demonstrates superior performance than BOA and ABC on the particular benchmark functions used in this study. The simulation results point towards the fact that BOA/ABC is competitive with the other algorithms and it has the ability to provide efficient results on real world complex problems.

## VIII. Conclusion

In the present work, a hybrid BOA/ABC is proposed for numerical optimization problems. Although BOA is an effective algorithm, however, it faces the problem of entrapment into local optima. In order to solve this problem, ABC is introduced in BOA. ABC enables the proposed BOA/ABC to improve the poor solutions and along with it is able to save the good solutions in order to maintain the diversity of the population. Ten benchmark functions are used to investigate the performance of BOA/ABC. The results demonstrated that BOA/ABC makes better use of exploration and exploitation of the butterflies' information, than original BOA. The convergence of BOA/ABC is faster than original BOA and it shows superior results on higher dimensional problems.

## Acknowledgement

The authors wish to acknowledge the Department of RIC, I.K. Gujral Punjab Technical University, Kapurthala, Punjab, India

## References

[1] Yang, Xin-She. *Nature-inspired metaheuristic algorithms*. Luniver press, 2010.

[2] Vasant, Pandian. *Handbook of Research on Artificial Intelligence Techniques and Algorithms, 2 Volumes*. Information Science Reference-Imprint of: IGI Publishing, 2015.

[3] Kennedy, James. "Particle swarm optimization." In *Encyclopedia of machine learning*, pp. 760-766. Springer US, 2011.

[4] Yang, Xin-She. "Firefly algorithm, stochastic test functions and design optimisation." *International Journal of Bio-Inspired Computation* 2, no. 2 (2010): 78-84.

[5] Yang, Xin-She, and Suash Deb. "Cuckoo search via Lévy flights." In *Nature & Biologically Inspired Computing, 2009. NaBIC 2009. World Congress on*, pp. 210-214. IEEE, 2009.

[6] Dorigo, Marco, Mauro Birattari, and Thomas Stutzle. "Ant colony optimization." *IEEE computational intelligence magazine* 1, no. 4 (2006): 28-39.

[7] Onwubolu, Godfrey C., and B. V. Babu. *New optimization techniques in*

*engineering*. Vol. 141. Springer, 2013.

[8] Gosavi, Abhijit. "Simulation-based optimization." Parametric Optimization Techniques and Reinforcement Learning. Kluwer Academic Publishers (2003).

[9] Arora, Sankalap, and Satvir Singh. "Butterfly algorithm with Lèvy Flights for global optimization." In *Signal Processing, Computing and Control (ISPCC), 2015 International Conference on*, pp. 220-224. IEEE, 2015.

[10] Arora, Sankalap, and Satvir Singh. "An improved butterfly optimization algorithm with chaos." *Journal of Intelligent &Fuzzy Systems* (2016)

[11] Karaboga, Dervis, and Bahriye Basturk. "A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm." *Journal of global optimization* 39, no. 3 (2007): 459-471.

[12] Karaboga, Dervis, Beyza Gorkemli, Celal Ozturk, and Nurhan Karaboga. "A comprehensive survey: artificial bee colony (ABC) algorithm and applications." *Artificial Intelligence Review* 42, no. 1 (2014): 21-57.

[13] Gong, Wenyin, Zhihua Cai, and Charles X. Ling. "DE/BBO: a hybrid differential evolution with biogeography-based optimization for global numerical optimization." *Soft Computing* 15.4 (2010): 645-665.

[14] Guo, Lihong, et al. "An effective hybrid firefly algorithm with harmony search for global numerical optimization." *The Scientific World Journal* 2013 (2013).

[15] Sahnehsaraei, M. Andalib, et al. "A hybrid global optimization algorithm: particle swarm optimization in association with a genetic algorithm." *Complex System Modelling and Control Through Intelligent Soft Computations*. Springer International Publishing, 2015. 45-86.

[16] Abdullah, Afnizanfaizal, et al. "A new hybrid firefly algorithm for complex and nonlinear problem." *Distributed Computing and Artificial Intelligence*. Springer Berlin Heidelberg, 2012. 673-680.

[17] El-Abd, Mohammed. "A hybrid ABC-SPSO algorithm for continuous function optimization." *Swarm Intelligence (SIS), 2011 IEEE Symposium on*. IEEE, 2011.

[18] Li, Li, Fangmin Yao, Lijing Tan, Ben Niu, and Jun Xu. "A novel DE-ABC-based hybrid algorithm for global optimization." In International Conference on Intelligent Computing, pp. 558-565. Springer Berlin Heidelberg, 2011.

[19] Fleurent, Charles, and Jacques A. Ferland. "Genetic and hybrid algorithms for graph coloring." *Annals of Operations Research* 63, no. 3 (1996): 437-461.

[20] Prosser, Patrick. "Hybrid algorithms for the constraint satisfaction problem." *Computational intelligence* 9, no. 3 (1993): 268-299.

[21] Kalra, Shifali, and Sankalap Arora. "Firefly Algorithm Hybridized with Flower Pollination Algorithm for Multimodal Functions." In *Proceedings of the International Congress on Information and Communication Technology*, pp. 207-219. Springer Singapore, 2016.

[22] Gupta, Samiti, and Sankalap Arora. "A Hybrid Firefly Algorithm and Social Spider Algorithm for Multimodal Function." In *Intelligent Systems Technologies and Applications*, pp. 17-30. Springer International Publishing, 2016.

[23] Fukuda, Sho, Yuuma Yamanaka, and Takuya Yoshihiro. "A Probability-based Evolutionary Algorithm with Mutations to Learn Bayesian Networks." *International Journal of Interactive Multimedia and Artificial Intelligence* 3.1 (2014): 7-13.

[24] Arora, Sankalap, and Satvir Singh. "A conceptual model of Butterfly algorithm" In *Latest initiatives and Innovations in Communication and Electronics (IICE), 2015 National Conference on*, pp. 69-72. , 2015.

[25] Blair, Robert B., and Alan E. Launer. "Butterfly diversity and human land use: Species assemblages along an urban grandient." *Biological conservation* 80, no. 1 (1997): 113-125.

[26] Zwislocki, Jozef J. *Sensory neuroscience: Four laws of psychophysics*. Springer Science & Business Media, 2009.

[27] Stevens, Stanley Smith. *Psychophysics*. Transaction Publishers, 1975.

[28] Karaboga, Dervis, and Bahriye Akay. "A comparative study of artificial bee colony algorithm." *Applied mathematics and computation* 214, no. 1 (2009): 108-132.

[29] Akay, Bahriye, and Dervis Karaboga. "A modified artificial bee colony algorithm for real-parameter optimization." *Information Sciences* 192 (2012): 120-142.

[30] Meza, Joaquín, Helbert Espitia, Carlos Montenegro, and Rubén González Crespo. "Statistical analysis of a multi-objective optimization algorithm based on a model of particles with vorticity behavior." *Soft Computing* (2015): 1-16

[31] Liang, J. J., B. Y. Qu, P. N. Suganthan, and Q. Chen. "Problem definitions and evaluation criteria for the CEC 2015 competition on learning-based real-parameter single objective optimization." *Technical Report201411A, Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou China and Technical Report, Nanyang Technological University, Singapore* (2014).

[32] Maesono, Yoshihiko. "Competitors of the Wilcoxon signed rank test." *Annals of the Institute of Statistical Mathematics* 39, no. 1 (1987): 363-375.

[33] Houck, Christopher R., Jeff Joines, and Michael G. Kay. "A genetic algorithm for function optimization: a Matlab implementation." *NCSU-IE TR 95*, no. 09 (1995).

[34] Lam, Albert YS, Victor OK Li, and J. Q. James. "Real-coded chemical reaction optimization." *IEEE Transactions on Evolutionary Computation* 16, no. 3 (2012): 339-353.

[35] Ho, Yu-Chi, and David L. Pepyne. "Simple explanation of the no-free-lunch theorem and its implications." *Journal of optimization theory and applications* 115, no. 3 (2002): 549-570.

**Sankalap Arora** was born on Nov 3, 1988. He received his Bachelor's degree (B.Tech.) and Master's degree (M.Tech.) from Lovely Professional University, Phagwara, Punjab (India) with specialization in Computer Science & Engineering. He is currently pursuing his Doctoral degree (Ph.D.) from Punjab technical University, Kapurthala, Punjab (India). He has more than 5 years' research and teaching experience. His fields of special interest include Nature Inspired Algorithms, Engineering Design problems and Wireless Sensor Networks. He has published nearly 20 research papers in reputed International Journals and Conferences.

**Satvir Singh** was born on Dec 7, 1975. He received his Bachelor's degree (B.Tech.) from Dr. B.R.Ambedkar National Institute of Technology, Jalandhar, Punjab (India) with specialization in Electronics & Communication Engineering in year 1998,Master's degree (M.E.) from Delhi Technological University (Formerly, Delhi College of Engineering), Delhi (India) with distinction in Electronics & Communication Engineering in year 2000 and Doctoral degree (Ph.D.) from Maharshi Dayanand University, Rohtak, Haryana (India) in year 2011. He has more than 13 years' research and teaching experience years of teaching experience. His fields of special interest include Evolutionary Algorithms, High Performance Computing, Type-1 & Type-2 Fuzzy Logic Systems and Wireless Sensor Networks for solving engineering problems. He is active member of an editorial board of International Journal of Electronics Engineering and published nearly 30 research papers in International Journals and Conferences

# A Solution to the N-Queens Problem Using Biogeography-Based Optimization

Ali Habiboghli and Tayebeh Jalali

*Computer Science & Engineering Department, Khoy Branch, Islamic Azad University, Khoy, Iran*

***Abstract* — Biogeography-based Optimization (BBO) is a global optimization algorithm based on population, governed by mathematics of biogeography, and dealing with geographical distribution of biological organisms. The BBO algorithm was used in the present study to provide a solution for the N-queens problem. The performance of the proposed algorithm has been evaluated in terms of the quality of the obtained results, cost function, and execution time. Furthermore, the results of this algorithm were compared against those of genetic and particle swarm algorithms.**

***Keywords* — N-Queens Problem, Biogeography-based Optimization Algorithm, Evolutionary strategy.**

## I. Introduction

Optimization is a process of making something better. During this process the primary conditions are examined by different methods and the obtained data are used to make improvement in an idea or a method. Optimization is a mathematical tool which includes providing answers for many questions on how to solve different problems [2, 3]. Optimization deals with finding the best possible answer for a problem. Meta-heuristic algorithms simulate natural processes through stochastic optimization methods. The shorter the time and the lower the required quality of the answer, the better applying meta-heuristic approaches will be. In fact, meta-heuristic algorithms are subsets of approximate optimization algorithms which have exiting solutions from local optimization and are applicable on a wide range of problems [4].

Biogeography-based algorithm (BBO) is an evolutionary algorithm that was originally introduced by Dan Simon in 2008 [1]. BBO is a novel global optimization algorithm based on biogeography theory within the domain of smart optimization dealing with geographical distribution of biological organisms. Mathematical models of biogeography describe speciation (the evolution of new species), the migration of species between islands, and the extinction of species. This algorithm has been used for single-target optimization of many criterion functions [7, 9] and solving a wide range of real-life optimization problems such as sensor selection for estimation of airplane engine efficiency [1] or categorization of satellite images [5].

Various approaches have been introduced so far as solutions to the N-queens problem, some of which are mentioned below:

In the paper [10] the heuristic algorithms have been used for solving the N-queens problem. Meta-heuristic algorithms such as genetic algorithm, simulated refrigeration algorithm, and forbidden search have been applied to solve this problem. The reference [11] describes application of the genetic algorithm with various patterns for solving the N-queens problem. The paper [12] deals with the effect of the particle swarm algorithm on generating an optimum solution for the problem.

Shuihua et al. [16] proposed two novel machine-learning based classification methods. Their developed system includes, wavelet entropy (WE), principal component analysis (PCA), feed forward neural network (FNN) trained by fitness-scaled chaotic artificial bee colony (FSCABC) and biogeography- based optimization (BBO), respectively. They show that their proposed method is effective for fruit classification.

Mehran Tamjidy et al. [17] present an evolutionary optimization algorithm based on geographic distribution of biological organism to deal with hole-making process problem. The aim of their study is to minimize the none-productive time, including tool travelling time and tool switching time, by using biogeography-based optimization algorithm. Their obtained results show that the proposed algorithm can efficiently improve the solution quality in words of minimizing none-productive time.

Vanitha and Thanushkodi [18] used an effective biogeography-based optimization algorithm to solve economic load dispatch (ELD) problem. Their proposed algorithm has been applied to ELD problem for verifying its feasibility and the convergence of EBBO is presented.

Saremi and Mirjalili [19] used biogeography-based optimization algorithm for integrating chaos. Their results show that chaotic maps are able to improve the performance of biogeography-based optimization algorithm.

There are different methods for solution of N-queen problem [13, 14, 15]. Also it has been applied different optimization methods for solution of N-queen problem such as particle swarm optimization (PSO), Genetic algorithms (GA) but BBO algorithm has not been used for this aim yet.

In this paper we used BBO algorithm for solution of N-queen problem and we compare performance of BBO algorithm with PSO and GA algorithms. Our comparisons are based on cost and execution time in mention of queen numbers.

The second section of the present paper deals with general principles of BBO evolutionary algorithm. The third section is focused on evaluation of the BBO algorithm and its comparison against GA and PSO algorithms. The fourth section is a conclusion to this study and ultimately some suggestions for future trends are given at the end of the paper.

## II. Biogeography-Based Optimization

BBO is an evolutionary algorithm based on population, which is inspired by animals and birds migration between islands. In fact, biogeography deals with geographical distribution of species. Islands that are hospitable to live have a high habitat suitability index (HSI). Features that influence HSI include rainfall, vegetative diversity, topographic characteristics, land area, and temperature. These features are called suitability index variables (SIVs). Islands with a high HSI have many species that emigrate to nearby habitats. In terms of habitability, SIVs are the independent variables and HSI is the dependent variable.



DOI: 10.9781/ijimai.2017.443

Islands with a high HSI have a low immigration rate because they already host many other species and cannot receive new members. Islands with a low HSI have a high immigration rate because of their low populations. Immigration of new species to habitats with high HSI may lead to increase in HSI in that area, since hospitability of a place depends on its geographical diversity. Application of biogeography for optimization was initially focused on using a natural process to solve an optimization problem. Similar to other evolutionary algorithms such as GA which are always associated with certain operators like mutation and selection, in BBO algorithm the operators of migration and mutation result in desirable changes in the trend of generations' creation. As the suitability of a habitat rises, the number of species and emigration rate increase, whereas immigration rate decreases [1].

Any habitat (solution) in BBO has an immigration rate (ג) and an emigration rate (μ) which is used in the form of possibility to share data between the solutions and are calculated by the following relations:

$$\lambda_i = I\left(1-\frac{k(i)}{n}\right) \quad (1)$$

$$\mu_i = E\left(\frac{k(i)}{n}\right) \quad (2)$$

where I and E represent maximum rates of immigration and emigration respectively, k(i) is the number of species in $i^{th}$ habitat which ranges from 1 to n that is the number of members in a certain population (n for the best solution and 1 for the worst solution).

Sudden mutations may change HSI in a habitat. Furthermore, they may make the number of species deviate from its balanced value. This issue is modeled in BBO as SIV mutation and the mutation rate is determined by probability of the number of existing species in the habitat.

$$m(s) = m_{max}\left(\frac{1-Ps}{Pmax}\right) \quad (3)$$

Where $m_{max}$(maximum rate of mutation) is defined by user. Ps represents the probability that the habitat supports exactly S species. Excellent and hostile habitats will tend to mutation and change. This pattern of mutation results in more diversity in population.

The stages of BBO algorithm can be summarized as follow:

1. Parameters initialization (assign initial values to the parameters);
2. Stochastic generation of primary solutions (habitats);
3. Obtaining the number of species, S, immigration rate, ג , and emigration rate, μ, related to each habitat from HSI;
4. Modification of non-elite habitat based on immigration and emigration rates, and recalculation of migration operator and HSI;
5. For each habitat the probability of the number of habitants is modified. Subsequently, each non-elite habitat is mutated, and then the mutation operator and the amount of HSI are calculated again for each habitat.
6. Iteration should start from step 3.
7. This cycle may end after predefined number of generations or when an acceptable solution is obtained.

TABLE I.
BBO ALGORITHM PERFORMANCE BY STEP IMPLEMENTATION OF 500 AND POPULATION OF 30

| Number of Queens | | 5×5 | 8×8 | 10×10 | 20×20 | 100×100 | 200×200 |
|---|---|---|---|---|---|---|---|
| Cost | Mutation = 0.02 | 05055/0 | 48750/0 | 88235/0 | 6491/8 | 157/63 | 874/214 |
| Execution time | | 146199/0 | 202645/0 | 263386/0 | 2181/0 | 8316/0 | 0146/2 |
| Cost | Mutation = 0.04 | 04978/0 | 37754/0 | 48995/0 | 3467/8 | 339/64 | 678/213 |
| Execution time | | 170528/0 | 222510/0 | 282753/0 | 2213/0 | 1184/0 | 9796/1 |
| Cost | Mutation = 0.06 | 05805/0 | 18492/0 | 58131/0 | 9745/8 | 398/63 | 870/213 |
| Execution time | | 152917/0 | 215785/0 | 288850/0 | 1986/0 | 8144/0 | 976/1 |
| Cost | Mutation = 0.08 | 05709/0 | 22720/0 | 29223/0 | 2190/8 | 673/63 | 564/214 |
| Execution time | | 145144/0 | 211193/0 | 280975/0 | 2419/0 | 8416/0 | 0033/2 |

TABLE II.
BBO ALGORITHM PERFORMANCE BY STEP IMPLEMENTATION OF 500 AND POPULATION OF 50

| Number of Queens | | 5×5 | 8×8 | 10×10 | 20×20 | 100×100 | 200×200 |
|---|---|---|---|---|---|---|---|
| Cost | Mutation = 0.02 | 04901/0 | 13687/0 | 64436/0 | 2285/8 | 239/63 | 786/213 |
| Execution time | | 260871/0 | 394331/0 | 458567/0 | 4436/0 | 6989/1 | 7981/4 |
| Cost | Mutation = 0.04 | 04978/0 | 10688/0 | 46647/0 | 0923/9 | 987/63 | 564/214 |
| Execution time | | 278102/0 | 388269/0 | 473595/0 | 5561/0 | 6927/1 | 2177/4 |
| Cost | Mutation = 0.06 | 04901/0 | 21222/0 | 42099/0 | 9823/8 | 765/63 | 574/213 |
| Execution time | | 266770/0 | 388476/0 | 460342/0 | 5986/0 | 6929/1 | 2627/4 |
| Cost | Mutation = 0.08 | 04940/0 | 11803/0 | 88235/0 | 4534/8 | 459/63 | 788/213 |
| Execution time | | 255299/0 | 380295/0 | 466751/0 | 5973/0 | 7208/1 | 9828/4 |

TABLE III
GA ALGORITHM PERFORMANCE BY STEP IMPLEMENTATION OF 500 AND POPULATION OF 30

| Number of Queens | | 5×5 | 8×8 | 10×10 | 20×20 | 100×100 | 200×200 |
|---|---|---|---|---|---|---|---|
| Cost | Mutation = 0.02 | 0008/0 | 454/0 | 9676/0 | 5664/6 | 541/48 | 369/101 |
| Execution time | | 010003/0 | 149606/0 | 260403/0 | 1080/0 | 1231/0 | 2098/0 |
| Cost | Mutation = 0.04 | 0008/0 | 5636/0 | 024/1 | 9936/5 | 363/47 | 482/103 |
| Execution time | | 013432/0 | 178321/0 | 266426/0 | 1060/0 | 1826/0 | 1859/0 |
| Cost | Mutation = 0.06 | 0008/0 | 3448/0 | 8716/0 | 9196/5 | 645/47 | 794/102 |
| Execution time | | 015204/0 | 126024/0 | 266711/0 | 1210/0 | 1718/0 | 2257/0 |
| Cost | Mutation = 0.08 | 0024/0 | 4272/0 | 9624/0 | 7888/5 | 126/47 | 091/104 |
| Execution time | | 017084/0 | 151551/0 | 26134/0 | 1345/0 | 1678/0 | 2299/0 |

TABLE IV.
GA ALGORITHM PERFORMANCE BY STEP IMPLEMENTATION OF 500 AND POPULATION OF 50

| Number of Queens | | 5×5 | 8×8 | 10×10 | 20×20 | 100×100 | 200×200 |
|---|---|---|---|---|---|---|---|
| Cost | Mutation = 0.02 | 0 | 2392/0 | 7548/0 | 9614/5 | 359/48 | 026/102 |
| Execution time | | 018133/0 | 121314/0 | 311989/0 | 1521/0 | 2087/0 | 3027/0 |
| Cost | Mutation = 0.04 | 0 | 2184/0 | 7956/0 | 8906/5 | 121/47 | 934/102 |
| Execution time | | 023361/0 | 115198/0 | 352058/0 | 1523/0 | 1817/0 | 3301/0 |
| Cost | Mutation = 0.06 | 0 | 208/0 | 7104/0 | 0487/6 | 063/48 | 825/102 |
| Execution time | | 018361/0 | 125868/0 | 301072/0 | 1574/0 | 2274/0 | 3763/0 |
| Cost | Mutation = 0.08 | 0 | 2692/0 | 7288/0 | 6512/5 | 019/47 | 985/103 |
| Execution time | | 0175/0 | 153405/0 | 366623/0 | 1611/0 | 2507/0 | 4107/0 |

TABLE V.
PSO ALGORITHM PERFORMANCE BY STEP IMPLEMENTATION OF 50 AND POPULATION OF 30

| Number of Queens | | 5×5 | 8×8 | 10×10 | 20×20 | 100×100 | 200×200 |
|---|---|---|---|---|---|---|---|
| Cost | W= 0.02 | 0004/0 | 50324/0 | 9344/0 | 567/4 | 149/39 | 536/96 |
| Execution time | | 023302/0 | 189087/0 | 305571/0 | 6545/0 | 9306/0 | 0916/1 |
| Cost | W= 0.04 | 0016/0 | 5578/0 | 0672/1 | 743/4 | 451/39 | 026/97 |
| Execution time | | 02245/0 | 243094/0 | 307073/0 | 7315/0 | 8441/0 | 1564/1 |
| Cost | W= 0.06 | 0004/0 | 5076/0 | 0594/1 | 941/4 | 226/39 | 984/96 |
| Execution time | | 025709/0 | 230874/0 | 309601/0 | 6371/0 | 7721/0 | 0802/1 |
| Cost | W= 0.08 | 0032/0 | 5252/0 | 1524/1 | 801/4 | 914/38 | 041/97 |
| Execution time | | 025855/0 | 23055/0 | 32764/0 | 5883/0 | 9941/0 | 0391/1 |

TABLE VI.
PSO ALGORITHM PERFORMANCE BY STEP IMPLEMENTATION OF 50 AND POPULATION OF 50

| Number of Queens | | 5×5 | 8×8 | 10×10 | 20×20 | 100×100 | 200×200 |
|---|---|---|---|---|---|---|---|
| Cost | W= 0.02 | 0 | 3092/0 | 601273/0 | 351/4 | 482/38 | 439/96 |
| Execution time | | 033574/0 | 206495/0 | 686414/0 | 8786/0 | 26302/1 | 6505/1 |
| Cost | W= 0.04 | 0 | 4744/0 | 842/0 | 466/4 | 026/38 | 540/96 |
| Execution time | | 032076/0 | 290509/0 | 566925/0 | 9998/0 | 1099/1 | 0183/2 |
| Cost | W= 0.06 | 0 | 3684/0 | 944/0 | 852/4 | 503/38 | 941/96 |
| Execution time | | 032369/0 | 30929/0 | 601719/0 | 9962/0 | 3091/1 | 9643/1 |
| Cost | W= 0.08 | 0 | 39/0 | 934/0 | 753/4 | 129/38 | 843/96 |
| Execution time | | 036943/0 | 298051/0 | 575719/0 | 9785/0 | 0723/1 | 8218/1 |

## III. EXPERIMENTAL RESULTS

The N-queens puzzle is one of the classic problems with a long history [6]. The N-queens puzzle is the problem of putting N chess queens on a ג chessboard such that none of them is able to capture any other using the standard chess queen's moves. In this section the results from implementing N-queens problem using BBO algorithm, and also comparison of its results with genetic (GA) and particle swarm optimization (PSO) algorithms will be discussed. Since all three algorithms are stochastic in nature, each of them has been repeated 500 times. The focus of the experiments was on the impact of variations in values of BBO parameters.

In this paper for study BBO algorithm performance in N-queen problem, we solve the problem with different queens. Our study has been done on 5, 8, 10, 20, 100, and 200. Due to BBO algorithm solution is produced such as most of other meta-heuristic algorithms stochastically so experiments are repeated 500 times and average of them is obtained in this 500 times.

Our experiments are based on effect of BBO algorithm change value. Condition of algorithm termination is reach to optimal answer and end of maximum repeat. BBO method is based on population. First, it is produced a random population for solution of N-Queen problem. This population is produced in mention of problem space (number of queens and number of population). After this main operation, BBO algorithm is performed. Then, the data is sorted for finding based possible solution. Finally, after doing these maximum solutions in population will been probability answer. Possible number of fulmination returns as cost in this method and it is written for cost function. If a solution breaks main condition of problem then it saves maximum value of fulmination as cost. For this reason, we are interested in finding optimal cost. Fig. 1 shows BBO algorithms cost function with 50 populations.

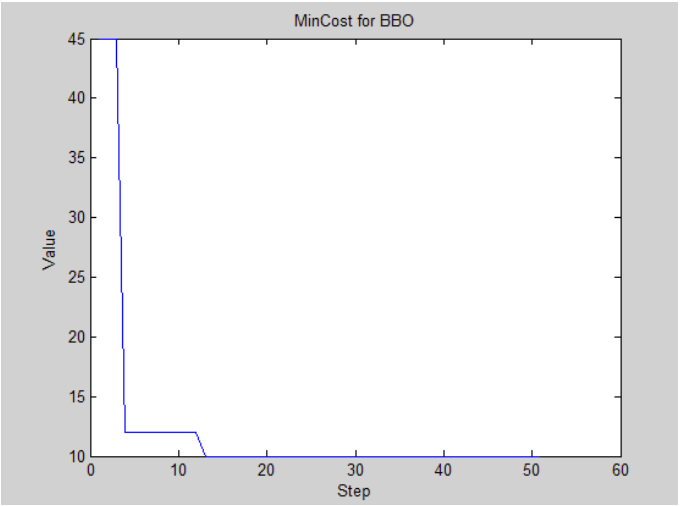



Fig 1. BBO cost function with 50 population.

Numbers of population selections are effective in optimal answer in population oriented meta-heuristic algorithm. A small number of population or unsuitable number of population selection has low variety answers so it decreases discovery power in search space and makes precocious convergence. On the other hand if number of population increases, then discovery power also increases. But more times it is required to get precocious convergence. So the number of population should be appropriate so that we can have desirable results in best time.

According to the obtained results from Tables 1 and 2, population growth leads to decrease in costs, but increase in implementation period. As the number of queens increases an appreciable increase is observed in both costs and implementation period. Variations in the mutation rate have no significant influence on costs. The following diagrams (Fig. 1) have been plotted taking 500 implementation steps and 10 queens into account.

In comparing BBO algorithm performance with GA and PSO algorithms, execution time and number of population are considered 500 and 30, 50, respectively in N-queen problem. Mutation rate are studied for 0.2, 0.4 0.6, 0.8. Results of GA and PSO algorithms are shown in tables 3, 4, 5, 6.

Fig. 2,3,4,5 show performance of BBO, GA and PSO algorithms with queen numbers and execution time for mutation rates 0.2, 0.4 0.6, 0.8. Mutation rates have little impact on performance of algorithm. But mutation rates 0.4 and 0.8 show best performance for the BBO algorithm.

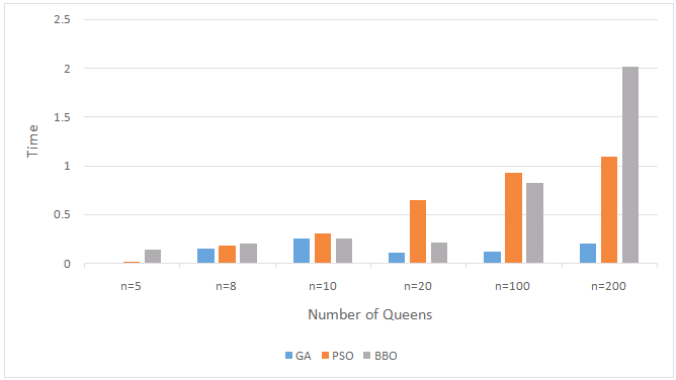



Fig. 2. Comparing the BBO, GA and PSO algorithms with mutation 0.02.

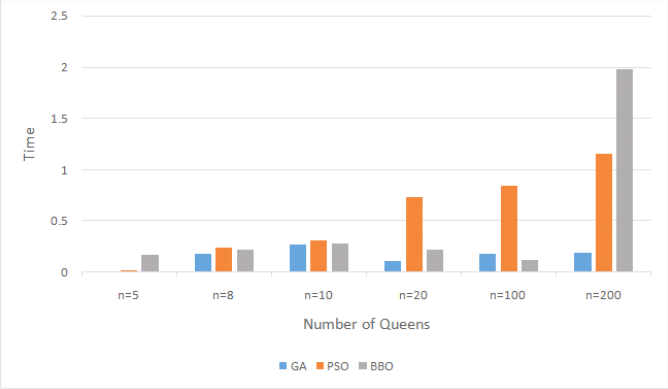



Fig. 3. Comparing the BBO, GA and PSO algorithms with mutation 0.04.

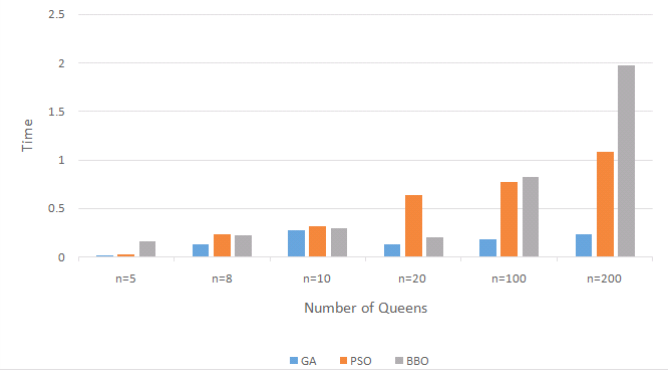



Fig. 4. Comparing the BBO, GA and PSO algorithms with mutatiom 0.06.

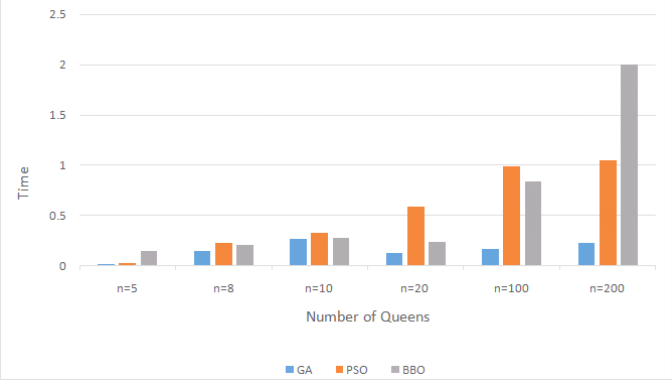



Fig. 5. Comparing the BBO, GA and PSO algorithms with mutatiom 0.08.

By comparing PSO, GA, and BBO algorithms and considering the diagrams, it can be noticed that as the number of queen's increases the performance of the BBO algorithm becomes better relative to PSO and GA algorithms, besides the implementation period will be shorter.

This algorithm has an advantage over others for solving the N-queens problem. However, this superiority can be held up to maximum 100 queens. As it can be seen from the diagrams, BBO shows a poor performance in the cases with more queens (200 and more). Moreover, the cost of the BBO algorithm is higher comparing with GA and PSO algorithms.

## IV. Conclusion

The obtained results from examination indicated capability of BBO for solving the N-queens problem. As the contrastive analysis of two algorithms showed, BBO algorithm is more accurate and quicker than PSO. Comparing with GA algorithm, the more queens are there, the better the performance of the BBO algorithm will be. The performance falls with more queens (200 and more).

In comparing BBO algorithm with GA algorithm, GA algorithm has better performance. GA algorithm is better than BBO and PSO algorithm because of their execution time. Both BBO, GA algorithms have fast convergence. If convergence of PSO algorithm is fast, then it has not exit mechanism of local optimum trap. So, it selects local optimum as global optimum while BBO algorithm escapes local optimum and gets to optimum answer.

Considering the fact that BBO is a relatively newly-developed algorithm, a wide range of potential research fields can be proposed for this algorithm. For example we can integrate this algorithm with others to achieve better results. In order to obtain more accurate answers and shorten the time of implementation, other algorithms such as exact algorithms can be used to generate better primary solutions, hence to improve the performance.

BBO algorithm can be used for the optimization problem solution of the real world as image restoration, image segmentation, video coding and wireless sensor network. We suggest that BBO algorithm is used in above problems in future work. Also this algorithm can be applied for problem solutions that are NP-hard as travelling salesman problem (TSP).

## References

[1] Dan Simon, "Biogeography-Based Optimization", IEEE Transactions on Evolutionary Computation, vol. 12, no. 6, Dec. 2008.

[2] R.L. Haupt and S.E. Haupt, "Practical Genetic Algorithms", 2nd Edition, John Wiky& Sons Inc, 2004.

[3] P. Pedregal. "Introduction to Optimization", Springer, New York Inc.,2004.

[4] Talbi.EL-Ghazali, "Metaheuristics : From Design to Implementation", John Wily and sons, 2009.

[5] N. Johal, S. Singh, and H. Kundra, " A hybrid FPAB/BBO algorithm for satellite image classification", International Journal of Computer Applications, vol. 6, no. 5, pp. 31- 36, Sep. 2010.

[6] B.Y. Qu, J.J. Liang, P.N. Suganthan, "Niching particle swarm optimization with local search for multi-modal optimization", Information Sciences 197 (2012) 131–143.

[7] M. Ergezer, D. Simon, D. Du, "Oppositional biogeography-based optimization", In Proceedings of the IEEE Conferencon Systems, Man, and Cybernetics, IEEE, San Antonio,TX, USA, pp. 1035-1040,2009.

[8] D. Du, D. Simon, M. Ergezer, "Biogeography-based optimization combined with evolutionary strategy and immigration refusal", In Proceedings of the IEEE Conference on Systems, Man, and Cybernetics,IEEE, San Antonio, TX,USA, pp. 1023-1028, 2009.

[9] H. Ma, S. Ni, M. Sun, " Equilibrium species counts and migration model tradeos for biogeography-based optimization", In Proceedings of the 48th IEEE Conference on Decision and Control, IEEE, Shanghai, China, pp. 3306-3310, 2009.

[10] Ivica Martinjak, Marin Golub, "Comparison of Heuristic Algorithms for the N-Queen Problem", July 2007.

[11] K. D. Crawford, "Solving the N-Queens Problem Using Genetic Algorithms," in ACM/SIGAPP Symposium on Applied Computing, 1992, pp. 1039-1047.

[12] Aftab Ahmed, Attique Shah, Kamran Ali Sani and Abdul Hussain Shah Bukhari, "Particle Swarm Optimization For N-Queens Problem", *Journal of Advanced Computer Science and Technology*, vol. 1, no. 2, pp. 57-63, 2012.

[13] Amooshahi A., Joudaki M., Imani M., and Mazhari N., "Presenting a New Method Based on Cooperative PSO to Solve Permutation Problems: A Case Study of n-Queen Problem," in Proceedings of the 3rd International Conference on Electronics Computer Technology, anyakumari, India, vol. 4, pp. 218-222, 2011.

[14] Martinjak I. and Golub M., "Comparison of Heuristic Algorithms for the n-Queen Problem," in Proceedings of the 29 th International. Conference on Information Technology Interfaces, Cavtat,Croatia, pp. 59-764, 2007.

[15] Draa A., Meshoul S., Talbi H., and Batouche M., "A uantum-Inspired Differential Evolution Algorithm for Solving the n-Queens Problem," the International Arab Journal of Information Technology, vol. 7, no. 1, pp. 21-27, 2010.

[16] Wang, Shuihua; Zhang, Yudong; Ji, Genlin; Yang, Jiquan; Wu, Jianguo; Wei, Ling. 2015. "Fruit Classification by Wavelet-Entropy and Feedforward Neural Network Trained by Fitness-Scaled Chaotic ABC and Biogeography-Based Optimization." *Entropy* 17, no. 8: 5711-5728. doi: 10.3390/e17085711

[17] Mehran Tamjidy , Shahla Paslar , B.T. Hang Tuah Baharudin , Tang Sai Hong , M.K.A. Ariffin, "Biogeography-based optimization (BBO) algorithm to minimise non-productive time during hole-making process" , International Journal of Production Research , Vol. 53, no. 6, 2015.

[18] Vanitha, M. and K. Thanushkodi, "An Effective Biogeography-Based Optimization Algorithm to Slove Economic Load Dispatch Problem", Journal of Computer Science, vol. 8, no 9, pp. 1482-1486, 2012 .

[19] Shahrzad Saremi and Seyedali Mirjalili, "Integrating Chaos to Biogeography-Based Optimization Algorithm ",International Journal of Computer and Communication Engineering, vol. 2, no. 6,November 2013.

**Tayebeh Jalali** received the B.Sc. and M.Sc. degree from department of computer Science & engineering, Khoy Branch, Islamic Azad University, Khoy, Iran in 2010, 2015 respectively. She is now with the Ministry of Education of Khoy, West Azarbayjan, Iran. Her research focuses on artificial intelligence, algorithms.

**Ali Habiboghli** received the B.S. degree from department of engineering of Islamic Azad University, khoy Branch in 2004. He received the M.Sc degrees from electronic, computer engineering and information technology from Islamic Azad University, Qazvin Branch, Iran in 2007. From 2007 to already is with the Islamic Azad University of KHOY Branch. His research focuses on artificial intelligence, algorithms, biometric and image processing.

# Modified Three-Step Search Block Matching Motion Estimation and Weighted Finite Automata based Fractal Video Compression

Shailesh D. Kamble[1], Nileshsingh V. Thakur[2], and Preeti R. Bajaj[3]

[1]*Computer Science & Engineering, Yeshwantrao Chavan College of Engineering, India*
[2]*Computer Science & Engineering, Prof Ram Meghe College of Engineering & Management, India*
[3]*Electronics Engineering, G. H. Raisoni College of Engineering, India*

***Abstract* — The major challenge with fractal image/video coding technique is that, it requires more encoding time. Therefore, how to reduce the encoding time is the research component remains in the fractal coding. Block matching motion estimation algorithms are used, to reduce the computations performed in the process of encoding. The objective of the proposed work is to develop an approach for video coding using modified three step search (MTSS) block matching algorithm and weighted finite automata (WFA) coding with a specific focus on reducing the encoding time. The MTSS block matching algorithm are used for computing motion vectors between the two frames i.e. displacement of pixels and WFA is used for the coding as it behaves like the Fractal Coding (FC). WFA represents an image (frame or motion compensated prediction error) based on the idea of fractal that the image has self-similarity in itself. The self-similarity is sought from the symmetry of an image, so the encoding algorithm divides an image into multi-levels of quad-tree segmentations and creates an automaton from the sub-images. The proposed MTSS block matching algorithm is based on the combination of rectangular and hexagonal search pattern and compared with the existing New Three-Step Search (NTSS), Three-Step Search (TSS), and Efficient Three-Step Search (ETSS) block matching estimation algorithm. The performance of the proposed MTSS block matching algorithm is evaluated on the basis of performance evaluation parameters i.e. mean absolute difference (MAD) and average search points required per frame. Mean of absolute difference (MAD) distortion function is used as the block distortion measure (BDM). Finally, developed approaches namely, MTSS and WFA, MTSS and FC, and Plane FC (applied on every frame) are compared with each other. The experimentations are carried out on the standard uncompressed video databases, namely, akiyo, bus, mobile, suzie, traffic, football, soccer, ice etc. Developed approaches are compared on the basis of performance evaluation parameters, namely, encoding time, decoding time, compression ratio and Peak Signal to Noise Ratio (PSNR). The video compression using MTSS and WFA coding performs better than MTSS and fractal coding, and frame by frame fractal coding in terms of achieving reduced encoding time and better quality of video.**

***Keywords* — Fractal Coding, Motion Wstimation, Three Step Search, Cross Hexagon Search, Encoding Time, Compression Ratio.**

## I. Introduction

Video compression techniques deal with the lossy or lossless data compression for the series of image sequences. Gray and color image sequences are the customers for the video compression techniques. From the preprocessing point of view different color spaces [1] can be used for the image sequences. With the recent development in the area of internet and multimedia technology with moving video, a color image sequences processing plays an important role. Each color plane is represented by 8 bits/pixel in a RGB color space therefore a RGB color image is represented by 24 bits/ pixel. The intra-frame and inter-frame coding is used to reduce the spatial and temporal data redundancy present in the image sequences.

Block matching motion estimation plays an important role in inter-frame coding technique to reduce temporal redundancy present in the series of image sequences. Therefore it is a most popular and efficient technique for computing the motion vectors which has been used by various coding standards. Video coding standards are related to the organization, such as ITU-T Rec. H.261, H.263, ISO/IEC MPEG-1, MPEG-2, MPEG-4, and recent progress is H.264/AVC. Different block searching approaches are available for motion estimation [2-7] between the successive frames. A block-matching algorithm [5-6] can be used to improve the quality of video and performance of coding process [7].

Block-matching algorithms are the most successful approaches for motion estimation in the video compression technology because of it easy to understand and with some efforts it can be implemented easily. The simplest full search (FS) block matching algorithm provides a full exhaustive search within the search window for searching the optimal solution. Therefore to reduce the computational cost of FS, many block matching motion estimation algorithms are proposed such as three step search [8], 2D logarithmic search [9], orthogonal search [10], cross search [11], binary search [12], new three-step search [13], the four-step search [14], the block-based gradient descent search [15], the diamond search [16], the cross-diamond search [17], efficient three step search [18], etc. All these search algorithms employs a rectangular and diamond search patterns having the center-biased motion vector distribution characteristics [19-20]. Hexagon-based search algorithm which employs a two hexagon-based pattern i.e. large and small, and results in fewer search points is proposed [21]. Novel Cross-Hexagon based Search algorithm [6] consists of two cross shape patterns and two hexagon-based patterns. Modified partial distortion criterion (MPDC) used a certain block of pixels which improve the computations [22]. New Cross-hexagon search (NHEXS) algorithm [23-24] consist of two cross search patterns and hexagon search patterns which is similar to fast block-matching Motion Estimation [6].

Block matching Motion estimation is widely used in the video application areas such that rain pixel recovery for videos, video compression, medical video processing [25-26], object tracking and



DOI: 10.9781/ijimai.2017.444

surveillance [27], etc. Motion estimation means displacement of pixels position from one frame to another frame which gives the best motion vector (MV) [28]. In Block matching algorithm, each frame is divided into a fixed sized macro blocks and each corresponding macro blocks in the current frame are then compared with the adjacent neighbor macro blocks in the previous frame to estimate a motion vector (MV). The estimated MV specifies the displacement of a macro block from one position to another position in a previous frame as shown in Fig. 1. Motion vector for the block Bf(x, y) is computed as (+1,-1) i.e. MV-Bf(x,y) = (+1,-1). Still there is a scope of improvement in this field for implementing the fast block matching motion estimation algorithm.

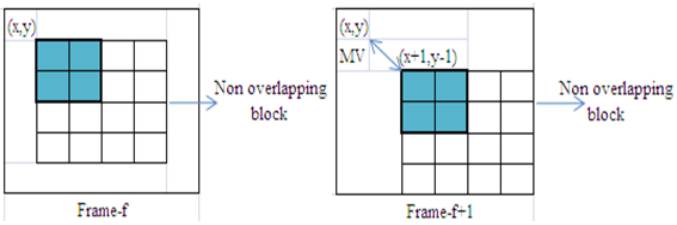



Fig.1. Block matching motion vector estimation.

Using Weighted Finite Automata (WFA) the weights are assigned to state transitions proposed by [29-30], WFA provide a powerful tool for image representation as a WFA and compress the image in term of a good compression ratio. The inference algorithm for WFA subdivides an image into a set of non-overlapping range blocks and then separately approximates each range block with a linear combination of the domain block. WFA coding techniques is based on the idea of fractal coding that an image has self-similarity in itself [31-34] i.e. WFA coding is similar to the fractal coding approach.

This paper discusses a modified three step search algorithm for block matching motion estimation and WFA coding approach for color video compression. The paper is organized as follows. Section 2 reviews Three Step Search algorithm. New cross hexagonal search algorithm discussed in section 3. The proposed MTSS presented in section 4 and explains the video compression process using MTSS and WFA coding. Section 5 presents experimental results of proposed MTSS approach in comparison with TSS, ETSS and NTSS block matching algorithm and further the results of proposed MTSS and WFA coding is compared with MTSS and fractal coding. Finally, Section 6 presents the conclusion and future scope based on proposed work.

## II. Three Step Search Algorithm

Three step search (TSS) proposed by Koga et al. [8] is one of the earliest attempt to implement a fast motion estimation algorithm. The computational cost of this algorithm is less in terms of MAD matching criterion, average search point and computation required as compared to the full search algorithm. In this algorithm first step is to define a fixed size of 16 ×16 macro block, searching parameter p pixels in all four directions and 9×9 search window in the central part of the 16 ×16 macro block for searching the best match. The TSS algorithm is summarized as follow:

Step 1: Plot 9 points in the search window at the equal distance of step size s=4 and checked 9 points on the 9×9 search window. .

Step 2: Step size is divided by 2 i.e. s=2 and check 8 points to generate 5×5 square shape pattern. If minimum block distortion measure (BDM) is one of the nine points of generated search window then this point consider as a center point in the step 3.

Step 3: Step size is divided by 2 i.e. s=1 and check 8 points to generate 3×3 square shape pattern and search will be terminated.

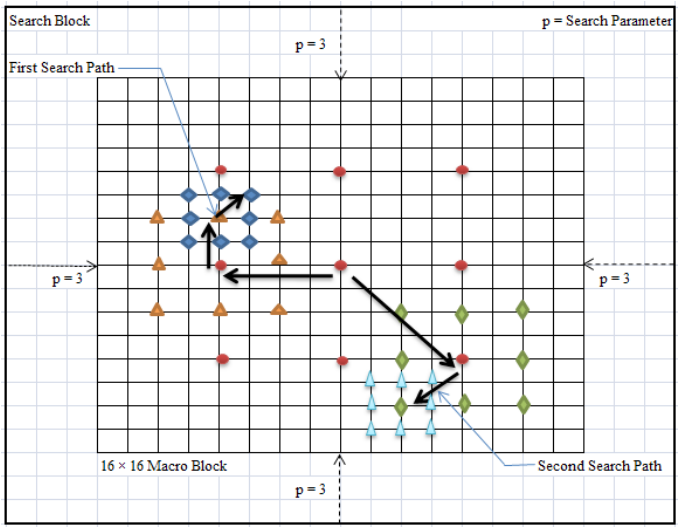



Fig.2. Two different search paths for three step search motion estimation.

The minimum BDM point found at the 3×3 square shape pattern is the final motion vector. Fig. 2 shows the two different search path of TSS for estimating a motion vector within search window. TSS can be easily extended up to n-step search for larger search window. The number of checking points required for TSS is 25.

## III. New Cross Hexagonal Search Algorithm

Kamel Belloulata et al. [24] proposed a novel fast block matching algorithm called new cross hexagon (NCHEXS) pattern based search using two small/large cross shape(SCSP/LCSP) pattern as a first three initial steps and two small/large cross hexagon shape pattern (SHSP/LHSP) as a subsequent steps of search. In the beginning of this algorithm, initialize SCSP by plotting 5 points at the center of 16×16 macro blocks [28]. The algorithm is summarized as follow:

Step 1: If minimum BDM point found at the center of the small cross shape pattern then stop searching otherwise go to step 2.

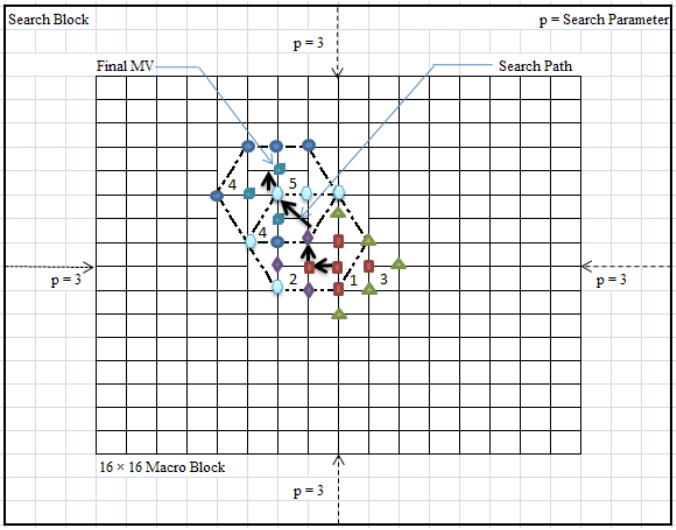
Step 2: If minimum BDM point found at the center of newly formed small shape pattern then stop searching otherwise go to step 3.

Step 3: Check the another 3 unchecked points of large cross pattern and 2 unchecked points of the square center biased to show the best possible direction for the hexagonal search.

Step 4: A new large hexagonal shape pattern is formed by considering a center point as minimum BDM point found in previous step. If minimum BDM point found at the center of new large hexagonal pattern then go to step 5 otherwise again form a new large hexagon pattern i.e., repeat step 4.

Step 5: If minimum point found at the center of large hexagonal pattern then large hexagonal pattern shifted to small hexagonal pattern and find best motion vector in small hexagon shape pattern

Fig. 3 shows the search path of NCHEXS. The number of checking point required for NCHEXS is from 5 to 8 for best case which is better than all the techniques available for estimating a motion vector.




*SCSP: 1st step:* , *2nd step:* , *LCSP: 3rd step:* , *LHSP: 4th step:* & ,*SHSP: 5th step:* *Final MV.*

Fig.3. Search path in New Hexagon search motion estimation.

## IV. Proposed Approach

The main objective of proposed work is to develop a mechanism for color video compression based fractal coding technique using Modified three-step search (MTSS) and weighted finite automata (WFA) coding. Therefore, fractal coding based color video compression using MTSS block matching algorithm and WFA coding is proposed.

In color images, Each R, G and B components contains 8 bit data and also every color image contains lots of redundancy. In this approach, each frame is converted into the $YC_bC_r$ color space i.e. most suitable color space for video processing and the first component of the frame in $YC_bC_r$ color space is treated as gray scale image. The first component of $YC_bC_r$ i.e. luminance component is sensitive to human eye and similar to grey scale image and remaining two chrominance component consist of color information and human eye is not that much sensitive for these two planes. Therefore, the focus is on the compression of the first luminance plane. Each gray scale frame is divided into a fixed/equal sized macro blocks. The MTSS approach is used for estimating a motion vector in the current frame with respect to the previous frame. The predicted frame for the current frame is then created from the previous frame with assigned motion vectors. Encoding and decoding of the previous frame is performed using WFA or FC. Encoding and decoding of the difference between the predicted frame and the current frame is, also, performed using WFA or FC. The first frame is encoded and decoded using WFA/FC and then the predicted frame for the second frame is formed from the decoded first frame using MTSS. Then, the predicted frame is encoded using WFA/FC and decoded to form the predicted frame for the third frame. This process is repeated for all the frame sequences. The process flow of proposed MTSS and WFA/FC coding approach for video compression is shown in Fig. 4.

### *A. Modified Three Step Search (MTSS) Algorithm*

The proposed MTSS approach is used to calculate the motion compensation prediction error (MCPE) between two consecutive frames. In MTSS, Two cross search pattern i.e. small and large cross-search pattern and two cross-hexagon search pattern i.e. large cross-hexagon and small cross-hexagon search pattern is used in the center part of search window to exploit central biased characteristic of Motion Vector (MV) in video sequences. The proposed MTSS approach

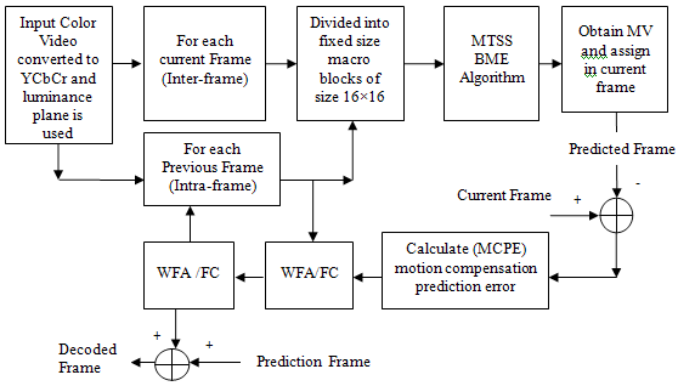



Fig.4. MTSS and WFA/FC video coding Process Flow.

consists of TSS and NCHEXS approach. TSS and NCHEXS employ square and hexagon based shape pattern of different sizes respectively. If minimum BDM point found at outer point of 9×9 search window then TSS will execute. Otherwise NCHEXS will execute at the central part of the search window. Fig. 5 shows the search pattern used in first step of MTSS proposed approach. The MTSS can be summarized as follows:

Step 1: Total 9+4 points are checked, if minimum BDM point found at the center of 9+4 points then stop searching otherwise go to step 2.

Step 2: If minimum BDM point found at the outer part of search window then search process is same as TSS discussed in section 2 otherwise go to step 3.

Step 3: If minimum BDM point found at the 4 outer points of small cross search pattern (SCSP) then search process is same as new cross hexagon search (NCHEXS) discussed in section 3.

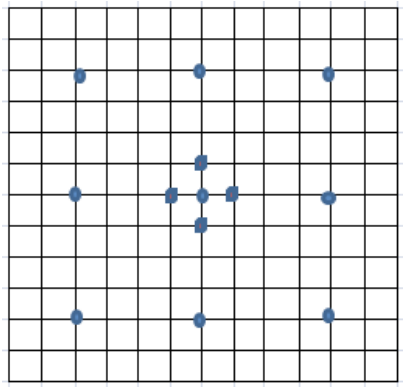


Fig.5. Search pattern used in first step of modified TSS (MTSS).

Block diagram and two search path of proposed MTSS approach are shown in Fig. 6 and 7 respectively. The numbers of checking points required for MTSS approach are 13 (9+4) points for best case i.e. stationary block is shown in Fig. 5, 13 to 29 (9+4+8+8) points for average case and more than 29 points for worst case is shown Fig. 7. While checking points are needed with the new three step search (NTSS) algorithm are 17 for best case, 17 to 33 points for average case and 33 checking points are required for worst case. In TSS, 25 checking points are needed in all the cases.

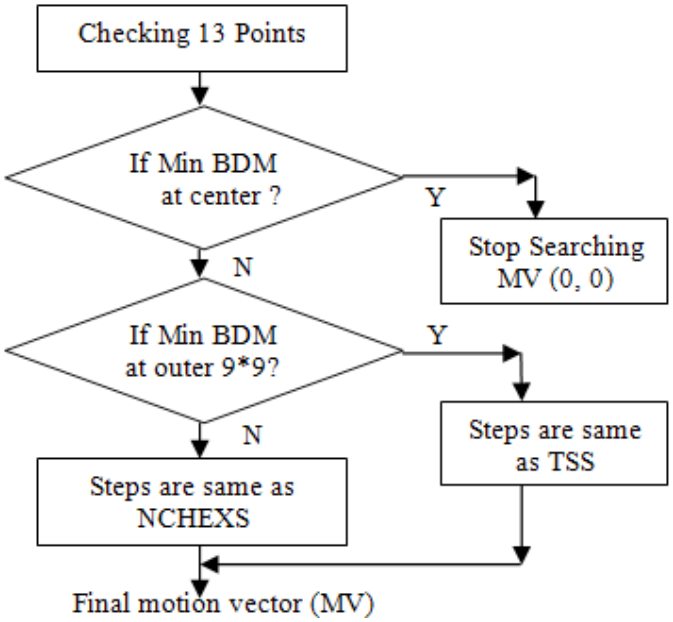



Fig..6. Block schematic of the Modified TSS (MTSS).

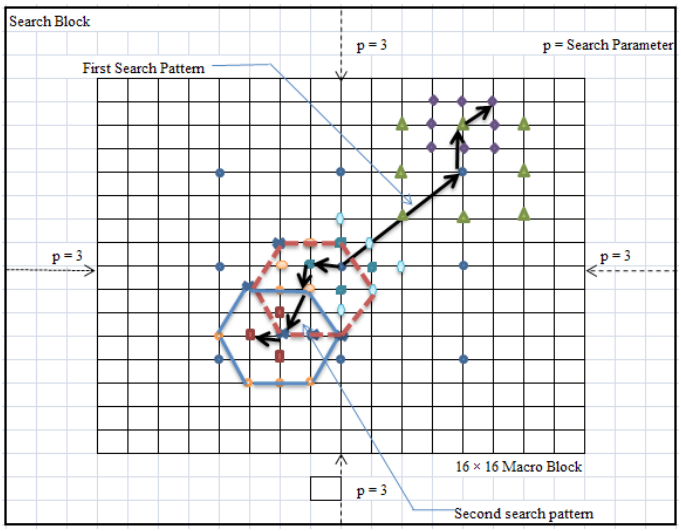



Fig.7. Two Search paths in modified TSS.

The concept of intra-frame and inter-frame coding are used for exploiting the self similarity present in the frame. The intra-frame coding is similar to the individual frame by frame coding i.e. individual frames are coded independently. The first frame in a video sequence is always an intra-frame because there is no previous data related to the first frame. It is not necessary in a video sequence always a first frame is an intra-frame. Intra-frame occurs anywhere in video sequences when the scene of video sequence is totally changed. On the other side in inter-frame coding, the previous frame and current frame is coded using proposed MTSS block matching motion estimation algorithm and WFA coding. The previous and current frames are extracted from the video sequence and apply the proposed MTSS algorithm on each fixed sized blocks. The obtained motion vectors (MV) are assigned to the current frame *i.e.,* the newly formed predicted frame. Then calculate the motion compensated predicted error (MCPE) frame by deleting the current frame from newly formed predicted frame. Finally apply the WFA encoding and decoding on them as shown in block schematic Fig. 4. The WFA decoded frame is considered as a previous frame for the new subsequent current frame and same process is repeated.

### *B. Proof-Outlines of the Parameters used in Modified Three-Step Search (MTSS) algorithm*

In general, videos consist of different number of frames. Consecutive frames have the huge spatial redundancy. Occurrence of the change in next frame is very less in comparison with the previous frame. Most of the part of next frame is similar to the previous frame. Therefore, the change, in general, may occur within the area of 3×3 or 5×5 or 7×7 or 9×9 neighborhood of pixels. Hence, to measure the similarity between two consecutive frames, it is better to use the bigger size search window to grab most of the area of macro block size (i.e. 16×16). So, initially, 9×9 search window (centre-biased) is selected. Now, 3 or 4 pixels are left on every side of 9×9 search window. As unexplored area consist of 3 or 4 pixels, the scope of next search window of size (3×2+1)×(3×2+1) i.e. 7×7 or (2×2+1)×(2×2+1) i.e. 5×5 exists. Here, we have used 5×5 search window (centre-biased). Now, unexplored area consist of 1 pixel, the scope of next search window of size (1×2+1)×(1×2+1) i.e. 3×3 exists. Hence, we have used 3×3 search window (centre-biased).

### *C. Weighted Finite Automata Representation*

A finite automaton is a mathematical model used to represent an image. It exploits the self-similarity, i.e. redundancy, in an image. WFA constitute an extended version of finite automata, with transitions labeled as the addresses of the sub-images and weights assigned from one state to another state. WFA are used to represent images in a 6-tuple form, i.e. M= (Q, ∑, W, I, F, $q_0$), where,

Q is a finite set of states, i.e. {$q_0$, $q_1$, $q_2$ ...$q_n$};

∑ is a finite set of input symbols/quadrant addresses, i.e.

∑= {0, 1, 2, 3} for quadtree WFA,

∑= {0, 1} for bintree WFA, and

∑= {0, 1, 2, 3, 4, 5, 6, 7, 8} for nona-tree WFA;

W is the weight function between two states, i.e.

W Є Q×Q → R for W ($q_0$, $q_1$) = R, where $q_0$, $q_1$ Є Q and R is a real number, i.e. the weight between states $q_0$ and $q_1$;

I is the initial configuration of states Q → R and indicates which states correspond to the entire image; I ($q_0$) = 1 and I($q_i$) = 0, $q_0 \neq q_i$, where $q_0$, $q_i$ Є Q and I =1, 2, 3....n;

F is the final configuration of states Q → R, e.g. F ($q_0$) = f (Є), where $q_0$ Є Q and f (Є) is the average intensity (greyness) of the entire image;

$q_0$ is the initial state of the WFA, i.e. the entire original image: $q_0$ Є Q

The transition from ($q_0$, 1) = $q_1$ Є Q×∑ →Q if W ($q_0$, $q_1$) = R or W ($q_0$, $q_1$) ≠ 0 implies that there is a transition from $q_0$ to $q_1$ on the input symbol labelled by 1. Further, $W_i$($q_0$, $q_1$) denotes a weighted transition from $q_0$ to $q_1$ on the input symbol, i.e. state or sub-image i.

The basic principles of the WFA approach are somewhat similar to those of the fractal image compression approach based on PIFS (partitioned iterated function systems). Both the approaches use the fact that images used in practice have a certain amount of self-similarity present in images to achieve compression. In other words, a sub-image of the image to be compressed may be similar to another sub-image of the same image, except perhaps for size, contrast, or brightness. The main difference between the two approaches is that the PIFS-based fractal compression uses affine transformations to find the self-similarity of sub-images, while on the other hand, WFA finds a sub-image as a weighted linear combination of other states/sub-images. The real value (weight) and quadrant address assigned with each transition along an edge in the WFA indicate how each state in the WFA is expressed as a linear combination of the other states:

$$(s_i)q = Wq(s_i, s_0)s_0 + Wq(s_i, s_1)s_1 + \cdots + Wq(s_i, s_n)s_n \quad (1)$$

Where $s_i$ is the image associated with state $s_i$ and $(s_i)q$ indicates the address of quadrant q of state $s_i$ .

WFA provide a powerful tool for image generation and compression. First, the image is subdivided into non-overlapping sub-images through a quadtree partitioning scheme. These sub-images are identical to those range blocks used in the PIFS-based fractal image compression approach. Next, one or more sub-images that are very similar or identical to the original image or to each range block/sub-image to be encoded separately are obtained from a domain block/sub-images present in domain pool, and a transition graph is constructed to describe the relationship between these sub-images and finally, the image to be encoded using WFA approach. The domain pool may contain all states or sub-images which could be generated from the given partitioning scheme. In general, the WFA uses an inference algorithm to construct a transition graph that is very similar to graphs used to represent finite state automata. The various states / sub-images of the finite state automata are then compressed to become the compressed image.

The process of decoding an encoded image with suitable example is discussed below. The image to be encoded at resolution 4×4 i.e. $2^{k=2} \times 2^{k=2}$ is given in Fig. 8 (a). The generated quadtree, WFA representation and its transition diagram representation are given in Fig. 8 (b), 8(c) & 8(d).

The WFA representations in the form of matrix are as follows.

$$Q=\{ S1 , S2 \} , \sum = \{ 0,1,2,3 \} , I = [1\ 0] , F = \begin{bmatrix} 2.2 \\ 2 \end{bmatrix},$$

$$W0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, W1 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, W2 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \text{ and } W3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

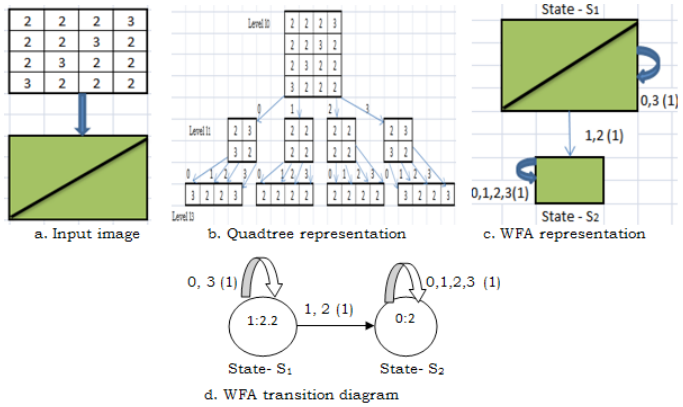



Fig.8. Process of WFA encoding/decoding.

*WFA Encoding Process*

The WFA encoding algorithm takes as input a grayscale image of size $2^k \times 2^k$ and gives as output the WFA representation of an image of size $2^k \times 2^k$ i.e. the initial configuration I, the final configuration F and weights W. In WFA all the four sub-image/ quadrants of an image are processed by approximating the quadrant/sub-image with the linear combination of all the existing states. If the quadrant of state/image is not a linear combination of the existing states then the quadrant is chosen as a new state and placed in the unprocessed state list; if not, then add transition and stores coefficients in the list. All the four quadrants/sub-images of a state/image are processed before moving to the next unprocessed state present in the list. Once all the new unprocessed states present in the unprocessed list have been processed, the algorithm terminates. The recursive WFA encoding algorithm is shown below.

*Input : Image M of size $2^k \times 2^k$*

*Output: Generated WFA M representing the image M of size $2^k \times 2^k$*

*Step-1: n =0, Number of states in WFA*

*$W_q (s_i, s_j)$ =0, indicate the weight associated with the transition from state $s_i$ to $s_j$ with real value $W_q$.*

*$(s_i)_q$ = quadrant indicating address q of state $s_i$.*

*i=1, The index of first unprocessed state.*

*$I(s_0)$=1,Initial configuration of entire image*

*$F(s_0) = f(\epsilon)$ is the average grayness or brightness of entire image*

*Step-2: For every q $\in \sum$ = {0, 1, 2, 3} do*

*Sub-image I = $(s_i)_q$*

*Approximate the sub-image/quadrant with linear combination of the existing states/ images*

*Approximate I= $W_q(s_i, s_0)s_0 + W_q(s_i, s_1)s_1 + \ldots + W_q (s_i, s_n) s_n$*

*add transition and store coefficient in the list*

*else n = n + 1*

*Create new state $s_n$ = I*

*Add transition from state $s_i$ to new state $s_n$ with labeled q and weight 1*

*set $F(s_n)$ = and $I(s_n)$ = 0*

*Append new state $s_n$ to WFA*

*Step-3: i= i+1*

*if i <= n then goto Step 2*

*WFA Decoding Process*

The decoding algorithm takes as input an encoded WFA represented by $n + n_0$ states, F and W created during encoding process as well as the resolution level k and returns as output decoded image of size $2^k \times 2^k$. An WFA decoding approach is discussed below.

*Input: An encoded EWFA represented by $n + n_0$ states, F, W and integer k for an image resolution $2^k \times 2^k$*

*Output: Decoded image M of size $2^k \times 2^k$*

*$n_0$ =1, Number of initial states in WFA*

*n = Number of non initial state created during execution of encoding algorith in WFA*

*$l_n$ = Level of quadtree generated from top $l_0$ to bottom $l_n$*

*$f(s_i, \epsilon) = F(s_i)$ return the value of average grayness or brightness of image associated with state $s_i$*

*$f(s_i, aw)$ is the function returns the value of state $s_i$ at quadrant a=0,1,2,3 and length of input string generated by $\sum^*$*

*To decode each state in WFA call routine built_decoder (WFA M, int n)*

```
{
f(s_i, ϵ)= F(s_i) where i= 1,2,...n
for (i = 1 to l_n – 1)
{
  for ( s = 1 to n + n_0)
  {
    for every input symbol or string w ∈ Σ^(i-1) do
    {
      for every quadrant in a image a ∈ Σ do
      {
        f(s_i, aw) =
      }
    }
  }
  Rearrange all values f(s_i, aw) calculated in matrix form
}
```

## V. Experimental Results and Discussion

The proposed approach is implemented using MATLAB® 2013a (8.1.0.604). The experiments were carried out on a system with an Intel® Core™ i3 CPU (2.40 GHz). The experiments were carried out on color videos (i.e. Soccer, Suzie, Bus, Football, Xylophone, Paris, Traffic, Akiyo, Ice, and Mobile sequences) obtained from online resources i.e., http://media.xiph.org/video/derf/ (see Fig. 9). Standard input streams with different frame rates, lengths of sequences, and frame widths/heights were considered to demonstrate the performance of the proposed approach (see Table I).

Fig.9. Uncompressed input video sequences.

TABLE I

Specification of standard databases

| Sequences | Frame Rate frame/s | Length of video (s) | Frame Width × Height |
|---|---|---|---|
| Soccer | 20 | 6 | 176 × 144 |
| Suzie | 15 | 6 | 176 × 144 |
| Bus | 20 | 3 | 176 × 144 |
| Football | 15 | 6 | 176 × 144 |
| Xylophone | 30 | 5 | 640 × 480 |
| Paris | 29 | 13 | 352 × 288 |
| Traffic | 15 | 8 | 160 × 120 |
| Akiyo | 29 | 10 | 325 × 288 |
| Ice | 15 | 6 | 176 × 144 |
| Mobile | 29 | 10 | 352 × 188 |

### A. Quality Measures

Inter-frame and intra-frame coding is used to eliminate the large amount of temporal and spatial redundancy exists in the video sequences and therefore, helps in compressing them. The matching of the one current frame macro block and previous frame macro block is based on the output of matching criteria. The macro block that results in the minimum value is the one that matches the closest to current block with respect to the corresponding previous frame macro block. The popular matching criteria used for block matching motion estimation are mean of absolute difference (MAD), mean squared error (MSE) and sum of absolute difference (SAD) given by equation 2 ,3 and 4 respectively.

$$\mathrm{MAD}(i,j) = \frac{1}{N\times N}\sum_{i=0}^{i=N-1}\sum_{j=0}^{j=N-1}|Cij - Rij| \tag{2}$$

$$\mathrm{MSE}(i,j) = \frac{1}{N\times N}\sum_{i=0}^{i=N-1}\sum_{j=0}^{j=N-1}(Cij - Rij)^2 \tag{3}$$

$$\mathrm{MSE}(i,j) = \frac{1}{N\times N}\sum_{i=0}^{i=N-1}\sum_{j=0}^{j=N-1}(Cij - Rij)^2 \tag{4}$$

Where, N×N is the row and column of the macro block, $C_{ij}$ and $R_{ij}$ are the pixels value compared in current macro block and previous macro block, respectively.

In Block matching algorithms, the size of macro block is the important parameter for motion estimation. Smaller the macro block size results in more motion vectors and more macro blocks per frame. Therefore, improves a quality of motion compensated prediction error (MCPE). Most video coding standards used a macro block of size 16×16 and 8×8. The best/single motion vector is computed for each macro block in reference frame. On the other hand, total number of search point to find motion vector per frame is one of the key parameter in block matching motion estimation algorithm. The performance of video coding is measured in terms of compression ratio (CR), quality of video i.e. PSNR, encoding time and decoding time. The compression ratio (CR) is given by equation 5.

$$\mathrm{CR} = \frac{\text{Size of original video}}{\text{Size of compressed file}} \tag{5}$$

Therefore, compression ratio in percentage is computed from 5 and given by equation 6.

$$\mathrm{CR}\,\% = \big(1 - (1/\mathrm{CR})\big)\times 100 \tag{6}$$

When measuring the quality of compressed video, the peak signal-to-noise ratio is used. Sometimes mean squared error (mse) is also used, is given by 7.

$$\mathrm{mse} = \frac{1}{n}\sum(X_i - Y_i) \tag{7}$$

From this, the PSNR for an 8-bit grayscale image is defined by equation (8), where 255 is the maximum value for 8-bit pixel can assume.

$$\mathrm{PSNR(dB)} = 10\log_{10}(255^2/\mathrm{mse}) \tag{8}$$

### B. Stepwise Results Obtained for Color Space Conversion and the Quad-tree Partitioning Scheme

The intermediate results based on color space conversion and Quadtree partitioning of first ten frames of "soccer" video are shown in Fig. 10. The input video "Soccer" was obtained from an uncompressed video database i.e., http://media.xiph.org/video/derf/. The first 10 frames are shown in Fig. 10(a). The $YC_bC_r$ color space i.e., Y- luminance component sequences of the first 10 frames are shown in Fig. 10(b). The Y- luminance component sequences of the first 10 frames are then converted into gray frames, as shown in Figures 10(c). Fig. 10(d) shows first 10 frames of quadtree partitioning of gray frames. Note that in experimentation the standard input frame size must be converted into square of size $2^n \times 2^n$. This conversion involves shrinking and replicating the pixels for finding out the frame of size $2^n \times 2^n$. The presented approach initially converts the RGB sequence into the $YC_bC_r$ color space sequence and finally convert Y- luminance component to grey scale for applying the quadtree partitioning.

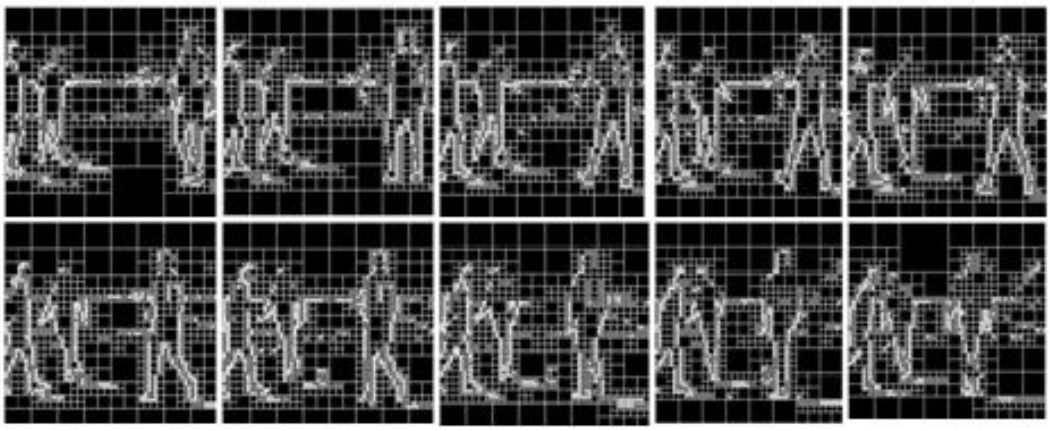
### C. Evaluation and Comparison of Modified Three Step Search (MTSS) Block Matching Estimation

The mean absolute difference (MAD) minimum cost function is used as the Block distortion measure (BDM) given by equation 2. The performance of the MTSS block matching motion estimation algorithm is compared with the existing block matching algorithms i.e., TSS, NTSS and ETSS. Performance evaluation parameters used for comparison are- MAD and total number of search points/ check points required per frame and the results are shown in Fig. 11 and 12 respectively.

a. First ten color frames of soccer video.

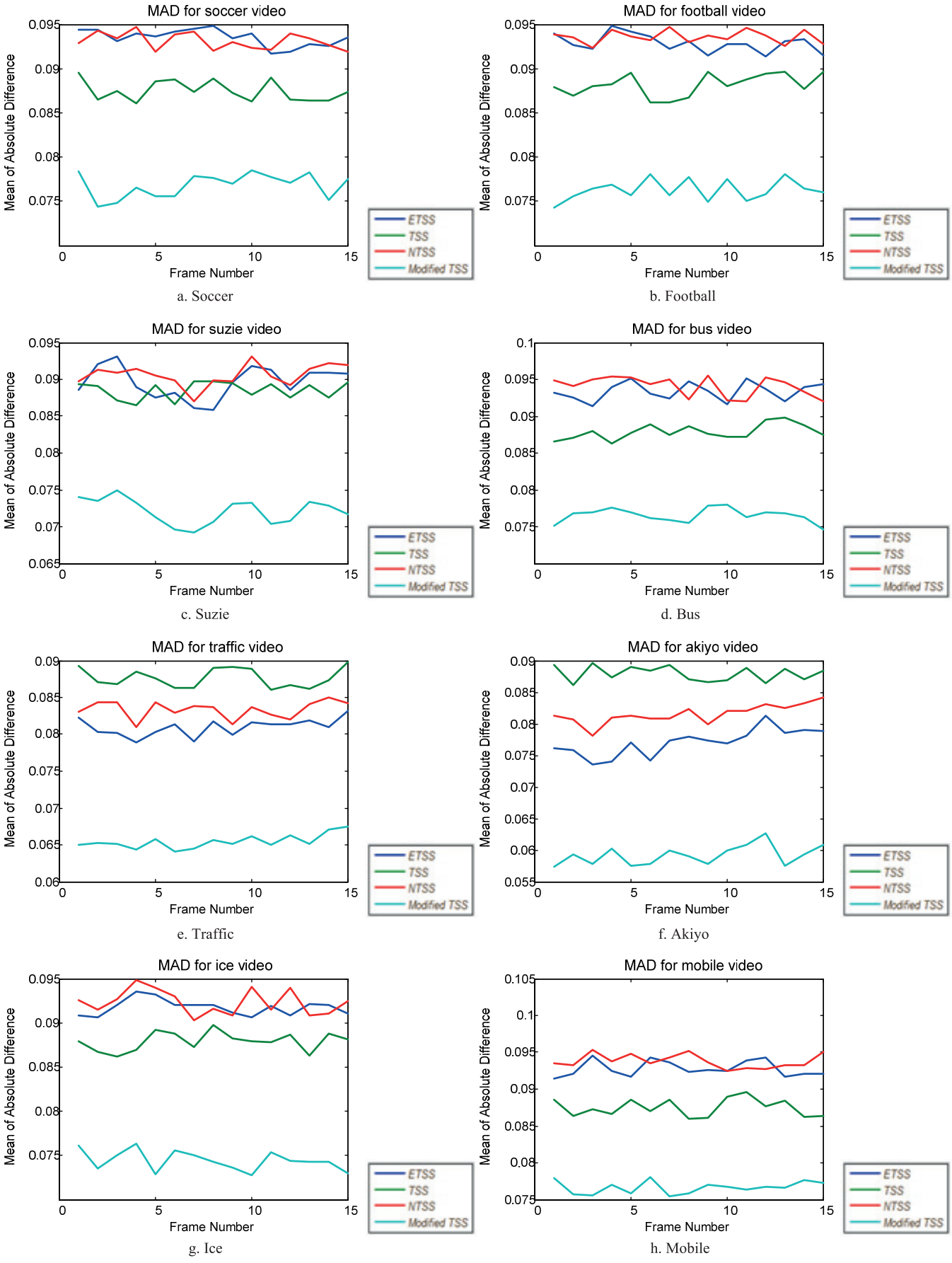
b. First ten YCbCr frames of soccer video.

c. First ten gray frames of soccer video.



d. First ten frames of quadtree partitioning

Fig.10. Intermediate results for the first ten frames of "soccer video".



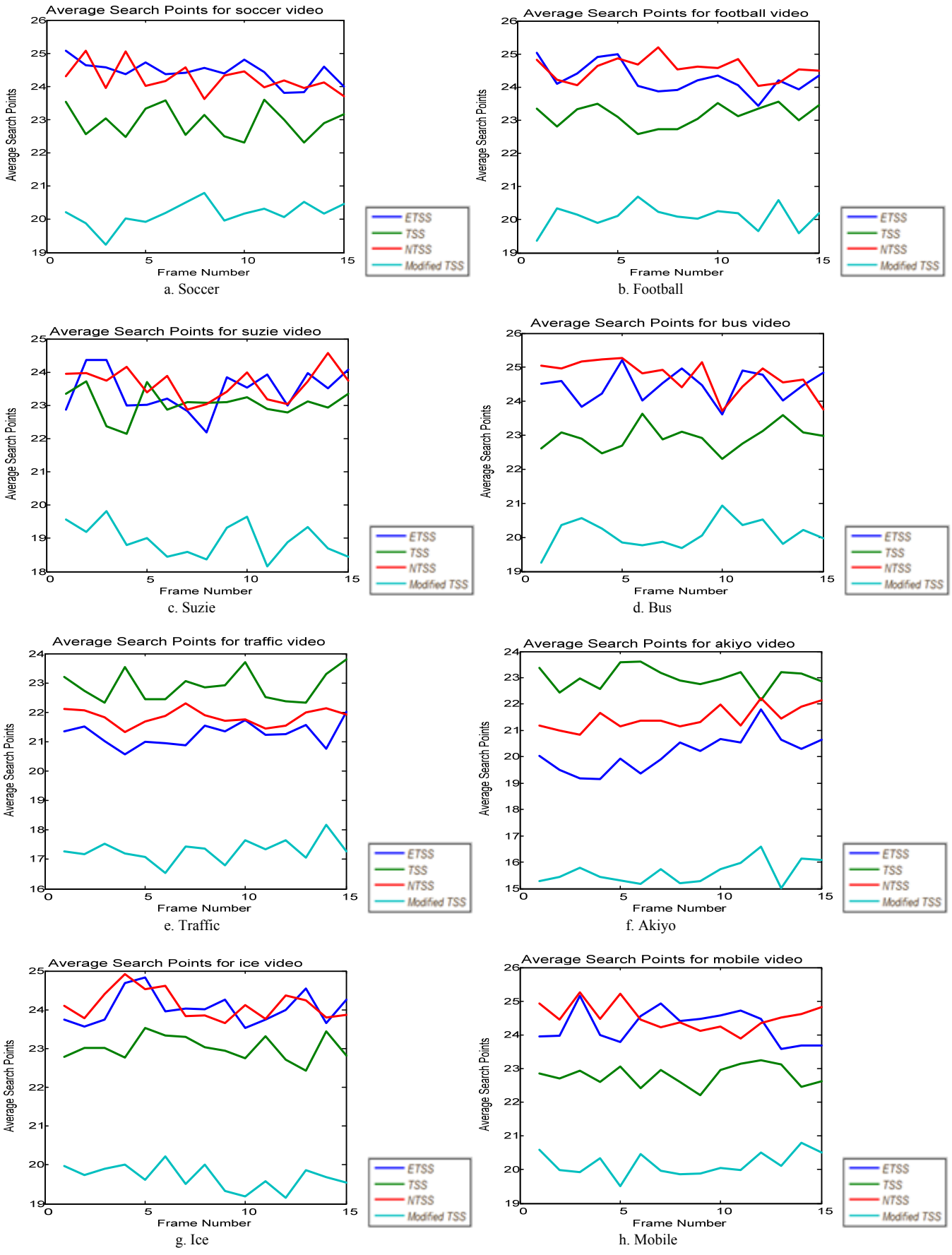

a. Soccer b. Football

c. Suzie d. Bus

e. Traffic f. Akiyo

g. Ice h. Mobile

Fig.11.Performance comparison of MTSS in terms of mean absolute difference (MAD).




a. Soccer b. Football c. Suzie d. Bus e. Traffic f. Akiyo g. Ice h. Mobile

Fig.12.Performance comparison of MTSS in terms of average search points/check points.

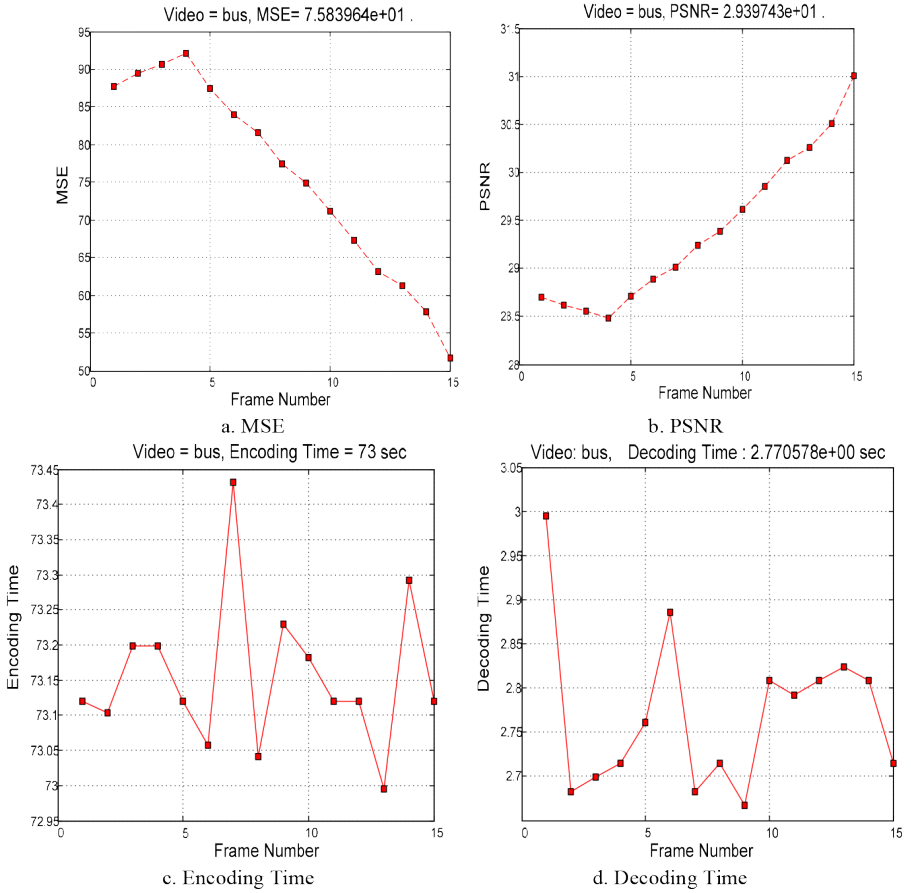



Fig.13. Performance of MTSS and WFA coding in terms of MSE, PSNR, Encoding and decoding Time.

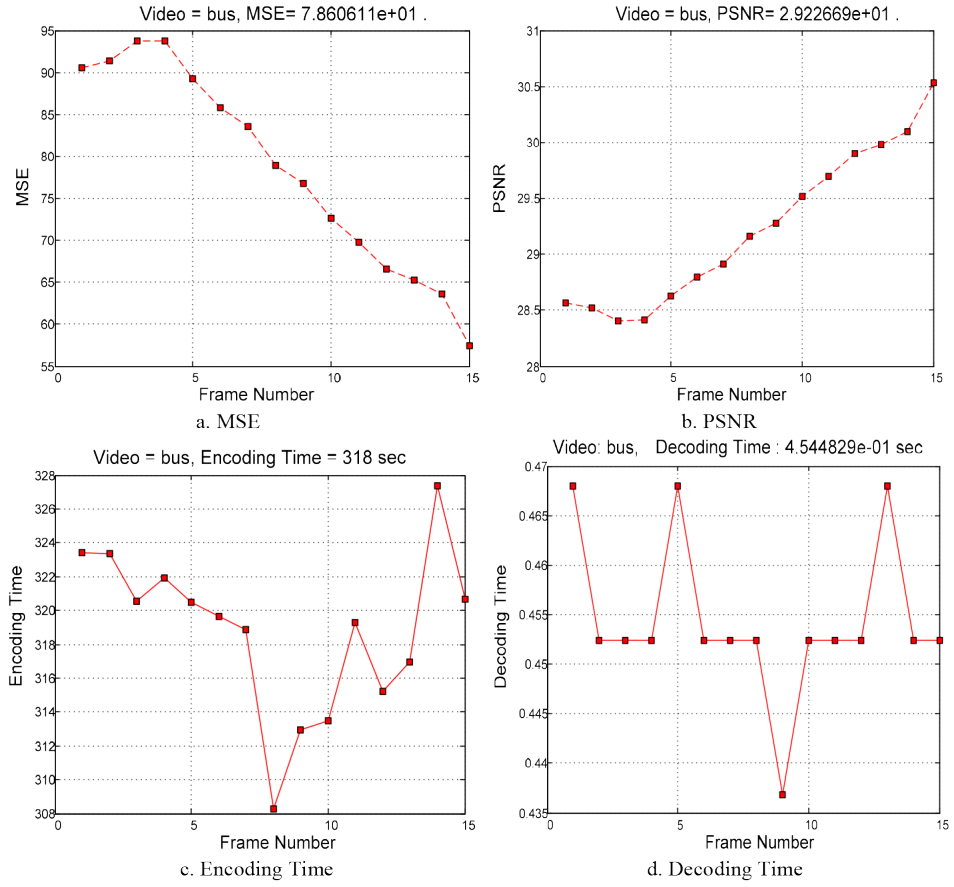



Fig.14. Performance of MTSS and Fractal Coding results for the first fifteen frames of "bus" sequences

During experimentation the search parameter p=3 pixels in both the horizontal and vertical directions for the macro block of size 16×16 are considered as the search area for a macro block.

From These Figures, the proposed MTSS coding gives less number of MAD and search points/check points required for each frame than TSS, NTSS and ETSS. The matching of one macro block with other macro block is based on the minimum output of cost function given in equation 2.

### *D. Evaluation and Comparison of MTSS Block Matching Estimation and Weighted Finite Automata Coding*

The performance of proposed approach measured on the evaluation parameters are quality of decoded video i.e. PSNR, MSE, compression ratio, decoding and encoding time at search parameter p=3 as shown in Table II.

Fig. 13 shows the performance measure of first 15 frames of bus sequence on MSE, PSNR, encoding time and decoding time.

### *E. Evaluation and Comparison of MTSS Block Matching Estimation and Fractal Coding*

In every image, there is a scope to have the self similarity. Fractal coding explores the self similarity existing in the images. For fractal encoding and decoding process the approach with range block of size 8×8 and domain block of size 16×16 is used. Table II to III gives the statistical performance comparison of first 15 frames of standard sequences. The proposed MTSS block matching algorithm and WFA coding approach is compared with MTSS and fractal coding approach given in Table III. Also both the approaches are compared with the frame by frame fractal coding approach. The fractal image compression approach is further implemented on image/video sequences based on intra-frame coding for comparison in terms of encoding time, compression ratio, and quality of video i.e. PSNR given in Table IV. Fig. 14 shows the performance measure in terms of MSE, PSNR and encoding time of first 15 frames of bus sequence using MTSS and fractal coding approach, and frame by frame based fractal coding approach respectively. Fig. 15 shows the 15$^{th}$ decoded frame of bus sequence in MTSS and WFA coding.

### *F. Evaluation and Comparison of Fractal Coding*

From the Table II to Table IV, we can find that the proposed video compression using MTSS and WFA coding performs better than MTSS and fractal coding, and frame by frame fractal coding in terms of achieving reduced encoding time and better quality of video. The performance measure of simple fractal coding (individual frame coding) for first 15 frames of video sequences in terms of MSE, PSNR, encoding time and decoding time is given in Table IV.

Fig.15. 15$^{th}$ decoded frame "bus" sequences.

TABLE II

PERFORMANCE COMPARISONS OF MTSS AND WFA CODING FOR FIRST 15 IMAGE SEQUENCES AT SEARCHING PARAMETER P=3: AVERAGE MSE, PSNR, ENCODING AND DECODING TIME

| Sequences | MSE | PSNR dB | Encoding Time (s) | Decoding Time (s) | Compression ratio |
|---|---|---|---|---|---|
| Soccer | 51.59 | 31.02 | 71 | 2.739 | 0.3025 |
| Suzie | 13.95 | 36.97 | 72 | 2.793 | 0.2779 |
| Traffic | 12.21 | 37.26 | 73 | 2.858 | 0.3825 |
| Bus | 75.83 | 29.39 | 73 | 2.770 | 0.3447 |
| Football | 60.27 | 30.33 | 72 | 2.747 | 0.3375 |
| Ice | 43.65 | 31.82 | 66 | 2.759 | 0.2689 |
| Akiyo | 7.94 | 39.14 | 73 | 2.831 | 0.0715 |
| Paris | 22.76 | 34.59 | 69 | 2.711 | 0.902 |
| Mobile | 40.88 | 32.06 | 65 | 2.760 | 0.095 |

TABLE III

PERFORMANCE COMPARISONS OF MTSS AND FRACTAL CODING FOR FIRST 15TH IMAGE SEQUENCES AT SEARCHING PARAMETER P=3: AVERAGE MSE, PSNR, ENCODING AND DECODING TIME

| Sequences | MSE | PSNR dB | Encoding Time (s) | Decoding Time (s) | Compression ratio |
|---|---|---|---|---|---|
| Soccer | 58.55 | 30.46 | 306 | 0.4451 | 0.3052 |
| Suzie | 23.95 | 34.35 | 312 | 0.430 | 0.2783 |
| Traffic | 14.17 | 36.62 | 263 | 0.400 | 0.3839 |
| Bus | 78.60 | 29.22 | 318 | 0.454 | 0.3418 |
| Football | 63.22 | 30.12 | 324 | 0.433 | 0.3416 |
| Ice | 52.69 | 30.98 | 305 | 0.426 | 0.2720 |
| Akiyo | 16.67 | 35.96 | 250 | 0.401 | 0.0721 |
| Paris | 39.86 | 32.18 | 303 | 0.404 | 0.0895 |
| Mobile | 54.95 | 30.73 | 337 | 0.427 | 0.1009 |

TABLE IV

RESULTS OF FRACTAL CODING FOR FIRST 15TH FRAMES ON STANDARD DATABASES: AVERAGE MSE, PSNR, ENCODING TIME AND DECODING TIME

| Sequences | MSE | PSNR dB | Encoding Time (s) | Decoding Time (s) | Compression ratio |
|---|---|---|---|---|---|
| Soccer | 44.65 | 31.64 | 4332 | 0.411 | 0.2848 |
| Suzie | 27.63 | 33.71 | 4330 | 0.408 | 0.2493 |
| Traffic | 27.05 | 33.81 | 4324 | 0.410 | 0.3612 |
| Bus | 58.88 | 30.43 | 4345 | 0.430 | 0.1410 |
| Football | 65.46 | 29.97 | 4349 | 0.416 | 0.2412 |
| Ice | 50.03 | 31.14 | 4311 | 0.409 | 0.2523 |
| Akiyo | 34.26 | 32.78 | 4328 | 0.407 | 0.1817 |

## VI. CONCLUSION AND FUTURE SCOPE

In this paper, a modified three step search (MTSS) algorithm and WFA coding approach is proposed to reduce the encoding time. A MTSS block matching motion estimation approach performs better in terms of small MAD and less average search points/check points required than the TSS, NTSS and ETSS at search parameter p=3 pixels. MTSS performs efficiently for the frames with slow and fast motions. Hence, the MTSS algorithm is suitable for video applications. The proposed MTSS and WFA coding approach performed better than MTSS and FC as well as frame-by-frame FC in terms of the encoding time. In MTSS and WFA, the encoding time is reduced by 70% to 80%

in comparison with MTSS and FC.

Developed block matching algorithms have scope for improvement through optimization of searching process by exploring the nearest neighborhood of pixels. The developed block matching algorithm can be combined with different coding mechanisms from video compression point of view.

## References


[1] Gonzalez, R. C., Woods, R. E. (2005) '*Digital Image Processing*', Second Edition, Pearson Education Asia.

[2] Acharjee, S., Dey, N., Biswas, D., Das, P., & Chaudhuri, S. S. (2012), "A novel Block Matching Algorithmic Approach with smaller block size for motion vector estimation in video compression", *12th* IEEE *International Conference on Intelligent Systems Design and Applications (ISDA)*, 2012, pp. 668-672.

[3] Acharjee, S., Biswas, D., Dey, N., Maji, P., & Chaudhuri, S. S. (2013), "An efficient motion estimation algorithm using division mechanism of low and high motion zone", IEEE *International Multi-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), 2013,* pp. 169-172.

[4] Acharjee, S., Pal, G., Redha, T., Chakraborty, S., Chaudhuri, S. S., & Dey, N. (2014), "Motion vector estimation using parallel processing", IEEE *International Conference on Circuits, Communication, Control and Computing (I4C), 2014,* pp. 231-236.

[5] Zhu, S., Tian, J., Shen, X. and Belloulata, K. (2009) '*A new cross-diamond search algorithm for fast block motion estimation*', In the IEEE Int. Conf. Image Processing, ICIP'09, Cairo, Egypt, Vol. I, pp. 1581-1584.

[6] Zhu, S., Tian, J., Shen, X. and Belloulata, K. (2009) '*A Novel Cross-Hexagon Search Algorithm Based on Motion Vector Field Prediction*', In the IEEE Int. Symposium on Industrial Electronics, ISIE'09, Seoul, Korea, pp. 1870-1874.

[7] Zhu, S., Hou, Y., Wang, Z. and Belloulata, K. (2010) '*A novel fractal video coding algorithm using fast block matching motion estimation technology*', In the International Conference on Computer Application and System Modeling, ICCASM'10,Taiyuan, China, Vol. 8, pp.360-364.

[8] Koga, T., Iinuma, K., Hirano, A., Iijima, Y. and Ishiguro, T. (1981) '*Motion compensated interframe coding for video conferencing*', In Proc. National Telecommunications Conf., New Orleans, LA, pp. G5.3.1 –G5.3.5.

[9] Jain, J. R. and Jain, A. K. (1981) '*Displacement measurement and its application in interframe image coding*', IEEE Transactions on Communications, vol. 29, pp. 1799-1808.

[10] Puri, A., Hang, H. M. and Schilling, D. L. (1987) '*An efficient block matching algorithm for motion compensated coding*', Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc., pp. 1063-1066.

[11] Ghanbar, M. (1990) '*The cross search algorithm for motion estimation*', IEEE Trans. Commun., Vol. COM-38, pp. 950-953.

[12] Urabe, T., Afzal, H., Ho, G., Pancha, P. and Zarki, M. E. (1994) '*MPEG Tool- an X window-based MPEG encoder and statistical tool*', Multimedia a syst. 1(5). pp. 220-229.

[13] R. Li, Zeng, B. and Liou, M. L. (1994) '*A new three-step search algorithm for block motion estimation*', IEEE Transactions on Circuits and Systems for Video Technology, vol. 4, no.4, pp. 438-443.

[14] Po, L. -M. and Ma, W.-C. (1996) '*A novel four-step search algorithm for fast block motion estimation*', IEEE Transactions on Circuits and Systems for Video Technology, vol. 6, no. 3, pp. 313–317.

[15] Liu, L. K. and Feig, E. (1996) '*A block based gradient descent search algorithm for block motion estimation in video coding*', IEEE Transactions on Circuits and Systems for Video Technology, vol. 6, no. 4 , pp. 419-422.

[16] Zhu, S. and Ma, K.K.(2000) '*A new diamond search algorithm for fast block-matching motion estimation*', IEEE Transactions on Image Processing, vol. 9, no. 2, pp. 287-290.

[17] Cheung, C. H. and Po, L. M. (2002) '*A novel cross-diamond search algorithm for fast block motion estimation*', IEEE Transactions on Circuits and Systems for Video Technology, vol. 12, no. 12, pp. 1168-1177.

[18] Jing, X. and Lap-Pui, C.(2004) '*An efficient three-step search algorithm for block motion estimation*', IEEE transactions on multimedia, Vol. 6, No.3, pp. 435-438.

[19] Tham, J.Y., Ranganath, S., Ranganath, M. and Kassim, A. A.(1998) '*A novel unrestricted center-biased diamond search algorithm for block motion estimation*', IEEE Transactions on Circuits and Systems for Video Technology, vol. 8, no. 4, pp. 369-377.

[20] Chen, H. M., Chen, P. H., Yeh, K. L., Fang, W. H., Shie, M. C. and Lai, F.(2007) '*Center of Mass-Based Adaptive Fast Block Motion Estimation*', EURASIP Journal on Image and Video Processing, vol. Article ID 65242, 11 pages.

[21] Zhu, C., Lin, X. and Chau, L-P.(2002) '*Hexagon-based search pattern for fast block motion estimation*', IEEE Transactions on Circuits and Systems for Video Technology, vol. 12, pp. 349-355.

[22] Cheung, C. H. and Po, L. M. (2000) '*Normalized partial distortion search algorithm for block motion estimation*', IEEE Transactions on Circuits and Systems for Video Technology, vol. 10, pp. 417-422.

[23] Li H., Liu M. (2009) '*Cross-Hexagon-based motion estimation algorithm using motion vector adaptive search technique*', International Conference on Wireless Communications & Signal Processing, Nanjing, 2009, pp. 1-4.

[24] Belloulata, K., Zhu, S., and Wang, Z. (2011) '*A Fast Fractal Video Coding Algorithm Using Cross-Hexagon Search for Block Motion Estimation*', International Scholarly Research Network Signal Processing, volume 2011, Article ID 386128.

[25] Acharjee, S., Ray, R., Chakraborty, S., Nath, S., & Dey, N. (2014), "Watermarking in motion vector for security enhancement of medical videos", IEEE *International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2014,* pp. 532-537.

[26] Ikeda, N., Araki, T., Dey, N., Bose, S., Shafique, S., El-Baz, A., ... & Suri, J. S. (2014), "Automated and accurate carotid bulb detection, its verification and validation in low quality frozen frames and motion video", *International angiology: a journal of the International Union of Angiology*, *33*(6), 573-589.

[27] Dey, N., Ashour, A., & Acharjee, S. (2017), "Applied Video Processing in Surveillance and Monitoring Systems", pp. 1-321, Hershey, PA: IGI Global. doi:10.4018/978-1-5225-1022-2.

[28] Kamble, S. D., Thakur, N.V., Malik, L. G. and Bajaj, P. R. (2014) '*Fractal Video Coding Using Modified Three-step Search Algorithm for Block-matching Motion Estimation*', Computational Vision and Robotics, Proceedings of International Conference on Computer Vision and Robotics, ICCVR'14, Advances in Intelligent Systems and Computing ,Vol. 332, pp 151-162 Springer-India.

[29] Culik, K. and Kari, J. (1995) '*Inference algorithms for WFA and image compression*', In Y. Fisher, editor, Fractal Image Compression, chapter 13, pages 243-258. Springer-Verlag.

[30] Kari, J. and Franti, P. (1994) '*Arithmetic coding of weighted finite automata*', Theoretical Informatics and Applications, 28(3-4):343-360.

[31] Hafner, U.(1996) '*Refining Image Compression with Weighted Finite Automata*', IEEE Data compression Conference, pp. 359-368.

[32] Katritzke, F. (2001) '*Refinements of Data Compression Using Weighted Finite Automata*', Ph.D. dissertation.

[33] Katritzke, F., Merzenich, W., and Thomas, M. (2003) '*Enhancements of partitioning techniques for image compression using weighted finite automata*', Elsevier, 2003.

[34] Kamble, S. D., Thakur, N.V., Malik, L. G. and Bajaj, P. R. (2015) '*Color video compression based on fractal coding using quad-tree weighted finite automata*', Information system design and intelligent application, Proceedings of Second International Conference INDIA 2015,vol.2, Advances in Intelligent System and Computing, Springer India, vol. 340, pp-649-658.

**Shailesh D. Kamble** received Bachelor of Engineering degree in Computer Technology from Yeshwantrao Chavan College of Enginering, Nagpur, India under the Rashtrasant Tukdoji Maharaj Nagpur University, Nagpur, India. He received Master of Engineering degree from Prof Ram Meghe Institute of Technology and Research, formerly known as College of Engineering, Badnera under Sant Gadge Baba Amravati University, Amravati, India. Presently, he is Ph.D. candidate in department of Computer Science and Engineering from G.H. Raisoni College of Engineering (An Autonomous Institution Affiliated to Rashtrasant Tukdoji Maharaj Nagpur University), Nagpur, India. He is working as the Assistant Professor in Department of Computer Science and Engineering at Yeshwantrao Chavan College of Enginering, Nagpur, India. He is having over 15 years of teaching and research experience. His current research interests include image processing, video processing and language processing. He is the author or co-author of more than 20 scientific publications in International Journal, International Conferences, and National Conferences. He is the life member of ISTE, India, IE, India.

**Nileshsingh V. Thakur** received Bachelor of Engineering degree in Computer Science Engineering from Government College of Engineering, Amravati, India and Master of Engineering degree from College of Engineering, Badnera under Amravati University and Sant Gadge Baba Amravati University in 1992 and 2005 respectively. He received Ph.D. degree in Computer Science Engineering under Department of Computer Science Engineering from Visvesvaraya National Institute of Technology, Nagpur, India on 1st February, 2010. His research interest includes image and video processing, sensor network, computer vision, pattern recognition, artificial neural network and evolutionary approaches. He is having over 24 years of teaching and research experience. Presently, he is Professor and Head in Department of Computer Science and Engineering and Dean, PG Studies at Prof Ram Meghe College of Engineering and Management, Badnera- Amravati, Maharashtra, India. He is the author or co-author of more than 70 scientific publications in refereed International Journal, International Conferences, National Journal, and National Conferences. He is a member of editorial board of over eight International journals; also, he is the life member of ISTE, India, IAENG and IAEME. He also worked as the reviewer for refereed international journals and conferences.

**Preeti R. Bajaj** is an Electronics Engineer Graduated in 1991, Post graduate in 1998 and awarded doctorate in Electronics Engg in 2004. Having 25 year of experience, currently she is Director of G.H. Raisoni College of Engineering, Nagpur. Her research interest includes Intelligent Transportation System, Soft Computing, Hybrid Intelligent Systems & Applications of Fuzzy logic in ITS. Her professional society affiliation includes Fellow-Institute of Engineers, Fellow IETE, Senior Member- IEEE, LM-ISTE, and LM-CSI, member ACM. She is presently Vice chair students activities- India council for IEEE. She is the first Indian to be selected as Technical Committee Chair on System Man and Cybernetics Society of the IEEE. She had been consultant for NHAI, Government of India. Six Doctoral students have been awarded PhD & 15 have done masters under her. She has authored & co-authored 100 plus reviewed Publications. She has published three book chapters. She is founder General Chair of ICETET series of IEEE Conferences. Under her leadership GHRCE has been awarded Autonomy, NBA accreditation to 21 programs, NAAC with A Grade, TEQIP World Bank project 1 and 2 leading to transform a private Engineering institution into India's premier, Engineering and Research Institute.

# Construction of a Benchmark for the User Experience Questionnaire (UEQ)

Martin Schrepp[1], Andreas Hinderks[2], Jörg Thomaschewski[2]

[1]SAP AG, Germany
[2]University of Applied Sciences Emden/Leer, Germany

***Abstract* — Questionnaires are a cheap and highly efficient tool for achieving a quantitative measure of a product's user experience (UX). However, it is not always easy to decide, if a questionnaire result can really show whether a product satisfies this quality aspect. So a benchmark is useful. It allows comparing the results of one product to a large set of other products. In this paper we describe a benchmark for the User Experience Questionnaire (UEQ), a widely used evaluation tool for interactive products. We also describe how the benchmark can be applied to the quality assurance process for concrete projects.**

***Keywords* — User Experience, UEQ, Questionnaire, Benchmark.**

## I. Introduction

In today's competitive market, outstanding user experience (UX) is a must for any product's commercial success. UX is a very subjective impression, so in principle it is difficult to measure. However, given the importance of this characteristic, it is important to measure it accurately. This measure can be used, for example, to check if a new product version offers improved UX, or if a product is better or worse than the competition [1].

There are several methods to quantify UX. One of the most widespread are usability tests [2], where the number of observed problems and the time participants need to solve tasks are quantitative indicators for the UX quality of a product. However, this method requires enormous effort: finding suitable participants, preparing tasks and a test system, and setting up a test site. Therefore typical sample sizes are very small (about 10-15 users).

In addition, it is a purely problem-centered method, i.e. it focuses on detecting usability problems. Usability tests are not able to provide information about users' impression of hedonic quality aspects, such as novelty or stimulation, although such aspects are crucial to a person's overall impression concerning UX [3].

Other well-known methods rely on expert judgment, for example, cognitive walkthrough [4] or usability reviews [5] against established principles, such as Nielsen's usability heuristics [6]. Like usability tests, these methods focus on detecting usability issues or deviations from accepted guidelines and principles. They do not provide a broader view of a product's UX.

A method that is able to measure all types of quality aspects and at the same time collect feedback from larger samples are standardized UX questionnaires. "Standardized" means that these questionnaires are not a more or less random or subjective collection of questions, but result from a careful construction process. This process guarantees accurate measuring of the intended UX qualities.

Such standardized questionnaires try to capture the concept of UX through a set of questions or items. The items are grouped into several dimensions or scales. Each scale represents a distinct UX aspect, for example efficiency, learnability, novelty or stimulation.

A number of such questionnaires exist. Questionnaires related to pure usability aspects are described, for example, in [8], [9]. Questionnaires covering the broader aspect of UX are, for example, described in [10], [11], and [12]. Each questionnaire contains different scales for measuring groups of UX aspects. So the choice of the best questionnaire depends on an evaluation study's research question, i.e. on the quality aspects to measure. For broader evaluations, it may make sense to use more than one questionnaire.

One of the problems in using UX questionnaires is how to interpret results, if no direct comparison is available. Assume that a UX questionnaire is used to evaluate a new program version. If a test result from an older version exists, the interpretation is easy. The numerical scale values of the two versions can be compared by statistical test to show whether the new version is a significant improvement.

However, in many cases the question is not "*Is UX of the evaluated product better than UX of another product or a previous version of the same product?" but "Does the product show sufficient UX?*" So there is no separate result to compare with. This is typically the case when a new product is released for the first time. Here it is often hard to interpret whether a numerical result, for example a value of 1.5 on the *Efficiency* scale, is sufficient. This is the typical situation where a benchmark, i.e. a collection of measurement results from a larger set of other products, is helpful.

In this paper we describe the construction of a benchmark for the User Experience Questionnaire (UEQ) [12], [13]. This benchmark helps interpret measurement results. The benchmark is especially helpful in situations where a product is measured with the UEQ for the first time, i.e. without results from previous evaluations.

## II. The User Experience Questionnaire (UEQ)

### A. Goal of the UEQ

The main goal of the UEQ is a fast and direct measurement of UX. The questionnaire was designed for use as part of a normal usability test, but also as an online questionnaire. For online use, it must be possible to complete the questionnaire quickly, to avoid participants not finishing it. So a semantic differential was chosen as item format, since this allows a fast and intuitive response.

Each item of the UEQ consists of a pair of terms with opposite meanings.

Examples:

*Not understandable* o o o o o o o *Understandable*

*Efficient* o o o o o o o *Inefficient*

Each item can be rated on a 7-point Likert scale. Answers to an item therefore range from -3 (fully agree with negative term) to +3 (fully



DOI: 10.9781/ijimai.2017.445

agree with positive term). Half of the items start with the positive term, the rest with the negative term (in randomized order).

### B. Construction process

The original German version of the UEQ uses a data analytics approach to ensure the practical relevance of the constructed scales. Each scale represents a distinct UX quality aspect.

An initial set of more than 200 potential items related to UX was created in two brainstorming sessions with two different groups of usability experts. A number of these experts then reduced the selection to a raw version with 80 items. The raw version was used in several studies on the quality of interactive products, including a statistics software package, cell phone address books, online collaboration software or business software.

In these studies, 153 participants rated the 80 items. Finally, the scales and the items representing each scale were extracted from this data set by principal component analysis [12], [13].

### C. Scale structure

This analysis produced the final questionnaire with 26 items grouped into six scales:

- *Attractiveness*: Overall impression of the product. Do users like or dislike it? Is it attractive, enjoyable or pleasing?

  6 items: *annoying / enjoyable, good / bad, unlikable / pleasing, unpleasant / pleasant, attractive / unattractive, friendly / unfriendly*.

- *Perspicuity*: Is it easy to get familiar with the product? Is it easy to learn? Is the product easy to understand and clear?

  4 items: *not understandable / understandable, easy to learn / difficult to learn, complicated / easy, clear / confusing*.

- *Efficiency*: Can users solve their tasks without unnecessary effort? Is the interaction efficient and fast? Does the product react fast to user input?

  4 items: *fast / slow, inefficient / efficient, impractical / practical, organized / cluttered.*

- *Dependability*: Does the user feel in control of the interaction? Can he or she predict the system behavior? Does the user feel safe when working with the product?

  4 items: *unpredictable / predictable, obstructive / supportive, secure / not secure, meets expectations / does not meet expectations*.

- *Stimulation*: Is it exciting and motivating to use the product? Is it fun to use?

  4 items: *valuable / inferior, boring / exciting, not interesting / interesting, motivating / demotivating*.

- *Novelty*: Is the product innovative and creative? Does it capture users' attention?

  4 items: *creative / dull, inventive / conventional, usual / leading-edge, conservative / innovative.*

Scales are not assumed to be independent. In fact, a user's general impression is captured by the *Attractiveness* scale, which should be influenced by the values on the other 5 scales (see Fig. 1).

*Attractiveness* is a pure valence dimension. *Perspicuity*, *Efficiency* and *Dependability* are pragmatic quality aspects (goal-directed), while *Stimulation* and *Novelty* are hedonic quality aspects (not goal-directed) [14].

Applying the UEQ does not require much effort. Usually 3-5 minutes are sufficient for a participant to read the instructions and complete the questionnaire. The UEQ can either be used in a paper-pencil form as part of a classical usability test (and this still is the most common application), but also as an online questionnaire.

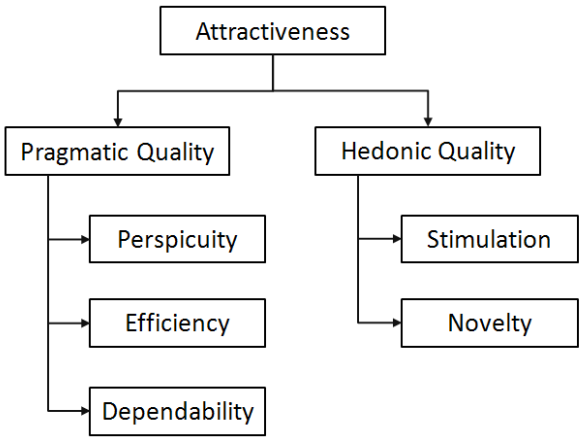



Fig. 1. Assumed scale structure of the User Experience Questionnaire (UEQ).

### D. Validation

The reliability (i.e. the consistency of the scales) and validity (i.e. that scales really measure what they intend to measure) of the UEQ scales was investigated in several usability tests with a total of 144 participants and an online survey with 722 participants. These studies showed a sufficient reliability of the scales (measured by Cronbach's Alpha). In addition, several studies have shown a good construct validity of the scales. For details see [12], [13].

### E. Availability and language versions

For a semantic differential like the UEQ, it is very important that participants can fill it out in their natural language. Thus, several contributors created a number of translations.

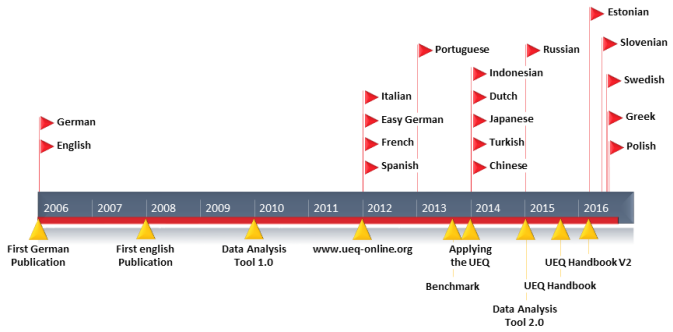



Fig. 2. Timeline of UEQ development.

The UEQ is currently available in 17 languages (German, English, French, Italian, Russian, Spanish, Portuguese, Turkish, Chinese, Japanese, Indonesian, Dutch, Estonian, Slovene, Swedish, Greek and Polish).

The UEQ in all available languages, an Excel sheet to help with evaluation, and the UEQ Handbook are available free of charge at www.ueq-online.org.

Helpful hints on using the UEQ are also available from Rauschenberger et al. [15].

## III. Why do we need a benchmark?

The goal of the benchmark is to help UX practitioners interpret scale results from UEQ evaluations.

Where only a single UEQ measurement exists, it is difficult to judge

whether the product fulfills the quality goals. See Fig. 3 as an example of an evaluation result.

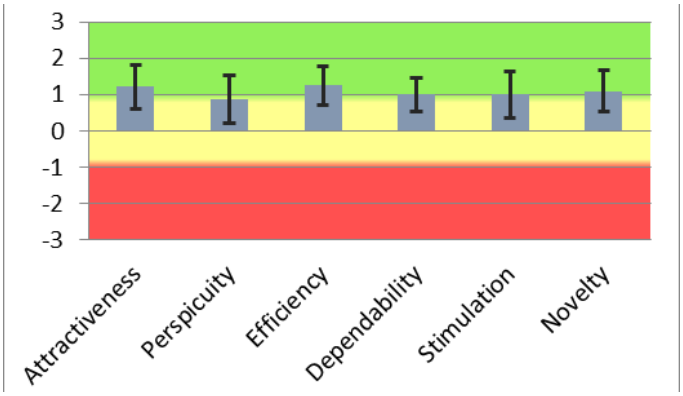



Fig. 3. Example chart from the data analysis Excel sheet showing the observed scale values and error bars for an example product.

Is this a good or bad result? Scale values above 0 represent a positive evaluation of the quality aspect; values below 0 represent a negative evaluation. But what does this actually mean? How do other products score?

If we have, for example, a comparison to a previous version of the same product or to a competitor product, then it is easy to interpret the results.

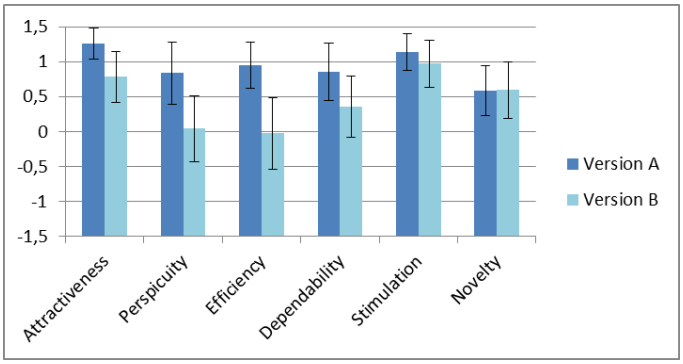



Fig. 4. Comparison between two different products. Here it is much easier to interpret the results, since the mean scale values can be directly compared.

A simple statistical test, for example a t-test, can be used to find out whether version A shows a significantly higher UX than version B.

But when a new product is launched, a typical question is whether the product's UX is sufficient to fulfill users' general expectations. Obviously no comparison to previous versions is possible in this case. It is also typically not possible to get evaluations of competitor products. The same is true for a product that has been on the market for a while, but is being measured for the first time.

Users form expectations of UX during interactions with typical software products. These products need not belong to the same product category. For example, users' everyday experience with modern websites and interactive devices, like tablets or smartphones, has also heavily raised expectations for professional software, such as business applications. So if a user sees a nice interaction concept in a new product, which makes difficult things easier, this will raise his or her expectations for other products. A typical question in such situations is: "*Why can't it be as simple as in the new product?*".

Thus, the question whether a new product's UX is sufficient can be answered by comparing its results to a large sample of other commonly used products, i.e. a benchmark data set. If a product scores high compared to the products in the benchmark, this can indicate that users will generally find the product's UX satisfactory.

## IV. Construction of the Benchmark

Over the last couple of years, such a benchmark was created for the UEQ by collecting data from all available UEQ evaluations. The benchmark was only made possible by a huge number of contributors, who shared the results of their UEQ evaluation studies. Some of the data comes from scientific studies using the UEQ, but most of the data comes from industry projects.

The benchmark currently contains data from 246 product evaluations using the UEQ. These evaluated products cover a wide range of applications. The benchmark contains complex business applications (100), development tools (4), web shops or services (64), social networks (3), mobile applications (16), household appliances (20) and a couple of other (39) products.

The benchmark contains a total of 9,905 responses. The number of respondents per evaluated product varied from extremely small samples (3 respondents) to huge samples (1,390 respondents). The mean number of respondents per study was 40.26.

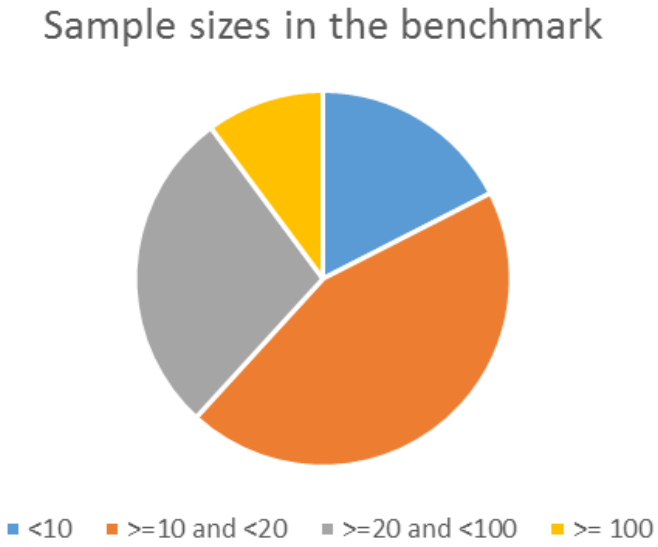



Fig. 5. Distribution of the sample sizes in the benchmark data set.

Many evaluations were part of usability tests, so the majority of the samples had less than 20 respondents (65.45%). The samples with more than 20 respondents were usually collected online.

Of course, the studies based on tiny samples with fewer than 10 respondents (17.07%) do not carry much information. It was therefore verified whether these small samples had an influence on the benchmark data. Since the results do not change much when studies with less than 10 respondents are eliminated, it was decided to keep them in the benchmark data set.

The mean values and standard deviations (in brackets) of the UEQ scales in the benchmark data set are:

- Attractiveness: 1.04 (0.64)
- Efficiency: 0.97 (0.62)
- Perspicuity: 1.06 (0.67)
- Dependability: 1.07 (0.52)
- Stimulation: 0.87 (0.63)
- Originality: 0.61 (0.72)

Nearly all of the data comes from evaluations of mature products, which are commercially developed and designed. Thus, it is no surprise that the mean value is above the neutral value (i.e. 0) of the 7-point Likert scale.

Since the benchmark data set currently contains only a limited number of evaluation results, it was decided to limit the feedback per scale to 5 categories:

- *Excellent*: The evaluated product is among the best 10% of results.
- *Good*: 10% of the results in the benchmark are better than the evaluated product, 75% of the results are worse.
- *Above average*: 25% of the results in the benchmark are better than the evaluated product, 50% of the results are worse.
- *Below average*: 50% of the results in the benchmark are better than the evaluated product, 25% of the results are worse.
- *Bad*: The evaluated product is among the worst 25% of results.

Table 1 shows how the categories relate to observed mean scale values.

TABLE I
BENCHMARK INTERVALS FOR THE UEQ SCALES

| | **Att.** | **Eff.** | **Per.** | **Dep.** | **Sti.** | **Nov.** |
|---|---|---|---|---|---|---|
| Excellent | ≥ 1.75 | ≥ 1.78 | ≥ 1.9 | ≥ 1.65 | ≥ 1.55 | ≥ 1.4 |
| Good | ≥ 1.52<br>< 1.75 | ≥ 1.47<br>< 1.78 | ≥ 1.56<br>< 1.9 | ≥ 1.48<br>< 1.65 | ≥ 1.31<br>< 1.55 | ≥ 1.05<br>< 1.4 |
| Above average | ≥ 1.17<br>< 1.52 | ≥ 0.98<br>< 1.47 | ≥ 1.08<br>< 1.56 | ≥ 1.14<br>< 1.48 | ≥ 0.99<br>< 1.31 | ≥ 0.71<br>< 1.05 |
| Below average | ≥ 0.7<br>< 1.17 | ≥ 0.54<br>< 0.98 | ≥ 0.64<br>< 1.08 | ≥ 0.78<br>< 1.14 | ≥ 0.5<br>< 0.99 | ≥ 0.3<br>< 0.71 |
| Bad | < 0.7 | < 0.54 | < 0.64 | < 0.78 | <0.5 | < 0.3 |

The comparison to the benchmark is a first indicator for whether a new product offers sufficient UX to be successful in the market. It is sufficient to measure UX by a large representative sample of users. Usually 20-30 users already provide a quite stable measurement. Comparing the different scale results to the products in the benchmark allows conclusions regarding the relative strengths and weaknesses of the product.

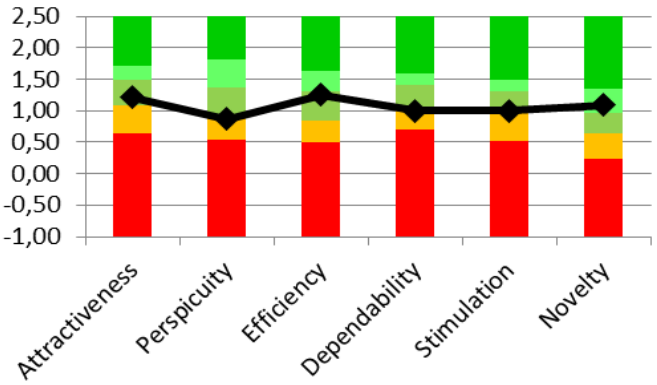



Fig. 6. Visualization of the benchmark in the data analysis Excel sheet of the UEQ. The line represents the results for the evaluated product. The colored bars represent the ranges for the scales' mean values.

It must be noted that the general UX expectations have grown over time. Since the benchmark also contains data from established products, a new product should reach at least the *Good* category on all scales.

## V. Benchmark as part of quality assurance

A UX benchmark can be a natural part of the quality assurance process for a new product. Assume that a new product is planned. The crucial quality aspects for a successful launch can easily be identified according to the product type and the intended market positioning. These identified quality aspects should reach a very good value in a later UEQ evaluation.

Let us assume that a new Web application should be developed. Users should be able to handle this application intuitively, without help or reading of documentation, to order services over the Web. The new application's design should be original and unconventional to grab users' attention. In addition, it should not be boring to use, so that users will come back.

In this example it is clear that *Perspicuity*, *Originality* and *Stimulation* are the most important UX aspects. So it would be a natural goal for the application to reach the *Excellent* category on these scales and at least an *Above Average* on the other UEQ scales. A benchmark – together with a clear idea of the importance of the UX quality aspects / UEQ scales – can help define clear and understandable quality goals for product development. These goals can easily be verified by using the UEQ questionnaire later on.

## VI. Conclusion

We described the development of a benchmark for the User Experience Questionnaire (UEQ). This benchmark helps interpret UX evaluations of products. It is currently available in 17 languages at www.ueq-online.org inside the "UEQ Data Analysis Tool" Excel file. The benchmark is especially helpful in situations where a product is measured for the first time with the UEQ, i.e. where no results from previous evaluations exist for comparison. In this article we also described how the benchmark can be used to formulate precise and transparent UX quality goals for new products.

A weakness of the current benchmark is that it does not distinguish between different product categories, i.e. there is only one benchmark data set for all types of products. Since most of the data in the benchmark comes from business applications or websites, it may be difficult to use for special applications or products, such as games, social networks or household appliances. The quality expectations for such types of products may simply be quite different from those expressed in the benchmark.

In the future we will try to create different benchmarks for different product categories. However, this requires collecting a larger number of data points per product category in UEQ evaluations and will therefore take some time.

## References

[1] Schrepp, M.; Hinderks, A. & Thomaschewski, J. (2014). Applying the User Experience Questionnaire (UEQ) in Different Evaluation Scenarios. In: Marcus, A. (Ed.): Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience. Lecture Notes in Computer Science, Volume 8517, S. 383-392, Springer International Publishing.

[2] Nielsen, J. (1994). Usability engineering. Elsevier.

[3] Preece, J., Rogers, Y.; Sharpe, H. (2002): Interaction design: Beyond human-computer interaction. New York: Wiley.

[4] Rieman, J., Franzke, M., & Redmiles, D. (1995, May). Usability evaluation with the cognitive walkthrough. In Conference companion on Human factors in computing systems (pp. 387-388). ACM.

[5] Nielsen, J. (1992, June). Finding usability problems through heuristic evaluation. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 373-380). ACM.

[6] Nielsen, J. (1994, April). Enhancing the explanatory power of usability heuristics. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (pp. 152-158). ACM.

[7] Nielsen, J. (1994): Heuristic Evaluation. In: J. Nielsen; R.L. Mack (Eds.): Usability Inspection Methods. New York: Wiley. S. 25-62.

[8] Brooke, J., 1996. SUS-A quick and dirty usability scale. Usability evaluation in industry, 189(194), 4-7.

[9] Kirakowski, J.; Corbett, M. (1993). SUMI: The Software Usability Measurement Inventory. British Journal of Educational Technology, Vol. 24, Nr. 3, S. 210–212.

[10] Hassenzahl, M.; Burmester, M.; Koller, F. (2003): AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. [AttrakDiff: A questionnaire to measure perceived hedonic and pragmatic quality] In: J.Ziegler; G. Szwillus (Eds.): Mensch & Computer 2003. Interaktion in Bewegung. Stuttgart: Teubner. S. 187-196.

[11] Visual Aesthetics of Websites Inventory (Moshagen, M. & Thielsch, M. T. (2010). Facets of visual aesthetics. International Journal of Human-Computer Studies, 68 (10), 689-709.).

[12] Laugwitz, B.; Schrepp, M. & Held, T. (2006). Konstruktion eines Fragebogens zur Messung der User Experience von Softwareprodukten. [Construction of a questionnaire for the measurement of user experience of software products] In: A.M. Heinecke & H. Paul (Eds.): Mensch & Computer 2006 – Mensch und Computer im Strukturwandel. Oldenbourg Verlag, S. 125 – 134.

[13] Laugwitz, B.; Schrepp; M. & Held, T. (2008). *Construction and evaluation of a user experience questionnaire.* In: Holzinger, A. (Ed.): USAB 2008, LNCS 5298, pp. 63-76.

[14] Hassenzahl, M. (2001). The effect of perceived hedonic quality on product appealingness. International Journal of Human-Computer Interaction, 13(4), pp. 481-499.

[15] Rauschenberger, M.; Schrepp, M.; Perez-Cota, M.; Olschner, S.; Thomaschewski, J. (2013): Efficient Measurement of the User Experience of Interactive Products. How to use the User Experience Questionnaire (UEQ). Example: Spanish Language Version. In: IJIMAI, 2(1), pp. 39-45.

**Martin Schrepp** has been working as a user interface designer for SAP AG since 1994. He finished his Diploma in Mathematics in 1990 at the University of Heidelberg (Germany). In 1993 he received a PhD in Psychology (also from the University of Heidelberg). His research interests are the application of psychological theories to improve the design of software interfaces, the application of *Design for All* principles to increase accessibility of business software, measurement of usability and user experience, and the development of general data analysis methods. He has published several papers in these research fields.

**Andreas Hinderks** holds a diploma in Computer Science and is Master of Science in Media Informatics by University of Applied Science Emden/Leer. He has worked as a Business Analyst and a programmer from 2001 to 2016. His focus then lay on developing user-friendly business software. Currently, he is a freelancing Business Analyst and Senior UX Architect. Also, he is a Ph.D. student at the University of Applied Science Emden/Leer. He is involved in research activities dealing with UX questionnaires, process optimization, information architecture, and user experience since 2011.

**Jörg Thomaschewski** was born in 1963. He received a PhD in physics from the University of Bremen (Germany) in 1996. He became Full Professor at the University of Applied Sciences Emden/Leer (Germany) in September 2000. His research interests are Internet applications for human-computer interaction, e-learning, and software engineering. Dr. Thomaschewski is the author of various online modules, e.g., “Human-Computer Communication,” which are used by the Virtual University (online) at six university sites. He has wide experience in usability training, analysis, and consulting.

# Handwritten Character Recognition Based on the Specificity and the Singularity of the Arabic Language

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

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

***Abstract* — A good Arabic handwritten recognition system must consider the characteristics of Arabic letters which can be explicit such as the presence of diacritics or implicit such as the baseline information (a virtual line on which cursive text are aligned and/join). In order to find an adequate method of features extraction, we have taken into consideration the nature of the Arabic characters. The paper investigate two methods based on two different visions: one describes the image in terms of the distribution of pixels, and the other describes it in terms of local patterns. Spatial Distribution of Pixels (SDP) is used according to the first vision; whereas Local Binary Patterns (LBP) are used for the second one. Tested on the Arabic portion of the Isolated Farsi Handwritten Character Database (IFHCDB) and using neural networks as a classifier, SDP achieve a recognition rate around 94% while LBP achieve a recognition rate of about 96%.**

***Keywords* — Handwritten Arabic Character Recognition, Feature Extraction, Texture Descriptor, Structural Feature.**

## I. Introduction

Today even with the emergence of new technologies, people still use the paper as a physical medium of communication and information storage. The collection and archiving of papers and historical documents is one of the greatest goals of nations, as these archives are an inexhaustible mine of valuable information.

Many documents are stored in their original form (as papers). Scanning might be enough to preserve these documents from degradation, but it is not good enough to allow for quick access to information using text queries.

Intelligent Character Recognition (ICR) systems allow the conversion of handwritten documents into electronic version, while Optical Character Recognition (OCR) systems deal with printed documents (produced by typewriter or computer). ICR is more difficult to implement because of the wide range of handwritten styles as well as image degradation.

There are several applications which use ICR, such as document digitization, storing, retrieving and indexing, automatic mail sorting, processing of bank checks and processing of forms. The importance of these applications has leads to intensive research for several years.

1. The architecture of an ICR system consists generally of five stages (fig.1):

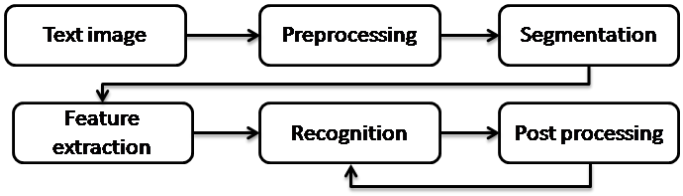



Fig. 1. Phases of character recognition process.

2. Preprocessing: contains techniques for image enhancement and normalization, such as: smoothing, noise reduction, slope normalization, contour detection, slant and skew correction …etc
3. Segmentation : the process of partitioning a document into homogeneous entities such as lines, words and characters;
4. Feature extraction: This deal with the extraction of some features from the image, which should permit discrimination between different classes (words or characters).
5. Learning/Recognition: which allows learning the classification rules based on the characteristics drawn from a training set.
6. Post-processing : It contains techniques for word verification (lexical, syntax and semantic)

Arabic language today is spoken by over 300 million people and it is the official language of many countries, and it contains a huge inheritance of documents to digitize. A robust system of recognition of Arabic handwriting documents can also serve other languages using the Arabic script such as Farsi, Urdu...

The Arabic script is written from right to left and is semi-cursive in both printed and handwritten versions. There are 28 letters, and the shape of these letters change depending on their position in the word, as they are preceded and/or followed by other letters or isolated, some letters can take four different forms: for example the letter 'Ain' (ع, عـ, ـعـ, ـع).

The diacritics play an essential role in reading, some letters have the same shape, and the only distinction is the number of points (diacritics) that can go up to three and their positions either on top or bottom of the letter. For example, three different letters 'ba','taa','thaa' (ب, ت, ث) have the same basic shape but the position and the number of points are different (fig.2).

All these facts make the Arabic script more challenging than other script like Latin.

Many collections of Arabic manuscripts are now in archives and libraries around the world, but unfortunately despite its importance, remained not exploited.

This paper focuses on the recognition of Arabic handwritten letters in their isolated form. We will investigate the efficiency of two feature extraction methods namely: Spatial Distribution of Pixels (SDP) and Local Binary Pattern (LBP) while taking into consideration the specificity of the Arabic script.

The rest of the paper is organized as follows: The next section, examines some related works on isolated Arabic character recognition. Section 3 describes the used dataset, the pre-processing and the classification stage. Section 4 provides a detailed overview and the results of the used feature extraction methods. Section 5 gives a comparative analysis of the proposed method with other works. Finally section 6 concludes the paper.



DOI: 10.9781/ijimai.2017.446

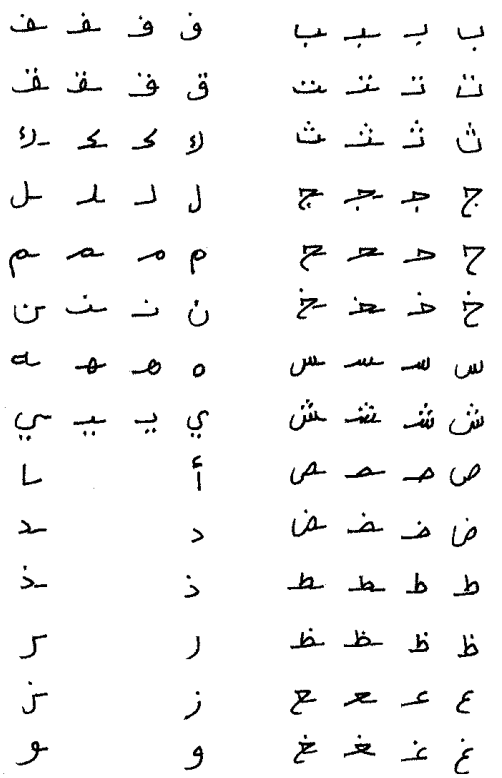


Fig. 2. Different form of the Arabic alphabets according to their positions.

All these facts make the Arabic script more challenging than other script like Latin.

Many collections of Arabic manuscripts are now in archives and libraries around the world, but unfortunately despite its importance, remained not exploited.

This paper focuses on the recognition of Arabic handwritten letters in their isolated form. We will investigate the efficiency of two feature extraction methods namely: Spatial Distribution of Pixels (SDP) and Local Binary Pattern (LBP) while taking into consideration the specificity of the Arabic script.

The rest of the paper is organized as follows: The next section, examines some related works on isolated Arabic character recognition. Section 3 describes the used dataset, the pre-processing and the classification stage. Section 4 provides a detailed overview and the results of the used feature extraction methods. Section 5 gives a comparative analysis of the proposed method with other works. Finally section 6 concludes the paper.

## II. Related Works

The challenge concerning the shape recognition problem such as handwritten character recognition remain in finding features that maximize the interclass variability while minimizing the intra-class variability [1]. Feature extraction methods can be categorized into two classes [2]:

- Structural features [3, 4], which extract geometrical and topological properties such as the number and position of dots, the presence of loops, the orientation of curves...etc.
- Statistical features [5], such as histograms of projection profile and transitions, moments, histograms of gray level distribution, Fourier descriptors and chain code...etc.

A Technique for recognizing hand printed Arabic characters using induction learning is presented in [6]. The method extracts structural features while tracing the path from the skeleton of the character where it finds an end or junction point. These features are primitives such as lines, curves and loops which are stored in a binary tree where the relationship between them is considered. The features such as the relation between primitives, the orientation of lines and curves, and the number of dots are used in an inductive learning program, in order to generate Horn clauses. As mentioned by the authors this can generalize over large degree of variation between writing styles. 30 samples from each character were selected for training and 10 samples were used for test. The average correct recognitions rate obtained using cross-validation was 86.65%.

The method in [7] uses preprocessing steps to remove noise, and then extract morphological and statistical features from the main body and secondary components. Using back propagation neural network on the CENPRMI [8] dataset, they report 88% of recognition rate.

Based on the extraction of normalized central and Zernike moments features from the main and secondary components, the work in [9], uses SVM as classifier with Non-dominated Sorting Genetic Algorithm for feature selection. The authors claim to reach 10% classification error on a dataset of isolated handwritten character of 48 persons.

The authors in [10] use neural network with the wavelet coefficients on a corpus of Arabic isolated letters from more than 500 writers. Using a network whose input vector contains 1024 input give 12% of error rate.

The work in [11] proposes a set of features to distinguish between similar Farsi letters. The first stage is based on the general shape structure to find the best match for a letter. In the second stage statistical feature such as distributive and concavity are extracted after partitioning a letter into smaller parts, this allows the distinction of structurally dissimilar letters. Vector quantization has been employed to test the features on 3000 letters and achieved 85.59% of accuracy.

The method proposed in [12], pre-processes the images in order to remove noisy points, and then uses moments, Fourier descriptor of the projection profile and centroid distance with Principal Component Analysis to reduce dimension of the feature vector. Using SVM on a database containing 1000 Arabic isolated characters, the authors achieved a 96.00% of recognition rate.

The paper in [13] presents a comparative study for window-based descriptors for the recognition of Arabic handwritten alphabet. Descriptors such as HOG, GIST, LBP, SIFT and SURF are first extracted from the entire image and evaluated using Support Vector Machine, Artificial Neural Network and Logistic Regression and then extracted from horizontal and vertical overlapped spatial portions of the image. The same paper introduces a dataset for Arabic handwritten isolated alphabet which contains about 8988 letters collected from 107 writers. The authors claims to reach 94.28% as the best recognition rate using SVM with RBF kernel, with SIFT features from the entire image.

In 2009 [14], a competition for handwritten Farsi/Arabic character and digit recognition, grouped four works. For character recognition the CENPARMI and IFHCDB [15] databases were used in training and testing steps. The best reported recognition rate is 91.85% that corresponds to a system based on hierarchy of multidimensional recurrent neural networks that works directly on raw input data with no feature extraction step.

The paper in [16] discusses the effectiveness of the use of Discrete Cosine Transform and Discrete Wavelet Transform to capture discriminative features of Arabic handwritten characters. On a dataset of 5600 characters, the coefficients of both techniques have been extracted in a zigzag fashion from the top-left corner of the image and used as input to the artificial neural network. Extracting 400 coefficients from the 128x128 resized image gives about 79.87% for DCT and 40.71% for DWT.

In [17] a two-stage SVM based scheme is proposed for recognition of Farsi isolated characters. After binarizing the image, the undersampled bitmaps and chain-code directional frequencies of the contour pixels are used as features. For the first stage the characters are grouped into 8 classes, and then the undersampled bitmaps feature is used to assign an input image to the class that belongs to. In the second stage the classifier are trained on those classes using this time chain-code feature in order to discriminate between the characters belonging to the same class. Using the IFHCDB database, the authors reach a recognition rate of 98% and 97%, respectively for 8-class and 32-class problems.

In [18], after applying some preprocessing technique and normalizing the image, zoning and crossing counts are combined to represent the feature set. Self-organized map is used to cluster the classes and for creation of binary decision tree. For each node the classifier among SVM, KNN and neural networks who gives the best recognition rate is considered as the main classifier of the node. Tested on the IFHCDB dataset, the authors claim the reach a recognition rates of 98.72, 97.3 and 94.82 respectively when considering 8, 20 and 33 clusters.

A method was proposed [19] for the recognition of Persian isolated handwritten characters, after preprocessing the images, the derivatives of the projection profile histograms in four directions is used as feature with a Hamming neural network. This classifier is implemented using CUDA on GPU in order to speed-up the classification time. On a subset of the Hadaf database, the author claim to reach a 94.5% of recognition rate while accelerating the algorithm 5 times.

## III. Dataset, preprocessing and classification stage

### A. Dataset

The database used in this work, is the same used in the ICDAR 2009 competition [14] under the name of Isolated Farsi Handwritten Character Database (IFHCDB) [15], it contains 52380 characters and 17740 numbers. The images are scanned in grayscale at a resolution of 300 dpi, and each character has a size of 95x77. The distribution of characters in this database is not uniform, which means that the number of samples is not the same for all characters.

In this work only of the Arabic portion (28 Arabic characters) is considered, which represents approximately 97% of all character set (Arabic and Farsi), and which is divided into 35989 images for learning and 15041 for the testing.

### B. Preprocessing

In the feature extraction phase the used LBP descriptor (see section IV) which allows to describe textures relies heavily on the distribution of the grayscale level in the image and since the characters in the dataset are extracted from specific areas; these areas sometimes contain noise and degradation of gray level. Extracting the LBP histogram from the entire image of the character poses a risk of capturing useless information and can thus have a negative impact on the recognition results.

To remedy to this problem; first the image is binarized using a global threshold (i.e. Otsu), then convolved with a “low pass” filter, which results in smoothing the contour of the character (fig.3).

This pretreatment allows LBP to capture the grayscale level of pixels that lie within the contour which will better discriminate between different characters.

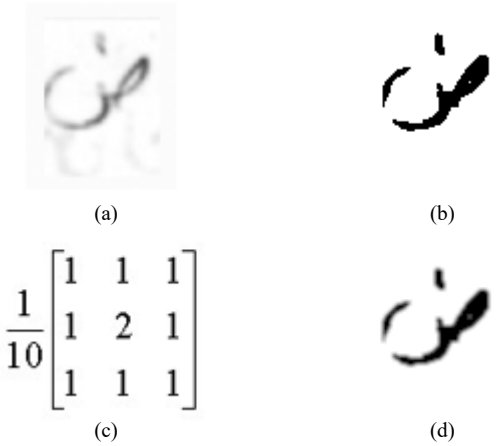



Fig. 3. Preprocessing steps: (a) The original image, (b) the binary image using Otsu, (c) the smoothed image using the filter in (d).

### C. Classification

The interest here is to investigate the feature extraction methods, in order to choose the one that works best. So initially we choose Artificial Neural Networks (ANN) as classifier. Afterwards, if the used classifier is replaced with another one that outperforms neural networks, it will certainly improve the results.

A feed-forward neural network with Scaled Conjugate Gradient as training algorithm is used here. After experimental tests, a configuration of 100 hidden layers gives the optimal recognition rates.

## IV. Feature extraction

To recognize the Arabic characters, a descriptor must be able to recognize two different characters that are written similarly (many characters have the same body shape and can only be differentiated using the information about the diacritical points), but at the same time has the capability to recognize the same character that is written differently as shown in the Fig. below.


Fig. 4. Some Arabic characters writing in a similar way. (a) samples of handwritten character, (b) printed form of the characters in (a).

Furthermore, the baseline information which is a special characteristic of the Arabic script has to be taken into consideration to further improve the recognition of similarly written characters.

### A. The concept of baseline in the Arabic script

The baseline is a virtual line on which cursive text are aligned and join. According to the shape of Arabic characters, the baseline can be either at the top, at the bottom or across the body of the character (the horizontal lines in Fig. 5).

Words of Arabic language are made up of characters written from right to left and linked with each other. These linking parts allow continuity and smoothness in the writing which in turn allows easy and fast reading (the red portions in Fig. 5).

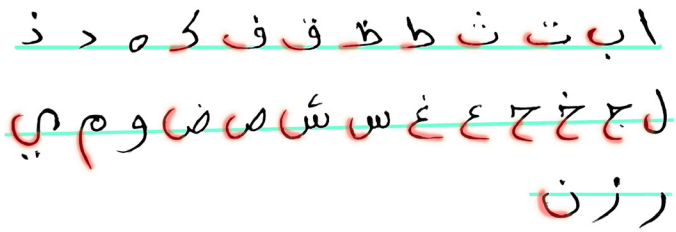
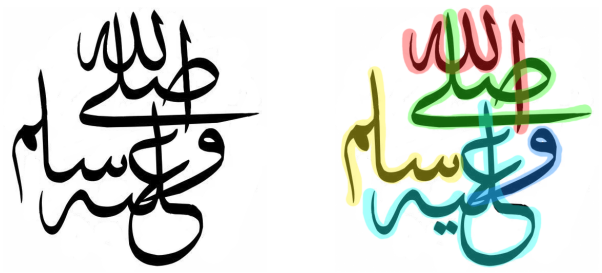


Fig. 5. The position of the baseline (bottom, across and top) according to different characters.

In fact most of isolated Arabic characters have two parts: one contains the identity of the character (the useful part) and the other contains the linking part with the next character. This gives Arabic script more flexibility in writing, such as the case of calligraphy as shown in Fig. 6, we steel read the words easily and this thanks to the presence of the baseline information which lies in useful part of the characters.



(a) (b)

Fig. 6. Example of Arabic calligraphy: (a) the original sentence, (b) detection of different word in the sentence according to the baseline information.

When dealing with the recognition of Isolated Arabic characters one must take into consideration that the linking parts in the characters can lead to confusion between those having similar ones. As a solution, in the feature extraction phase, the character must not be considered as one part; rather it must be segmented into different parts allowing us to distinct the useful one from the linking one. So we seek a unique way to divide all the characters in a way that approximates their baseline.

Inspired by these observations and in order to cope with the problem of similarity of Isolated Arabic characters, the problem can be seen from two different perspectives:

- The shape of character is a set of pixels that are spatially distributed. The way these pixels are arranged informs us about the structural features of such character and gives us idea about the location of the useful parts. Here, techniques that capture the distribution of pixels from different regions can be used.
- The way how the character is written influences on it shape i.e. the texture in the beginning of the character (the useful part) is not the same as in its ending. So statistical techniques can be used to better capture this information.

In the following, we introduce each of the chosen feature extraction method namely Spatial Distribution of Pixels and Local Binary Patterns corresponding to the above perspectives.

### B. Spatial Distribution of Pixels (SDP)

#### 1) Definition

The image of the character can be seen as a set of pixels that are linked together and have a certain position and dispersion in the space. Capturing this information will allow to better distinguish between the Arabic letters.

To measure this dispersion of pixels, the rectangle enclosing the character and its diacritical points is divided using a grid, which will allow for a better comparison between different characters even if they could have different sizes.

As explained below, partitioning of the image into an odd number will be adequate to measure the symmetry and the arrangement of the Arabic characters:

- The division of the character into four equal and symmetrical parts allows the distinction between symmetrical characters such as (ب, ت, ن…) from asymmetrical ones (ص, ش, و, م…).
- The location of diacritics in columns C2, C3 and C4 (fig.7), allows to differentiate characters that have similar shapes but differ in the number and position of diacritics such as (س, ش, ر, ز, ت, ث).
- The location of the baseline in rows R2, R3 and R4 (fig.7), provides additional information to differentiate between characters such as (ب, ن, ا, ر…)
- The division of the character in four regions: up, down, left and right, allows it to be precisely located in the bounding box i.e. (و, م, ل, ن…).

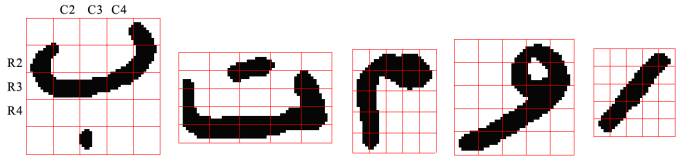



Fig. 7. Partitioning the rectangle enclosing the character in 25 regions of the same size.

The reason we choose to divide the character into a 5x5 grid is justified by the fact that 5 is the smallest number that permits the above points while reducing the computation time.

In what follows, certain blocks in the grid that will allow us to capture the information mentioned above are explained.

#### 2) SDP application

To extract feature vector of spatial distribution of pixels, four configurations are considered:

##### a) The first configuration

The number of black pixels in the four blocks of the grid is divided by the area of these blocks (each block contains four regions as shown in fig.8.a). In the same way, the percentage of black pixels in the middle column and the middle row of the grid is computed as shown in fig.8.b and fig.8.c.

Finally these percentages are used to differentiate between characters.

Using the feed-forward neural network, we have got a 81.22% of recognition rate.

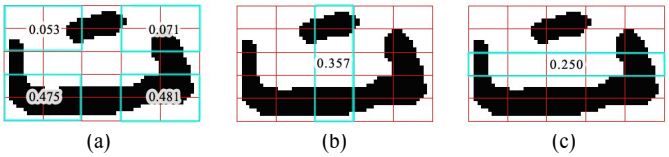



Fig. 8. The calculation of percentages according to the first configuration.

*b) The second configuration*

In order to integrate the information about the location of the character's body, the following percentages are calculated:

- The percentage of black pixels in the two rows at top (fig.9.a).
- The percentage of black pixels in the two rows at down (fig.9.a).
- The percentage of black pixels in the two columns at left (fig.9.b).
- The percentage of black pixels in the two columns at right (fig.9.b).

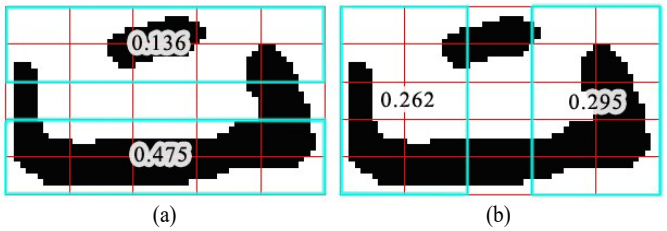



Fig. 9. The calculation of percentages according to the second configuration.

In addition to the previous configuration, the feature vector in this one contains the four percentages about the location of the character.

Using the feed-forward neural network with this new configuration, we have achieved a rate of 87.65%.

*c) The third configuration*

After splitting the rectangle enclosing the character and its diacritics in 25 regions of the same dimensions, the percentage of black pixels in each region is calculated by dividing their number by the total number of black pixels in the rectangle. In this case the overall shape of the image is the one of interest (like zoom out).

These percentages are inserted line by line in a vector which will be considered as descriptor of the character (fig.10).

Using the feed-forward neural network, we found 92.99% of recognition rate.

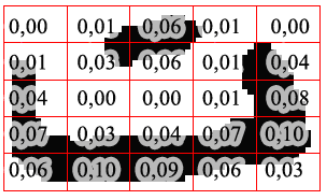



Fig. 10. The calculation of percentages according to the third configuration.

*d) The fourth configuration*

In this configuration, the percentage of black pixels in each region is calculated by dividing their number by the area of the region where they are located. In this case, the interest is the information about the presence and the fill of local pixels in each region (like zoom in).

These percentages are inserted line by line in a vector which is used to describe the character (fig.11).

Extracting the vector according to this configuration has achieved a rate of 93.84%.

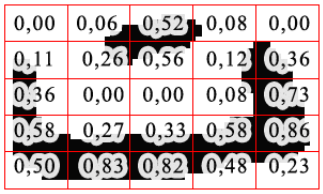



Fig. 11. The calculation of percentages according to the fourth configuration.

*3) Summary*

Table I shows the TOP-5 measures of recognition rates of the four configurations for extracting SDP feature using the feed-forward neural network described above.

For instance the TOP2 and TOP4 are respectively the percentages of samples that the true class is among the two first and the four positions in the list of candidates.

TABLE I

TOP-5 MESURES FOR DIFFERENT CONFIGURATIONS OF SDP EXTRACTION

| Configuration of SDP extraction | TOP1 | TOP2 | TOP3 | TOP4 | TOP5 |
|---|---|---|---|---|---|
| First configuration | 81.22 | 91.52 | 95.76 | 97.50 | 98.37 |
| Second configuration | 87.65 | 95.39 | 97.83 | 98.82 | 99.25 |
| Third configuration | 92.99 | 97.61 | 98.80 | 99.21 | 99.43 |
| **Fourth configuration** | **93.84** | **97.76** | **98.84** | **99.27** | **99.44** |

The first configuration includes the concept of symmetry, location of the baseline and diacritical points. In addition to the first configuration, the second one includes information about the location of the character in the grid.

In the third configuration, the percentage of black pixels in all the cells of the grid, allows to encompass all the concepts of the first two configurations, which results by an increase in the recognition rate.

Finally, in the fourth configuration, the calculation of the percentages of pixels with respect to the regions they are located in, which also helps to make the descriptor more insensitive to the size of the character, and therefore improves the recognition rate evermore.

## C. Local Binary Pattern

*1) Definition*

Developed by Ojala et al [20], LBP for Local Binary Pattern is a descriptor designed firstly to analyze textures in terms of local spatial patterns and gray level contrast.

In its basic form, a 3x3 window is used to produce labels for each pixel by thresholding to neighboring pixels with the center pixel in the window and considers the result as a binary number. The histogram of the $2^8 = 256$ different labels is used to describe a texture.

In [21], LBP has been improved to be used with a circular neighboring with a radius and a number of neighborhoods.

The decimal value of a pixel is calculated as:

$$LBP_{P,R} = \sum_{p=0}^{P-1} f(g_p - g_c)2^p \qquad f(x) = \begin{cases} 1, & f \quad x \geq 0; \\ 0, otherwise \end{cases}$$

With, $P$ stands for the number of neighbors, $R$ is the radius of the circle, and $g_p$ is the gray level of the pixel in the position $p$.

Fig. 12 shows an example of LBP calculation. The neighboring pixels that are greater or equal to the value of the central pixel are replaced by 1 and the lower pixels are replaced by 0. The LBP label corresponds to the binary value taken in a circular order (11101101) which is then converted to decimal (237).

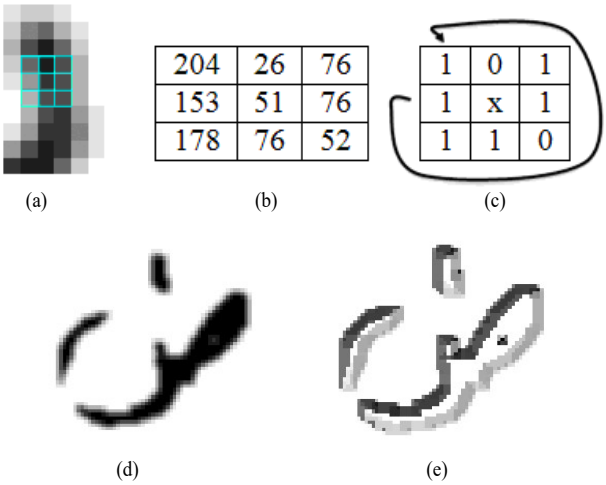



Fig. 12. Example of LBP computation: (a) a 3x3 window centered on a pixel, (b) the gray levels in the window, (c) thresholding the neighborhood compared to pixel x, (d) an image after preprocessing, (e) result of LBP application with a configuration $LBP_{8,1}$

Another extension of LBP is called Uniform Pattern [21], it allow reducing the size of the vector while keeping the same performance. The idea comes from the fact that there are patterns that occur frequently compared to others and must be grouped in a single pattern.

A pattern is called uniform if it contains at most two bit transitions from 0 to 1 and vice versa. For example 00000000 (0 transitions), 01110000 (two transitions) and 11001111 (two transitions) are Uniform, but the patterns 11001001 (4 transitions) and 01010010 (6 transitions) are not. For a neighborhood of 8 pixels, there are 256 patterns, 58 of them are uniform.

After the calculation of the label of each of pixel $f_{lbl}(x,y)$, the LBP histogram can be defined by:

$$H_i = \sum_{x,y} I\{f_{lbl}(x,y) = i\}, i = 0,...,n-1$$

*N* is the number of labels, *I{A}* equal 1 if A is true, otherwise it equal 0.

In addition to its success in several areas such as: face recognition [22], writer identification [23], and handwritten and digits recognition [24], the motivation behind the use of LBP as a descriptor is justified by the fact that it has proved its discriminating power for textures through the combination of information of spatial local patterns and intensity which we believe can be useful in the shape recognition problem.

To detect the variation in the strokes of the character, we can look at the neighborhood of the pixels in the contour. We found out that it is more appropriate to use LBP method which offers technical possibilities to implement this vision.

In this work, the uniform version of LBP with a configuration of $LBP_{8,1}$ is used.

### 2) LBP application

The natural idea is to extract the LBP histogram from the entire image (fig.13). Using the feed-forward neural network, we have got an 85.26% of recognition rate.

To improve this recognition rate, the useful part of the image is considered, i.e. the part containing only the character, by cropping it.

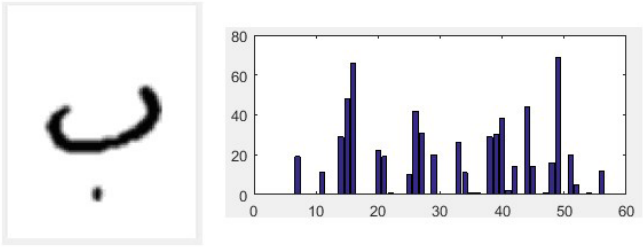



Fig. 13. (a) The input image, (b) the LBP histogram extracted from the entier image.

### a) The first configuration

In this configuration, the character and its diacritical point are included within the smallest rectangle (fig.14).

Extracting LBP histogram in this configuration has achieved a rate of 88.47%, an improvement of about 3.21% compared to the extraction of LBP from the entire image.

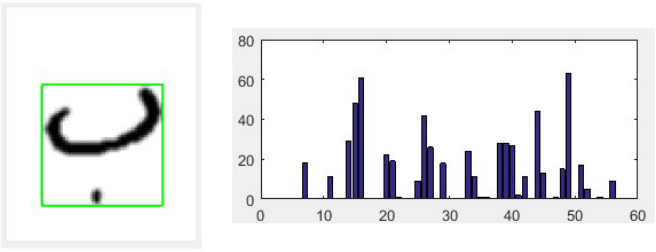



Fig. 14. (a) In green the rectangle enclosing the character, (b) the LBP histogram extracted from the rectangle.

### b) The second configuration

In their isolated forms several Arabic characters have similar endings or terminations (ج, ع, س, ص). If we only extract the LBP histogram from the whole image, it does not effectively allow differentiating between different characters that are written almost in the same way.

To remedy to this problem, we try to approximate the baseline of the character by dividing the rectangle enclosing the character into four regions from the center of gravity of the body of the character, in order to break it into the four quadrants while highlighting areas which do not look the same. As shown in fig.15, the portions a2 and a4 respectively look similar to the portions b2 and b4, on the other hand portions a1, a3, b1 and b3 are different.

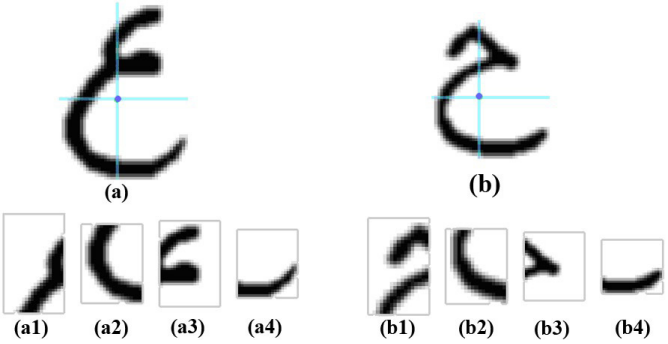



Fig. 15. Different regions after spliting the character into four regions from its centroid.

After splitting the bounding box into four regions (quadrants) from the centroid of the character, LBP histograms are extracted from each region and concatenated to form one feature vector of 236 elements (fig.16).

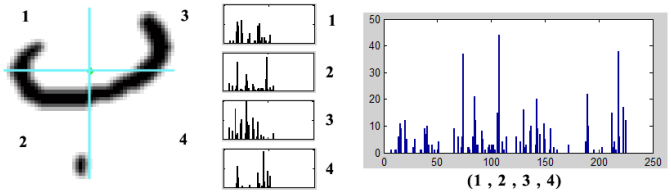



Fig. 16. Concatenation of the histograms extracted from the four regions from the centroid of the character.

This configuration has significantly improved the recognition rate, until we achieve a rate of 96.10%.

*c) The third configuration*

In the previous configuration we were interested on the body shape of the character, but in the Arabic language, there are more characters that are distinguishable only by the number and position of diacritics, for example (ح, خ, ب, ت, ن). In addition the histogram of LBP of the entire image does not capture the information of the existence of diacritical points.

To address this problem and to better approximate the baseline, the rectangle enclosing the character is divided into four regions from the point situated halfway between the centroid of the body and the centroid of diacritics. This will highlight the region that contains diacritical point(s). As shown in fig.17, the a2 and b3 portions contain information of diacritical points.

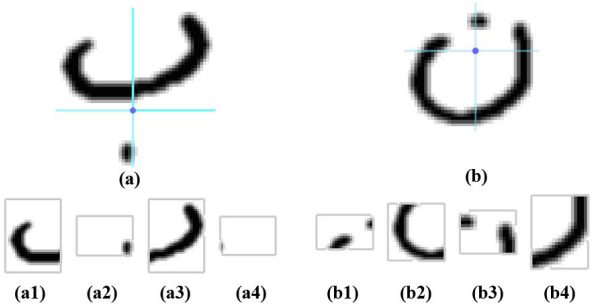



Fig. 17. Different regions after spliting the character into four regions from the centroid of both the character and its diacritics.

After splitting the bounding box in four regions from the center point of the centroid of the character and the centroid of diacritics, LBP histograms are extracted from each region and concatenated to form a vector of 236 elements (fig.18).

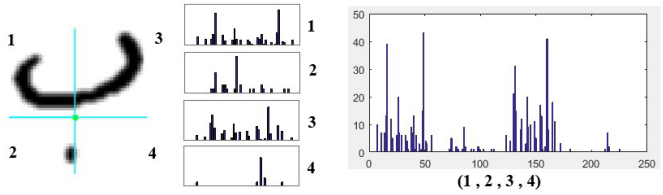



Fig. 18. Concatenation of the histograms extracted from the four regions from the centroid of of the character and its diacritics.

This configuration enables to further improve the recognition rate and achieve a rate of 96.31%.

*3) Summary*

Table II illustrates the top-5 recognition rate measures of different ways of extracting LBP features.

TABLE II

TOP-5 MESURES FOR DIFFERENT CONFIGURATIONS OF LBP EXTRACTION

| Configuration of LBP extraction | TOP1 | TOP2 | TOP3 | TOP4 | TOP5 |
|---|---|---|---|---|---|
| The entire image | 85.26 | 93.26 | 96.28 | 97.72 | 98.43 |
| First configuration | 88.47 | 94.98 | 97.18 | 98.26 | 98.84 |
| Second configuration | 96.10 | 98.88 | 99.45 | 99.68 | 99.80 |
| **Third configuration** | **96.31** | **98.91** | **99.44** | **99.65** | **99.78** |

It is clear that it is important to use the bounding box enclosing the character instead of the whole image. The concatenation of LBP histograms extracted according to the third configuration gives the best results.

These results are obtained because the third configuration process respected the nature and the characteristics of the Arabic language:

- Using the centroid information to get close to the baseline of the character, which is in fact within the character itself.
- Dividing the bounding box into four regions, allows differentiating different characters having resembling shapes (ع, ح).
- Integrating the centroid information of diacritics, allows to differentiate between similar letters shapes but differs in the diacritics (ن, ب, ث, ت).

## V. COMPARATIVE ANALYSIS

From the results it is clear that LBP descriptor outperforms SDP, in what follows a comparative analysis of the proposed method (the third configuration of LBP) with the existing systems working on IFHCDB dataset is performed.

The dataset used here contains the Farsi letters (The known 28 Arabic letters plus the letters پ, ژ, چ, گ, which makes it 32 letters). So in order to test the robustness of the proposed method it will be interesting to experiment it on the Farsi set too.

In some works the 32 Persian characters are grouped in 8 or 20 classes containing character with similar shapes. Therefore in order to be able to compare the proposed method with the existing ones we consider the same classes in those studies to construct the 8-class, 32-class and 33-class problems.

Table III shows both versions adopted to build the 8-class:

TABLE III

THE TWO USED VERSION OF 8-CLASS PROBLEM

| | 8-class version 1 (8-v1) | 8-class version 2 (8-v2) |
|---|---|---|
| Class 1 | ظ – ط – آ - ا | ه – د – ر – و – ا |
| Class 2 | ث - ت – پ – ب – ن – ل – ق – ف | م – آ |
| Class 3 | غ – ع – خ – ح – چ – ج | پ – ب |
| Class 4 | و – ژ – ز – ر – ذ – د | ذ – ز – ژ |
| Class 5 | ي – ض – ص – ش – س | ش – س – ض – ص |
| Class 6 | ك – گ | چ – خ – ح – ج – غ – ع |
| Class 7 | ه | ظ – ط – ك – گ – ل |
| Class 8 | م | ث – ت – ق – ف – ن – ي |

It is to be noted that the authors in [17], have implemented the work in [27] for the problem of 8-class in order to compare the results.

Table IV shows the results of our method compared to others when considering 8, 32 and 33 class problems.

TABLE IV

COMPARISON OF THE PROPOSED SYSTEM WITH OTHER METHODS ON IFHCDB DATASET

| Algorithms | Train size | Test size | Number of classes | Recognition rate (%) |
|---|---|---|---|---|
| Dehghan and Faez [26] | 36682 | 15338 | 8 – v1 | 81.47 |
| Alaei et al [17] | 36682 | 15338 | 8 – v1 | 98.10 |
| Rajabi et al [18] | 36000 | 13320 | 8 – v2 | 98.72 |
| Alaei et al [17] | 36682 | 15338 | 32 | 96.68 |
| Rajabi et al [18] | 36000 | 13320 | 33 | 94.82 |
| Proposed system | 36437 | 15233 | 8 – v1 | 95.63 |
| Proposed system | 36437 | 15233 | 8 – v2 | 95.56 |
| Proposed system | 36437 | 15233 | 32 | 95.87 |
| Proposed system | 36437 | 15233 | 33 | 96.04 |

As can be seen from the results, the proposed method is competitive even if we can see that assigning some character in the same group can lead to misclassification. It is to be noted that our focus was only on the Arabic portion of the IFHCDB dataset.

Table V shows some characters in IFHCDB dataset that have been correctly recognized despite their similarity.

TABLE V

SAMPLES OF SIMILAR ARABIC HANDWRITTEN CHARACTERS WHICH WERE CORRECTLY RECOGNIZED

| | | | | | |
|---|---|---|---|---|---|
| Character image |  |  |  |  |  |
| Character class | ب | ت | ث | ن | ن |
| Character image |  |  |  |  |  |
| Character class | ن | ف | ق | ك | ا |
| Character image |  |  |  |  |  |
| Character class | د | ذ | ر | ر | م |
| Character image |  |  |  |  |  |
| Character class | م | ي | ي | ل | ه |
| Character image |  |  |  |  |  |
| Character class | س | ش | ح | خ | ع |

## VI. CONCLUSION

This paper deals with the recognition of Arabic handwritten characters. The goal is to find the best way of features extraction from two perspectives (textural and structural) while taking into consideration the specificity and the nature of the Arabic script.

SDP and LBP feature extraction methods are configured and used to implement these two perspectives. Tested on IFHCDB dataset and using ANN as classifier, we have found that the percentages of pixels extracted after dividing the image into a 5x5 grid, allows to reach a recognition rate around 94%, while extracting LBP histograms after dividing the image into four regions from the centroid of the character's body and its diacritics allows to achieve a recognition rate around 96%.

Since the recognition rate in TOP2 is about 99%, keeping two results in the recognition of a character belonging to a word is beneficial in the stage of word recognition, since we will have more than one possibility.

In future work, we will investigate the collaboration between what we call the word agent (the entity responsible to recognize a word) and the character agent.

## ACKNOWLEDGMENT

The authors are thankful to S. Mozaffari for providing the dataset for the experiment. The authors are also thankful to the Maxware Technology stuff for their moral support and professional help, without forgetting faculty of science, University Ibn Tofail for providing the infrastructural facilities that helped to complete this work.

## REFERENCES

[1] P.A. Devijver, J. Kittler, Pattern Recognition: A Statistical Approach, 1982.

[2] V. Govindan, A. Shivaprasad, Character recognition—a review, Pattern Recognit. (1990).

[3] L. Heutte, T. Paquet, J. V Moreau, Y. Lecourtier, C. Olivier, A structural/statistical feature based vector for handwritten character recognition, Pattern Recognit. Lett. 19 (1998) 629–641.

[4] M. Parvez, S.A. Mahmoud, Arabic handwriting recognition using structural and syntactic pattern attributes, Pattern Recognit. 46 (2013) 141–154.

[5] J. Cai, Integration of structural and statistical information for unconstrained handwritten numeral recognition, IEEE Trans. Pattern Anal. Mach. Intell. 21 (1999) 263–270. doi:10.1109/34.754622.

[6] A. Amin, Recognition of hand-printed characters based on structural description and inductive logic programming, in: Proc. Int. Conf. Doc. Anal. Recognition, ICDAR, 2001: pp. 333–337.

[7] A. Sahlol, C. Suen, A Novel Method for the Recognition of Isolated Handwritten Arabic Characters, arXiv Prepr. arXiv1402.6650. (2014).

[8] H. Alamri, J. Sadri, Ching Y Suen, N. Nobile, C.Y. Suen, A Novel Comprehensive Database for Arabic Off-Line Handwriting Recognition, Elev. Int. Conf. Front. Handwrit. Recognit. (2008) 664–669, Montreal.

[9] G. Abandah, N. Anssari, Novel moment features extraction for recognizing handwritten Arabic letters, J. Comput. Sci. 5 (2009) 226–232. doi:10.3844/jcssp.2009.226.232.

[10] A. Asiri, M. Khorsheed, Automatic Processing of Handwritten Arabic Forms, in: Proc. World Acad. Sci. Eng. Technol., 2005: pp. 313–317.

[11] J. Shanbehzadeh, H. Pezashki, A. Sarrafzadeh, Features Extraction from Farsi Hand Written Letters, Image Vis. Comput. (2007) 35–40.

[12] R.Hamdi, F.Bouchareb, M.Bedda, Handwritten Arabic character recognition based on SVM Classifier, in: 3rd Int. Conf. Inf. Commun. Technol. From Theory to Appl., 2008: pp. 1–4.

[13] M. Torki, M.E. Hussein, A. Elsallamy, M. Fayyaz, S. Yaser, Window-Based Descriptors for Arabic Handwritten Alphabet Recognition: A Comparative Study on a Novel Dataset, arXiv Prepr. arXiv1411.3519. (2014).

[14] S. Mozaffari, H. Soltanizadeh, ICDAR 2009 handwritten Farsi/Arabic character recognition competition, in: Proc. Int. Conf. Doc. Anal. Recognition, ICDAR, 2009: pp. 1413–1417. doi:10.1109/ICDAR.2009.283.

[15] S. Mozaffari, K. Faez, F. Faradji, M. Ziaratban, S.M. A Golzan, A Comprehensive Isolated Farsi/Arabic Character Database for Handwritten OCR Research, Tenth Int. Work. Front. Handwrit. Recognit. (2006) 385–389.

[16] A. Lawgali, A. Bouridane, Handwritten Arabic Character Recognition: Which feature extraction method, Int. J. Adv. Sci. Technol. 34 (2011) 1–8.

[17] A. Alaei, P. Nagabhushan, U. Pal, A new two-stage scheme for the recognition of persian handwritten characters, in: Proc. - 12th Int. Conf. Front. Handwrit. Recognition, ICFHR 2010, 2010: pp. 130–135.

[18] M. Rajabi, N. Nematbakhsh, S. Amirhassan Monadjemi, A New Decision Tree for Recognition of Persian Handwritten Characters, Int. J. Comput. Appl. 44 (2012) 52–58. doi:10.5120/6271-8433.

[19] M. Askari, M. Asadi, A.A. Bidgoli, Isolated Persian/Arabic handwriting characters: Derivative projection profile features, implemented on GPUs, J. AI. (2016).

[20] T. Ojala, M. Pietikäinen, D. Harwood, A comparative study of texture

measures with classification based on featured distributions, Pattern Recognit. 29 (1996) 51–59. doi:10.1016/0031-3203(95)00067-4.

[21] T. Ojala, M. Pietikäinen, T. Mäenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Trans. Pattern Anal. Mach. Intell. 24.7 (2002): 971-987. doi:10.1109/TPAMI.2002.1017623

[22] T. Ahonen, A. Hadid, M. Pietikainen, Face Description with Local Binary Patterns: Application to Face Recognition, TPAMI. 28 (2006) 2037–2041.

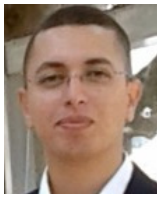
[23] Y. Hannad, I. Siddiqi, M.E.Y. El Kettani, Writer identification using texture descriptors of handwritten fragments, Expert Syst. Appl. 47 (2016) 14–22. doi:10.1016/j.eswa.2015.11.002.

[24] M. Biglari, F. Mirzaei, J. Neycharan, Persian/Arabic Handwritten Digit Recognition Using Local Binary Pattern, Int. J. Digit. (2014).

[25] S. Tulyakov, S. Jaeger, V. Govindaraju, D. Doermann, Review of classifier combination methods, Stud. Comput. Intell. 90 (2008) 361–386. doi:10.1007/978-3-540-76280-5_14.

[26] Poh, N., & Bengio, S. Database, protocols and tools for evaluating score-level fusion algorithms in biometric authentication. Pattern Recognit, (2006) 39(2), 223-233.

[27] M. Dehghan, Farsi handwritten character recognition with moment invariants, in: Digit. Signal Process. Proc., 1997: pp. 507–510.



**Youssef Boulid** received his M.S. degree in Decision Support Systems and Project Management in 2012 from University Ibn Tofail, Faculty of science, Kenitra- Morocco. Currently he is preparing a PhD 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 in Rabat - Morocco. Now he is a Professor at university Ibn Tofail in Kenitra - Morocco. His research interests include CAE, CAD 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 in Kenitra - Morocco. His research interests include Multivariate statistics and Image recognition algorithms.

# A Depth Video-based Human Detection and Activity Recognition using Multi-features and Embedded Hidden Markov Models for Health Care Monitoring Systems

Ahmad Jalal[1] , Shaharyar Kamal[2] and Daijin Kim[1]

[1]Department of Computer Science and Engineering, POSTECH, Pohang, 790-784, Korea
2Department of Electronics and Radio Engineering, KyungHee, Suwon, 446-701, Korea

***Abstract* — Increase in number of elderly people who are living independently needs especial care in the form of healthcare monitoring systems. Recent advancements in depth video technologies have made human activity recognition (HAR) realizable for elderly healthcare applications. In this paper, a depth video-based novel method for HAR is presented using robust multi-features and embedded Hidden Markov Models (HMMs) to recognize daily life activities of elderly people living alone in indoor environment such as smart homes. In the proposed HAR framework, initially, depth maps are analyzed by temporal motion identification method to segment human silhouettes from noisy background and compute depth silhouette area for each activity to track human movements in a scene. Several representative features, including invariant, multi-view differentiation and spatiotemporal body joints features were fused together to explore gradient orientation change, intensity differentiation, temporal variation and local motion of specific body parts. Then, these features are processed by the dynamics of their respective class and learned, modeled, trained and recognized with specific embedded HMM having active feature values. Furthermore, we construct a new online human activity dataset by a depth sensor to evaluate the proposed features. Our experiments on three depth datasets demonstrated that the proposed multi-features are efficient and robust over the state of the art features for human action and activity recognition.**

***Keywords* —Depth camera, Embedded Hidden Markov Models, Feature Extraction, Human Activity Recognition.**

## I. Introduction

MONITORING human activities of daily living is an essential way of describing the functional and health status of a human [1]. Therefore, human activity recognition (HAR) is one of genuine components in personalized life-care and healthcare systems, especially for the elderly and disabled [2]. To monitor daily activities of the elderly people, video-cameras can be deployed in smart environments, such as smart homes or smart hospitals to acquire time-series activity video clips. According to the world health organization survey, the population of older people is rapidly increasing all over the world and their healthcare needs become more complex which consume more resources (i.e., human and healthcare expenditures). Thus, healthcare monitoring services are needed to overcome the extensive resource utilization and improve the quality of life of elder people [3]. Indeed, several studies support that personalized life-care and healthcare services can decrease the mortality rate especially for the elderly people. For instance, in the European countries, it is estimated that the survival proportion of older people is increasing while receiving the personalized healthcare services instead of receiving institutional care or nursing homecare [4]. Thus, the aim of this study is to propose an efficient depth video-based HAR system that monitors the activities of elder people 24 hours/day and provides them an intelligent living space which comfort their life at home.

To improve the personalized healthcare services of elderly people, automated HAR systems are required to monitor the elderly daily activities and provide safe and independent life at home [5], [6]. In automated HAR system, various sensors are used to extract signal or video data, provide continuous health monitoring by observing their daily routine activities and generate an alert in case of emergency to authorized person (i.e., doctor, nurse and relatives). Many researchers in the field of automated HAR frequently use wearable sensors or vision-based sensors to extract features data. In wearable sensor HAR technology, subjects are asked to wear sensing devices (i.e., accelerometer, magnetometer and gyroscope) at different locations of human body to capture sequence of data [7], [8]. However, HAR system built by wearable sensors have certain difficulties faced by elderly people to perform daily activities such as discomfort to wear sensors to their body parts for long time, elderly people often forget to wear proper suit equipment's of wearable sensors, directional control issues causes unreliable data recording during complex subject's movements (i.e., especially in smart phones and wrist bands/watches) and there is a relatively difficulty in terms of energy consumptions and device settings. On the other hand, video sensors have certain issues such as privacy, fixed devices at predetermined positions and pre-processing complexity. From the above literature, we have observed that video sensors have mostly solvable issues, richer information and have wider scope, therefore, we motivated to use vision-based HAR approach to recognize activities of elderly people at home.

Vision-based HAR system is a challenging research topic in the field of computer vision and pattern recognition. It enables the development of various practical applications such as security systems, elderly healthcare systems, video surveillance and smart homes systems [9], [10] which provide personal security, cost-effectiveness, friendly services and efficient health care for elderly people. In recent years, most of the researches of vision-based HAR are focused on daily activity monitoring due to the fact that less medical aid, not due attention by their relatives and loss of independence causes lives risks and injuries in elderly people. This paper annotated a novel set of continuous online daily routine activities consisting of sitting down, eating, falling-down, exercising, and taking-medicine and reading an article. To get realistic and natural scenes faced by elderly people in their life, activities are defined and selected after visiting healthcare medical centers, dealing with doctors/nurses and reviewing medical research papers [11]-[13].



DOI: 10.9781/ijimai.2017.447

## II. Related Work

In this section, we will review related works from two different aspects including elderly healthcare applications and automatic activity recognition over multiple datasets.

### A. Elderly Healthcare Applications

Elderly healthcare monitoring has a direct link with HAR systems, where the sensor devices capture and examine the indoor behaviors and activities of elderly at hospitals, home or offices. In smart hospitals [14], automatically estimating hospital-staff-patients activities are examined and empowered HAR to boost our vision for the hospital as a smart environment. In smart homes [15], home activities of elderly people are recognized based on invariant features characteristics. While, in smart offices [16], authors proposed comfort management system along with activity recognition solutions that handles multiple-user and rapidly recognize office activities.

### B. Automatic Activity Recognition

Many existing studies have applied HAR utilizing video sensors technologies (i.e., RGB cameras) for human detection, tracking and activity recognition. In [17], a new model is proposed for activity recognition that combines a powerful mid-level representation, in the form of HoG and BoW poselets, with discriminative key frame selection based on conventional videos. In [18], an epitomic representation for modeling is introduced where the video activity sequence is divided into segments to extract moving objects and short-time motion trajectories. This information is further processed by Iwasawa matrix decomposition to represent the effect of rotation, scaling and projective action on the state vector and used for activity recognition. In [19], a view-specific approach is proposed for representation of movements as temporal templates. These templates indicate the presence of motion in binary values and the function of the motion in a sequence. Then, a matching algorithm is used to construct a recognition system. However, these cameras have certain limitations such as technical infeasibility to differentiate between near and far parts of human body, limited information (binary or RGB intensity values), highly sensitive with lighting conditions and unreliable for postures having self-occlusions.

To improve HAR capabilities and human silhouettes representation, depth sensors [20]-[22] have been released to facilitate the human detection, feature extraction and activity recognition tasks. Compared with the digital RGB cameras, depth cameras provide additional human body parts information, insensitivity to light changes and easily normalized during body orientation/size changes. In addition [21]-[26], several research articles have used depth maps information to explore their features extraction from two basic types such as depth silhouettes features and skeleton-based features for recognizing human activities using depth sequences. In the depth silhouettes features, many researches used a set of depth pixels of the depth images or human shape silhouettes to extract the features. In [20], action graph is used to model explicitly the dynamics of human motion and a bag of 3D points to characterize a set of salient postures that corresponds to the nodes in the action graph for recognizing actions/activities. In [21], a new descriptor is proposed for activity recognition using a histogram capturing the distribution of the surface normal orientation in the 4D space of time, depth, and spatial coordinates. To build the histogram, they created 4D projectors, which quantize the 4D space and represent the possible directions for the 4D normal. In [22], semi-local features called random occupancy pattern (ROP) features are proposed which employed a novel sampling scheme and extracted from randomly sampled 4D subvolumes with different sizes and locations using depth images.

Instead of relying on depth silhouettes features, many researchers have explored features based on skeleton joints information. In [23], a set of features such as body pose, hand position, motion information and point-cloud features are proposed having three dimensional Euclidean coordinates and the orientation matrix of each joint to recognition activities using RGBD images. In [24], an effective method is proposed that consists of a new type of features based on position differences of 3D joints and Eigenjoints. The Eigenjoints are able to capture the properties of posture, motion and offset of each frame. Then, they used frame descriptors of Eigenjoints without quantization and classify different actions based on Naïve Bayes Nearest Neighbor (NBNN). In [25], an actionlet ensemble model is developed to represent each action, capture the intra-class variance and recognize various actions and activities using benchmark depth datasets. Although, depth video-based HAR systems are quite feasible for recognizing activity, however, it is still difficult using just depth silhouettes features or joint point's information especially during self-occlusion. Therefore, our research work is focused on utilizing depth data based on merging both silhouettes and joint points information for feature representation, activity training and recognition.

Our main contribution of this paper is to propose a real-time body parts tracking method that has the ability to track the self-occluded human body parts especially in case of torso rotation. Also, the proposed method detects and controls the fast moving human body parts and has invariant characteristics with respect to body size and human position which strengthen our contributions in HAR. Combination of depth silhouettes and joint information features cover various factors such as robust to noises, missing joints and capture the local dependencies over the embedded HMM acting as a novel methodology in order to enhance the recognition rate over all three depth datasets. In addition, we provide a new online depth human activity dataset, which becomes a benchmark in HAR.

In this work, a novel approach is proposed for HAR by considering multi-features approach that is extracted in 3D coordinates along with time space (i.e., 3D human silhouettes and spatiotemporal joints values) and embedded HMMs. These features deal with intensity differentiation profiles, directional angular values, local motion of active body parts and temporal frame movements which provide compact and sufficient information for human action and activity recognition. All these features are concatenated together and converted into discrete symbols by considering vector quantization algorithm. Meanwhile, active regions of depth silhouettes (i.e., moving body regions, arms, legs and interest body joints) make specific classes for training and recognition using embedded HMMs. In order to determine the recognition performance, we build a new online depth activity dataset that contains segmented video sequences for training phase and unsegmented video sequences for testing phase, which will be a benchmark for activity recognition based on depth data. In addition, we evaluate our system according to the standard experimental protocols definition on three challenging depth datasets. Our results outperform all published state of the art feature extraction methods.

The rest of the paper is organized as follows. In Section III, we describe the system architecture including problem statement, dataset generation and proposed HAR system model. Section IV presents the detail description of HAR. Section V describes the experimental results and comparisons using proposed and state of the art methods. Finally, Section VI presents the conclusion of the paper.

## III. System Architecture and Methodology

### A. Problem Statement

Due to recognition of natural scenes of continuous human activities without any instructions to subjects, video based HAR systems faces

various problems such as dynamic backgrounds changes in a scene along with self-occlusion, object hurdles or body parts rotation (i.e., especially torso) of human subjects and body orientation or sizes changes frequently due to different subjects performing activities at different distances from cameras. Also, similar postures of different activities (i.e., falling-down and taking an object, take-medicine and eating and clapping and exercise) causes reduction in recognition performance and processing time consumption during testing of activities especially online datasets.

We proposed an online depth HAR system that utilized person-tracking system, multi-features and embedded HMMs algorithms to solve the above mentioned problems. For the first problem, we developed our real-time body tracking system to control self-occlusion and provide freely human movements in a scene. For the second problem, we normalized skeleton models and applied invariant features to manage body size changes. While, third problem is solved by considering multi-features having depth silhouettes and joints information to identify difference in between different activities having similar postures. For the fourth problem, we introduced embedded HMMs concept to overcome the redundant data usage during testing time, improve the computational processing and increase recognition performance.

### B. Dataset Generation

During daily routine activities, elderly people are mainly involved in a mixture of static sequences (i.e., minor movements of body parts) and dynamic sequences (i.e., major movements of different body parts) of activities. Therefore, our dataset provided both types of activity sequences having natural routine behaviors and continuous recognition of elderly people such as eating, taking-medicine, sitting down, exercising, falling-down and cleaning. Due to monitoring of continuous elderly activities, detecting starting and ending times of all occurring activities are controlled by a sliding window approach. In Fig. 1 we annotated an online continuous depth dataset by considering the daily life activities of elderly people at home or offices.

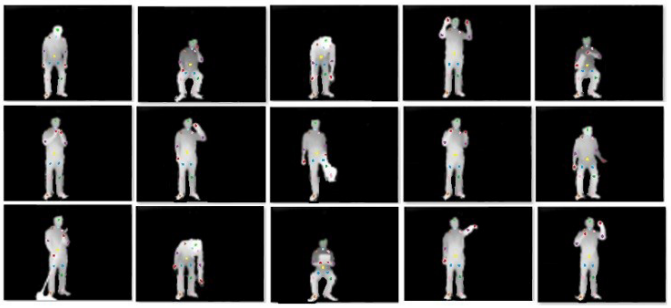


Fig. 1. Samples of human depth silhouettes along with joint point's location in our depth annotated dataset. Top row: sitting down, taking-medicine, falling-down, both hands waving, eating; Middle row: clapping, phone conversation, walking, exercising, stand up; Bottom row: cleaning, taking an object, reading an article, pointing as object and hand waving.

### C. Proposed HAR System Model

The proposed framework of HAR system consists of the following processes namely as, (1) depth imaging acquisition, (2) human silhouettes segmentation and tracking, (3) feature representation and extraction based on multi-features, (4) clustering algorithm and vector quantization, and (5) human activity modeling, training and recognition. Fig. 2 shows the overall architecture of our proposed activity recognition system.

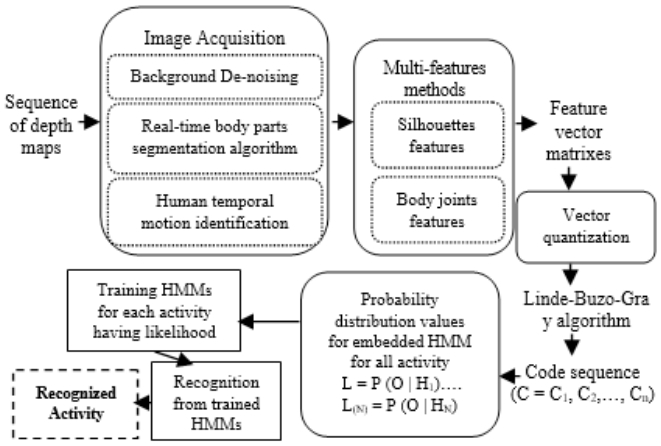



Fig. 2. System architecture of the proposed activity recognition system.

## IV. Image Acquisition, Feature Extraction and Activity Recognition

In this section, we describe depth imaging acquisition, feature extraction via multi-features approach, symbol representation via Linde-Buzo-Gray (LBG) clustering algorithm and activity training/recognition using embedded HMMs.

### A. Depth Imaging Acquisition

To capture 3D information, we utilized a RGB-D camera (i.e., Kinect) to acquire a pair of RGB images and depth maps. These depth maps contain a considerable amount of noise data. However, noise reduction is an important process before extracting multi-features. Therefore, to remove the noisy background areas from the depth map, we applied pixel differentiation method as

$$P_d^t(x,y,z) = b_t(x,y,z) - d_t(x,y,z) > T_{values} \quad (1)$$

Where $b_t(x, y, z)$ and $d_t(x, y, z)$ are the background and depth intensity pixel values at time $t$ and $T_{values}$ is a positive threshold value. Meanwhile, to apply floor removal mechanism, we simply ignore ground line (i.e., $y$ parameters) which acts as lowest value (i.e., equal to zero) corresponding to a given pair of $x$ and $z$ axis. However, to extract accurate human silhouette region [26] from the scene, we calculate depth intensity center values from the scenes using connected component labeling technique (see Fig. 3). In component labeling technique, the variation of pixel intensity in an image is observed using raster scanning where every depth pixel $d_n$ of the connected component has depth value, intensity values of two neighboring pixels are within a threshold $\delta_n$ and each object/subject has its depth center $d_c$ values is assigned as $| d_c - d_n | \leq \delta_t$. Due to depth center values, we monitored the pixel-neighboring intensity variation in between the consecutive frames which remove the unnecessary objects (i.e., cupboard, doors or chairs) from the scenes.

Lastly, we applied human movement detection $D(f)$ by considering temporal continuity constraints between consecutive frames.

$$D(f) = \sqrt{\left(f_{t-1}^x - f_t^x\right)^2 + \left(f_{t-1}^y - f_t^y\right)^2 + \left(f_{t-1}^z - f_t^z\right)^2} \quad (2)$$

As a result, human silhouettes regions are enclosed within the rectangular bounding box having specific parametric values (i.e., height and width) based on motion detection. Fig. 3 describes the overall scenario of real time body tracking system including (a) depth maps having noisy information, (b) temporal human motion using ridge data, (c) human identification and (d) depth human silhouettes.

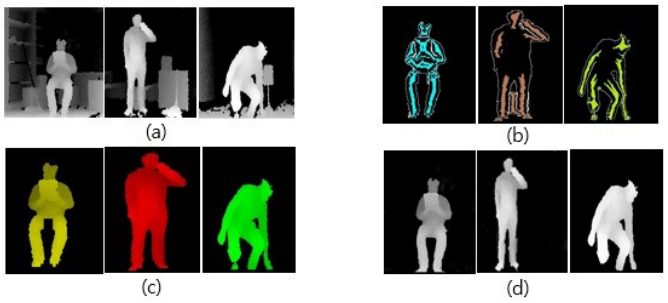



Fig. 3. Overall description of human silhouettes identification as (a) depth maps having noisy information, (b) temporal human motion using ridge data, (c) labeled human body silhouette and (d) depth human silhouettes identification of reading an article, phone conversation and falling-down activities.

### B. Feature Extraction via Multi-Features Approach

In this section, we extract features from human body silhouettes and joints information (i.e., include 15 joint points) via depth images. However, a set of feature extraction techniques provide a compact representation of image content by describing invariant characteristics of local body parts, multi-view differentiation and spatiotemporal body joints motion which are derived to merge together having spatial and temporal depth silhouettes characteristics.

#### 1) Invariant Features

To process the human depth silhouettes, we compute the total pixel intensities along lines of different locations (i.e., 0 to 180 degrees) to identify specific view directions and the center points of human silhouettes are selected as the reference point, which is defined as

$$R_D(\rho,\theta)\int_{\infty}^{-\infty}\int_{\infty}^{-\infty} f(x,y)\,\delta(\rho - x\cos\theta - y\sin\theta)\,dxdy \tag{3}$$

$R_D(\rho,\theta)$ is the line integral of the 2D radon function along a line between positive to negative infinite value. These 2D information are passed through sum of the squared Radon transform values [27] to create a 1D profiles as shown in Fig. 4. These 1D profiles are strong candidates to provide translation and scaling invariant features. Finally, a feature vector with 180 dimensions is extracted, instead of the 2D shape matrix.

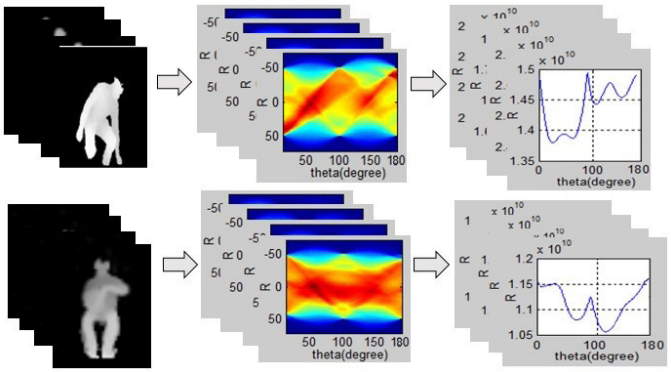



Fig. 4. Invariant features performed over various depth activities.

#### 2) Multi-view Differentiation Features (MDF)

Due to similar postures, object occlusions and missing body parts from the frontal view, extra views (i.e., side and top) are used for the feature vectors to improve the accuracy of the classifier. Therefore, we applied Cartesian planes over the human depth silhouettes to get the 2D images of side and top views. These images are passed through a frame differential mechanism where current frame is compared with the next frame to get pixel intensity information for feature processing. Fig. 5 explains the multi-view differentiation features having (a) side and (b) top views of forward kick and bend actions using MSR Action3D dataset.

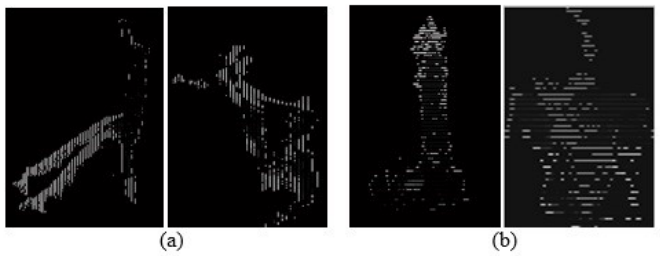



Fig. 5. Multi-view differentiation features based on (a) side and (b) top views in case of forward kick and bend actions using MSR Action3D dataset.

#### 3) Spatiotemporal Body Parts Motion Features (BMF)

To consider the discriminative information for determining how a person has moved (i.e., spatiotemporally) during the activity sequence, we considered the shape information having specific motion region of body parts in an activity. Therefore, we calculated the gradient orientation, pixel intensity in between initial frame till final frame and Mahalanobis distance for matching the input activity from the stored templates is defined as

$$D_{seq} = \sum_{j=1}^{N} \left| I_j^t(x,y,z) - I_j^{t-1}(x,y,z) \right| > \tau \tag{4}$$

where $I_j$ is image sequence along with $t$ and $t$-$1$ to evaluate the temporal sequential human motion of overall activity images. However, those regions having maximally confident patches of temporal values (i.e., greater than specific threshold values) are tracked based on optical flow mechanism which are enclosed by rectangular box as shown in Fig. 6.

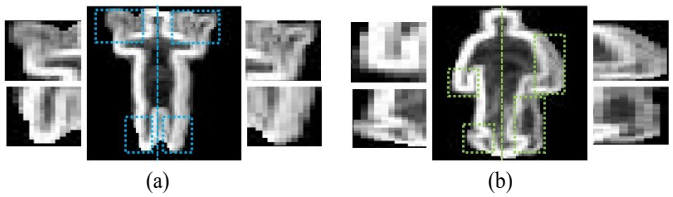



Fig. 6. Spatiotemporal body parts motion features based on maximally confident patches in case of (a) both hands waving and (b) walking activities using our depth annotated dataset.

However, both MDF and BMF features spaces produce a higher number of features dimension, thus, principal component analysis (PCA) is used here to extract global information [28] from all activities data and approximate the higher features dimension data into lower dimensional features. In this work, 200 and 150 principal components (PCs) are used for MDF and BMF features to process the activity data and are expressed $PC = m_i e_{top}$ where $PC$ is the PCA projection of feature vectors, $m_i$ is the zero mean vector and $e_{top}$ is the top eigenvectors indicating higher variance among overall eigenvalues.

#### 4) Temporal Joints Difference Features (TDF)

In addition to depth human silhouettes, the multi-features also provided skeleton motion joints features. In order to make use of the additional motion information from joints information, we applied current frame differentiation method to calculate joint points difference

between the current frame $f_i$ and all the respective frames $f_d$ of activity sequence can be represented as

$$f_{current}^{diff} = \left\{ f_i^j - f_r^j \mid i = 1; r = 2, \ldots, N \right\} \quad (5)$$

where 3D skeleton joints $j$ having all three coordinates axis at frame $t$ to $t+1,\ldots,t+N$ and the size of feature vector become 15x1, respectively.

### 5) Pairwise Joints Distance Features (PJF)

To consider the pairwise joint distance feature, we measure the joints distance between the active $a$ body joints with the inner $i$ body joints at each frame $t$. Here, the active body joints consist of head, shoulders, hands and feet, while, inner body joints include torso, neck, elbows, hips and knees. Thus, pairwise joints distance features $D_{pjf}$ is represented as

$$D_{pjf}(t) = \sqrt{\left(j_a^x - j_i^x\right)^2 + \left(j_a^y - j_i^y\right)^2 + \left(j_a^z - j_i^z\right)^2} \quad (6)$$

where $D_{pjf}$ becomes a vector of 54 dimensions. Fig. 7 shows 2D plots of TDF and PJF feature values using falling-down activity.

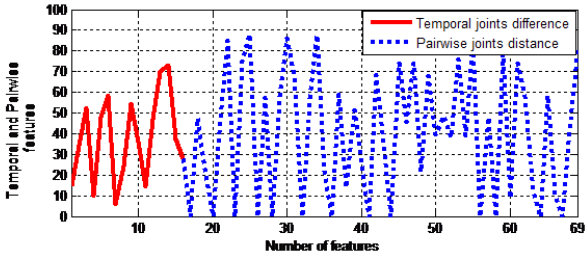



Fig. 7. Temporal joints difference and pairwise joints distance features are applied over falling-down activity.

### 6) Spatiotemporal Joints Angular Features

Moreover, the human body parts variation in terms of postures, direction, height and angular dimensions have a huge impact on the performance of the activity recognition. Therefore, we identify the gradients representation with respect to angles of each body joints in-between $t$ and $t-1$ consecutive frames of each sequence can be expressed as

$$\theta_{tan} = \arctan\left(C_{(1)}^t - C_{(1)}^{t-1} / C_{(2)}^t - C_{(2)}^{t-1}\right) \quad (7)$$

$$\varphi_{cos} = arccos\left(C_k - C_k^{'} / \left|C_k\right|\left|C_k^{'}\right|\right), k = x, y, z \quad (8)$$

where $C_{(1)}$ and $C_{(2)}$ are the pair-coordinates of all three respective axis. Both equations are examined for angular and sinusoidal features characteristics. As a result, the size of the feature vector of joints angular features representation of each activity frame become 45 x 2, respectively (See Fig 8.)

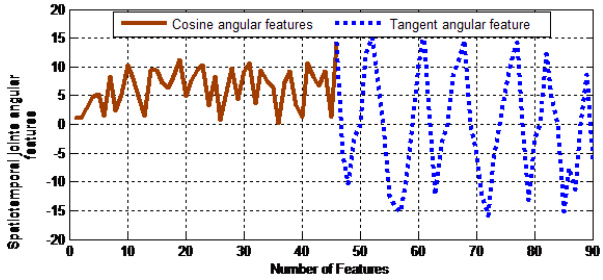



Fig. 8. 2D plot of spatiotemporal joints angular features using falling-down activity from our depth annotated dataset.

### C. Symbol Representation and Code Matrix Selection

All these sub-features are merged together to make a multi-feature vector size of 689x1. These multi-features are symbolized from the activity frames and generated from Linde-Buzo-Gray (LBG) clustering algorithm. Here, LBG used a splitting mechanism where the centroid for the training activity sequence is calculated and split into two nearest vectors. Each partition is restricted into specific centroid value. Therefore, each vector is then split into two vectors and the iteration is repeated until N-level centroid values are obtained. Finally, for each cluster, samples are assigned to the same class (i.e., activity label) that is the one of the closest cluster centroid. Fig. 9 shows the internal concept of feature dimensional structure and code matrix selection of proposed features.

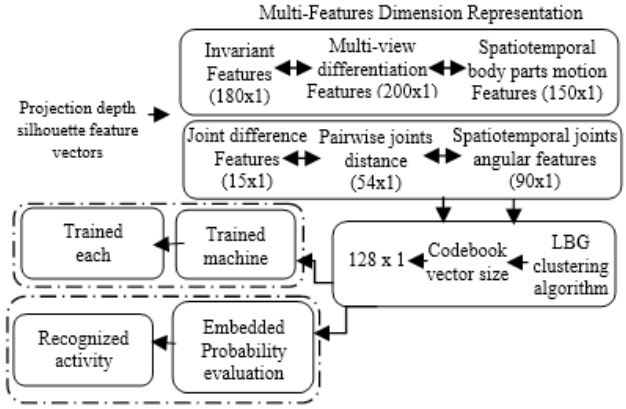



Fig. 9. Symbol representation and code matrix selection.

### D. Embedded Hidden Markov Model

To model, train and recognize different activities using depth data, we introduced a new concept of embedded HMM method. Therefore, embedded HMM is introduced which focused specifically at active feature regions of human body joints such as hands, head, feet and shoulders.

Also, it includes the overall human silhouettes information or all joints information which contain redundant information such as static body regions (i.e., torso, chest and forearms) and inactive body joints (i.e., elbows, neck and hips). These kinds of unnecessary information causes reduction during performance of recognition accuracy results.

Also, full-body silhouettes contain specific or active feature regions (i.e., moving body parts areas) which are augmented together to build a single HMM having $O_a$ observation probabilities of $M$ active feature regions of each activity as $a$.

$$A_l = \sum_{l=1,2,\ldots,N,a=1}^{M} P\left(O_a | h_{la}\right) \quad (9)$$

where $A_l$ indicates the likelihood of $l$ HMM with respect to $N$ number of activities. Fig. 10 shows active feature regions of overall human silhouettes to calculate specific likelihood of each activity. Finally, recognized activity $R$ is chosen as desired activity having maximum likelihood values [29] among all activities during testing.

$$R_{act} = argmax\left\{A_l / l\right\} \quad (10)$$

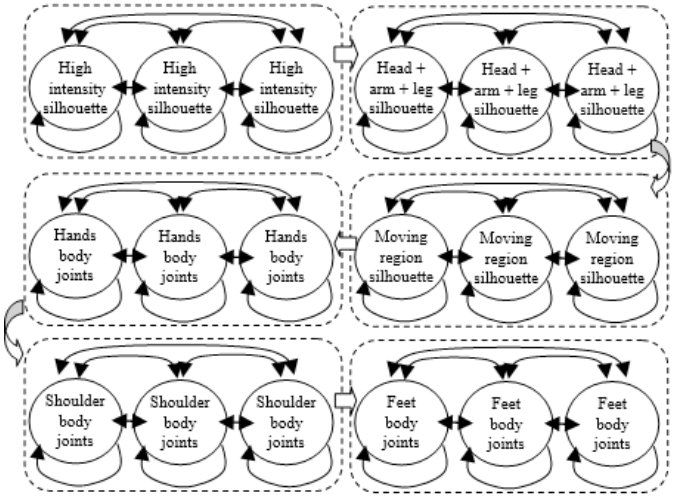



Fig. 10. Embedded HMM structure for all activities or actions using specific full-body silhouettes and joints information.

## V. Experimental Results and Discussion

In this section, we divide our experimental results into two different research aspects such as 1) elderly healthcare applications and 2) automatic activity recognition datasets which evaluate the performance of proposed and the state of the art methods.

### A. Elderly Healthcare Applications

In this subsection, we evaluate our method by considering three benchmark datasets [15] performed by elderly people at multiple environments (i.e., hospital, home and office). Table I compares the recognition accuracy of proposed and state of the art methods based on same settings as [15].

TABLE I

Recognition Accuracy Comparison between Proposed and State of the Art Methods using Three Healthcare Dataset

| Healthcare Applications | Invariant features [30] | Motion features [15] | Proposed method |
|---|---|---|---|
| Smart hospital activities | 86.09 | 90.33 | 94.82 |
| Smart home activities | 88.68 | 92.33 | 95.15 |
| Smart office activities | 89.43 | 93.58 | 95.97 |

Similarly, we compare our detection activities along with active frames (i.e., frames that contain given activity types) having selected human silhouettes and obtained precision, recall and F-measure. Table II compares the performance of proposed and state of the art methods. It is quite obvious that our method is effective to encode relevant activity information.

TABLE II

Precision, Recall and F-measure Comparisons on the Smart Home Dataset for Activity Localization

| Methods | Precision | Recall | F-measure |
|---|---|---|---|
| Invariant features [30] | 78.3 | 80.6 | 78.9 |
| Motion features [15] | 80.9 | 82.4 | 81.3 |
| Proposed method | 84.7 | 87.3 | 85.8 |

### B. Experiments on Automatic Activity Recognition Datasets

In this subsection, we conduct experiments on three challenging depth-based activity and action datasets such as online self-annotated dataset [31], MSRDailyActivity3D and MSRAction3D for recognition purpose using multi-features and embedded HMM. However, the content of each dataset, experimental setting and results are described in the following sub-subsections.

#### 1) Online Self-Annotated Dataset

In this experiment, the depth-video activity dataset [32] is collected for the fifteen different activities based on daily life healthcare monitoring scenarios often encountered by elderly people. These activities include: *sitting down, taking-medicine, falling-down, both hands waving, eating, clapping, phone conversation, walking, exercising, stand up, cleaning, taking an object, reading an article, pointing an object and hand waving,* respectively. All activities are captured in labs and halls. The total dataset consists of 705 video sequences performed by sixteen different subjects.

Its training datasets include 675 segmented video sequences and its testing phase contains 30 unsegmented video sequences having time duration of two to four minutes. From Table III, we illustrate that the proposed method achieved recognition rate of 71.6% for all fifteen activities.

TABLE III

Comparison of Recognition Accuracy between Proposed and State of the Art Methods using Online Self-annotated Dataset

| Methods | Accuracy (%) |
|---|---|
| Dynamic temporal warping [33] | 38.7 |
| Multi-part bag-of-poses [34] | 47.6 |
| HOJ3D [35] | 49.6 |
| Multimodal approach [36] | 51.6 |
| Depth silhouettes context features [32] | 57.6 |
| Multi-Features method | 71.6 |

While, Table IV summarizes the comparison of our method with the state of the art methods by considering a decision of recognition results based on 100 frame sliding window approach.

TABLE IV

Mean Recognition Accuracy of Proposed Multi-features Method using Online Self-annotated Dataset

| Activities | Accuracy (%) | Activities | Accuracy (%) |
|---|---|---|---|
| Sitting down | 38.7 | Hand waving | 67.7 |
| Falling-down | 47.6 | Taking-medicine | 78.1 |
| Eating | 51.6 | Both hands waving | 58.2 |
| Phone conversation | 89.9 | Clapping | 78.5 |
| Exercising | 54.8 | Walking | 81.3 |
| Cleaning | 73.4 | Stand up | 63.7 |
| Reading an article | 58.2 | Take an Object | 69.6 |
| | | Pointing an Object | 84.4 |

In addition, Table V shows the performance evaluation of online self-annotated activity recognition dataset from all three indicators such as precision, recall and F-measure. It is clearly justified that the proposed features are significantly better than conventional ones.

TABLE V

Performance Evaluation using Precision, Recall and F-measure Parameters for Online Self-annotated Dataset

| Methods | Precision | Recall | F-measure |
|---|---|---|---|
| [30] | 53.7 | 56.2 | 55.8 |
| [15] | 58.3 | 61.8 | 60.5 |
| [36] | 61.6 | 62.1 | 62.8 |
| [32] | 63.9 | 65.3 | 64.2 |
| Proposed method | 68.4 | 67.9 | 67.3 |

### 2) MSRDailyActivity3D Dataset

The MSRDailyActivity3D dataset [25] was collected with human daily activities captured by the Kinect sensor. This dataset consists of sixteen activities including: drink, eat, read book, call cellphone, write, use laptop, vacuum cleaner, cheer up, sit still, toss paper, play game, lie down, walk, play guitar, stand up and sit down. All subjects perform activities into both standing and sitting on sofa poses. The number of activity video sequences is 320. While, dataset is quite challenging due to human object interactions. Some sample images of MSRDailyActivity3D dataset are shown in Fig. 11.

Fig. 11. Sample depth images used in MSRDailyActivity3D dataset.

However, this dataset includes one sample per subject, therefore, we applied leave-one-subject-out (LOSO) cross validation process in our experiment. Table VI presents the recognition accuracy of our proposed multi-features method.

TABLE VI
Recognition Accuracy Comparison between Proposed and State of the Art Methods using MSRDailyActivity3D Dataset

| Methods | Accuracy (%) |
|---|---|
| Eigenjoints [24] | 58.1 |
| Joint position features [25] | 68.0 |
| Graph based genetic programming [37] | 72.1 |
| Moving Pose[38] | 73.8 |
| Integrating Joints features [39] | 76.0 |
| Motion features [40] | 79.1 |
| Actionlet ensemble[25] | 85.7 |
| Super normal vector [41] | 86.2 |
| Depth Cuboid Similarity features[42] | 88.2 |
| Multi-Features method | 92.2 |

Also, we compare the recognition performance using MSRDailyActivity3D dataset where the proposed method achieved a superior mean recognition rate of 92.2% over the state of the methods [24], [25], [37]-[42] as shown in Table VII.

TABLE VII
Recognition Performance Results of Proposed Multi-features Method using MSRDailyActivity3D Dataset

| Activities | Accuracy (%) | Activities | Accuracy (%) |
|---|---|---|---|
| Drink | 89.6 | Eat | 96.2 |
| Read Book | 93.4 | Call cell phone | 97.7 |
| Write | 87.5 | Use laptop | 89.6 |
| Vacuum Cleaner | 98.8 | Cheer up | 96.4 |
| Sit still | 87.3 | Toss paper | 88.1 |
| Play game | 89.3 | Lie down | 98.3 |
| Walk | 95.7 | Play guitar | 87.3 |
| Stand up | 90.9 | Sit down | 89.4 |

### 3) MSRAction3D Dataset

The MSRAction3D dataset [20] was captured with a depth sensor (i.e., Kinect device) by the Microsoft Researcher team. It includes 20 different action types as: *high arm wave, horizontal arm wave, hammer, hand catch, forward punch, high throw, draw x, draw tick, draw circle, hand clap, two hand wave, side boxing, bend, forward kick, side kick, jogging, tennis swing, tennis serve, golf swing and pick up & throw.* The dataset consists of 567 depth map sequences performed by 10 subjects. Also, the background of this dataset is clean and the human silhouettes are available in each frame. This dataset is quite challenging due to similar postures of different actions especially hands and legs movements. Several samples of MSRAction3D dataset are shown in Fig. 12.

Fig. 12. Sample depth images of MSRAction3D dataset.

Also, we follow the same experimental setting as [25] and obtained the recognition accuracy of 93.1% as shown in Table VIII.

TABLE VIII
Recognition Accuracy Comparison using MSRAction3D Dataset

| Methods | Accuracy (%) |
|---|---|
| Dynamic temporal warping [33] | 54.0 |
| Bag of 3D points [20] | 74.7 |
| HOJ3D [35] | 79.0 |
| Motion and Shape features [43] | 82.1 |
| Eigenjoints [24] | 82.3 |
| Semi Supervised learning [44] | 83.5 |
| Grassmannian manifold [45] | 86.2 |
| HON4D [21] | 88.3 |
| Pose Set [46] | 90.0 |
| HOD Descriptor [47] | 91.2 |
| Euclidean group algorithm [48] | 92.4 |
| Multi-Features method | 93.1 |

In addition, we compare our method with the state of the art methods [31], [20], [35], [24], [21], [43]-[48] on the cross subject test setting and obtained a significantly improved recognition performance over existing works as shown in Table IX.

TABLE IX
Mean Recognition Rate of Proposed Multi-features Method using MSRAction3D Dataset

| Activities | Accuracy (%) | Activities | Accuracy (%) |
|---|---|---|---|
| High arm wave | 89.5 | Horizontal arm wave | 90.9 |
| Hammer | 98.6 | Hand catch | 89.8 |
| Forward punch | 96.4 | High throw | 93.6 |
| Draw x | 94.1 | Draw tick | 95.5 |
| Draw circle | 98.8 | Hand clap | 87.7 |
| Two hand wave | 97.2 | Side boxing | 98.8 |
| Bend | 98.7 | Forward kick | 83.9 |
| Side kick | 94.1 | Jogging | 88.7 |
| Tennis swing | 88.5 | Tennis serve | 86.8 |
| Golf swing | 93.6 | Pick up and throw | 97.4 |

To evaluate the recognition performance based on various codebook sizes, Fig. 13 shows the recognition accuracies of all three depth datasets having different codebook sizes. We determine the codebook size as 128 experimentally using LBG clustering algorithm because the greater codebook size makes minor changes in HAR accuracy.

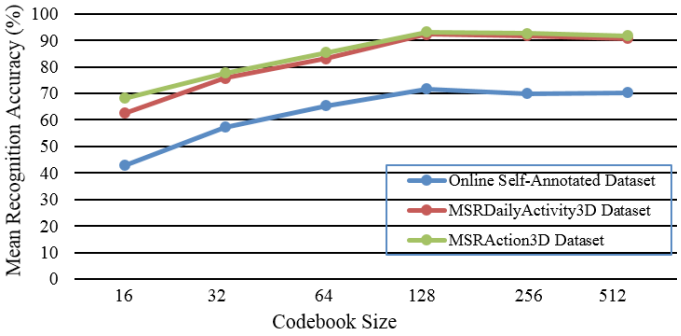



Fig. 13. Recognition accuracies versus different codebook sizes of all three depth datasets.

However, three-state HMM model is selected for the training/testing of all three depth activity/actions datasets after experimenting with different number of states HMMs models as shown in Fig. 14.

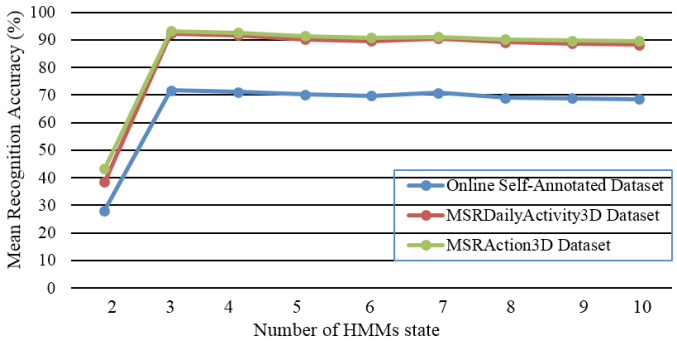



Fig. 14. Recognition accuracies versus different number of states for HMMs using all three depth datasets.

## VI. Conclusion

In this paper, a novel approach has been proposed for robust HAR system utilizing multi-features along with embedded HMMs from depth video sensor. The HAR framework contains novel characteristics as (1) a novel real-time body parts tracking system is introduced to extract human silhouettes from noisy background, (2) a robust spatio-temporal multi-features obtained from the full-body human depth silhouettes and joints body parts information, (3) development of new online depth dataset which become a benchmark for video-based HAR systems, and (4) a new concept of embedded HMMs. Our experimental results on three challenging depth datasets have shown the significant recognition performance of our features over the state of the art features extraction techniques. The proposed system is directly applicable to any e-health monitoring systems, such as monitoring healthcare problems for elderly and sick people, or examines the indoor activities of people at home or hospitals.

In the future work, we will exploit the effectiveness of our features by merging the RGB features along with multi-view invariant characteristics over more complex activities datasets including human-to-human interactions and human-object interactions. Also, some discriminative/generative models are used to robust training/recognition phase to strengthen our HAR algorithm.

## References


[1] P. Petersen, D. Kandelman, S. Arpin, and H. Ogawa, "Global oral health of older people-Call for public health action," *Community dental health*, vol. 27, no. 4, pp. 257–268, Dec. 2010.

[2] V. Osmani, S. Balasubramaniam, and D. Botvich, "Human activity recognition in pervasive health-care: Supporting efficient remote collaboration," *Journal of Network and Computer Applications*, vol. 31, no.4, pp. 628–655, Nov. 2008.

[3] J. Dunn, M. Rudberg, S. Furner, and C. Cassel, "Mortality, disability, and falls in older persons: the role of underlying disease and disability," *American Journal of Public Health*, vol. 82, no. 3, pp. 395–400, Mar. 1992.

[4] U. Reinhardt, "Does the aging of the population really drive the demand for health care?," *Health Affairs*, vol. 22, no. 6, pp. 27–39, Nov. 2003.

[5] E. Mynatt and W. Rogers, "Developing technology to support the functional independence of older adults," *Ageing International,* vol. 27, no. 1, pp. 24–41, Dec. 2001.

[6] A. Jalal, Y. Kim, Y. Kim, S. Kamal, and D. Kim, "Robust human activity recognition from depth video using spatiotemporal multi-fused features," *Pattern recognition*, vol. 61, pp. 295–308, Jan. 2017.

[7] W. Xu, M. Zhang, A. Sawchuk, and M. Sarrafzadeh, Co-recognition of human activity and sensor location via compressed sensing in wearable body sensor networks, in *Proc. 2012 Ninth international Conf. Wearable and Implantable Body Sensor Networks*, London, 2012, pp. 124–129.

[8] M. Kreil, B. Sick, and P. Lukowicz, "Dealing with human variability in motion based, wearable activity recognition," in *Proc. IEEE international Conf. Pervasive Computing and Communications Workshops*, Budapest, 2014, pp. 36–40.

[9] S. Kamal and A. Jalal, "A hybrid feature extraction approach for human detection, tracking and activity recognition using depth sensors," *Arabian Journal for Science and Engineering*, vol.41, no. 3, pp. 1043–1051, Mar. 2016.

[10] S. Kamal, A. Jalal and D. Kim, "Depth Images-based Human Detection, Tracking and Activity Recognition Using Spatiotemporal Features and Modified HMM," *Journal of electrical engineering and technology*, vol.11, no. 6, pp. 1857–1862, Nov. 2016.

[11] D. Chen, A. Bharucha, and H. Wactlar, "Intelligent video monitoring to improve safety of older persons," in *Proc. IEEE Conf. Engineering in Medicine and Biology Society*, Lyon, 2007, pp. 3814–3817.

[12] A. Jalal, N. Sharif, J. Kim, and T. Kim, "Human activity recognition via recognized body parts of human depth silhouettes for residents monitoring services at smart homes," *Indoor and Built Environment*, vol.22, pp. 271–279, Aug. 2013.

[13] A. Jalal, S. Kamal, and D. Kim, "Depth Map-based Human Activity Tracking and Recognition Using Body Joints Features and Self-Organized Map," in *Proc. IEEE Conf. on computing, communication and networking technologies*, China, 2014, pp. 1–6.

[14] D. Sanchez, M. Tentori, and J. Favela, "Activity recognition for the smart hospital," *IEEE Intelligent systems*, vol. 23, no.2, pp. 50–57, Mar. 2008.

[15] A. Farooq, A. Jalal, and S. Kamal, "Dense RGB-D map-based human tracking and activity recognition using skin joints features and self-organizing map," *KSII Transactions on internet and information systems*, vol. 9(5), pp. 1856-1869, 2015.

[16] A. Jalal, S. Kamal, and D. Kim, "A depth video sensor-based life-logging human activity recognition system for elderly care in smart indoor environments," *Sensors*, vol. 14(7), pp. 11735-11759, 2014.

[17] M. Raptis and L. Sigal, "Poselet key-framing: A model for human activity recognition," in *Proc. IEEE Conf. Computer vision and pattern recognition*, Oregon, 2013, pp. 2650–2657.

[18] N. Cuntoon and R. Chellappa, "Epitomic representation of human activities," in *Proc. IEEE Conf. Computer vision and pattern recognition*, Minneapolis, 2007, pp. 1–8.

[19] A. Bobick and J. Davis, "The recognition of human movement using temporal templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 23, no. 3, pp. 257–267, Mar. 2001.

[20] W. Li, Z. Zhang, and Z. Liu, "Action recognition based on a bag of 3d points," in *Proc. IEEE Conf. Computer vision and pattern recognition workshops*, San Francisco, 2010, pp. 9–14.

[21] O. Oreifej and Z. Liu, "Hon4d: Histogram of oriented 4d normal for activity recognition from depth sequences," in Proc. *IEEE Conf. Computer vision and pattern recognition*, Oregon, 2013, pp. 716–723.

[22] J. Wang, Z. Liu, J. Chorowski, Z. Chen, and Y. Wu, "Robust 3d action recognition with random occupancy patterns," in *Proc. European Conf. Computer Vision*, Florence, 2012, pp. 872–885.

[23] J. Sung, C. Ponce, B. Selman and A. Saxena, "Unstructured human

activity detection from rgbd images," in *Proc. IEEE conf. Robotics and Automation*, Saint Paul, 2012, pp. 842–849.

[24] X. Yang and Y. Tian, "Eigenjoints-based action recognition using naive-bayes-nearest-neighbor," in *Proc. IEEE Conf. Computer vision and pattern recognition workshops*, Providence, 2012, pp. 14–19.

[25] J. Wang, Z. Liu Y. Wu, and J. Yuan, "Mining actionlet ensemble for action recognition with depth cameras," in *Proc. IEEE Conf. Computer vision and pattern recognition*, Providence, 2012, pp. 1290–1297.

[26] A. Jalal and Y. Kim, "Dense depth maps-based human pose tracking and recognition in dynamic scenes using ridge data," in *Proc. IEEE Conf. Advanced video and signal based surveillance*, Seoul, 2014, pp. 119–124.

[27] A. Jalal, S. Kamal, and D. Kim, "Human depth sensors-based activity recognition using spatiotemporal features and hidden markov model for smart environments," *Journal of computer networks and communications*, vol. 2016, pp. 1–11, Sep. 2016.

[28] V. Deepu, S. Madhvanath, and A. Ramakrishnan, "Principal component analysis for online handwritten character recognition," in *Proc. IEEE Conf. Pattern recognition*, Cambridge, 2004, pp. 327–330.

[29] A. Jalal, Y. Kim, S. Kamal, A. Farooq and D. Kim, "Human daily activity recognition with joints plus body features representation using Kinect sensor," *in Proc. IEEE conference on Informatics, electronics and vision*, Japan, 2015, pp. 1–6.

[30] Y. Wang, K. Huang and T. Tan, "Abnormal activity recognition in office based on R transform," *in Proc. IEEE conference on image processing*, San Antonio, 2007, pp. 341–344.

[31] A. Jalal, "IM-DailyDepthActivity dataset," imlab.postech.ac.kr/ databases. htm, 2015, [Online; accessed 19 August- 2016].

[32] A. Jalal, S. Kamal and D. Kim, "Individual detection-tracking -recognition using depth activity images," *in Proc. IEEE conference on Ubiquitous robots and ambient intelligence*, Korea, 2015, pp. 450–455.

[33] M. Muller and T. Roder, "Motion templates for automatic classification and retrieval of motion capture data," *in Proc. ACM SIGGRAPH/ Eurographics symposium on computer animation*, Austria, 2006, pp. 137–146.

[34] L. Seidenari, V. Varano, S. Berretti, A. Bimbo, and P. Pala, "Recognizing actions from depth cameras as weakly aligned multi-part bag-of-poses," in *Proc. IEEE Conf. computer vision and pattern recognition workshops*, Portland, 2013, pp. 479–485.

[35] L. Xia, C. Chen, and J. Aggarwal, "View invariant human action recognition using histograms of 3d joints," in *Proc. IEEE Conf. Computer vision and pattern recognition workshops*, Providence, 2012, pp. 20–27.

[36] A. Keceli and A. Can, "A multimodal approach for recognizing human actions using depth information," in *Proc. IEEE Conf. Pattern recognition workshops*, Stockholm, 2014, pp. 421–426.

[37] L. Liu and L. Shao, "Learning discriminative representations from RGB-D video data," in *Proc. Inter. Conf. on artificial intelligence*, Beijing, 2013, pp. 1493–1500.

[38] M. Zanfir, M. Leordeanu, and C. Sminchisescu, "The moving pose: an efficient 3d kinematics descriptor for low-latency action recognition and detection," in *Proc. IEEE Conf. Computer Vision*, Sydney, 2013, pp. 2752–2759.

[39] Q. Li, Y. Zhou and A. Ming, "Integrating joint and surface for human action recognition in indoor environments, in *Proc. IEEE Conf. security, pattern analysis and cybernetics*, China, 2014, pp. 100–104.

[40] A. Jalal and S. Kamal, "Real-Time Life Logging via a Depth Silhouette-based Human Activity Recognition System for Smart Home Services," *in Proc. of the IEEE International Conference on Advanced Video and Signal-based Surveillance*, Korea, 2014, pp. 74-80.

[41] X. Yang and Y. Tian, "Super normal vector for activity recognition using depth sequences, in *Proc. IEEE Conf. Computer vision and pattern recognition*, Columbus, 2014, pp. 804–811.

[42] L. Xia and J. Aggarwal, "Spatio-temporal Depth cuboid similarity feature for activity recognition using depth camera, in *Proc. IEEE Conf. Computer vision and pattern recognition*, Portland, 2013, pp. 2834–2841.

[43] A. Jalal, S. Kamal and D. Kim, "Shape and motion features approach for activity tracking and recognition from Kinect video camera, in *Proc. IEEE Conf. advanced information networking and applications workshops*, Korea, 2015, pp. 445–450.

[44] M. Mabrouk, N. Ghanem and M. Ismail, "Semi supervised learning for human activity recognition using depth cameras, in *Proc. IEEE Conf. on machine learning and applications*, US, 2015, pp. 681–686.

[45] R. Slama, H. Wannous and M. Daoudi, "Grassmannian representation of motion depth for 3D human gesture and action recognition, in *Proc. Inter. Conf. on Pattern recognition*, Sweden, 2014, pp. 3499–3504.

[46] C. Wang, Y. Wang, and A. Yuille, "An approach to pose based action recognition," in *Proc. IEEE Conf. Computer vision and pattern recognition*, Portland, 2013, pp. 915–922.

[47] M. A. Gowayyed, M. Torki, M. E. Hussein and M. El-Saban, "Histogram of oriented displacements (HOD): describing trajectories of human joints for action recognition," in *Proc. Inter. Conf. on artificial intelligence*, China, 2013, pp. 1351–1357.

[48] R. Vemulapalli, F. Arrate and R. Chellappa, "Human action recognition by representing 3D skeletons as points in a lie group," in *Proc. IEEE Conf. computer vision and pattern recognition*, Columbus, 2014, pp. 588–595.

**Ahmad Jalal** received his B.S. degree in Computer Science from Iqra University, Peshawar, Pakistan and M.S degree in Computer Science from Kyungpook National University, Republic of Korea. He received his Ph.D. degree in the Department of Biomedical Engineering at Kyung Hee University, Republic of Korea. Now, he is working as Post-doctoral Research fellowship at POSTECH (Pohang University of Science and Technology). His research interest includes human computer interaction, image processing and computer vision.

**Shaharyar Kamal** He received his B.S. degree in Software Engineering from City University of Science and Information Technology, Peshawar, Pakistan and M.S. degree in Computer Engineering from Mid Sweden University, Sweden. He is currently enrolled in Ph.D. in the Department of Radio and Electronics Engineering at Kyung Hee University, Republic of Korea. His research interest includes wireless communication, image and signal processing.

**Daijin Kim** He received the B.S. degree in electronic engineering from Yonsei University, Seoul, Korea, in 1981, and the M.S. degree in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Taejon, 1984. In 1991, he received the Ph.D. degree in electrical and computer engineering from Syracuse University, Syracuse, NY. During 1992–1999, he was an Associate Professor in the Department of Computer Engineering at Donga University, Pusan, Korea. He is currently a Professor in the Department of Computer Science and Engineering at POSTECH, Pohang, Korea, a vice president of office of academic information affairs and Director of BK21 þ POSTECH CSE Institute. His research interests include computer vision, human computer interaction, and intelligent systems.

# Robust Artificial Immune System in the Hopfield network for Maximum *k*-Satisfiability

Mohd Asyraf Bin Mansor, Mohd Shareduwan Bin Mohd Kasihmuddin, and Saratha Sathasivam

*School of Mathematical Sciences, Universiti Sains Malaysia*

***Abstract* — Artificial Immune System (AIS) algorithm is a novel and vibrant computational paradigm, enthused by the biological immune system. Over the last few years, the artificial immune system has been sprouting to solve numerous computational and combinatorial optimization problems. In this paper, we introduce the restricted MAX-*k*SAT as a constraint optimization problem that can be solved by a robust computational technique. Hence, we will implement the artificial immune system algorithm incorporated with the Hopfield neural network to solve the restricted MAX-*k*SAT problem. The proposed paradigm will be compared with the traditional method, Brute force search algorithm integrated with Hopfield neural network. The results demonstrate that the artificial immune system integrated with Hopfield network outperforms the conventional Hopfield network in solving restricted MAX-*k*SAT. All in all, the result has provided a concrete evidence of the effectiveness of our proposed paradigm to be applied in other constraint optimization problem. The work presented here has many profound implications for future studies to counter the variety of satisfiability problem.**

***Keywords* —Artificial Immune System, Brute Force Algorithm, Hopfield Neural Network, Maximum *k*-satisfiability.**

## I. Introduction

The astonishing power of the biological systems has been a core impetus for the researcher to enhance and create a computational paradigm [1]. One of the brand new bio-inspired metaheuristic techniques is the artificial immune system (AIS) that enthused from the robust vertebrate immune system [2]. The artificial immune system (AIS) has been transformed into a rapid growing heuristics method in parallel computation and optimization problem. In point of fact, most of AIS practitioners have dedicated on the learning and memory domains of the immune system in order to resemble it with the human's immune system [3]. The breakthrough work by Farmer *et al*. (1986) [4] was the eye-opener for the other researchers to venture and improve the artificial immune system algorithm. After a few years, we can deduce that the main advances within AIS have revolved on three important immunological forte namely, the immune clonal selection, immune networks, and negative selection [5]. Thus, it was proven as a notable searching technique by various researchers. The AIS algorithm has been applied to wide range problems such as global optimization [6], pattern recognition [7], multiple sequence alignment [8] and shop scheduling conundrum [9]. The related work by Layeb *et al*. [10] has demonstrated the robustness of the artificial immune system in tackling the general maximum satisfiability problem. Hence, we will improve the work by taking the advantage of clonal selection power in artificial intelligence together with the Hopfield network to solve maximum *k*-satisfiability problem.

The main motivation of this paper is to compare the effectiveness of the searching techniques incorporated with the Hopfield neural network in solving MAX-*k*SAT problem. Since a few decades ago, Boolean satisfiability has emerged from a classical problem into a bunch of various hard problems [11]. Hence, MAX-*k*SAT is a counterpart of classical Boolean satisfiability that has captured the attention of many researchers in the optimization field [12]. Specifically, the MAX-*k*SAT can be delineated as the maximum number of satisfied clauses achieved by any complete assignment [13]. The integration of artificial immune system and Hopfield neural network has been proven in solving some real life or industrial computational problems [43, 44]. The Hopfield neural network is vital due to its resemblance to the human biological brain [14]. Thus, the output we obtained after performing the searching technique will be stored as a content addressable memory (CAM). Thus, it is vital in identification or pattern recognition problem [45]. Hence, the Hopfield neural network will train the output systematically in order to drive towards the final output (global minima).

The idea of applying Hopfield neural network to hunt solution has been supported by numerous notable works [15, 16, 17]. Hopfield neural network comprises of a simple recurrent network that has an efficient associative memory and resembles the biological brain [14]. The important property of the Hopfield neural network is the minimization of energy whenever there is any change in inputs [15]. Moreover, the Hopfield neural network minimizes Lyapunov energy by utilizing the physical Ising spin of the neuron states. On top of that, the network produces global output by minimizing the network energy. Gadi Pinkas [16] and Wan Abdullah [17] described a bi-directional mapping between logic and energy function of symmetric neural network. Besides, both methods are the building blocks for a corresponding logic program. Due to effectiveness of energy changes in Hopfield neural network, several researchers have combined the idea of logic programming with Hopfield neural network. Several renowned models were developed by Sathasivam [18] and Wan Abdullah [17].

The main work of this paper is to propose a hybrid computational model by incorporating artificial immune system (AIS) and Hopfield neural network (HNN-MAX*k*SATAIS) in solving maximum *k*-satisfiability (MAX-*k*SAT) problem. Specifically, the work will focus on the capability of our proposed hybrid network in obtaining the maximum satisfied clauses in any MAX-*k*SAT formula. Therefore, the novelty can be found in the proposed hybrid technique since most of the researchers are only focusing on the standalone Hopfield neural network or metaheuristic to solve any maximum *k*-satisfiability problem. To frame the novelty, we will compare the HNN-MAX*k*SATAIS with conventional method, via integrating brute force algorithm with Hopfield neural network (HNN-MAX*k*SATBF).

The remaining frameworks of the paper are organized as follows. In Section II, the noteworthy concept of Boolean satisfiability (SAT) and maximum *k*-satisfiability (MAX-*k*SAT) problem are introduced. After that, we discuss in brief about the neuro-searching techniques namely Brute force (BF) search and artificial immune system (AIS) in Section III. Section IV discusses the Hopfield model, Wan Abdullah's



DOI: 10.9781/ijimai.2017.448

method and Sathasivam's relaxation method. Furthermore, we present the theory implementation of our proposed network in Section V. Moving on, Section VI encloses the experimental results for HNN-MAX*k*SATAIS and HNN-MAX*k*SATBF in terms of global minima ratio, ratio of satisfied clause, fitness energy landscape value and the computation time. As a final point, we conclude with some remarks in Section VII.

## II. Maximum *k*-Satisfiability Problem

### A. *k-Satisfiability Problem*

The *k*-SAT problem can be defined as a problem to determine the satisfiability of a particular Boolean formula containing *k* literals per clause [36]. *k*-SAT is generally expressed in terms of *k* – CNF (Conjunctive Normal Form) or Krom formula [37]. *k*-SAT has been a special NP problem that represents various optimization problems such as circuit and pattern recognition. Solutions of the *k*-SAT are corresponding to the solution of the various optimization problems. *k*-SAT consists of a set of *m* variables where $x_1, x_2, \ldots\ldots, x_m$ with a set of literal or negation. These variables accommodate the *n* clauses $C_1, C_2 \ldots\ldots C_n$ . The finest definition of *k*-SAT Boolean formula is as followed:

$$P = \bigwedge_{i=1}^{n} C_i \tag{1}$$

where $\wedge$ is a logical AND connector and *P* denotes the entire Boolean formula for *k*-SAT. $C_i$ is a clausal form of DNF with *k* variables. Boolean value or truth value of the variable only consists of 1 or -1. These two values impersonate the information of True and False respectively. The goal of *k*-SAT problem is to decide whether there exists an assignment that makes *P* become satisfiable. In this paper, randomized 2-SAT and 3-SAT will be the benchmark to our satisfiability problem. Randomized 2-SAT and 3-SAT clauses are as followed:

$$C_i = \bigvee_{j=1}^{k} \left( x_{ij}, y_{ij} \right), \ k = 2 \tag{2}$$

$$C_i = \bigvee_{j=1}^{k} \left( x_{ij}, y_{ij}, z_{ij} \right), \ k = 3 \tag{3}$$

### B. *Restricted Maximum k-Satisfiability*

Restricted maximum *k*-Satisfiability problem (MAX-*k*SAT) is a variant of Boolean SAT problem. Given a Boolean formula *P* in conjunctive normal form (CNF) with *n* clauses containing *k* number of variables per clause and positive integer $g$ where $g \leq n$ . MAX-*k*SAT can be defined implicitly as a pair of $(\lambda, \theta)$ [10] where $\lambda$ is the set of all possible solution $\{1, -1\}^n$ bit string and $\theta$ is a mapping $\lambda \rightarrow \xi$ which denotes the score of the assignments. $\xi$ is scored based on correct clauses. Therefore, MAX-*k*SAT problem consist of defining the best bipolar/binary assignments to the variables in *P* that simultaneously satisfy at least *g* of the *n* clauses. The aim of the system is to decide the "optimized" assignment that can satisfy the maximum number of clauses containing *k* variables. It was proven that MAX-*k*SAT is NP-complete problem. There are numerous classifications of the MAX-*k*SAT namely, weighted MAXSAT [33] and Partial MAXSAT [34]. In our case, restricted value of *k* in MAX-*k*SAT only allowed $k = 2$ and $k = 3$. Restricted MAX-*k*SAT can be included as combinatorial problem that works in parallel with logic programming. Equation (4) is an example of MAX-*k*SAT formula:

$$P = (x \vee y) \wedge (x \vee \neg y) \wedge (\neg x \vee y) \wedge (\neg x \vee \neg y) \wedge (z \vee w) \tag{4}$$

Equation (4) consists of variables $x, y, z \ and \ w$ and formula *P* is impossible to satisfy since there are no assignments that make the above constraint become true. As the number of clauses increases, finding satisfying assignment will be terribly complex.

## III. Neuro- Searching Paradigm

Neuro-searching techniques are staple in the computational method in order to hunt for the optimal solutions. The implementation of Hopfield neural network in logic programming has been successfully done by a few researchers. However, we require a solid neuro-searching paradigm to facilitate and foster the training process especially when the problem is getting complex. The standalone Hopfield network has limitations especially in terms of complexity [18]. The complexity will determine whether our models able to withstand more complex problems or not. Thus, we embedded the conventional brute force search as a heuristic method to the Hopfield neural network in solving MAX-*k*SAT problem (HNN-MAX*k*SAT). Conversely, we proposed the artificial immune system (AIS) algorithm as searching technique for the Hopfield neural network (HNN-MAX*k*SATAIS). In this paper, neuro-searching techniques were implemented in computing the maximum number of clauses for maximum 2-satisfiability (MAX-2SAT) and maximum 3-satisfiability problem (MAX-3SAT).

### A. *Brute Force (BF) algorithm*

Brute-force (BF) algorithm can be delineated as a straight forward local search method for an element with a specific property between combinatorial domains including probabilities, permutations, combinations, logics, satisfiability or arrays of a set [23]. It is a conventional technique and the simplest algorithm, extensively used in pattern searching problem [24, 37]. The brute force searching paradigm is based on "generate and test" principle to enumerate all possible solutions in a search space. In our context, we will implement the brute force algorithm as a searching technique in the Hopfield network for solving maximum *k*-satisfiability problem (HNN-MAX*k*SATBF).

The brute force algorithm is a generalized problem-solving paradigm that usually enumerates all possible solutions and check whether the solutions satisfies the given formula [24]. It will return the satisfying assignment for the formula, if such exists. In essence, MAX*k*SAT formula with a specific number and combinations of clauses is addressed as a combinatorial optimization problem to be solved by our proposed technique. Furthermore, the brute force algorithm will explore any combinations of assignments and directly compute the total number of satisfied clauses. Technically, the brute force algorithm will allow our paradigm to hunt for the total satisfied clauses brutally, even in tremendous search dimension [25]. Specifically, the brute force search will evaluate the candidate solutions clause by clause in order to obtain the feasible solution. The feasible assignment will be stored in the Hopfield memory as content addressable memory (CAM) [30].

The brute force algorithm (exhaustive search) has been widely used by the researchers in solving numerous satisfiability problems, including the maximum satisfiability problem [23, 24]. Hence, it can be considered as the primitive searching tool for the standalone Hopfield neural network. Therefore, the brute force searching method is practically easy to implement by the researchers [25].

The underlying reason we venture this conventional method is to ascertain the degree of effectiveness of HNN-MAX*k*SATBF. Theoretically, the brute force algorithm devours more computation time in searching for the maximum satisfied assignment completely. Hence, the computation complexity is represented as $O(2^n)$. This hybrid

network will encounter higher complexity as the program attempts large and more constrained clauses. Thus the computation time will be intensified and might end up in infinite loop. In this paper, we will generate random bit strings (represented the assignment of the MAX-*k*SAT logic) and straight away record the number of satisfied clauses. Hence, it is not guaranteed that the bit strings are not converging to global maxima during the first iteration of BF.

The brute force algorithm is given as follows:

**Step 1**

Generate the candidate bit strings.

**Step 2**

Test the candidate bit strings and compute the number of satisfied clauses.

**Step 3**

**If**

Return the assignment with the maximum number of satisfied clauses as an output.

**Else**

Repeat Step 1 and 2.

### *B. Artificial Immune System (AIS)*

The artificial immune system is a population based paradigm where every individual represents a potential solution to the problem [6]. The artificial immune system is popularized by Framer *et al*. (1986) [4] by modelling the Jerne's Immune network theory. On top of that, the artificial immune system (AIS) algorithm can be illustrated as a distributed network and able to do parallel processing. We proposed the binary artificial immune system according to the immune clonal selection perspective. Technically, the binary artificial immune system has been implemented by several researchers for binary optimization and pattern recognition [6, 7].

Furthermore, the complex interactions between entities within each level will ensure the immune system to shield the body after any harmful entity and exogenous agent, known as antigen. A particular form of cell, identified as B-cells, is in charge for the destruction of the antigen. Hence, the B-cell produces antibodies that bind with the antigens and mark them for damage [3]. In addition, the power of the antibody or antigen binding is called antigenic affinity [8]. Robust features of that immune system have fortified their adaptation to information technology in solving numerous problems. In this paper, we only focused on clonal selection that will be implemented in our binary AIS.

The remarkable feature in our biological immune system is the capability to build antibodies to combat the new antigens or pathogens [1, 3]. Hence, the immune clonal selection process depicts the fundamental structures of an immune response towards an antigenic stimulus. By all means, it inaugurates the idea that only those cells can identify the antigen proliferate; thus, being nominated against those that do not [5]. Specifically, the B-cells will produce the antibodies if any incoming antigen is discovered. Then, the particular B-cells distinguish the antigen proliferate via cloning process. Significantly, the main event during clonal mutation is called somatic hypermutation, whereby the genetic maturation and variation can be improved [7, 9]. The B-cells with higher affinity will be differentiated into plasma and memory cell, whereby the worst one will be destroyed [35].

In our paper, we developed a hybrid paradigm by implementing the Hopfield network and binary AIS to do a MAX-*k*SAT logic programming HNN-MAX*k*SATAIS. In our exploration on binary AIS, the binary strings were illustrated as the B-cells. Thus, the bit string of 1 and -1 will be represented as "true" and "false" respectively.

The artificial immune system algorithm is given as follows:

**Stage 1: Initialization**

Generate and initialize the 100 B-cells (bit strings). Generally, any massive and diverse population of B-cells may represent a massive space search of solutions that can lead the program to global solutions. On contrary, a smaller population size can contribute to local minima solutions.

**Stage 2: Affinity evaluation**

Compute the affinity of every B-cells (each of candidate solutions). The affinity measure can be delineated as the total number of satisfied clauses in any particular MAX-*k*SAT formula. The affinity evaluation of the artificial immune system is given as follows:

$$aff = c_1(x) + c_2(x) + c_3(x) ..... + c_N(x) \tag{5}$$

where $c_1, c_2, c_3 ..... c_N$ are the number of clause checked by artificial immune system and $N$ is the number of clauses present in the formula. Specifically, the role of the fitness function is to evaluate the candidate bit strings.

**Stage 3: Selection**

Select the best 5 B-cells with the highest affinity. The selected B-cells will stand the chance to perform the cloning process.

**Stage 4: Cloning**

Clone and duplicate the selected B-cells by implementing the roulette wheel mechanism [42]. Therefore, the newly produced B-cells population will comprise of 200 cloned B-cells. We need to consider the initial affinity ( $aff_i$ ) and the total affinity of the population to check the number of possible clone. The $\beta$ is the number of population clone that the program want to introduce to the search space. In our study, we set $\beta = 200$ to be punched into formula (6).

$$\begin{pmatrix}\text{The number of}\\ \text{clone allowed}\end{pmatrix} = \frac{aff_i}{\sum aff} \times \beta \tag{6}$$

**Stage 5: Normalization**

Normalize the B-cells $(aff\ N_i)$ via equation (7). Thus, the antibodies that exist in a memory response will achieve a higher average affinity than those of the initial primary response. It is called the maturation of the immune response process.

$$aff\ N_i = \frac{aff_i - \min aff}{\max aff - \min aff} \tag{7}$$

**Stage 6: Somatic Hypermutation**

The mutation process in AIS is vital in order to improve the quality of B-cells. The process is enriched by the "somatic" principle whereby the nearer the match, the more disruptive the mutation [6]. In order to avoid possible local maxima in terms of affinity (non-improving B-cell), the somatic hypermutation might be very useful. Mutation for each B-cell works by flipping the variable from 1 to -1 or vice versa. The flipping process will improve the B-cells (bit strings) to achieve the best affinity value.

$$\begin{pmatrix}\text{Number of}\\ \text{Mutation (Nb)}\end{pmatrix} = \left(\frac{1}{NN}\right)\left(aff\ N_i\right) + \left(1 - aff\ N_i\right)\left(0.01\right) \tag{8}$$

Thus, we can compute the affinity value for the newly formed B-cells (matured population). NN denotes the number of variables in

one particular randomized MAX-*k*SAT. If the B-cells are able to reach the maximum affinity, the solution will exit the program. On contrary, if any of the B-cell did not manage to achieve maximum affinity, the program will reset the affinity of all B-cells and repeat stage 1 until 6.

In this paper, we hybridized AIS algorithm with the Hopfield neural network as a network based on logic programming to solve MAX-*k*SAT problems (HNN-MAX2SATAIS and HNN-MAX3SATAIS).

## IV. Neuro- Logic In Hopfield Neural Network

### A. The Hopfield Neural Network

One of the milestone for neural network was the associative model proposed by Hopfield at the beginning of 1980s. This model has been celebrated by many optimization problems such as computer network, pattern recognition and scheduling problem since this network focused the ensembles of computing unit.

Hopfield design consist of interconnected unit called neurons, forming a network. Computation in Hopfield network is executed by collections of activated neurons [15, 38]. Most of the literature suggest Hopfield network contains good properties including parallel execution for fast computation and exceptional stability [39]. In particular, we choose Hopfield network to do MAX-*k*SAT problem because the truth value of the problem can be well presented in Ising model, well distributed, integration of CAM [41], simple implementation and easy to synthesis with other algorithm.

The units in Hopfield nets are called binary threshold unit [30] which can only consider bipolar values of 1 and -1. The paramount definition for unit I's activation, $a_i$ are given:

$$a_i = \begin{cases} 1 & if \sum_j W_{ij} S_j > \xi_i \\ -1 & Otherwise \end{cases} \tag{9}$$

where $W_{ij}$ is the synaptic weight from unit $j$ to $i$. $S_j$ is the state of unit $j$ and $\xi_i$ is the threshold of unit $i$. The network's architecture comprises of $N$ recognized neurons, each was described by an Ising spin variable model. The connections in Hopfield net contain no connection with itself $W_{ii} = W_{jj} = 0$. This gave Hopfield the symmetrical features in terms of architecture [18]. Neuron state is basically bipolar $S_i \in \{1, -1\}$ and the spin points follow in the direction of magnetic field. This causes each neuron to flip until the equilibrium is reached. Thus, it follows the dynamics $S_i \rightarrow \mathrm{sgn}(h_i)$ where $h_i$ is the local field of the neurons connection. When dealing with higher order neurons connection, the sum of the field induced by each neuron are as followed:

$$h_i = .... + \sum_j W_{ijk}^{(3)} S_j S_k + \sum_j W_{ij}^{(2)} S_j + W_i^{(1)} \tag{10}$$

Since the synaptic weight in Hopfield network is constantly symmetrical and does not contain zero diagonal, the updating rule maintains as follows [31]:

$$S_i(t+1) = \mathrm{sgn}\left[h_i(t)\right] \tag{11}$$

Equation (10) is necessary in order to ensure the energy decrease monotonically with the network. The generalized Lypunov energy equation that accommodates more neurons connections is as followed:

$$E = -\frac{1}{3}\sum_i \sum_j \sum_k W_{ijk}^{(3)} S_i S_j S_k - \frac{1}{2}\sum_i \sum_j W_{ij}^{(2)} S_i S_j - \sum_i W_i^{(1)} S_j \tag{12}$$

This energy function is significant because it establishes the degree of convergence of the network [41]. The energy value obtained from the equation will be verified as global or local minimum energy. For our case, the network will produce the correct solution when the induced neurons state reached global minimum energy. The process of obtaining global minimum energy is always associated with the correct model of synaptic weight. For MAX-*k*SAT, we implemented Wan Abdullah's updating technique to obtain the synaptic weight for our network [17].

### B. Wan Abdullah's Method in Learning MAX-kSAT Clauses

MAX-*k*SAT can be treated as one of the constrained optimization problem that is being carried out on Hopfield neural network. Wan Abdullah's method became the pioneer in synaptic weight extraction based on logical inconsistencies [17]. Cost function that corresponds to MAX-*k*SAT clauses is the minimization of logical inconsistencies.

$$\min_{i \in (0,\infty),\, C_i = 1} \neg C_i \tag{13}$$

As the number of "wrong" assignment decreases, the number of satisfied MAX-*k*SAT clauses will increase. For example, consider the following randomized MAX-2SAT and MAX-3SAT problem with $\alpha$ and $\phi$

$$\alpha = (A \vee B) \wedge (A \vee \neg B) \wedge (\neg A \vee B) \wedge (\neg A \vee \neg B) \wedge (C \vee D) \tag{14}$$

$$\begin{aligned} \phi = {} & (P \vee Q \vee R) \wedge (\neg P \vee Q \vee R) \wedge (P \vee \neg Q \vee R) \\ & \wedge (P \vee Q \vee \neg R) \wedge (\neg P \vee \neg Q \vee R) \wedge (\neg P \vee Q \vee \neg R) \\ & \wedge (P \vee \neg Q \vee \neg R) \wedge (\neg P \vee \neg Q \vee \neg R) \wedge (S \vee T \vee U) \end{aligned} \tag{15}$$

Cost function $f_{\mathrm{cost}}$ for both equations (14) and (15) are as followed:

$$\begin{aligned} f_{\mathrm{cost}\,\alpha} = {} & \frac{1}{2}(1-S_A)\frac{1}{2}(1-S_B) + \frac{1}{2}(1-S_A) + \frac{1}{2}(1+S_B) \\ & + \frac{1}{2}(1+S_A)\frac{1}{2}(1-S_B) + \frac{1}{2}(1+S_A)\frac{1}{2}(1+S_B) \\ & + \frac{1}{2}(1-S_C)\frac{1}{2}(1-S_D) \end{aligned} \tag{16}$$

$$\begin{aligned} f_{\mathrm{cost}\,\phi} = {} & \frac{1}{2}(1-S_P)\frac{1}{2}(1-S_Q)\frac{1}{2}(1-S_R) + \frac{1}{2}(1+S_P)\frac{1}{2}(1-S_Q)\frac{1}{2}(1-S_R) \\ & + \frac{1}{2}(1-S_P)\frac{1}{2}(1+S_Q)\frac{1}{2}(1-S_R) + \frac{1}{2}(1-S_P)\frac{1}{2}(1-S_Q)\frac{1}{2}(1+S_R) \\ & + \frac{1}{2}(1+S_P)\frac{1}{2}(1+S_Q)\frac{1}{2}(1-S_R) + \frac{1}{2}(1+S_P)\frac{1}{2}(1-S_Q)\frac{1}{2}(1+S_R) \\ & + \frac{1}{2}(1-S_P)\frac{1}{2}(1+S_Q)\frac{1}{2}(1+S_R) + \frac{1}{2}(1+S_P)\frac{1}{2}(1+S_Q)\frac{1}{2}(1+S_R) \\ & + \frac{1}{2}(1-S_T)\frac{1}{2}(1-S_U)\frac{1}{2}(1-S_V) \end{aligned} \tag{17}$$

By comparing equation (16), (17) with equation (12), we obtained synaptic weight for $\alpha$ and $\phi$. The synaptic weights are shown in Table 1 and Table 2.

TABLE I

SYNAPTIC WEIGHT FOR $\alpha$ BASED ON WAN ABDULLAH'S METHOD

| $W$ | $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ |
|---|---|---|---|---|---|
| $W_A^{(1)}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $-\frac{1}{4}$ | $-\frac{1}{4}$ | 0 |
| $W_B^{(1)}$ | $\frac{1}{4}$ | $-\frac{1}{4}$ | $\frac{1}{4}$ | $-\frac{1}{4}$ | 0 |
| $W_{AB}^{(2)}$ | $-\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $-\frac{1}{2}$ | 0 |
| $W_C^{(1)}$ | 0 | 0 | 0 | 0 | $\frac{1}{4}$ |
| $W_D^{(1)}$ | 0 | 0 | 0 | 0 | $\frac{1}{4}$ |
| $W_{CD}^{(1)}$ | 0 | 0 | 0 | 0 | $-\frac{1}{2}$ |

TABLE II

SYNAPTIC WEIGHT FOR $\phi$ BASED ON WAN ABDULLAH'S METHOD

| $W$ | $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ | $C_6$ | $C_7$ | $C_8$ | $C_9$ |
|---|---|---|---|---|---|---|---|---|---|
| $W_P^{(1)}$ | $\frac{1}{8}$ | $-\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $-\frac{1}{8}$ | $-\frac{1}{8}$ | $\frac{1}{8}$ | $-\frac{1}{8}$ | 0 |
| $W_Q^{(1)}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $-\frac{1}{8}$ | $\frac{1}{8}$ | $-\frac{1}{8}$ | $\frac{1}{8}$ | $-\frac{1}{8}$ | $-\frac{1}{8}$ | 0 |
| $W_R^{(1)}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $-\frac{1}{8}$ | $\frac{1}{8}$ | $-\frac{1}{8}$ | $-\frac{1}{8}$ | $-\frac{1}{8}$ | 0 |
| $W_{PQ}^{(2)}$ | $-\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $-\frac{1}{8}$ | $-\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $-\frac{1}{8}$ | 0 |
| $W_{QR}^{(2)}$ | $-\frac{1}{8}$ | $-\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $-\frac{1}{8}$ | $-\frac{1}{8}$ | 0 |
| $W_{PR}^{(2)}$ | $-\frac{1}{8}$ | $\frac{1}{8}$ | $-\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $-\frac{1}{8}$ | $\frac{1}{8}$ | $-\frac{1}{8}$ | 0 |
| $W_{PQR}^{(3)}$ | $\frac{1}{16}$ | $-\frac{1}{16}$ | $-\frac{1}{16}$ | $-\frac{1}{16}$ | $\frac{1}{16}$ | $\frac{1}{16}$ | $\frac{1}{16}$ | $-\frac{1}{16}$ | 0 |
| $W_S^{(1)}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $\frac{1}{8}$ |
| $W_T^{(1)}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $\frac{1}{8}$ |
| $W_U^{(1)}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $\frac{1}{8}$ |
| $W_{ST}^{(2)}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $-\frac{1}{8}$ |
| $W_{TU}^{(2)}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $-\frac{1}{8}$ |
| $W_{SU}^{(2)}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $-\frac{1}{8}$ |
| $W_{STU}^{(3)}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $\frac{1}{16}$ |

Hebbian learning is another possible alternative to imprint the adequate synaptic weight for neurons connections. Sathasivam has shown that the synaptic weight obtained by using Wan Abdullah's method are similar due to clausal MAX-*k*SAT similarity. Although both method is expected to produce the similar synaptic weight information, Wan Abdullah's method is proven to minimize the spurious minima produced by clauses compared to Hebbian learning [16, 17]. Spurious minima can lead to local minima solutions.

### *C. Network Relaxation*

The quality of the solutions obtained by Hopfield network can be affected by multiple factors. Prolonged firing and receiving information among neurons can reduce the quality of the solutions. As the number of neurons increased, more interconnected neurons involved in firing and receiving information. In this case, network relaxation helps the neurons to "pause" before continuing the information exchange. "Network relaxation" is a series of relaxation loop in the program after local field has been obtained. Without proper relaxation mechanism, network tends to produce more local minima solution. Since MAX-*k*SAT contains more clausal constrained, we applied Sathasivam's relaxation technique [18] to ensure the network relaxed to its final state. Information exchange between neurons will be updated based on the following equation:

$$\frac{dh_i^{new}}{dt} = R\frac{dh_i}{dt} \tag{18}$$

Where $R$ denotes the relaxation rate and $h_i$ refers to the local field of the network as listed in equation (10). For our case, constant relaxation $R = 2$ will be used in our program.

## V. IMPLEMENTATION OF NEURO-HEURISTIC METHOD

The simulations for HNN-MAX*k*SATAIS and HNN-MAX*k*SATBF were executed on Microsoft Visual C++ 2013 for Windows 10. The main task of the program is to find "models" that can solve MAX-*k*SAT problem. The following algorithms shows how we can implement HNN-MAX*k*SATAIS and HNN-MAX*k*SATBF.

a) Translate MAX-*k*SAT clauses into Boolean algebra (if any).
b) Assign neuron to each variable in MAX-*k*SAT clauses
c) Initialize all synaptic weight to zero.
d) Derive MAX-*k*SAT cost function by assigning $X = \frac{1}{2}(1 + S_X)$ and $\overline{X} = \frac{1}{2}(1 - S_X)$. The state of the neuron reads true when $S_X = 1$ and reads false when $S_X = -1$
e) By using Wan Abdullah's method, compare the derived cost function with equation (16) or (17).
f) Check clauses satisfaction by using both BF and AIS. Each satisfied clause will be stored to the network.
g) Randomize the state of the neurons.
h) Compute the corresponding local field by using equation (10). Find the final state and run it for 5 times. If the state remains unchanged, we considered it as stable state.
i) Apply Sathasivam's relaxation method to the network via equation (14).
j) Find the corresponding final energy by using equation (12). Verify whether final energy is a local minimum energy of global minimum energy.
k) Find the corresponding global minima ratio, ratio of satisfied clauses, fitness energy landscape value and computation time.

Each simulation runs 100 trials with 100 combinations of neurons in order to reduce statistical error. According to Sathasivam [18], 0.001 is chosen as tolerance value since it gives us a better filtering mechanism to separate global or local minima solution.

## VI. Result and Discussion

The performance of simulated program model for both HNN-MAX*k*SATBF and HNN-MAX*k*SATAIS will be tested by a few tests. Both models will be compared in terms of global minima ratio, ratio of satisfied clauses, fitness energy landscape value and the computation time. In order to make a fair comparison for all tests, all simulations will be performed in the same device with the similar processor.

### A. Global Minima Ratio

Global minima ratio is elucidated as the ratio between the total global minimum energy over total number of runs [18]. Since the program pursues the global minimum energy for every neuron state in MAX-*k*SAT, it will be wise enough to indicate the simulation by checking it's the ratio of global minima. In our case, each simulation will produce 10000 bit strings each consist of final state of the neurons. 0.9500 global minima ratio value indicates 9500 bit strings solutions are global minimum and 500 bit strings are local minimum.

TABLE III
Global Minima Ratio

| NN | HNN-MAX2SATBF | HNN-MAX2SATAIS | HNN-MAX3SATBF | HNN-MAX3SATAIS |
|---|---|---|---|---|
| 10 | 0.9724 | 0.9999 | 0.9702 | 0.9993 |
| 20 | 0.9645 | 0.9904 | 0.9549 | 0.9822 |
| *30* | 0.9487 | 0.9767 | 0.9300 | 0.9606 |
| *40* | 0.9133 | 0.9632 | 0.8801 | 0.9399 |
| *50* | 0.8740 | 0.9408 | 0.8554 | 0.9230 |
| *60* | 0.8346 | 0.9225 | 0.8246 | 0.9117 |
| *70* | - | 0.9123 | - | 0.9045 |

Table III depicts the global minima ratio obtained by HNN-MAX*k*SATAIS and HNN-MAX*k*SATBF respectively. Hence, the capability of both hybrid networks can be measured by taking into consideration their global minima ratio for different variants of complexity. Sathasivam [18] explained the correlation between the global minima ratio with the nature of energy achieved at the end of the computation process. Theoretically, if the global minima ratio of the proposed hybrid network is close to one, almost all solutions in the network reached global minimum energy (global solution). According to Table 3, HNN-*k*MAXAIS has the capability to retrieve more accurate state compared to MAX-*k*SATBF. This is due to the effectiveness in the neuro-searching technique employed by the HNN-*k*MAXAIS. The ability of the B-cells in fighting the pathogens and improving affinity of the bit string in artificial immune system algorithm helps the program to search the solution optimally compared to the traditional brute force search algorithm. Hence, more solutions had successfully achieved the global minimum energy compared to the local minima (non-improving solution). Somatic hypermutation helps the program to avoid local maxima in terms of affinity (non improving affinity). Variation of mutation rate for each group of B-cells helps the program to spread the potential search space. This property causes AIS algorithm to find the correct solution and converge to global minima compared to traditional BF.

On the other hand, the limit for HNN-MAX-*k*SATBF is 60 neurons. After 60 neurons, the network in HNN-MAX-*k*SATBF trapped in trial and error state and consume more time to find the solution. On contrary, HNN-*k*MAXAIS is able to withstand number of neurons up to 70 neurons. The artificial immune system has been proven in reducing the complexity of the network. As the number of neurons increased, the complexity of the network increased, since the size of the constraint will enlarge indefinitely. Besides, AIS algorithm was able to sort the possible candidate solution (B-cells) effectively and can cope with more constraints compared to BF algorithm. The problem with HNN-MAX*k*SATBF was the nature of the brute-force search that deployed an intensive training process in hunting the correct neuron states. Therefore, the updating rule for HNN-MAX*k*SAT-BF will generate additional abrupt energy surfaces and more solutions are not improving (local minima). Based on the obtained global minima ratio, HNN-MAX*k*SATAIS has outperformed the HNN-MAX*k*SATBF in generating the global solutions.

### B. Ratio of Satisfied Clauses

As we deal with the MAX-*k*SAT problem, we are required to compute the possible satisfied clauses for each of the problems. Ratio of satisfied clauses can be defined as the total number of satisfied clauses over the total number of clauses [13].

TABLE IV
Ratio of Satisfied Clause

| NN | HNN-MAX2SATBF | HNN-MAX2SATAIS | HNN-MAX3SATBF | HNN-MAX3SATAIS |
|---|---|---|---|---|
| 10 | 0.725 | 0.750 | 0.805 | 0.875 |
| 20 | 0.708 | 0.750 | 0.796 | 0.870 |
| *30* | 0.700 | 0.748 | 0.784 | 0.852 |
| *40* | 0.694 | 0.742 | 0.753 | 0.838 |
| *50* | 0.680 | 0.735 | 0.718 | 0.825 |
| *60* | 0.643 | 0.731 | 0.696 | 0.800 |
| *70* | - | 0.726 | - | 0.788 |

NN=Number of neurons.

Table IV portrays the ratio of the satisfied clauses over total clause obtained HNN-MAX2SATAIS, HNN-MAX2SATBF and HNN-MAX3SATAIS, HNN-MAX3SATBF. In maximum satisfiability problem, MAX-2SAT and MAX-3SAT clauses will never be fully satisfied. Hence, the searching method will be able to return the maximum number of satisfied clauses as the output. We can further deduce that, the higher the ratio obtained, the more clauses will be satisfied in any MAX-*k*SAT problem. According to Table 4, HNN-MAX*k*SATAIS is proven to obtain more satisfied clauses in MAX-*k*SAT compared to conventional Brute force method. As the number of neurons increased, the HNN-MAX*k*SATAIS is still able to maintain the quality of the ratio. The ability of B-cells in artificial immune system algorithm to perform hyper mutation will drive into optimal solutions. Thus, it will be able to cope the higher number of neurons. Conversely, HNN-MAX*k*SATBF will produce a slightly lower ratio of satisfied clauses since most of the solutions obtained trapped at local minima. The generate and test procedure in the brute force search will cause the bit string to reach local maxima easily. This is the higher energy barrier needed during searching the optimum assignment. Thus, high stability in HNN-MAX*k*SATAIS will reduce the spurious minima which will cause the retrieved solutions become local minima solutions.

### C. Fitness Energy Landscape Value

Basically, the effectiveness of a paradigm can be measured by taking into consideration the fitness energy landscape value. Hence, the fitness energy landscape value is linked with the pattern or input storing competency. The ruggedness of the energy model depicts the performance of each of our algorithms. Thus, the fitness energy landscape value is computed by using Kauffman's model [29]. Fig. 1 and 2 represent the obtained fitness energy landscape values for HNN-MAX*k*SATAIS and HNN-MAX*k*SATBF respectively.

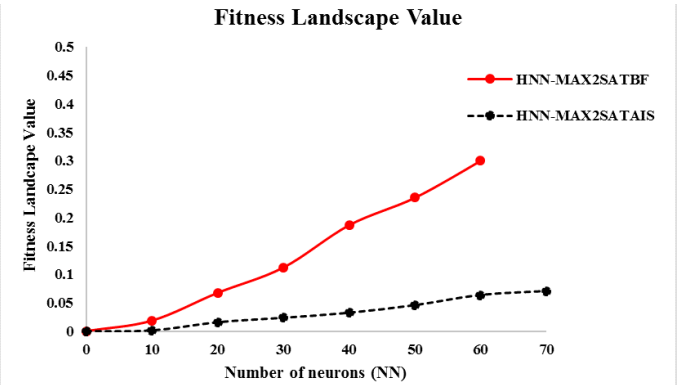



Fig. 1. Fitness energy landscape value for HNN-MAX2SATBF and HNN-MAX2SATAIS.

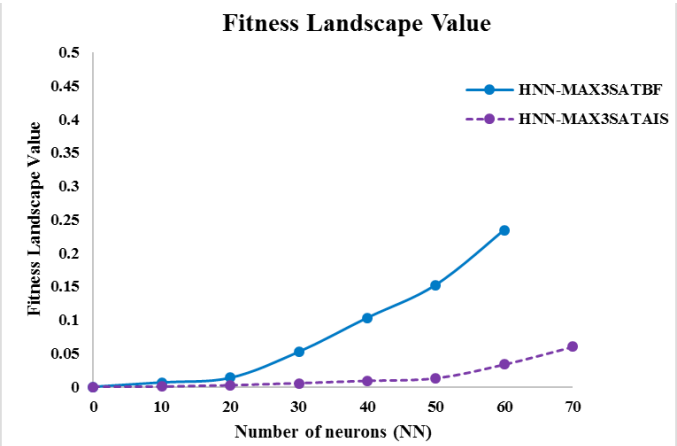



Fig. 2. Fitness energy landscape value for HNN-MAX3SATBF and HNN-MAX3SATAIS.

As shown in the Fig. 1 and Fig. 2, the difference in energy for HNN-MAX*k*SATAIS is very nearly flat (zero) compared to HNN-MAX*k*SATBF. Yet another observation that can be made from the graph is the capability of the artificial immune system algorithm to optimize the energy required throughout the exploration compared to the brute force search algorithm. Also, it was observed that the fitness energy landscape for HNN-MAX*k*SATBF can be observed as more rugged than the HNN-MAX*k*SATAIS. One thing to ponder is the MAX-*k*SAT clauses always allied to the ruggedness of the energy landscape. The more rugged the energy landscape, the harder it will be to achieve feasible solution [21]. Since the complexity of the solution searching has been reduce drastically by AIS in HNN-MAX*k*SATAIS, additional relaxation time was added before the network retrieve the final states. Consequently, HNN-MAX*k*SATAIS has a remarkable ability to store MAX*k*SAT pattern compared to HNN-MAX*k*SATBF. As a result, more global minimum energy produced that will drive to more feasible solutions. Thus, we are able to generate and retrieve the maximum satisfied clauses accurately. Other than that, HNN-MAX*k*SATBF was able to withstand up to 60 neurons due to the complexity.

### *D. Computation Time*

The robustness of our algorithms can be approximately proved by looking at the effectiveness of the entire computation process. In our case, the computation time refers to the expanse of time for which our network was used to complete the whole computation process [31]. Thus, the computation process involves the training and generating the maximum satisfied clauses by our proposed paradigms. In this study, we measure the computation time by using the standard unit of second for the simplicity.

TABLE V

COMPUTATION TIME

| NN | HNN-MAX2SATBF | HNN-MAX2SATAIS | HNN-MAX3SATBF | HNN-MAX3SATAIS |
|---|---|---|---|---|
| *10* | 20 | 1 | 29 | 2 |
| *20* | 95 | 11 | 126 | 19 |
| *30* | 316 | 53 | 440 | 64 |
| *40* | 2645 | 242 | 3003 | 303 |
| *50* | 10835 | 498 | 12657 | 685 |
| *60* | 86102 | 1325 | 98368 | 1502 |
| *70* | - | 3049 | - | 3166 |

NN=Number of neurons.

Table V depicts the overall computation time recorded for our proposed algorithms, HNN-MAX*k*SATAIS and the conventional algorithm HNN-MAX*k*SATBF. Since we deal with MAX-2SAT and MAX-3SAT clauses, the training process consumes more time to minimize the logical inconsistencies than the randomized *k*-SAT problem. For instance, as the number of neurons increased, the computation time to generate the maximum number of clauses also increased. This is due to the fact that maximum *k*-satisfiability problem will never be fully satisfied, but we can possibly calculate the maximum number of clauses that will be satisfied. Hence, the states retrieved from the network can improve the global solutions that maximize the number of satisfied clauses. Thus, the whole process incurs more computation time. Generally, MAX-3SAT requires more time than MAX-2SAT due to complexity as the number of literals entrenched in the formula also higher.

According to Table V, HNN-MAX2SATAIS and HNN-MAX3SATAIS require less computation time compared to the other counterparts, HNN-MAX2SATBF and HNN-MAX3SATBF. The undoubted evidence beyond that results are due to more neurons being forced to jump the energy barrier to relax into global solutions during the training process [38]. Additionally, the training process by using brute force requires more computational time due to "generate and test" process in hunting the maximum number of satisfied assignments. On the contrary, when we applied AIS algorithms, the CPU time was reported faster due to the efficiency of the B-cells to improve towards the desired solution. Before the network was fed by the solution, B-cells with high and low affinity will be considered in finding the best B-cells [7]. Hence, HNN-MAX*k*SATAIS experienced less computation burden during the training processes as compared to HNN-MAX*k*SATBF.

## VII. CONCLUSION

A robust paradigm is developed by taking the advantages of a brand new immune inspired heuristic method, known as artificial immune system algorithm and the power of Hopfield network to solve MAX-*k*SAT problem. Hence, the performance comparison was made between our proposed paradigm with the conventional brute force method integrated with Hopfield neural network. In comparison of the algorithms, the HNN-MAX*k*SATAIS has shown more robustness to produce the global output and feasible solution compared to HNN-MAX*k*SATBF. The exploration showed solid performances of HNN-MAX*k*SATAIS in terms of the global minima ratio, ratio of satisfied clause, fitness landscape value and the computation time. In our future work, we would like to introduce our proposed algorithm to solve the other variants of satisfiability problems such as Weighted maximum satisfiability, quantified maximum satisfiability and unique satisfiability problem.

## Acknowledgment

The authors fully acknowledged Ministry of Higher Education (MOHE) and School of Mathematical Sciences, Universiti Sains Malaysia for the support which makes this important research viable and effective.

## References

[1] R. B. Xiao and L. Wang, "Artificial immune system: principle, models, analysis and perspectives", *CHINESE JOURNAL OF COMPUTERS-CHINESE EDITION*, vol. 25, no. 12, pp. 1281-1293, 2002.

[2] E. Hart and J. Timmis, "Application areas of AIS: The past, the present and the future", *Applied soft computing*, vol. 8, no. 1, pp. 191-201, 2008.

[3] D. Dasgupta, Z. Ji and F.A González, "Artificial immune system (AIS) research in the last five years" in *IEEE Congress on Evolutionary Computation,* vol. 1, pp. 123-130, 2003.

[4] J. D. Farmer, N.H. Packard and A. S. Perelson, "The immune system, adaptation, and machine learning", *Physica D: Nonlinear Phenomena*, vol. *22*, no. 1, pp. 187-204, 1986.

[5] J. Timmis and M. Neal, "A resource limited artificial immune system for data analysis", *Knowledge-Based Systems*, vol. 14, no. 3, pp. 121-130, 2001.

[6] N. Cruz-Cortés, D. Trejo-Pérez and C. A. C. Coello, "Handling constraints in global optimization using an artificial immune system", in *International Conference on Artificial Immune Systems*, pp. 234-247, 2005.

[7] L. N. de Castro, and J. Timmis, "Artificial immune systems: a novel paradigm to pattern recognition", *Artificial Neural networks in pattern Recognition*, no. 1, pp.67-84, 2002.

[8] A. Layeb and A. H. Deneche, "Multiple sequence alignment by immune artificial system", in 2007 IEEE/ACS International Conference on Computer Systems and Applications, pp. 336-342, 2007, May.

[9] O. Engin and A. Döyen, "A new approach to solve hybrid flow shop scheduling problems by artificial immune system", *Future generation computer systems*, vol. 20, no. 6, pp.1083-1095, 2004.

[10] A. Layeb, A. H. Deneche and S. Meshoul, "A new artificial immune system for solving the maximum satisfiability problem" in *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, pp. 136-142, 2010, June.

[11] P. Hansen and B. Jaumard, 1990. "Algorithms for the maximum satisfiability problem". *Computing*, *44*(4), pp.279-303.

[12] P. Crescenzi, and V. Kann, 1997, July. "Approximation on the web: A compendium of NP optimization problems". In *International Workshop on Randomization and Approximation Techniques in Computer Science* (pp. 111-118). Springer Berlin Heidelberg.

[13] A. Z. Broder, A. M. Frieze and E. Upfal, "On the satisfiability and maximum satisfiability of random 3-CNF formulas", in *SODA*, vol. 93, pp. 322-330, 1993, January.

[14] R. Rojas, Neural Networks: A Systematic Introduction. Berlin: Springer, 1996.

[15] J.J. Hopfield, D. W. Tank, "Neural computation of decisions in optimization problem", Biological Cybernatics, vol. 52, 141-152, 1985.

[16] G. Pinkas, "Symmetric neural networks and propositional logic satisfiability", *Neural Computation*, vol. 3, no. 2, pp.282-291, 1991.

[17] W. A. T. W. Abdullah, "Logic Programming on a Neural Network", *Malaysian Journal of Computer Science*, vol. 9, no. 1, 1-5, 1993.

[18] S. Sathasivam, "Upgrading Logic Programming in Hopfield Network, Sains Malaysiana", vol. 39, 115-118, 2010.

[19] W. A. T. W. Abdullah, "The logic of neural networks". *Physics Letters A*, vol. 176, no. 3, pp. 202-206, 1993.

[20] J. R. M. Fernández, J. D. L. C. B Vidal, "Improved Shape Parameter Estimation in K Clutter with Neural Networks and Deep Learning", *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 3, 2016.

[21] S. Sathasivam, W. A. T. W. Abdullah, "Flatness of the Energy Landscape for Horn Clauses", *SSJ*, vol 1, no. 2, pp. 2, 2008.

[22] U. Aiman, and N. Asrar, "Genetic Algorithm Based Solution to SAT-3 Problem", *Journal of Computer Sciences and Applications*, vol. 3, no. 2, pp. 33-39, 2015.

[23] D. Verma, N. Kakkar & N. Mehan, "Comparison of Brute-Force and KD Tree Algorithm", *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 3, no. 1, pp. 5291-5294, 2014.

[24] R. A. Abdeen, "An Algorithm for String Searching Based on Brute-Force Algorithm", *International Journal of Computer Science and Network Security*, vol. 11, no. 7, pp. 24-27, 2011.

[25] I. Zinovik, D. Kroening, and Y. Chebiryak, "Computing binary combinatorial gray codes via exhaustive search with SAT solvers", *Information Theory, IEEE Transactions,* vol. 54, no. 4, pp. 1819-1823, 2008.

[26] K. Haddouch, K. Elmoutaoukil, and M. Ettaouil, "Solving the Weighted Constraint Satisfaction Problems Via the Neural Network Approach", *International Journal of Interactive Multimedia and Artificial Inteligence*, *4* (Special Issue on Artificial Intelligence Underpinning), 2016.

[27] B. Borchers and J. Furman, "A two-phase exact algorithm for MAX-SAT and weighted MAX-SAT problems", *Journal of Combinatorial Optimization*, vol. 2, no. 4, pp. 299-306, 1998.

[28] R. A. Kowalski, "The early years of logic programming", Communications of the ACM, vol. 31, no. 1, pp. 38-43, 1988.

[29] A. Imada and K. Araki, "What does the landscape of a Hopfield associative memory look like?" in *Evolutionary Programming VII*, Springer Berlin Heidelberg, pp. 647-65, 1998.

[30] S. Haykin, *Neural Networks: A Comprehensive Foundation*, New York: Macmillan College Publishing, 1999.

[31] M. Velavan, Z. R. Yahya, M. N. A. Halif, & S. Sathasivam, "Mean Field Theory in Doing Logic Programming Using Hopfield Network", *Modern Applied Science*, vol. 10, no. 1, p. 154, 2015.

[32] E. Elbeltagi, T. Hegazy and D. Grierson, Comparison among five evolutionary-based optimization algorithms. *Advanced engineering informatics*, vol. 19, no. 1, pp. 43-53, 2005

[33] B. A. Madsen and P. Rossmanith, "Maximum exact satisfiability: NP-completeness proofs and exact algorithms", BRICS Report Series, vol. 11, no.19, 2004.

[34] B. Borchers and J. Furman, "A two-phase exact algorithm for MAX-SAT and weighted MAX-SAT problems", *Journal of Combinatorial Optimizatio*n, vol. 2, no. 4, pp. 299-306, 1998.

[35] A. Baldominos Gómez, N. Luis Mingueza, and M. C. García del Pozo, "OpinAIS: An Artificial Immune System-based Framework for Opinion Mining", *International Journal of Interactive Multimedia and Artificial Inteligence, 2015*.

[36] V. Raman, B. Ravikumar and S. S. Rao, "A simplified NP-complete MAXSAT problem", *Information Processing Letters*, vol. 65, no. 1, pp. 1-6, 1998.

[37] L. Bordeaux, Y. Hamadi, and L. Zhang, "Propositional satisfiability and constraint programming: A comparative survey", *ACM Computing Surveys (CSUR)*, vol. 38, no. 4, p.12, 2006.

[38] S. Sathasivam, P.F. Ng, N. Hamadneh, "Developing agent based modelling for reverse analysis method", *Research Journal of Applied Sciences, Engineering and Technology*, vol. 6, no. 22, pp. 4281-4288, 2013.

[39] X. G. Ming and K. L. Mak. "A hybrid Hopfield network-genetic algorithm approach to optimal process plan selection." *International Journal of Production Research* 38, no. 8, pp. 1823-1839, 2000.

[40] U. P. Wen, K. M. Lan & H. S. Shih, "A review of Hopfield neural networks for solving mathematical programming problems", *European Journal of Operational Research*, vol. 198, no. 3, pp. 675-687, 2009.

[41] L. M. Ionescu, A. G. Mazare, & G. Serban, "VLSI Implementation of an associative addressable memory based on Hopfield network model", *IEEE Semiconductor Conference*, vol. 2, pp. 499-502, 2010.

[42] D. E. Goldberg, and K. Deb, "A comparative analysis of selection schemes used in genetic algorithms.", *Foundations of genetic algorithms*, vol. 1, pp. 69-93, 1991.

[43] V. B. Semwal, K. Mondal, and G. C. Nandi, "Robust and accurate feature selection for humanoid push recovery and classification: deep learning approach", *Neural Computing and Applications*, pp. 1-10, 2015.

[44] J. Singha, and R. H. Laskar, "Hand gesture recognition using two-level speed normalization, feature selection and classifier fusion", *Multimedia Systems*, pp.1-16, 2016.

[45] P. Kumari, and A. Vaish, "Feature-level fusion of mental task's brain signal for an efficient identification system", *Neural Computing and Applications*, vol. 27, no. 3, pp.659-669, 2016.

**Mohd Asyraf Mansor** obtained his MSc (2014) and BSc(Ed) (2013) from Universiti Sains Malaysia. He is currently pursuing Ph.D degree at School of Mathematical Science, Universiti Sains Malaysia. His current research interests include evolutionary algorithm, satisfiability problem, neural networks, logic programming and heuristic method especially Artificial Immune System.

**Mohd Shareduwan bin M. Kasihmuddin** received his MSc (2014) and BSc(Ed) (2013) from Universiti Sains Malaysia. He is currently pursuing Ph.D degree with School of Mathematical Science, Universiti Sains Malaysia Penang Malaysia. His current research interests include neuro-heuristic method especially Artificial Bee Colony algorithm, constrained optimization problem, neural network and logic programming.

**Saratha Sathasivam** is a lecturer in the School of Mathematical Sciences, Universiti Sains Malaysia. She received her MSc and BSc(Ed) from Universiti Sains Malaysia. She received her Ph.D at Universiti Malaya, Malaysia. Her current research interest are neural networks, agent based modeling and constrained optimization problem.

# Exploring the Relevance of Search Engines: An Overview of Google as a Case Study

Ricardo Beltrán-Alfonso. Andres Torres-Tautiva. Paulo Alonso Gaona-Garcia. Carlos Enrique Montenegro-Marin

*Universidad Distrital Francisco José de Caldas, Bogotá, Colombia*

***Abstract* — The huge amount of data on the Internet and the diverse list of strategies used to try to link this information with relevant searches through Linked Data have generated a revolution in data treatment and its representation. Nevertheless, the conventional search engines like Google are kept as strategies with good reception to do search processes. The following article presents a study of the development and evolution of search engines, more specifically, to analyze the relevance of findings based on the number of results displayed in paging systems with Google as a case study. Finally, it is intended to contribute to indexing criteria in search results, based on an approach to Semantic Web as a stage in the evolution of the Web.**

***Keywords* — Browsers, Google, Internet, Search, Algorithms.**

## I. Introduction

The constant growing of web-related services and the amount of data Internet handles, bring a new challenge for all information users associated to its veracity and quality. However, it is not an easy task, considering that the contents on Internet increase each second regardless their use or efficacy [1], which requires a deep and clear analysis about the quality of the search results.

Internet, and its fast evolution during years, has become a transforming factor for social, political, cultural, and especially scientific structures [1]. Also, this evolution has transformed the way in which the information has been stored and distributed by Internet, regarding internet significant size that hinders finding relevant results. It makes necessary structuring an architecture based in the Semantic Web that allows a more efficient search, analyzing it according to its results [2].

There are precursors in the field of quality assessment in search results on Internet, like [3, 4] and [5]; authors who have published results from their investigations about Web pages' evaluation. Among the most representative instruments, they highlight the use of control lists and propose a list of useful indicators to evaluate the quality of the digital information. This proves that the necessity to evaluate the information is not a new subject but it is a prevalent topic in the implementation of the Semantic Web: a relevant aspect for the adoption of strategies that allow an increased use inside the field of agents and multi-agent systems about search processes [2].

The purpose of this study is to analyze the relevance of the results obtained in search engines on Internet regarding the real content of each indexed resource. For that, the search engine Google will be used as a case study. This study has identified the following considerations in order to obtain results according subjective impressions of searches by students based on the use of keywords. Therefore, the study has focused on the following conditions: 1) tests with undergraduate students located in Bogotá - Colombia, 2) participants aged 18 to 25 years and 3) the use of keywords according to the frequency of keywords identified by the Colombian region. Based on these considerations, the study aims to analyze the relevance of indexing results according to keywords most frequently in undergraduate students within the region of Bogota - Colombia. To carry out this study we develop an algorithm that allows scrolling through the results of searches obtained by Google to identify the way in which these results are indexed, and validation criteria of the obtained results. Finally, it is intended to analyze these results with the aim of proposing methods that will improve the indexing of search results based on strategies and initiatives defined by the Semantic Web, and from these approaches generate strategies of linking data through Linked Data [7] and intelligent agents.

This document is structured as follows: In section 2, all the key aspects associated with resource search, search engines and valid methods of analysis for our study will be addressed. Section 3 shows the implementation of the algorithm necessary to scroll through the results detailing its use and showing a guide to apply it in other related researches. In section 4, preliminary results analysis that the algorithm implementation produces will be presented making punctual emphasis in the actual results vs. the shown results. Section 5, we present discussion of our study. As a final point, conclusions and the future applications of this research are presented.

## II. Background

Visual Analytics is a field that arises from the visualization of information and its practice has as a target to support the analytic reasoning through interactive visual interfaces [8]. Therefore, Visual Analytics combines techniques of automated analysis with interactive visualizations for effective understanding, reasoning and decision-making over data sets that are too big and complex [9]. Meanwhile [10] defines this analysis as the science of analytic reasoning facilitated by interactive visual interfaces. In summary, visual analytics is a tool preceding automated research; it is a tool that, with aid of interactive visual interfaces, allows us to have a deep analysis, providing information for reasoning and decision-making.

### *A. The relevance of visual analytics*

Visual analytics includes in its work area several scientific branches such as scientific & information visualization, data mining, data management, human computer interaction, and perception cognition, that work inside its scope. For example, the integration of scientific visualization and information include works in the fields of space-time data, data analysis and management, and human perception and cognition [9].

Web analytics arises as the analysis and presentation of data gathered from Internet with the purpose of helping the company in the management and optimization of its digital strategy [8]. It is necessary to mention that visual analytics would focus on Web analytics, which we can say it is centered in gathering info given by the users after visiting a website, and the further analysis of the obtained data [9].



DOI: 10.9781/ijimai.2017.449

Meanwhile, it may be set that the Semantic Web has as main purpose to lay out the information in such a way that it can be accessible by any user regardless his level of experience [2], and therefore solve the usual problems of information search. Even now that the Semantic Web has not as a purpose to become a reality and a primordial approach anymore, and spite of all the several technological efforts, it can only be partially used, whence intellectual and technological resources have to be used in a way that may meet these needs and contributes to build the Web with meaning by automating tasks like information search.

### B. Contents search

With the huge growth of the Web, the availability of electronic academic resources has also grown. Nowadays, the users face the need of finding more relevant academic resources for their necessities and their personal characteristics and cultural aspects based on suitable visual search interfaces [11-13]; nonetheless, locating resources in an infrastructure like Internet, that does not stop growing, is a complex activity. For this reason, there are strategies proposed to optimize location processes. One of the most common strategies is the use of a search engine to classify web files that contain interesting material. The most known and useful in terms of the precision in the last years is Google [14], the most popular search engine in the world, even though most of them work in the same way, and only differ in classification and relevance criteria [15]. Google was designed with the purpose of giving priority to the amount of visits a page has, and in this way it assigns priority to each page [16]. This model is known as PageRank [17, 18].

When a search through Google is made, the engine makes a classification and then shows a fragment of each page (or snippet), but since there is so much information, there are occasions when the desired info is not found (generally associated to a bad search). This is the reason for which Google offers several tools to make more effective searches. Among them Advanced Search, the use of symbols like quote marks or "site", and the versions of Google allow classification in a zone manner or by specific territories [19].

### C. Related work

The study of information metric sciences has been addressed by different authors, based on the study subject of the informative disciplines. There are different areas such as focused crawling [20], distributed crawling [21], extraction based on Domain Ontology [22], based on PageRank [18, 23], among other described in [24]; were we can extract information from Internet. Mathematics, applied in different ambits in social sciences, allows problem resolution and contributes to the development of mathematical models.

There are also the bibliometric and scientometric methods that are very similar, and sometimes called identical. These allow the study of growth, size and organization of the information, allowing planning in the organization of the scientific documentation, etc. They are based on the use of statistical models [25]. Once demonstrated the effort made to conduct a search in terms of quality and optimization of resources, it is established that the amount of irrelevant responses is proportional to the high amount of information circulating in the web of webs; also the use of additional indicators to improve the search or decreasing the irrelevant results is not enough. Authors like [26] propose a weighted architecture based on the number of clicks or [27] from combinational searches and listing the results in a metric manner, it is both a promising and applicable solution.

Another study, conducted by [28] proposes methods of dynamic data and also includes the sensible testing of changes in the website, and the dynamic recovery of the linked web pages. Also, authors like [20] implemented a tracking application based on Java, and its performance, in comparison to the conventional static approaches, was relatively good. Among the outstanding results of the experiment, it was shown a 59% improvement in the performance with the static tracking method.

In the study conducted by [29], it is addressed the question of how to build a practical system at great scale that can use the information in the hypertext. This gives an approximation of the importance and quality of a page certified by PageRank. Another study conducted by [30], proposes a series of heuristic criteria to identify trackers in the access registry of web servers. Later, the impact of the requests from robots in web caches was assessed and the behavior of the trackers was compared with the automatic buyers 'one. In the context of their work, the authors studied the time distribution between arrivals of crawler petitions.

There is another crawler known as Nutch, which is applied as a tracking mechanism in charge of analyzing documents published in Intranet [16]. This web crawler is free and with an open code, where the description of its management and configuration is found, and which is made through the console mode of the Linux operative system.

Meanwhile, a combination between the information recovery agents with the visual analysis is the work made by [31] through the development of a tool called WebTheme. This theme helps the users to understand rapidly big amounts of information of the web and to deepen into the interesting section of such information. Several Crawlers have already been recognized, as is the case of UbiCrawler [32], a tool made in Java, under an independent architecture and lineal scalability through the use of functions based in hash operations to divide the tracking domain.

Based on these referents, the next section presents the proposed methodology of the study to present a systematic analysis of search results given by Google as a case study based on a visual analysis tool.

## III. Work Methodology

With the objective of building the Crawler for the analysis, a series of stages is proposed where the structure and methodology of the work to be done are shown. In Fig. 1, a diagram with the methodology, stages, and activities to carry out is shown.

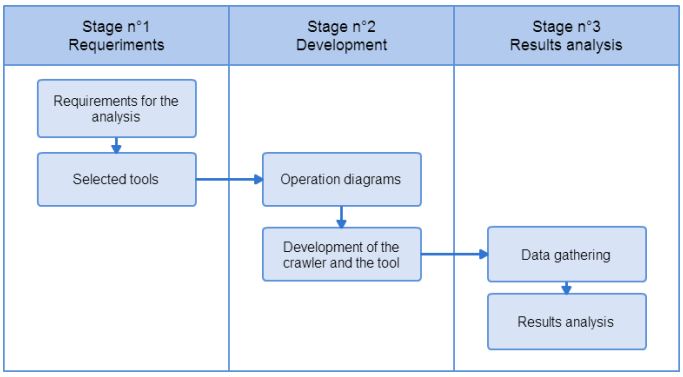



Fig. 1. Work methodology (own elaboration).

### A. Stage 1: requirements

This type of tool is used with multiple purposes, one of them described by [30] where a growth of the number and variety of active robots operating in the web are presented, and the impacts on the traffic on Internet and the activity in the web servers are analyzed. This condition is necessary to establish the requirements and limits of the Crawler to develop. Next, the requirements for both types are described.

### 1) Requirements for the analysis

Given that the proposed algorithm is a focused Crawler, it is necessary first to consider the page in which the analysis will be conducted, since it will do the search and show the results [30].

Since the objective of this study is to conduct an analysis of the results displayed in the Google search engine, the first thing to be analyzed is how Google conducts a search and how it changes the URL to move through the result pages. This activity is verified making a search in the page and observing the URL change.

As a result of this process, Google makes the search based on the URL https://www.google.com/search?q=”word_to_look_for” and advances through the pages adding &start=”number_multiple_of_10_based_on_which_results_are_shown”, to include all the results, “&filter=0” is added. It is worth knowing that Google shows the results in groups of 10 and it does it with the header h3 with the attribute class=r.

In the same way, it is important to stand out the outputs of the algorithm to write for its development, whereby it is defined that the algorithm must show the number of results indicated by Google and the number of actual results displayed given that “Google is designed to escalate” [29], so it optimizes the results shown and it is intended to analyze this action. Additionally, it must show the pages, it went through to get coincidences among them. Therefore, the methodology to extract information is based on the work made by [32] in terms of data extraction and their type.

### 2) Selected tools

The algorithm was developed in Java programming language according to the recommendations made by [28]. For the open development environment, NetBeans was used, which allows the construction of the algorithm using the Jsoup library, a library that allows working with HTML documents to analyze them and extract information from them. Likewise, for the complete tool JFreeChart library will be implemented, a library in charge of graphic management.

## B. Stage 2: development and elaboration of the algorithm

In this stage a Crawler will be constructed for data gathering and the complete tool for the analysis. For that, a diagram of operation and the necessary methods for the development are defined considering the architecture of a query on Internet proposed by [33].

### 1) Operation diagram

The operation of the algorithm is shown in Fig. 2 where the tool connects to Internet, specifically to the Google search engine and then it shows a series of results for further analysis.

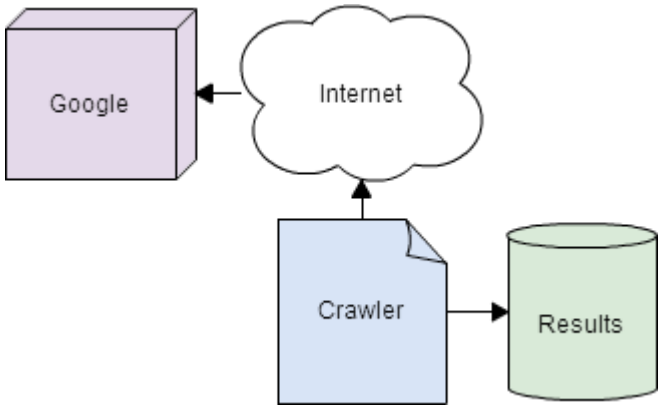



Fig. 2. Algorithm operation (own elaboration).

### 2) Development of the crawler

The Crawler is in charge of going through the generated results, creating a list of titles, a list of links, and a list of descriptions, all generated by the search engine. The implemented Crawler has two principal methods:

The first and most important method in the Crawler is the search method. This method has the task of conducting a search in a HTML document, filtering the results and setting them apart from the total amount of links in each page. This will generate relevant data for the research. The code of this method is shown in Fig. 3.

```
public void search(String s) throws IOException {//search desired items in a html document
    doc = Jsoup.connect(s + "&filter=0").userAgent("Mozilla").
            ignoreHttpErrors(false).timeout(0).get();

    Elements title = doc.select("h3[class=r]");
    Elements links = doc.getElementsByClass("s");

    for (Element link : title) {//write titles for search
        write = new FileWriter(titles, true);
        write.write(link.text() + "\r\n");
        write.close();
    }
    for (Element link : links) {//write titles for search
        write = new FileWriter(this.links, true);
        write.write(link.getElementsByTag("cite").text() + "\r\n");
        write.close();
    }
    for (Element link : links) {//write description for search
        write = new FileWriter(description, true);
        write.write(link.getElementsByTag("span").text() + "\r\n");
        write.close();
    }
    results += title.size();
    if (links.size() < 3) {
        flag = false;
    }
}
```

Fig. 3. Method that searches the results of a page (own elaboration).

The search method performs a connection with an URL using the Jsoup library, and then stores all the HTML files in the document doc. To do this connection, it is necessary to identify itself as a browser, in this case Mozilla in order that the engine will allow the connection and the page scanning. After having the document stored, the extraction and filtering of relevant data for research continue. In this case the titles, the links, and the descriptions will be the relevant data. For this filters are made by class or identifiers. These filters can be made through methods included in the Jsoup library. Next, data is stored in text document for a later analysis.

The second most important method is in charge of going through the pages and sending to the first method each page so it can conduct the link search. This method is observed in Fig. 5, where, by means of a for cycle, the generated page is changing by displaying ten results each time. The method implementation is shown in Fig. 4.

```
public void results() {//stores the number of results reported
    totalResults = doc.getElementById("resultStats").text();
}

public void follow(String s) {//following pages
    for (int i = 10; i < 2000; i = i + 10) {
        try {
            if (flag == true) {
                search(s + "&start=" + i);
            } else {
                break;
            }
        } catch (IOException ex) {
            Logger.getLogger(CrawlerGoogle.class.getName()).
                    log(Level.SEVERE, null, ex);
        }
    }
}
```

Fig. 4. Method to browse the pages (own elaboration).

The results method conducts a search inside the document doc looking for the results displayed by Google. On the other hand, the follow method has the task of going through the result pages. It is

important to point out that the method advances from 10 to 10 since this is the number of results by page that Google shows.

### 3) Final architecture of the crawler

The final architecture of the Crawler is shown in Fig. 5, where the Crawler class implemented by the library Jsoup is seen. This class executes a search in the Google search engine, goes through the generated pages, lists the number of results and generates a file with the paths for a later analysis. This type of architecture is proposed in studies made by [33,34], showing good results.

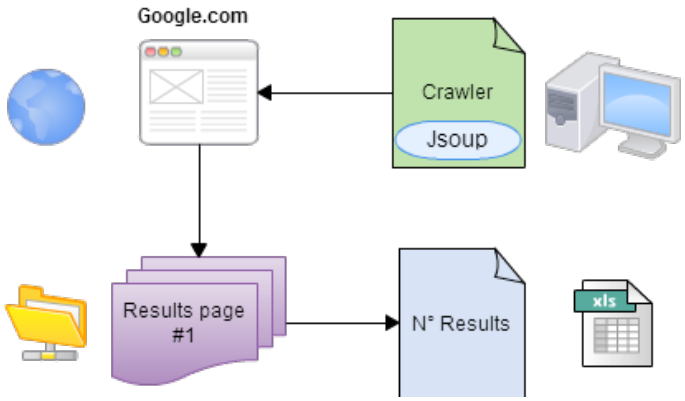



Fig. 5. Architecture of the Crawler (own elaboration).

### 4) Graphic tool

For the implementation and visualization of the results, in the same Crawler project a library called JFreeChart is implemented, which allows to visualize the results by means of graphics. In Fig. 6, there is a diagram showing its operation.

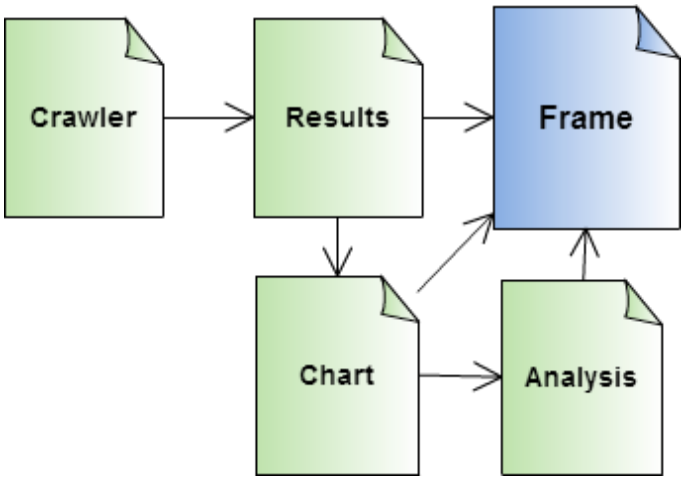



Fig. 6. Operation of the graphic tool (own elaboration).

Based on the Crawler, the results are generated in text documents that become graphics in the Chart class and are then visualized in the Frame class. This first part only includes the number of actual results vs. the number of results given by the search engine. Then, an analysis is made in its respective class where there are identified key axis, and a deep analysis of the results of the search, and an analysis of the related results.

## IV. Results

To show the behavior of the results of the searches, two phases of analysis were defined: 1) analysis of the effectiveness of the tool, for which related aspects with the number of obtained results are analyzed, as the number of unique results, and 2) objective analysis based on the obtained results with users through a definition of a study. Controlled experiments are carried out under controlled conditions, not necessarily in a real context, but in a "laboratory" or "experimental" setting. Participants were 50 undergraduate students between 18 to 25 years old and with a good level of English. Given the nature of the study, it was not necessary that users exhibit extremely specific knowledge of subject search, and neither strategy of advanced searches.

### A. Analysis of the effectiveness of the tool

This study includes keywords of common usage defined through a previous study where, based on a list of candidate words chosen at random, words whose meaning can be given in a conceptual manner, that can show different types of results like images, videos or reviews were filtered. The last criterion to select the keywords is the number of results given by the search engine. With these criteria, it is desired that the amount of possible results will be variable; this with the goal of showing the results for different volumes of information.

For the analysis of results, there were proposed five queries associated to knowledge fields of general culture, through the use of different terms, which were: fossil, museum, stadiums, London, and Rome. Based on these terms, in first instance the number of actual displayed results vs. the number of results displayed by the search engine was analyzed and then the traveled web pages were listed to analyze them. To be able to do the necessary data collection, the developed Crawler is used from which two initial data were obtained.

Table I show the results of two data types obtained. The first column corresponds to the number of related results displayed by the search engine. That is, the number of resources that the search engine has classified as related even when these cannot be verified in its totality. The second column corresponds to real number of results displayed by the search engine, that is, the real number of results related to the keywords of search.

TABLE I.
Words preliminary analysis

| Word | Results given by the Google search engine | Real results displayed based on Crawler |
|---|---|---|
| Fossil | 2,160,000 | 830 |
| Stadiums | 11,000,000 | 710 |
| London | 163,000,000 | 660 |
| Museum | 211,000,000 | 950 |
| Rome | 583,000,000 | 600 |

As it is presented in table I, the results displayed by the search engine surpass by great numbers the real results of available resources that actually have a relation with the search made. For example, for the case of the Fossil, from the 2,160,000 results given by the search engine, there are only 830 real resources. This same process was made with several iterations per word (20), and the results, in average, were associated with the ones indicated in Table I.

### B. Case study

To complement these results, a case study was conducted with the purpose of: 1) analyzing the behavior of undergraduate students at the moment of conducting a search process and 2) collating the results from the algorithm made in the Crawler with results obtained by people with experience in processes of information search on Internet. In summary, there were 5 proposed scenarios with different searches, and each scenario was evaluated with ten (10) different participants, that add up a total of 50 participants for the whole test.

For the selection of participants, the recommendations of applications tests proposed by [35] were considered where the selection

of participants is made in a random manner and the results are evaluated in a general way given that the subjects have an habitual character and are non-specific to a field, that is, a group of users is defined in relation to a documental analysis of a topic as it is done in [36].

For the case study, the participants were asked for looking for information related to a knowledge field in several scenarios like:

1. The first case consisted in a search where the user would simulate being a primary school teacher and his experience was required to teach a class about mammal animals.
2. In the second case, the participant assumed the role of a student and he was asked to conduct a study about the first computer in the history.
3. In the third case a user was defined as a young person with the desire of finding the most adequate definition of the word "computing".
4. A participant was defined with interest in knowing the benefits of a four-step engine for a motorbike.
5. Finally, the last case is presented to the user as a participant of an informal discussion where it is debated the World Ranking of best National Football Teams, for which it is necessary to consult Internet and have the best available information.

The data was gathered through an Internet survey performed to undergraduate students from different knowledge fields where the specific search case and a series of questions associated with daily situations were set where a necessity to make a low complexity search was presented. After conducting the search, they were asked to fill a form where the user provides information related to the obtainment of selected information. In this way, the participants show what their preferences are when they conduct a search and in which way the classification of relevance or veracity of the resources found is made.

The questionnaire starts by asking what was the search engine used for the research. In this way the relevance of this article was assessed when the Google search engine is used. The questionnaire continues with the query section, where it is established whether the user is using filters, the number of used keywords, and the time employed for the query. The last section of the questionnaire is related to the result assessment.

This section allows gathering information about the number of resources, the examined pages, source verification, and responsibility delegation to the search engine. A resume of volume dispersion for the first case is present in Fig. 7.

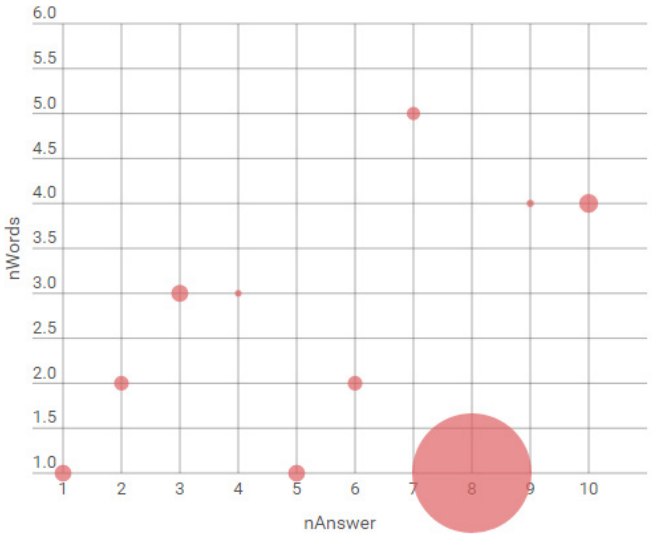



Fig. 7. Results volume dispersion for the first case (own elaboration).

Fig. 7 presents the results obtained for the first case study. In this occasion, there are responses from ten participants where a variation of the number of used keywords and the volume of results displayed by the search engine is shown. It is important to point out that each participant conducts a search with the same number of words, but with semantic differences that are not accounted for the study. Given that fig. 7 shows an unstable behavior, it is proposed to make a correlation statistical analysis in order to establish whether or not there is a relation between the number of used keywords and the dimension of the results.

In order to made the study a statistical hypothesis is proposed as follows.

H1 = There is a relation between the number of keywords used by a user and the number of results displayed by the search engine.

Nonfulfillment of the previous hypothesis is taken as a null hypothesis, having an allowed error of 5%.

Based on this hypothesis, we perform a Pearson correlation analysis for the first case as presented in Table II.

TABLE II.
PEARSON CORRELATION ANALYSIS FOR THE FIRST CASE

| Correlation | | | |
|---|---|---|---|
| | | **nWords** | **nResults** |
| nWords | Pearson Correlation | 1 | -0,399 |
| | Sig. (bilateral) | | 0,254 |
| | N | 10 | 10 |
| nResults | Pearson Correlation | -0,399 | 1 |
| | Sig. (bilateral) | 0,254 | |
| | N | 10 | 10 |

In the first proposed case study, there is a Pearson correlation index R of -0.399 with a significance value of 0.254, higher than the one set for 5% error. Having into account a negative value, the hypothesis which indicates that there is no relation between the number of keywords used by a user and the magnitude of results given by the search engine for the first case is rejected. From the results presented for case 2 (R = -0.01; SD= -0.997), case 3 (R = -0.212; SD= -0.613), case 4 (R = -0.465; SD= -0.176) and case 5 (R = -0.289; SD= -0.418), it can be concluded that the number of results given by the search engine is not associated to the number of keywords.

To go a little deeper into the user's behavior regarding the number of used keywords, a Kolmogorov Normality Analysis may be established given that there is a number of data over 30. Table III presents these results.

TABLE III.
USER'S BEHAVIOR REGARDING THE NUMBER OF WORDS

| Normality Analysis | | | | |
|---|---|---|---|---|
| | | **Kolmogorov-Smirnov** | | |
| | **nWords** | **Statistical** | **gl** | **Sig.** |
| nResults | 1,0 | 0,354 | 7 | 0,008 |
| | 2,0 | 0,399 | 10 | 0,000 |
| | 3,0 | 0,459 | 17 | 0,000 |
| | 4,0 | 0,169 | 9 | 0,200 |
| | 5,0 | 0,368 | 4 | 0,000 |

As it is presented in table III, the data is not in a normal distribution given that their significance values are under 5%. Nevertheless, in this case study it is possible to establish the users' preference regarding the use of keywords, as it is presented in Fig. 8.

Fig. 8 presents, in percentage, that in over 30% of the cases, the participants used three words for a search, and approximately 75% of the participants used from two to four words to conduct a search.

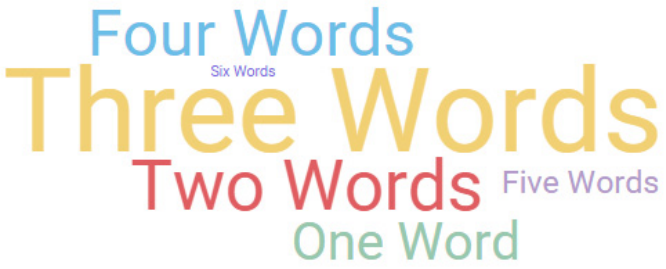



Fig. 8. Use of keywords by users (own elaboration).

### C. *Analysis of the participants' behavior*

The results of the case study give relevant information with respect to the number of visited pages, the number of consulted results and other user's behaviors at the moment of conducting a search. These results are shown in Fig. 9a,9b and 10.

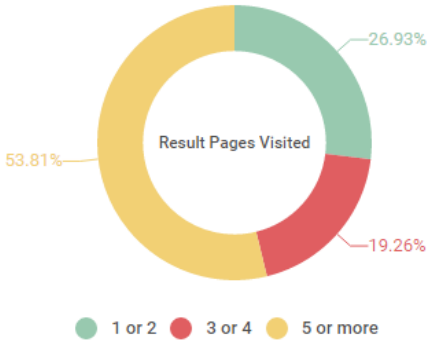



Fig. 9a. Analysis of results pages visited by participants (own elaboration).

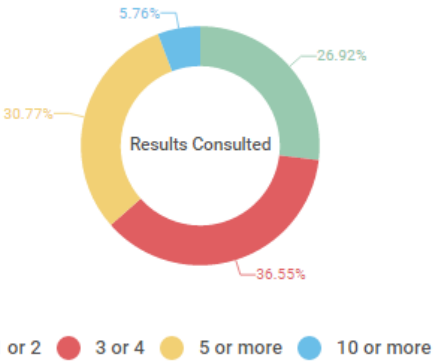



Fig. 9b. Analysis of pages consulted by participants (own elaboration).

Fig. 9a and 9b, shows two initial graphics about the behavior of the participant regarding the number of average results that are evaluated to find the desired information. The first graphic shows the amount of pages that are visited in the information search. In it, it is noted that a little over half of the people (53.81%) navigate between 5 and 10 pages to complete the solicited information; while the other half is divided between the first and fourth consulted page. With these results it can be deduced (for the particular case study) that the ideal number of pages to find valid information is near to five.

The second graphic shows the information sources consulted and evaluated by the user, and it is noted that the number of consulted results is less than 10 sources (5.76%), while the ideal number of sources in average for the search process is (36.55%) for most participants. Fig. 10 shows, in a classified manner, the behavior of the participants when doing the search process.


Fig. 10. Analysis of pages consulted by participants (own elaboration).

In the conducted study, it can be seen that only 2% of people use filter to conduct a search, which shows that the task of filtering the information is completely relied on the search engine. Another important conclusion is that 7 of each 10 participants (70%) found at least one source with an attractive title but with irrelevant content or not related to the research topic. Also, 55% of people consulted only well-known or previously consulted sources, so, their initial criteria was based on their own experience. These sources, like Wikipedia, YouTube, or WordReference are consulted by over half of the people without previously verifying their references or their pertinence to the query made. In the same proportion, 55% of people repeat the search specifying the query with greater detail, adding keywords, dates or type of resource.

### D. *Analysis of the results obtained by the tool*

Associated to the tool, from the obtained results by the Crawler different conclusions may be obtained. First, there is a minimal existent relationship between the displayed data and the actual data displayed after a search process. Among actual displayed data it may be classified by the type of displayed source and their respective quantities. This data is shown in Fig. 11. From the actual displayed results, 55% of the titles are directly related with the search keyword, while 20% are linked to related searches or similar ones. With the same percentage, there are results shown from a same page (Wikipedia, in this case). Lastly, 5% of the results are not related since they are advertisement or others.

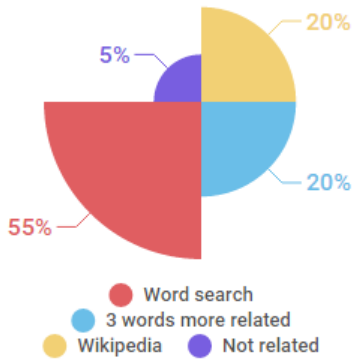



Fig. 11. Analysis of consulted sources (own elaboration).

Among the results displayed by the search engine and recovered by the Crawler, data can also be studied regarding the contents of the displayed sources. This data is shown in Fig. 12, where the first graphic shows that 60% of the displayed results are content generated after the year 2000 and half of the results are from 2015 (30%). This

could indicate that the Ranking in the search engine privileges the more recent and consulted results. In the next graphic in the same figure, it is shown the types of generated content: in 39% of the cases, there are shown pages from private companies (.com) followed by pages with domain of the country where the search is conducted. In this case: Colombia (.co) with a 50%.

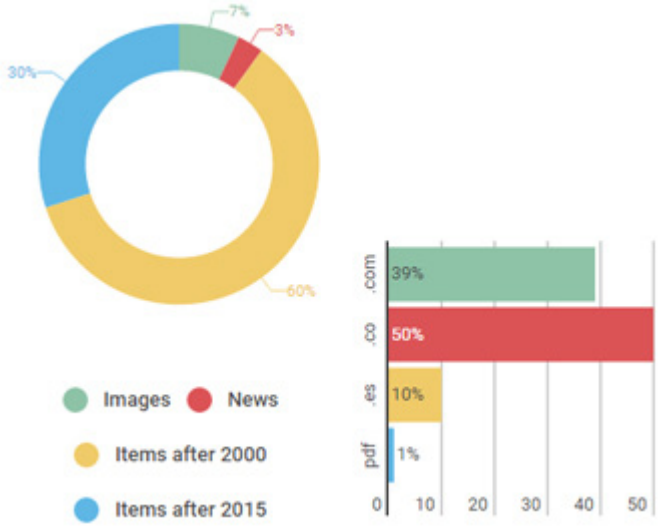



Fig. 12. Analysis of the types of consulted resources (own elaboration).

## V. Discussions

It is important to perform a previous evaluation of all the tools that will be used to provide the resources and their qualities in an appropriate manner. That being said, it is important to point out that the used search engine blocks periodically these tracking algorithms, and so the final tool needs waiting times to be employed in an experimental manner, but its results fulfill the objective of the research and provide enough bases for the next related studies.

Besides these aspects, it can also be said that the classification of resources the user does to prioritize a result, is based in his experience with the consulted page and he fully trusts on the ranking presented by the search engine, so much that the results supposed as trustable ones are not superior to five and the rigidness of the query is minimal, in barely the first five pages of displayed results.

In the results analysis a great volume of results may be identified in the displayed resources that is not relevant to the solicited information by participants in a search process. Therefore, the construction of tools that allow to catalog and explore information in the Web is essential and fundamental in the process of analysis of great volumes of data. For that reason, a good classification of the information is fundamental for the development based on Semantic Web for the linking of metadata [37, 38], that allows a classification of digital resources in a way that optimizes the search processes.

## VI. Conclusions

The implementation of a Crawler type tool to systematically browse over the web allows a fast classification of the results in a way that can be contrasted with a desired number of results to make an evaluation of a search. In this way relevant results are obtained like the lack of organization and distribution of the information in the Web.

The exhibited case study about the behavior of the users that conduct Internet searches was fundamental in the results of this study, since it permits the modeling of the behavior of the participants regarding their preferences or practical modes to conduct a search. According to these results, different aspects of behavior are evident like: i) the high grade of detachment when confirming the sources of the resources, ii) the minimum use of elements to intensify the searches like filters or related words, and iii) the unconditional acceptation of keywords suggested by the search engine.

As future work, it is intended to propose strategies that integrate agent systems or adaptive multi-agents, equipped with an initial experience defined by keywords or characteristics categorization [39], and needed by conventional users typified. In this strategy it is necessary to carry out a characterization of the user types based on the search needs and supply an agent system with criteria obtained based on the quality and quantity of given keywords. In the same way, it is proposed as a guide the use of open educational resources to conduct linking processes through technologies like Linked Data [7]. This way, the inter-operability of digital resources through semantic models that classify them may be used based on related knowledge fields, using information visualization and classification methods like the ones defined by [12]. For this, the use and classification of educative repositories by related topics is proposed, so that a web of repositories with academic content and specialized according to the needs of users using tools previously mentioned like Linked Data [40], categorization or agent systems, will be available to implant new methods of search based on specialized search engines.

## Acknowledgment

The authors gratefully acknowledged to Universidad Distrital Francisco José de Caldas trough GIIRA Research group.

## References

[1] N. Askitas and K. F. Zimmermann, “The internet as a data source for advancement in social sciences,” Int. J. Manpow., vol. 36, no. 1, pp. 2–12, 2015. doi: 10.1108/IJM-02-2015-0029

[2] G. Stumme, A. Hotho, and B. Berendt, “Semantic Web Mining. State of the art and future directions,” Web Semant., vol. 4, no. 2, pp. 124–143, 2006. doi:10.1016/j.websem.2006.02.001

[3] J. Zhao, X. Lu, X. Wang, and Z. Ma, "Web Information Credibility: From Web 1.0 to Web 2.0." International Journal of u-and e-Service, Science and Technology, vol. 8, pp. 161-172, 2015. doi: http://dx.doi.org/10.14257/ijunesst.2015.8.7.16

[4] J.W. Fritch, "Heuristics, tools, and systems for evaluating Internet information: helping users assess a tangled Web". Online Information Review, vol. 27, pp. 321-327, 2003. doi:10.1108/14684520310502270.

[5] B. Frazier, "Niche Search Engines: Expanding Information Discovery." The Reference Librarian, vol. 54, pp. 168-174, 2013. doi:10.1080/02763877.2013.755440

[6] M. Suárez-Barón and K. Salinas-Valencia, “An approach to semantic indexing and information retrieval,” Rev. Fac. Ing. Univ. Antioquia, no. 48, pp. 174–187, 2009.

[7] C. Bizer, T. Heath, and T. Berners-Lee, “Linked data-the story so far,” Semant. Serv. Interoperability Web Application Emerging Concepts, pp. 205–227, 2009. doi:10.4018/jswis.2009081901

[8] N. Andrienko and G. Andrienko, “Visual analytics of movement: An overview of methods, tools and procedures,” Information Visualization., vol. 12, no. 1, pp. 3–24, 2013. doi:10.1177/1473871612457601

[9] D. Keim, G. Andrienko, J.D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon, “Visual analytics: Definition, process, and challenges,” Lect. Notes Comput. Science, vol. 4950 LNCS, pp. 154–175, 2008. doi:10.1007/978-3-540-70956-5_7

[10] J. Kohlhammer, D. Keim, M. Pohl, G. Santucci, and G. Andrienko, “Solving problems with visual analytics,” in Procedia Computer Science, 2011, vol. 7, pp. 117–120. doi:10.1016/j.procs.2011.12.035

[11] P. A. Gaona-García, D. Martín-Moncunill, S. Sánchez-Alonso, and A. Fermoso, “A usability study of taxonomy visualisation user interfaces in

digital repositories," Online Inf. Rev., vol. 38, no. 2, pp. 284–304, 2014. doi:10.1108/OIR-03-2013-0051

[12] P. A. Gaona-García, S. Sánchez-Alonso, and C. E. Montenegro Marín, "Visualization of information: a proposal to improve the search and access to digital resources in repositories," Ingeniería e Investigación, vol. 34, pp. 83–89, 2014. doi: 10.15446/ing.investig.v34n1.39449

[13] P. A. Gaona-García, G. Stoitsis, S. Sánchez-Alonso, and K. Biniari, "An Exploratory Study of User Perception in Visual Search Interfaces Based on SKOS." Knowledge Organization, vol. 43, no. 4, 2016.

[14] J. Edosomwan and T. O. Edosomwan, "Comparative analysis of some search engines." S. Afr. J. Sci., vol. 106, no. 11–12, 2010. doi:10.4102/sajs.v106i11/12.169

[15] D. Tümer, M. A. Shah, and Y. Bitirim, "An empirical evaluation on semantic search performance of keyword-based and semantic search engines: Google, yahoo, msn and hakia," in Proceedings - 2009 4th International Conference on Internet Monitoring and Protection, ICIMP 2009, 2009, pp. 51–55. doi:10.1109/ICIMP.2009.16

[16] P. Gupta, S. K. Singh, D. Yadav, and A. K. Sharma, "An improved approach to ranking web documents," J. Inf. Process. Syst., vol. 9, no. 2, pp. 217–236, 2013. doi:10.3745/JIPS.2013.9.2.217

[17] T.H. Haveliwala, "Efficient Computation of PageRank," Stanford Univ. Technical Report, 1999.

[18] T. H. Haveliwala, "Topic-sensitive PageRank: a context-sensitive ranking algorithm for Web search," IEEE Transactions on Knowledge and Data Engineering, vol. 15, no. 4. pp. 784–796, 2003. doi:10.1109/TKDE.2003.1208999

[19] C. Kliman-Silver, A. Hannak, D. Lazer, C. Wilson, and A. Mislove, "Location, location, location: The impact of geolocation on web search personalization," in Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC, 2015, vol. 2015-October, pp. 121–127. doi:10.1145/2815675.2815714

[20] W. Zhang and Y. Chen, "Bayes topic prediction model for focused crawling of vertical search engine," in Proceedings - 2014 IEEE Computers, Communications and IT Applications Conference, ComComAp 2014, pp. 294–299. doi:10.1109/ComComAp.2014.7017213

[21] L. Zhang, H, Song, S. Yu, F. Ma. "Design and implementation of a high-performance distributed web crawler," Shanghai Jiaotong Daxue Xuebao/Journal of Shanghai Jiaotong University, Vol 38, pp. 59-61, 2004.

[22] L. Liu, T. Peng, "Post-processing of Deep Web Information Extraction Based on Domain Ontology," Advances in Electrical and Computer Engineering, vol.13, no.4, pp.25-32, 2013. doi:10.4316/AECE.2013.04005

[23] S. Brin, "The anatomy of a large-scale hypertextual Web search engine 1," Comput. Networks, vol. 30, no. 1–7, pp. 107–117, 1998.

[24] M. A. Kausar, V. S. Dhaka, and S. K. Singh, "Web crawler: a review," Int. Journal Computer Application, vol. 63, no. 2, 2013. doi: 10.5120/10440-5125

[25] A. Martínez-Rodríguez, "Cybermetric indicators: New proposals to measure information in the digital environment" ACIMED, vol. 14, no. 4, 2006.

[26] F. Ahmadi-Abkenari and A. Selamat, "An architecture for a focused trend parallel Web crawler with the application of clickstream analysis," Inf. Sci. (Ny)., vol. 184, no. 1, pp. 266–281, 2012. doi:10.1016/j.ins.2011.08.022

[27] V. Raval and P. Kumar, "SEReleC (Search Engine Result Refinement and Classification) - A Meta search engine based on combinatorial search and search keyword based link classification," in IEEE-International Conference on Advances in Engineering, Science and Management, ICAESM-2012, 2012, pp. 627–631.

[28] K. S. Kim, K. Y. Kim, K. H. Lee, T. K. Kim, and W. S. Cho, "Design and implementation of web crawler based on dynamic web collection cycle," in International Conference on Information Networking, 2012, pp. 562–566. doi:10.1109/ICOIN.2012.6164440

[29] S. Brin and L. Page, "Reprint of: The anatomy of a large-scale hypertextual web search engine," Comput. Networks, vol. 56, no. 18, pp. 3825–3833, 2012. doi:10.1016/j.comnet.2012.10.007

[30] M. D. Dikaiakos, A. Stassopoulou, and L. Papageorgiou, "An investigation of web crawler behavior: Characterization and metrics," Comput. Commun., vol. 28, no. 8, pp. 880–897, 2005. doi:10.1016/j.comcom.2005.01.003

[31] M. A. Whiting and N. Cramer, "WebThemeTM: Understanding web information through visual analytics," Lecture Notes Comput. Sci. (including Subser. Lect. Notes Artificial Intelligent Bioinformatics), vol. 2342 LNCS, pp. 460–468, 2002. 10.1007/3-540-48005-6_41

[32] P. Boldi, B. Codenotti, M. Santini, and S. Vigna, "UbiCrawler: A scalable fully distributed Web crawler," Software - Practice Exp., vol. 34, no. 8, pp. 711–726, 2004. doi:10.1002/spe.587

[33] N. Singhal, A. Dixit, R. P. Agarwal, and A. K. Sharma, "Regulating frequency of a migrating web crawler based on users interest," International Journal Engineering Technology, vol. 4, no. 4, pp. 246–253, 2012.

[34] R. Shettar and G. Shobha, "Web crawler on client machine," in Proceedings of the International MultiConference of Engineers and Computer Scientists, 2008, vol. 2.

[35] J. Nielsen, "Usability inspection methods," in Conference on Human Factors in Computing Systems - Proceedings, 1994, vol. 1994-April, pp. 413–414. doi:10.1145/259963.260531

[36] M. Maguire, "Methods to support human-centred design," Int. J. Hum. Comput. Stud., vol. 55, no. 4, pp. 587–634, 2001. doi:10.1006/ijhc.2001.0503

[37] M. A. Marzal, J. Calzad-Prado, M. J. C. Ruiz, and A. C. Cerveró, "Development of a controlled vocabulary for learning objects' functional description in an educational repository," in Proceedings of the International Conference on Dublin Core and Metadata Applications, 2006.

[38] D. Manjula and T. V Geetha, "Semantic search engine," J. Inf. Knowl. Manag., vol. 3, no. 1, pp. 107–117, 2004. doi:10.1142/S0219649204000729

[39] C. Cobos, E. León, and M. Mendoza, "A harmony search algorithm for clustering with feature selection," Rev. Fac. Ing. Univ. Antioquia, no. 55, pp. 153–164, 2010.

[40] F. J. Serón and C. Bobed, "VOX system: a semantic embodied conversational agent exploiting linked data," Multimedia Tools Application, vol. 75, no. 1, pp. 381–404, 2016. doi:10.1007/s11042-014-2295-5.

**Ricardo Beltrán-Alfonso**: Student of computer science at the engineering Faculty of the Universidad Distrital Francisco Jose de Caldas.

**Andres Torres-Tautiva**: Student of computer science at the engineering Faculty of the Universidad Distrital Francisco Jose de Caldas.

**Paulo Alonso Gaona-Garcia**: PhD in computer science and professor at the engineering Faculty, Universidad Distrital Francisco Jose de Caldas.

**Carlos Enrique Montenegro-Marin**: PhD in computer science and professor at the engineering Faculty, Universidad Distrital Francisco Jose de Caldas.

# Simple MoCap System for Home Usage

M. Magdin

*Constantine the Philosopher University in Nitra, Faculty of Natural Sciences , Department of Computer Science, Tr. A. Hlinku 1, 949 74 Nitra, Slovakia*

***Abstract* — Nowadays many MoCap systems exist. Generating 3D facial animation of characters is currently realized by using the motion capture data (MoCap data), which is obtained by tracking the facial markers from an actor/actress. In general it is a professional solution that is sophisticated and costly. This paper presents a solution with a system that is inexpensive. We propose a new easy-to-use system for home usage, through which we are making character animation. In its implementation we paid attention to the elimination of errors from the previous solutions. In this paper the authors describe the method how motion capture characters on a treadmill and as well as an own Java application that processes the video for its further use in Cinema 4D. This paper describes the implementation of this technology of sensing in a way so that the animated character authentically imitated human movement on a treadmill.**

***Keywords* — Multimedia, MoCap, Animations, LED Sensors.**

## I. Introduction

CONCEPT of real animation realized by hand drawing was created by J.Stuart Blackton. His most famous film is "Humorous Phases of Funny Faces" from the year 1906. This film has only three minutes, but it is considered the first animated film ever and Blackton as the first true animator - creator.

Gradually developed technologies led to improved animations and more approach to reality. Their foundation was interactivity - an inherent feature of multimedia. Therefore, we can conclude that the use of interactive and multimedia elements has a big impact. To create the impression of reality, a method using the movement of real people, which is transmitted to an animated character, was started to be used [36].

According to [24], the Animation Society is currently discussing the definition of animation. The way in which the introduction of new techniques interferes with the creative process of traditional animation production calls into question the concept of animation. The introduction of computers brought a "new pencil", with animation now under the influence of digital manipulation, currently recognizing the fascination with the effect of animation [37]. Today, films that are structured around visual effects, such as the Wachowski brothers' The Matrix (Warner Bros., 1999), point towards a path where today's film could be a sub-genre of animation [25].

According to [1], over the last three decades researchers from vision, computer graphics and robotics had been presenting methods for automatically generating camera control paths, which can be used in a large set of 3D applications. This problem is considered hard because of the large configuration space, as well as the huge number of factors that can affect the camera control [8, 27, 32, 14, 13]. The universal popularity of 3D games, which are based on human animation and 3D virtual environments, poses the camera control algorithms of an even more difficult problem, the capturing of human action scenes [1]:

1. The human character is an articulated object with many degrees of freedom. As a result, human motions are complex and its analysis is considered hard.
2. Human actions generally include several participating limbs, which have a prominent role in expressing the actual action. As a result, unlike in movement of simple geometric objects, the selection of expressive viewpoints should be affected by analyzing the motion and visibility of the body parts.
3. Small changes in the character pose often imply significant changes in the desired viewpoints for capturing this motion well.
4. The significance of the human actions is non-uniform over time. For example, routine actions such as a walk are visually less significant than a high action karate kick.

An efficient procedure is realized by using the motion capture data (MoCap data), which is obtained by tracking the facial markers from an actor/actress. In some cases of artistic animation, the MoCap actor/actress and the 3D character facial animation show different expressions. For example, from the original facial MoCap data of speaking, a user would like to create the character facial animation of speaking with a smirk. In this paper, we propose a new easy-to-use system for home usage, through which we are making the character animation [16].

According to [18], Motion capture (MoCap) data is extensively used in many multimedia applications such as in entertainment e.g., dance recording and training [5], in gaming e.g., cloud gaming [4], online gaming [22], etc., streaming animation [31] and online physical rehabilitation [34], [21] et al.

The first mention of using motion capture of human is from 1970, when the army used this technology for determining the position of the head of pilots when starting and landing on an aircraft carrier. A few years later the technology has started to be used for medical purposes in rehabilitation facilities. From the late 80s this technology is used in the entertainment industry. The definition of motion capture appears in 1995 in the work of Scott Dyer: "Capturing motion involves measurement of the position of the object as well as its orientation in physical space, in a way appropriate for work in computer. Objects of interest are human and other bodies, facial expressions, position cameras and lights and other elements in the scene" [8].

According to [19] MoCap gives some great opportunities to animators such as timesaving, simplifying the process of animation and others. MoCap can be applied in many other fields: sport (analyzing movements), military, robotics, game industry, medicine and others. This report is focused on using accelerometers in MoCap computer animation.

Hardware technology of Motion Capture equipments can be divided into several basic groups. In general, we can not tell which group is better or has higher quality. All have their advantages and disadvantages, and each has its ideal area for realization. According to a particular situation it is ideal to consider which version to use but in many cases it is difficult to combine two different hardware systems [7].

The main disadvantage of mechanical motion capture is lack of coordination with the floor, so that the human body can not completely



DOI: 10.9781/ijimai.2017.4410

implement jump (or otherwise to separate from the floor). The disadvantages are a slight limitation of movement and to option uses this system only for humanoid characters. While optical sensors can be used on any creature, mechanical exoskeleton captures and transmits only human skeleton, for which it was created [9].

The main drawback of optical motion capture is the interaction of light, side light sources and reflections, which are causing loss of data [9].

In the acoustic systems, there is no overlapping of sensors, but the system also has a number of disadvantages. Space suitable for scanning is highly restricted by placing receivers. Scan rate is relatively low, as it is limited by the speed of sound. If you use medium or smaller room, the sound pulses are reflected from objects and create data noise and inaccuracies [15].

According to [19] the important parameters of MoCap systems are:

- Accuracy: the error of measuring the movements of the actor. This is a basic problem to solve.
- Animation in real-time: the possibility of computer model to follow the actor's movements with a very short (negligible) time delay. Then the actor could experiment with different movements and make corrections.
- Freedom of movements: The freedom of actor movements (acting) gives an opportunity for rich and interesting animation.
- Frames per second: the recording speed. The greater the speed the better the system will capture fast movements like running, jumping, waving etc.
- Interruptions: recording the movement with no interruption in the process.
- Identification: the ability of the system to recognize body parts.
- External influence: the independence of the system from external influence like magnetic fields or light.
- Price: the less the better. As one can guess the goal is to make high quality MoCap system with affordable price.
- Training: training the actors to work with the system.
- Portability: the ability of the system to be moved.
- Software complexity: the complexity of the software developed for data analyzing and character animating.

Motion Capturing (MoCap) is in the last decades the standard used process by which we can record the motion and translate it into mathematically-usable signals. These signals correspond to the tracking of a number of key points in space over time. The increasing demand for rendering smooth and plausible 3D motion is fueling the development of motion capture (MoCap) systems. In this paper we will describe new approach of MoCap and its implementations. Our goal is to provide simple solution for home usage - automatically generate a 3D animation that describes in detail human actions by using static camera. Our methods extend the various existing concepts.

Related previous work is reviewed in Section II. Our advanced motion classification techniques are described in Section III. Experimental results are presented and analyzed in Section IV and V. Finally, concluding remarks, discussion, comparison of results of our and other systems and possible future extensions are summarized in Section V.

## II. Related work

For real motion capture, abbreviated MoCap, there are a number of professional and amateur solutions. Generally speaking, professional solutions (VICON, Gypsy7, Qualisys) are used in film industry and by gaming companies. They are economically very costly, but the resulting effect is high. Amateur solutions are cost-efficient, but they are inaccurate and contain many software and hardware deficiencies [8].

MoCap systems are presently standard in computer graphics, particularly in creation of 3D or 4D animations. Since there are different types MoCap systems, in this section, we describe only those studies which closely relate to our work.

The work of [12] and [13], suggests the implementation of a camera engine within a game pipeline, for the purpose of generating better viewpoint selection, and game summarization. Their system is based on solving constraints which consider a set of viewpoint quality attributes, as well as the camera control path quality. This work extends [1] by introducing attributes which relate to the subject's action. This addition was sufficient to significantly improve the resulting camera control path for such human animation based games. While work of [13] is focused on generating results in real-time, it lacks the ability to analyze the data and search for global solutions, method of [1] suggests this solutions. However, both solutions do not offer the possibility of creating 3D or 4D animations in real time.

A similar solution as used by us appeared already in 2008. Authors [20] use in their work various static camera positions which are then extended into a camera path by interpolation. In this solution the main problem is in camera path planning occlusions between objects. This problem is described in [2], [32] and [13]. In addition to such static scene related occlusions, here we focus on examining self occlusions of the various human limbs. This extension of standard occlusion constraint methods [13] was addressed in the context of camera control only method which suggested [1].

In last decade, various compute algorithms have been proposed (simple heuristics [30], [3], minimizing degenerated projection of polygon [17], [10]). These algorithms were categorized first time by [33]. In our work we will use some of the descriptors, which have been proved as effective.

In Slovakia, professional as well as amateur solutions are used. Examples include projects resulting from the Private School of Animation. Since 2013 this school has its own animation system based on motion capture. The system is composed of twenty cameras, which is an unique worldwide level. The system handles multiple actors, what the students can try. This school provides system for private animation to companies, but again it is a costly affair.

Between amateur solutions we can include these following five projects that were implemented in the 2004, 2005, 2006, 2010, and 2011.

In 2004, the [26] proposed a partial solution to the problem using 4 cameras that are placed perpendicular to each other and the more reflective elements located on the lower limbs. Analysis of points was resolved in the program VideoSQC however animation failed to create [26].

In 2005, [23] develops the idea of the previous solution. Their conception of the problem was much more complex. They used a system of 4 cameras and reflex by using points with light reflecting directly into the camera. Like in the previous case was used the program VideoSQC for analysis of coordinates and to the animation program BlueBot in which was animated walking part of the lower limbs [23].

Project of Krocka et al. in 2006 has not added many innovations to the topic. In general, they have improved the animation but again used the same programs as the projects before them.

The team of [28] developed a tool for modeling interactive animation of humanoid character. Authors solved the problem of importing data into a graphical modeling program. Data were imported using the BVH format, which is a classic file containing the position changes of individual points of the object in each frame. This plugin has a name Animik. Plugin uses the kinematics, the position of the root object and the relative rotation of the joints [28].

The last amateur solution resolved the motion sensing of human instructor. For shoot of the actor was used semi-professional system of six cameras. Using this system in normal room created a problem of restriction of movement. With the right setup of cameras was created a limited space of movement. This limited space of movement is defined as a cube with an edge of about 2 meters. A movement exceeding this space created problems in the analysis of data and caused loss of data. The pretreatment data was realized with program Vicon IQ. As an intermediary for further data processing and for creation of basic skeleton the program MotionBuilder was used. The problem arose from differences skeletons in the various programs. The construction of bone in Motion Builder is easier than construction in Poser, by converting data were again some data lost. Therefore, it happened that the limb is moved unrealistically, or some joints achieve the impossible angles, such as hand bent at the elbow backwards. Also, sometimes occurred overlapping limbs [29].

## III. Design Of System For Animation And Analysis Of Motion

Previous researches used data processing programs created by commercial companies. After careful evaluation of the problems we have chosen the solution of creating a custom program that will analyze and export data into the required format.

We decided to use optical technology MoCap, because this variant is the cheapest financially and in establishing its own system, it does not need complex electrical circuits or special devices. For motion capture of characters several cameras or compact cameras and a set of reflective elements mounted on the actor are needed.

### *A. Technical Part - Motion Capture*

For the motion capture of a walking man we chose a scene on a treadmill. We decided to use three compact cameras, one for left and one right leg and one for front that will capture the movement of both feet. By creating separation points of different limbs when shooting from the front, we solved software analysis of coordinates. For recognizing the points we used white LEDs. This solution, however, proved to be incorrect, since the movement of the actor points often overlapped and therefore there was the loss of data similarly as in case of previous amateur solutions. For additional sensing of figures we swapped white LEDs for color ones, for ease of distinction at a later stage of analysis. We did it on the assumption that the program will use the classic RGB model. Thus, we used red, blue and green ultra-bright LEDs. Deployment of LEDs was determined intuitive. Synchronization of motion capture using cameras (3x Canon EOS 5D Mark II) has solved the light element. At some point in time, when all three cameras (Fig. 1) panned motion, we turned off/on the light in the room. This point is therefore called time zero and is a crucial element for the beginning of the analysis of coordinates.

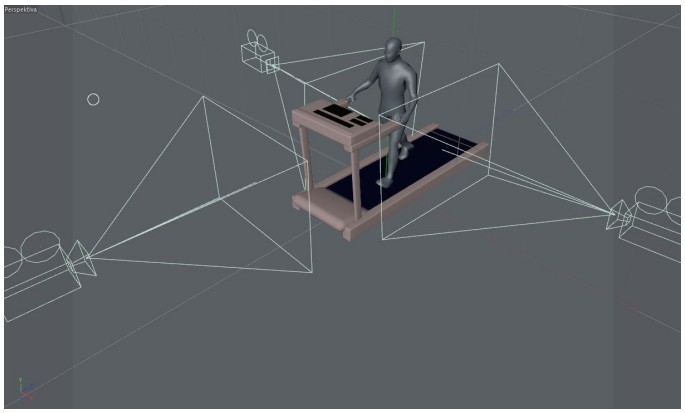


Fig. 1. Setup of cameras.

To eliminate reflections that create ultra-bright LEDs, we used the most black (matte), light absorbing substance. The actor was dressed in black, tight-fitting clothing (Fig. 2). We have created a prerequisite for later filtering quality in the software part.

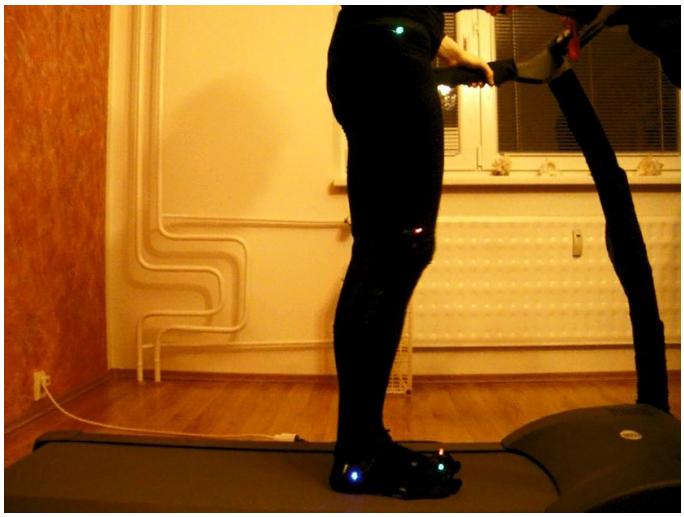


Fig. 2. The scanned Physique.

The last element that had to be solved was to eliminate non-essential elements in the captured video. This noise is in the case of optical MoCap technology everything that not is reflective elements. We therefore muted light in the room to a minimum, we filtered out other light sources which cameras catched and we left only faint, dim light necessary for orientation of the actor and the camera operator.

As can be seen in following Fig. (Fig. 3), by the movement of the feet back and forth, at the moment when this movement is the fastest, remains for the light elements a visual clue. This effect is caused by the imperfection of technique. By running the video we do not see this effect, but if we created the print screen from the video it is clear that the points are not circular cluster of pixels, but will have an ellipsoidal shape.

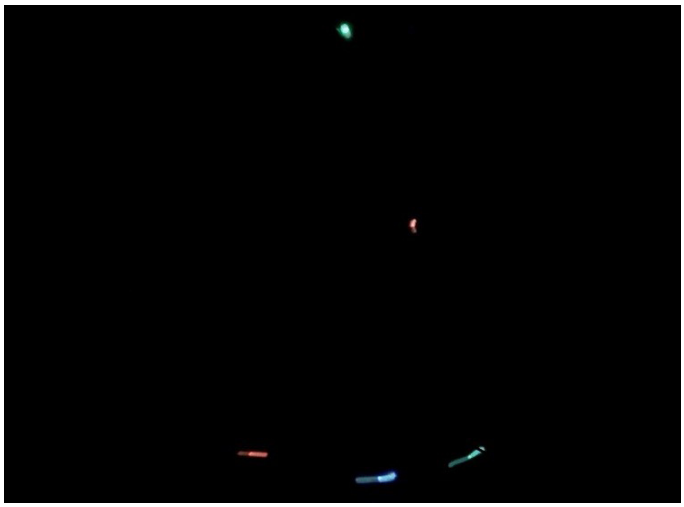


Fig. 3. Movement of reflex points.

Despite these imperfections, video sequences were taken approved as sufficient for our purposes. Therefore, we proceeded to the next phase, which was their modification.

### *B. Software part - processing and modification of scanned movement*

In this phase of work, we also decided to resolve the mirror effect

when shooting from two opposite sides. As seen in Fig. 4, the left camera captures the image so that if the actor moves the left foot, the foot appears to be closer to the camera and is closer to the left edge of the image, namely is cca 260 pixels. When viewed from the right camera at the moment the actor stepped left, this limb away as 260 pixels, but from the right edge images. From the perspective of the right camera is thus rotated view mirror. When analyzing the coordinates of this, it could have caused problems with the identification of coordinates.

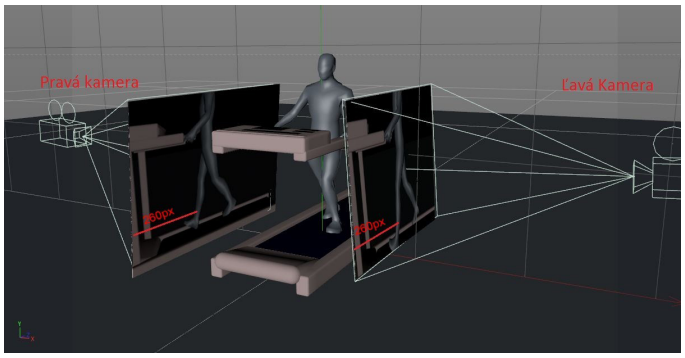



Fig. 4. Mirror effect.

On the video we used a scaling transformation with a matrix of horizontal rotation (1).

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

Analysis of the current state showed that at this stage of problem solving previous solvers had the most problems. Therefore, we decided to create a custom program that will identify and analyze the coordinates. We did not build on any existing program, we have created a program in NetBeans, build from scratch. However, we used the generated library (xuggle-xuggler), which operates with our program.


Fig. 5. Correctly identified cluster.

In the program even when shooting motion figures we worked with these three colors: red (R), green (G) and blue (B). By using these colors it is clear that it is possible to use the RGB model analysis. Its use, however, also introduced even more shortcomings. We found that we do not need to know how many blue, red and green are located around diodes. Fully sufficient was to specify only one color for each cluster. Therefore is the best using the HSV color model. This model meets our requirements almost perfectly. If we look, for example, blue color, so simply enter H than 240 ° and we are confident that we can find only blue. But even this was not a 100% solution because actually blue LEDs are off to a perfect shade of blue all the time. Pacing, we found that the color intensity of the admixture of gray was gently changing over time. Definition of the HSV model says that perfectly red color identified with H = 0 °, Green H = 120 ° and blue with H = 240 °, which represents a 120 degree interval for each basic color. So we decided to develop a tolerance for the search. Blue, therefore was procured from H = 240 °, but such an interval: H = <180 °, 300 °>, which we in the programming treated by calculating the angle = H / 360 °. This conversion no longer finds color in the interval <0,1>, which means NetBeans programming environment.

The problem is, when we identify 4 clusters (e.g. heel of the foot). This is as well as our case (shooting right and left sides).

At the idea that one video has 30 frames per second and we should manually set the HSV parameters for each cluster, it would take us weeks. Also, it could happen that we would merge some clusters. This was also one of the main reasons why we have created an application JFrame PointPaint. This subroutine deals with all settings under which is realized the export of the coordinates.

PointPaint (Fig. 6) can be divided into five main components:

1. Screen.
2. Components for handling with video.
3. Sets of settings.
4. Selection of videos and view.
5. Setting of HSB values.

Each part is designed for something else and includes several components and functions, by which we can identify the coordinates as accurately as possible.

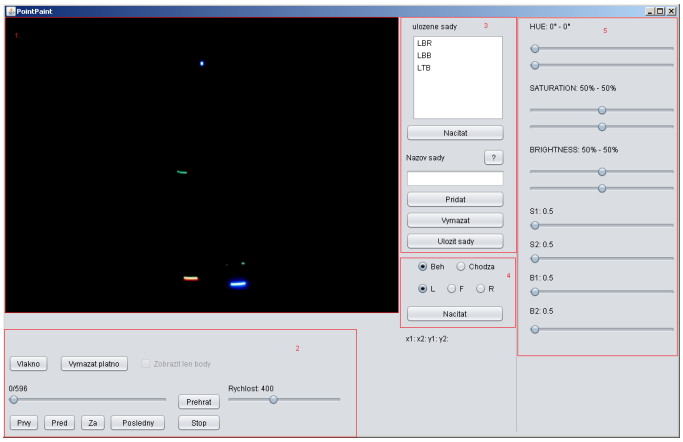



Fig. 6. Program PointPaint.

As the first step, the program determines the values $x_1$, $x_2$ and $y_1$, $y_2$ of LS_HSV. These parameters ensure that we were looking for a point in the correct area.

The second step is to define $x_{Min}$, $x_{Max}$, $y_{Min}$ and $y_{Max}$. Rewriting these values takes place according to comparisons based on HSV model, i.e. comparing the value of LS_HSV loaded with the appropriate value of the designated area.

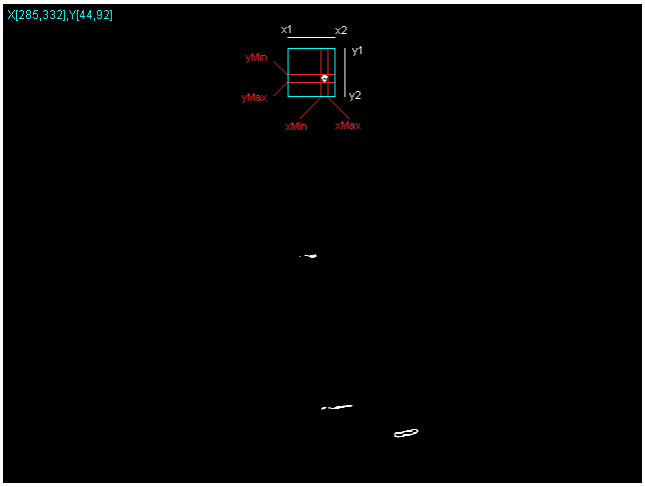



Fig. 7. Identification of the specific point of the cluster.

Finally, we use a simple algorithm in which these minimums and maximums are averaged by the formulas:

$$S_x = \frac{x_{Min} + x_{Max}}{2} \quad (2)$$

$$S_y = \frac{y_{Min} + y_{Max}}{2} \quad (3)$$

Position of specific points represents the mean value of the individual clusters, which for us is the best solution if the cluster has a circular or ellipsoidal form.

## IV. Modeling Characters in Cinema 4D

With the completion of setup of the HSV kit we achieved quite satisfactory results when identifying the clusters and specific points that cluster identified. These points were, however intended, only for each view separately, this means that each view was determined for all clusters of reflective elements, plus the corresponding grid co-ordinates. The aim of our work is but to combine several 2D coordinates into 3D coordinates for each specific reflex point. So we devised an algorithm, which converts the individual spatial coordinates.

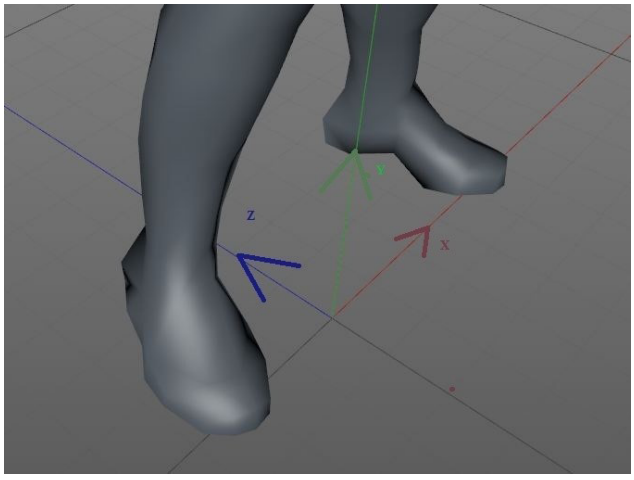



Fig. 8. Affine coordinate system.

In first step we created and defined the zero points. As seen in Fig. 8, we have created our own system of coordinates, for which the points will be recalculated. Point 0, i.e. the beginning of the coordinate system, we have determined as the point exactly between the feet. From this point were then calculated coordinates for the export to the visualization program Cinema4D.

From this perspective it is clear that the coordinates for this coordinates system and Affine coordinate system (Fig. 8) will be completely different as are coordinates for reflex points in different views. For example, actor which standing in a room as following the Fig. (Fig. 9), has when viewed L toe of the right foot at position [405,432]. The same position also reflects previous figure. For affine coordinate system, this toe has Sx and Sy negative; Sz near around zero.

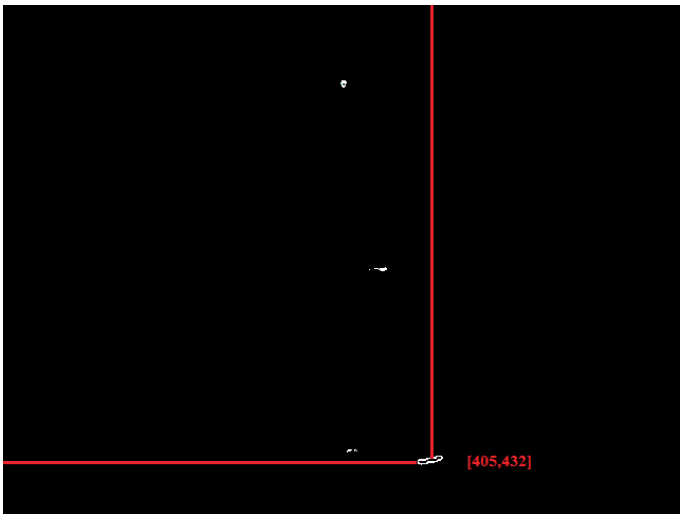



Fig. 9. Actor standing in at rest.

The outputs from the application PointPaint are the positions of all points for each image. Export was solved by generating part of the code that is inserted into an existing xml file. With the format XML we can work with program Cinema4D.

We have also programmed a number of calculations required for compatibility with the program Cinema4D. We created export of the earliest for the left leg, then for all the joints at once, and then the right foot. In first step we have to define the length of the generated code with the coordinates, representing the number of video frames, i.e. 597 (<= long in '597' />).

After these adjustments, we have to export the coordinates in the form <lreal v = 'Coordinate' />, which is a format for coordinates in Cineme4D. We wrote the coordinates in the order of *z*, *y*, and *x*.

After generating all coordinates we generated the coordinates of the beginning and end of the XML file for the process to Cinema4D.

As a model we used a standard Fig. from the collection of objects for Cinemu4D. As shown in fig. 10, we chose rectangular polygonal network. Quadrilaterals well deform and have not created as sharp spike like triangles and do not require complex calculations such as polygons.

This object, however, in this state we can only edit with the moving points, which would be very laborious. By changing the coordinates of a concrete point are changing just four corresponding polygons. Model Fig. consists of 927 independent points and calculating the coordinates for each of them is unreasonable and unrealistic.

We therefore used another method, and thus the "rigging". It is a method to manually create human skeleton inside the object. The creation of the skeleton can also be used for non humanoid figure, but there will not be such a natural movement, such as for the human figure.

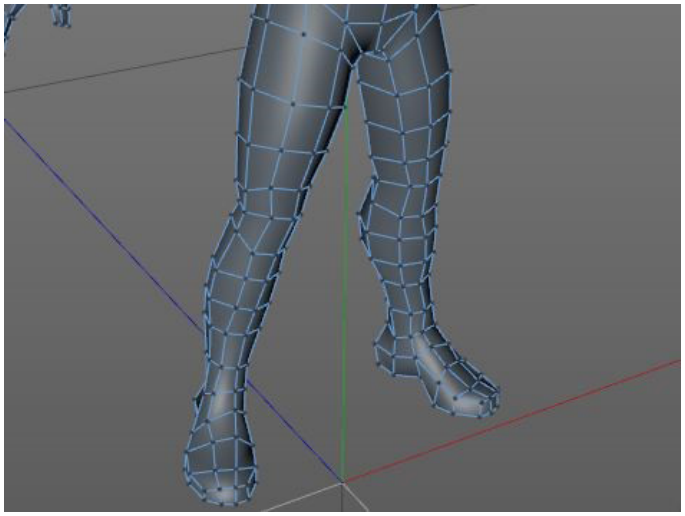


Fig. 10. Model Fig. of actor.

## V. Animated Movement Of Figure

There are several ways to create animation in Cinema4D, but the easiest and most practical is to identify the different positions of objects at different times and let the program calculate the intermediate state with using simple interpolation (keyframing).

Standard determination of positioning of individual objects unfortunately works only for simple objects. The complexity of method rigging makes it impossible to easily adjust the position of a particular joint, as this could cause breakage and separation of individual skeletal parts. Animation of movement of Fig. is therefore normally created do using only rotations.

Normally these are sufficient rotation, but for us it is important to determine the exact position of each joint and this is without giving rise to tearing or separation of the bones in the joints, what happens if we artificially specify the position of one of the joints.

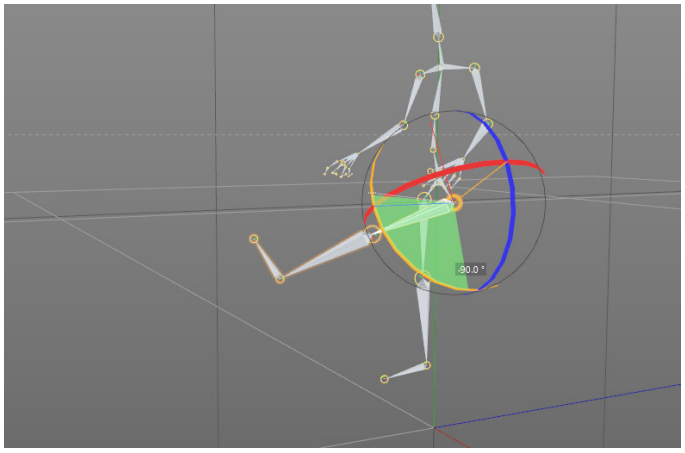



Fig. 11. Rotation of joints.

Therefore, it is necessary to add to the model several effectors which represent the forces acting between the joints of humanoid figure. Thus, the humanoid Fig. will behave realistically while maintaining keyframing techniques.

These forces we have expressed their as so-called IK objects. For each leg, we added one effector between the hip and heel, what is causing that the knee is not possible to bend to an unrealistic angle, i.e. greater than 180 °. Next effectors we added between the knee and the heel.

One effector we also included between hip joints, because they represent the hip bone, which is fixed.

Several effectors have also been added to the top of the humanoid Fig. though we did not animate them. Two effectors between the arms and wrists were added for realistic movement of the upper limbs. A two effectors directly on the wrist served for the character to "caught" the handles of the treadmill. So, even if the entire model is moved slightly to either side, as it is also in normal walking, the result is that the model will hold fast to the hand and is very slightly bent at the elbows and shoulders.

The thus prepared model is suitable for direct input of coordinates for each joint. The main advantage of IK effectors is that it does not determine the absolute position of individual parts, but only determines where this part was located.

The program according to the specified coordinates attempts to approximate the spatial data and the joint moves to the position closest to the point that we have identified.

But we must not violate any of the laws of physics, when we identify the model. For example, if an error occurs in the data, which would mean that is the tip of a 80 cm distance from the heel, toe there really does not move, but try to approximate the given coordinates. The foot will then be directed to a place where we identify faulty data, but no deformation of the foot occurs.

For easy keyframing we can only set position at a certain time on a certain percentage of the curve. But that would mean that we could split the curve at the most one hundred parts, as well as specifically indicated by the percentage might not hit the individual points. The captured frames of movement had more than five hundred points, so the division into hundredths would certainly be insufficient.

Fortunately, this complex modeling can be solved using the system programming tool Xpresso. We have created a system in which we have time animation transformed to a number of animated frames. Number of frames determines the index point of the curve and the position of this point, we transferred to the object of joint pertaining to a given curve.

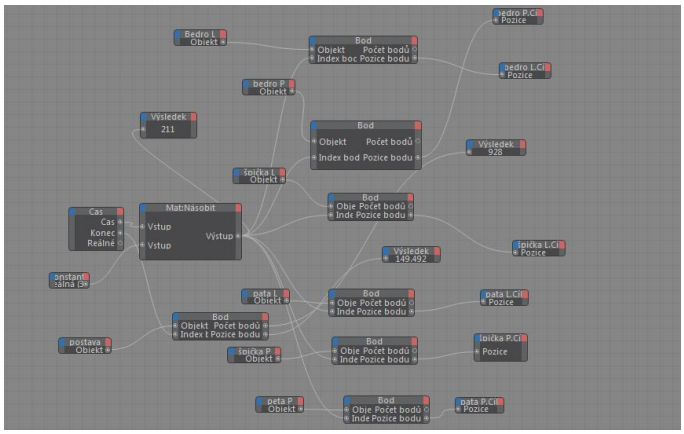


Fig. 12. The diagram in XPRESS.

From this moment, our character began to move in the desired manner and all joints copied movements for the data matrix.

## VI. Discussion - comparison of our and other MoCap LED sensors systems

For real motion capture, abbreviated MoCap, there are a number of professional and amateur solutions. Generally speaking, professional solutions (VICON, Gypsy7, Qualisys) are used in film industry and by gaming companies. They are economically very costly, but the

resulting effect is high. Amateur solutions are cost-efficient, but they are inaccurate and contain many software and hardware deficiencies [8].

**Vicon** is a professional solution of motion capture that is based on optical technology. The company supplies high-quality hardware such as cameras with the possibility of recording up to 250 frames per second. Vicon uses six various cameras for equally spaced system. The price of equipment Vicon depends from count and quality of the cameras.

TABLE I

Specifications of System Vicon

| Specification | Vicon |
|---|---|
| Camera model | MX-F20 |
| Lenses | 6x4/16mm |
| Sensor type | CMOS |
| Sensor dimensions | 1600x1280 |
| Price ($) | 15.000-250.000 |

**Qualisys** - output is animation of a real movement. Biomechanics approach opens the possibility of analysis of movement for rehabilitation purposes, such as analysis of the lower legs during running and walking.

TABLE II

Specifications of System Qualisys

| Specification | Qualisys low-end | Qualisys high-end |
|---|---|---|
| Camera model | Miqus M1 | Oqus 7+ |
| Lenses | 58x40/16mm | 54x42/16mm |
| Sensor type | CMOS | CMOS |
| Sensor dimensions | 1216x800 | 4096x3072 |
| Price ($) | 5.000$ | 20.000$ |

**OptiTrack** offers a blend of performance and usability that produces high-precision, biomechanically-relevant motion capture data via workflows that are unprecedented in their simplicity. Motive supports numerous biomechanics marker sets, with a focus on scientific and anatomically-valid options. Whether your analysis takes place in Visual3D, The MotionMonitor, MATLAB, or another third-party tool, these marker sets provide a robust marker tracking and auto-labeller.

TABLE III

Specifications of System Natural Point - Optitrack

| Specification | Natural Point |
|---|---|
| Camera model | Optitrack Flex V100R2 |
| Lenses | 12x4.5mm |
| Sensor type | CMOS |
| Sensor dimensions | 640x480 |
| Price ($) | 15.000$ |

**Gypsy7** - this MoCap system from company Meta Motion consists of a robust structure and is fitted on the human body. The construction is durable, made of light metal in order does not obstruct the movement.

Gypsy7 is not possible to be compared with these systems because it uses a completely different method of sensing the position of an object (system price is 14.000$ to 27.000$). Qualisys offers various solutions of cameras. In the Table 2 is stated the low and high-end solution. There exist other, relatively inexpensive solutions, e.g. Microsoft Kinect - price 200$. They are not intended to create 3D animation of man walking in real time.

The basis comparison of system Vicon and NaturalPoint and determination of their accuracy did [35] in 2013. The first phase of the study investigated the static and dynamic linear accuracy of the Vicon MX and Natural Point systems. In both assessments, a standard reference object was used (the Vicon "T" calibration object). The distance between two markers was chosen as the reference length and measured as 0.120 m using a 3 m Faro Fusion Arm (spatial accuracy = 0.104 mm). For the static test of linear accuracy, the reference object was positioned in the center of the calibrated volume. Three 10-s trials were recorded. The mean measured length of the 0.120 m object was calculated over the entire trial as the three-dimensional vector magnitude. Gait data were collected from a single participant, with an age of 22 years, height of 1.78 m and body mass of 75 kg. The linear accuracy tested identified a maximum absolute percentage error of 0.84% for the dynamic test in the mid volume of the Natural Point system. In all conditions, the Natural Point system produced higher errors than the Vicon system. No absolute percentage errors were found to exceed 1% deviation from the known length [35].

The main advantage of the PointPaint system is its accuracy. When you export the coordinates and their subsequent use, we recorded only one error in the coordinates from the five outputs (export of program). Each export contained 14328 values.

The resulting error rate is therefore only 1.39%. The advantage is also that if a system error occurs, it does not mean senseless of movement, but occurs raising the smooth movement due to physical changes kinetics model.

TABLE IV

Specifications of our system PointPaint

| Specification | Our system (PointPaint) |
|---|---|
| Camera model | Canon EOS 5D mark 2 |
| Lenses | 24.0x36.0/16mm |
| Sensor type | CMOS |
| Sensor dimensions | 5634x3753 |
| Price ($) | 1.500$ |

## VII. Conclusion

On the basis of the analysis on current situation, we chose the use of optical technology MoCap. While we have achieved the desired objective, it was necessary to solve a number of problems associated with the formation of a complete MoCap system.

The first big problem was capturing movement of actor, and set up of the scene. It turned out that the best choice is to use multi-color LEDs while filming. It is also necessary follow-tuning videos using video filters.

The main problem was to analyze coordinates and transformation of two-dimensional views into a three-dimensional affine coordinate system. These transformations we have programmed as part of the program PointPaint. The program also served as a mediation linking technical section and section of animation. PointPaint provides import video into the Java environment and export the resulting coordinates into animation program Cinema4D.

## References

[1] J. Assa, D. Cohen-Or, I. Yeh, & T. Lee. *Motion overview of human actions*. Paper presented at the *ACM SIGGRAPH Asia 2008 Papers, SIGGRAPH Asia'08*

[2] W. H. Bares, S. Thainimit, S.A. McDermott. *A model for constraint-based camera planning*. In Smart Graphics, Papers from the 2000 AAAI Spring Symposium, AAAI Press, vol. 4, 84–91, 2000.

[3] V. Blanz, M.J. Tarr, M.J., H. H. Bulthoff. *What object attributes determine canonical views*. Perception 28, 5, 575–600, 1999.

[4] W. Cai, V. C. M. Leung, L. Hu. *A cloudlet-assisted multiplayer cloud gaming system.* MONET 19(2):144–152, 2014. doi:10.1007/s11036-013-0485-4

[5] W. Choensawat, M. Nakamura, K. Hachimura. *Genlaban: a tool for generating labanotation from motion capture data*. Multimedia Tools Appl 74(23):10,823–10,846, 2014. doi:10.1007/s11042-014-2209-6

[6] M. Christie, P. OLivier. *Camera control in computer graphics*. In Eurographics 2006 Star Report, 89–113, 2006.

[7] L. Dickson, *Motion Capture Uses Meta Motion*. 2008, on-line at http://www.metamotion.com/motion-capture-uses/content-creation.html.

[8] S. Dyer, J. Martin, J. Zulauf, *Motion Capture White Paper*, 12 Dec 1995, on-line at http: //reality.sgi.com/employees/jam sb/mocap/MoCapWP_v2.0.html.

[9] M. Furniss, "Motion capture," *MIT communications forum*. 2000, on-line at: <http://web.mit.edu/comm-forum/papers/furniss.html>.

[10] F. Gómez, F. Hurtado, J.A. Sellares. Toussaint, G. T. *Nice perspective projections*. Journal of Visual Communication and Image Representation 12, 4, 387–400, 2001.

[11] Grafika.SK. *Motion Capture už aj na Slovensku*. 2014, on-line at: <http://grafika.sk/clanok/motion-capture-uz-aj-na-slovensku/>.

[12] N. Halper, N., Helbing, R., Strothotte, T. *A camera engine for computer games: Managing the trade-off between constraint satisfaction and frame coherence*. In EG 2001 Proceedings, Blackwell Publishing, vol. 20(3), 174–183, 2001.

[13] N. Halper, P. Olivier. *Camplan: A camera planning agent*. In AAAI 2000 Spring Symposium on Smart Graphics, AAAI Press, 92–100, 2000.

[14] L. W. He, M. F. Cohen, D. H. Salesin. *The virtual cinematographer: a paradigm for automatic real-time camera control and directing*. In SIGGRAPH 1996 Conference Proceedings, ACM, 217–224, 1996.

[15] T. Hermann, *AcouMotifor Acoustic Moon Syson – An Interaactive Sonification System for Acoustic Motion Control*. 2006. E-ISBN: 978-3-540-32625-0

[16] K. Hirose, T. Higuchi. Creating facial animation of characters via MoCap data. *Journal of Applied Statistics, 39*(12), 2583-2597, 2012. doi:10.1080/02664763.2012.724391.

[17] T. Kamada, S. Kawai. *A simple method for computing general position in displaying three-dimensional objects*. Computer Vision, Graphics, and Image Processing 41, 1, 43–56, 1988.

[18] M. A. Khan. Multiresolution coding of motion capture data for real-time multimedia applications. *Multimedia Tools and Applications,* , 1-16, 2016. doi:10.1007/s11042-016-3944-7

[19] N. T. Kostov, S. M. Yordanova, Y. D. Kalchev. *MoCap - the advantages of accelerometers and accuracy improvement*. International Journal of Mathematics and Computers in Simulation*, 9*, 60-64, 2015.

[20] J.-Y. Kwon, I. K. Lee. *Detemination of camera parameters for character motions using motion area*. The Visual Computer 24, 475–483, 2008.

[21] J. S. Lin, D. Kulic. *Online segmentation of human motion for automated rehabilitation exercise analysis*. IEEE Trans Neural Syst Rehabil Eng 22(1):168–180, 2014. doi:10.1109/TNSRE.2013.2259640

[22] L. Liu, A. Jones, N. Antonopoulos, Z. Ding, Y. Zhan. *Performance evaluation and simulation of peerto- peer protocols for massively multiplayer online games*. Multimedia Tools Appl 74(8):2763–2780, 2015. doi:10.1007/s11042-013-1662-y

[23] Ľ. Lučenič, "Animácia a vizuálna analýza chôdze človeka," 2005, *Team project FIIT STU*

[24] F. Luz, *Digital Animation: Repercussions of New Media on Traditional Animation Concepts*. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Vol. 6249 LNCS., 2010, doi:10.1007/978-3-642-14533-9_57.

[25] L. Manovich. *Image Future*. Animation 1(1), 2006, 25–44

[26] J. Mareták, "Animácia a vizuálna analýza chôdze človeka," 2004, *Team project FIIT STU*

[27] H. McCabe, J. Kneafsey. *A virtual cinematography system for first person shooter games*. In Proceedings of International Digital Games Conference, 25–35, 2006.

[28] M. Mihalovič, "Animik – nástroj na interaktívne modelovanie animácií humanoidných postáv," 2010, *Diploma thesis*.

[29] J. Nováček, "MOCAP – Snímání pohybu lidské postavy,", 2011, *Bachelor thesis*.

[30] S. Palmer, E. Rosch, P. Chase. *Canonical perspective and the perception of objects*. Attention and Performance IX, 135–151, 1981.

[31] M. Patoli, P. Gkion, M. Newbury. White. *Real time online motion capture for entertainment applications*. In: 3rd IEEE international conference on digital game and intelligent toy enhanced learning (DIGITEL), 139–145, 2010. doi:10.1109/DIGITEL.2010.39

[32] J. H. Pickering. *Intelligent Camera Planning for Computer Graphics*. PhD thesis, University of York, 2002.

[33] O. Polonsky, G. Patané , S. Biasotti, C. Gotsman, M. Spagnuolo, M. *What's in an image: Towards the computation of the "best" view of an object*. The Visual Computer 21, 8-10, 840–847, 2005.

[34] M. A. Rahman. *Multimedia environment toward analyzing and visualizing live kinematic data for children with hemiplegia*. Multimedia Tools Appl 74(15):5463–5487, 2014. doi:10.1007/s11042-014-1864-y

[35] D. Thewlis, C. Bishop, N. Daniell, G. Paul. *Next-generation low-cost motion capture systems can provide comparable spatial accuracy to high-end systems*. Journal of Applied Biomechanics*, 29*(1), 112-117, 2013.

[36] M. Vozár, P. Kuna, "Evaluation of acquired knowledge of mathematical subjects in the field of applied informatics," *ICETA 2013 - 11th IEEE International Conference on Emerging eLearning Technologies and Applications*, Proceedings, 2013, 405-410.

[37] P. Wells. *The Fundamentals of Animation*. AVA Publishing, 2006, Lausane

**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.

# Distributed Search Systems with Self-Adaptive Organizational Setups

Friederike Wall

*Universitaet Klagenfurt, Department of Controlling and Strategic Management, Austria*

***Abstract* — This paper studies the effects of learning-induced alterations of distributed search systems' organizations. In particular, scenarios where alterations of the search-systems' organizational setup are based on a form of reinforcement learning are compared to scenarios where the organizational setup is kept constant and to scenarios where the setup is changed randomly. The results indicate that learning-induced alterations may lead to high levels of performance combined with high levels of efficiency in terms of reorganization-effort. However, the results also suggest that the complexity of the underlying search problem together with the aspiration level (which drives positive or negative reinforcement) considerably shapes the effects of learning.**

***Keywords* — Agent-based Simulation, Complexity, NK Fitness Landscapes, Reinforcement Learning.**

## I. Introduction

The organizational setup of distributed search systems is a topic that is investigated in many disciplines, such as control theory, complex systems science or computational organization theory (for extensive reviews cf. [1], [2], [3]). The *coherence* of and the *coordination* within distributed search systems are among the predominant issues in this line of research, where the former is defined in terms of some of the systems' properties (e.g., solution quality) and the latter is concerned with actions and interactions of agents collaborating in a distributed search system [4], [5]. Thus, the key topics of the organizational setup of distributed search systems addressed refer to the appropriate segmentation of the overall search problem into sub-tasks, the way sub-tasks are assigned to agents, and the mechanisms to consolidate the (partial) solutions to sub-tasks into an overall solution. The overall solution should be as satisfactory as possible where its quality is determined on the basis of coherence metrics (e.g., [4], [5], [6]).

Hence, feasible consensus mechanisms, performant algorithms for search and optimization, and the appropriate assignment of tasks are of particular interest in this line of research [1] – in order to contribute to improving results with respect to coherence metrics of relevance. However, this line of research (mostly implicitly) assumes that the designer of a distributed search system decides which of these mechanisms, algorithms and ways of assignment are employed in the organizational setup of the search system. This paper follows an approach that, in a way, can be regarded as complement to the aforementioned line of research: not the designer of a search system is allowed to (exogenously) decide on the systems' organizational setup but the search system's organizational setup evolves endogenously. In particular, we allow for self-adaptation of the search systems' organizational setup, i.e., while searching for better solutions for the overall search problem (during "run-time") the search system is allowed to change its organization, where changes are based on feedback [7].

The idea of self-adaptive distributed search systems builds on prior studies which provide evidence that distributed search processes could remarkably benefit with respect to solution quality obtained from inducing organizational dynamics while searching for better solutions – may it be in the organizational setup of collaborating robots or "swarms" of unmanned aerial vehicles or in the organizational design of a firm where managers search for higher levels of firm performance [8], [9], [10], [11]. Apparently, organizational change per se tends to enhance the performance of a search system by inducing a shift towards more exploration, i.e., discovery of new solutions, and less exploitation, i.e., stepwise improvement. However, it is worth emphasizing that these studies employ *merely random-driven* organizational change in the sense that the search systems do not learn which organizational setups are more successful than others.

By investigating the effects of learning-based organizational dynamics, this paper goes a step beyond research studies that employ random-driven organizational changes. *In particular, this paper studies the effects of endowing distributed search systems with some capabilities to learn about their organization's performance and to adapt the organizational setup according to the search systems' performance.* This paper is an extended version of [12] which was presented at the 13th International Conference on Distributed Computing and Artificial Intelligence (DCAI). The extensions predominantly relate to the dimensionality of the search problems under investigation, to the time horizon of simulations, and to a sensitivity analysis with respect to the number of search agents.

It appears to be of particular interest to investigate whether search systems which employ learning-based organizational change outperform systems which make use of random changes in their organizational setup or systems which do not change their setup changes at all. This study intends to provide findings on the relative potential benefits of learning-based organizational dynamics. Since, it is well known that the task environment (in terms of the task complexity) tends to affect the performance of search, this paper particularly controls for the complexity of the search problems by employing an agent-based simulation model which is based on the framework of fitness landscapes [13], [14]. The next section introduces the key elements of the simulation model. Section III gives an overview of the performed simulation experiments. The results are presented in Section IV where, first, an in-depth analysis of some baseline scenarios of organizational change modes for different levels of complexity are provided. Second, a sensitivity analysis is presented which puts particular emphasis on the need for coordination within the search system where this need is considerably affected by the number of search agents who carry out sub-tasks.

## II. Outline of the Simulation Model

The study employs an agent-based simulation model which captures two intertwined adaptive processes: In (1) the short-term, search agents seek to find superior solutions for the search problem. The quality of a solution is measured on the basis of system's overall performance level



DOI: 10.9781/ijimai.2017.4411

achieved. We model search agents to operate on NK fitness landscapes [13], [14]. In (2) the mid-term, the search systems are allowed to adapt major features of their organizational setup. Changes are driven by reinforcement-learning, which is based on performance enhancements achieved. A schematic flow-chart of key features of the simulation model is displayed in Figure 1.

### *A. Short-Term Adaptive Search for Higher Levels of Performance*

The study employs the framework of NK-fitness landscapes, which were originally introduced in the domain of evolutionary biology [13]. An advantage of NK fitness landscapes is that they easily allow for controlling the complexity of the underlying search problem [15].

#### *1) Search Problem*

In each time step *t* of the observation period *T*, the search systems face an *N*-dimensional binary search problem, i.e., they seek for a superior configuration $\mathbf{d_t} = (d_{1t},\ldots,d_{Nt})$ (with $d_{it} \in \{0,1\}$, $i = 1,\ldots,N$ ) out of a set of possible solutions, which is given by $2^N$ different binary vectors. Each of the two states $d_{it} \in \{0,1\}$ contributes with $C_{it}$ to fitness $V(\mathbf{d_t})$ of the search system or – in other words – to the performance achieved by the search system. According to the NK framework, $C_{it}$ is randomly drawn from a uniform distribution with $0 \le C_{it} \le 1$.

The parameter *K* captures the complexity of the underlying search problem [15]: In particular, fitness contribution $C_{it}$ might not only depend on the single choice $d_{it}$ but also on a number of other choices where *K* indicates the number of $d_{jt}, j \neq i$ that affect $C_{it}$ in addition to $d_{it}$. In case of no interactions *K* is 0, and *K* equals $N-1$ for the case of maximum interdependence. Hence, contribution $C_{it}$ results from

$$C_{it} = f_i(d_{it}; d_{i_1 t},..d_{i_K t}) \tag{1}$$

with $\{i_1,...i_K\} \subset \{1,...,i-1,i+1,...,N\}$.

The overall performance $V_t$ achieved by the search system in period *t* is computed as the normalized sum of contributions $C_{it}$:

$$V_t = V(\mathbf{d_t}) = \frac{1}{N}\sum_{i=1}^{N} C_t \tag{2}$$

#### *2) Agents and their Choices*

The search for higher levels of performance $V_t$ is collaboratively performed by *M* search agents. In particular, the *N*-dimensional search problem is partitioned into *M* disjoint partial problems, and each of these sub-problems is exclusively delegated to one search agent *r,* $r = 1,\ldots,M$. The partial search problems are of equal size with $N_r = \frac{N}{M} \forall r = (1,\ldots,M)$. From the perspective of search agent *r,* the search problem is segmented into a partial search vector $\mathbf{d}_t^{\mathbf{r}}$ which contains the choices which are under the search agent *r*'s primary control and into a partial vector $\mathbf{d}_t^{\mathbf{r},res}$, which captures the residual choices that the other search agents $q \neq r$ are in charge of. However, with cross-segment interactions among the sub-problems, choices of agent *r* might affect the contribution of the other agents' choices to overall performance, and vice versa.

In each time step *t*, a search agent seeks to identify the best configuration for the "own" choices $\mathbf{d}_t^{\mathbf{r}}$ assuming that the other agents do not alter their choices made in *t*-1. Each agent *r* randomly discovers two alternatives in addition to the status quo choice $\mathbf{d}_{t-1}^{\mathbf{r}*}$, where, as compared to the status quo, one alternative (named *a*1) differs in one and the other one (labelled *a*2) differs in two single choices $d_{it}$. In consequence, in time step *t*, agent *r* has three options to choose from, i.e., keeping the status quo or switching to $\mathbf{d}_t^{\mathbf{r},a1}$ or $\mathbf{d}_t^{\mathbf{r},a2}$. Which of these options appears to be favorable from the search agent's perspective depends on the agent's "objective" $P_t^r$. This objective is shaped by parameter $\alpha^r$ which defines the extent to which the residual part of the decision problem is considered in addition to the "own" partial search problem. The objective function can be formalized by

$$P_t^r(\mathbf{d_t^r}) = P_t^{r,own}(\mathbf{d_t^r}) + \alpha^r \cdot P_t^{r,res} \tag{3}$$

with $V_t = V(\mathbf{d_t}) = \frac{1}{N}\sum_{i=1}^{N} C_{it}$ and $P_t^{r,res} = \sum_{q=1,q\neq r}^{M} P_t^{q,own}$ where $p = \sum_{s=1}^{r-1} N_s$ for $r > 1$ and $p = 0$ for $r = 1$.

However, the agents' ex ante evaluations of alternatives do not necessarily need to be perfect. In particular, agents might misjudge the options' contributions to objective $P_t^r(\mathbf{d_t^r})$. This may not only be an unintentional shortcoming of, e.g., agents' information processing capabilities but may also be intentionally induced: Some evidence suggests that imperfect information on the fitness (performance) of options could increase the effectiveness of search processes (e.g., [16], [17]). Previous research shows that false-positive evaluations of options increase the diversity of search. As a consequence, there is a chance to end situations of inertia induced by sticking to a local peak and to reach basins of attractions for higher levels of fitness. Hence, intentionally or not, our agents may eventually be endowed with slightly distorted information about the performance of options. Distortions are captured by adding error terms as exemplarily shown in Eq. (4):

$$\tilde{P}_t^{r,own}(\mathbf{d_t^r}) = P_t^{r,own}(\mathbf{d_t^r}) + e^{r,own}(\mathbf{d_t^r}) \tag{4}$$

For the sake of simplicity, distortions are modeled as relative errors added to the true performance (for other functions see [16]). The error terms follow a Gaussian distribution $N(0;\sigma^2)$ with expected value 0 and standard deviations $\sigma^{r,own}$ and $\sigma^{r,res}$. Variances are assumed to be the equal for all search agents $r = 1,\ldots,M$ and stable over time (if not altered by self-adaptation as described subsequently); all error terms are assumed to be independent from each other.

Apart from the search agents, the model captures a kind of "central agent" whose role is a twofold: (1) In the short-termed adaptive search, the central agent could – depending on the particular mode of coordination – intervene in the selection of choices. (2) In the mid-term, the central agent assesses performance enhancements and "learns" about successful organizational setups by reinforcement. The next section provides more details on the central agent's roles.

### *B. Mid-Term Adaptation of the Organizational Setup based on Reinforcement Learning*

The very core of this study is related to learning on the performance contributions of a search system's organization and, eventually, altering the organizational setup accordingly. The following two subsections describe the modelled mode of reinforcement learning as well as

the dimensions of the organizational setup which may be subject to organizational change.

*1) Mode of Reinforcement Learning.*

In each *T**-th time step, the central agent faces an *L*-dimensional decision problem related to the *L* dimensions of the organizational setup which can be altered. In particular, the central agent chooses a setup $\ddot{\mathbf{O}}_{\mathbf{t}} = (a_1(t),...,a_L(t))$ of alternatives $a_l \in A_l$ for all $l = 1,...,L$ and with $|A_l|$ giving the number of alternatives $a_l$ in set $A_l$.

The model employs a simple mode of reinforcement learning (for overviews see [18], [19]) based on statistical learning, i.e., a generalized form of the Bush-Mosteller model [20], [21]: In every *T**-th period, the propensities of choices are updated based on the stimuli resulting from the evaluation of the outcome (performance effects) achieved under the regime of prior choices of the organizational setup. The outcome $\omega$ of configuration $\ddot{\mathbf{O}}_{\mathbf{t}}$ is given by the maximal relative performance enhancement which is achieved within the last *T** periods of the adaptive walk, i.e.,

$$\varpi(\ddot{\mathbf{O}}_{\mathbf{t}}) = \max\left[\frac{V_{t-\tilde{t}} - V_{t-T^*}}{V_{t-T^*}},\ \tilde{t} = 1,...,(T^*-1)\right] \tag{5}$$

The evaluation of the outcome can be regarded positive (1) or negative (-1), where the assessment of ω depends on whether or not it, at least, equals an aspiration level *v*. Hence, the stimulus $\tau(t)$ results as

$$\tau(t) = \begin{cases} 1 \text{ if } \varpi(\mathbf{\Phi}_{\mathbf{t}}) \geq v \\ -1 \text{ if } \varpi(\mathbf{\Phi}_{\mathbf{t}}) < v \end{cases} \tag{6}$$

Let $p(a_l,t)$ denote the probability of an alternative within dimension *l* of organizational setup to be chosen at time *t* (where $0 \leq p(a_l,t) \leq 1$ and $\sum_{a_l \in A_L} p(a_l,t) = 1$); $a_l(t)$ denotes the option out of set $A_l$ which is implemented at time-step *t*. The probabilities of options $a_l \in A_l$ are updated based to the following rule, where l ($0 \leq \lambda \leq 1$) reflects the reinforcement strength [21]:

$$p(a_l,t+1) = p(a_l,t) + \begin{cases} \lambda \cdot \tau(t) \cdot (1 - p(a_l,t)) & \text{if } a_l = a_l(t) \wedge \tau(t) = 1 \\ \lambda \cdot \tau(t) \cdot p(a_l,t) & \text{if } a_l = a_l(t) \wedge \tau(t) = -1 \\ -\lambda \cdot \tau(t) \cdot p(a_l,t) & \text{if } a_l \neq a_l(t) \wedge \tau(t) = 1 \\ -\lambda \cdot \tau(t) \cdot \dfrac{p(a_l,t) \cdot p(a_l(t),t)}{1 - p(a_l(t),t)} & \text{if } a_l \neq a_l(t) \wedge \tau(t) = -1 \end{cases} \tag{7}$$

After the probabilities are updated according to Eq. 7, the organizational setup $\ddot{\mathbf{O}}_{\mathbf{t}}$ to be implemented from time steps $t+1$ to $t+T^*$ is determined randomly based on the updated probabilities.

*2) Organizational Setup.*

The vector of the organizational setup $\ddot{\mathbf{O}}$ is modelled to be three-dimensional, i.e., $L = 3$. Within each dimension, three options are given (i.e., $|A_l| = 3 \ \forall \ l \in \{1,2,3\}$). These dimensions relate to (see also Table I):

1. The objective of the search agents as controlled by parameter $\alpha^r$ in Eq. 3. For the sake of simplicity, $\alpha^r$ is modelled to be the same for all search agents $r = 1,...,M$. In the following we skip index *r*.
2. The precisions of ex ante-evaluations made by search agents and by the central agent as given by $\sigma^{r,own}$ and $\sigma^{r,res}$, and $\sigma^{cent}$, respectively.
3. The mode of coordination as selected out of three alternatives: (a) "decentralized": without intervention from the central agent or any coordination which other search agents, each search agent autonomously decides on the "own" partial choices $\mathbf{d}_t^{\mathbf{r}}$; (b) "lateral veto": the search agents inform each other about their preferences and are endowed with mutual veto power; (c) "centralized": each search agent informs the central agent about the two most preferred options from $\mathbf{d}_{t-1}^{\mathbf{r}*}$, $\mathbf{d}_t^{\mathbf{r},a1}$ and $\mathbf{d}_t^{\mathbf{r},a2}$; the central agent chooses that combination of preferences which promises the highest overall performance *V*.

## III. Simulation Experiments and Parameter Settings

In the simulation experiments, after a performance landscape is generated, the initial organizational setup (i.e., vector $\ddot{\mathbf{O}}_{t=0}$) of a search system is determined randomly with uniform probabilities $p(a_l, t=0)$ out of the options in each dimension *l* as introduced above and summarized in Table I. Next, the search systems are placed randomly in the performance landscape. Then, over an observation time *T* of 500 periods, the search systems are observed while searching for higher levels of performance. In each *T**-th period, probabilities are updated and organizational configurations are (eventually) altered (cf. Sec. B). Fig. 1 displays the key events during a simulation experiment capturing learning-based adaptation of the organizational setup.

TABLE I
Parameter Settings

| Parameter | Values / Types |
|---|---|
| Observation period | $T = 500$ |
| Number of choices | $N = 12$ |
| Interaction structures | block-diagonal ($K = 2$);<br>full interdependent ($K = 11$) |
| Number and scope of search agents | *Baseline scenarios:*<br>$M = 4$, with d1 = ($d1$,...,$d3$), d2=($d4$,…,$d6$), d3 = ($d7$,…,$d9$),<br>d4 = ($d10$,…,$d12$)<br>*Sensitivity analysis:*<br>$M = 2$ with d1 = ($d1$,...,$d6$), d2=($d7$,…,$d12$)<br>$M = 6$ with d1 = ($d1$,...,$d3$) to d6=($d10$,…,$d12$) |
| Number of organizational dimensions | $L = 3$ |
| Dimension *l*=1: Agents' objective | $\alpha \in \{0, 0.5, 1\}$ |
| Dimension *l*=2: Precision of evaluation | (*sr,own*, *sr,res*, *scent*) $\in$ {(0, 0, 0),<br>(0.1, 0.15, 0.125),<br>(0.05, 0.25, 0.125)} |
| Dimension *l*=3: Coordination mode | decentralized; lateral veto;<br>centralized |
| Interval of change | $T^* = 25$ and<br>for contrasting to "no change": $T^* > T$ |
| Reinforcement strength | $l \in \{0, 0.5\}$ |
| Aspiration level | $v \in \{0, 0.01\}$ |

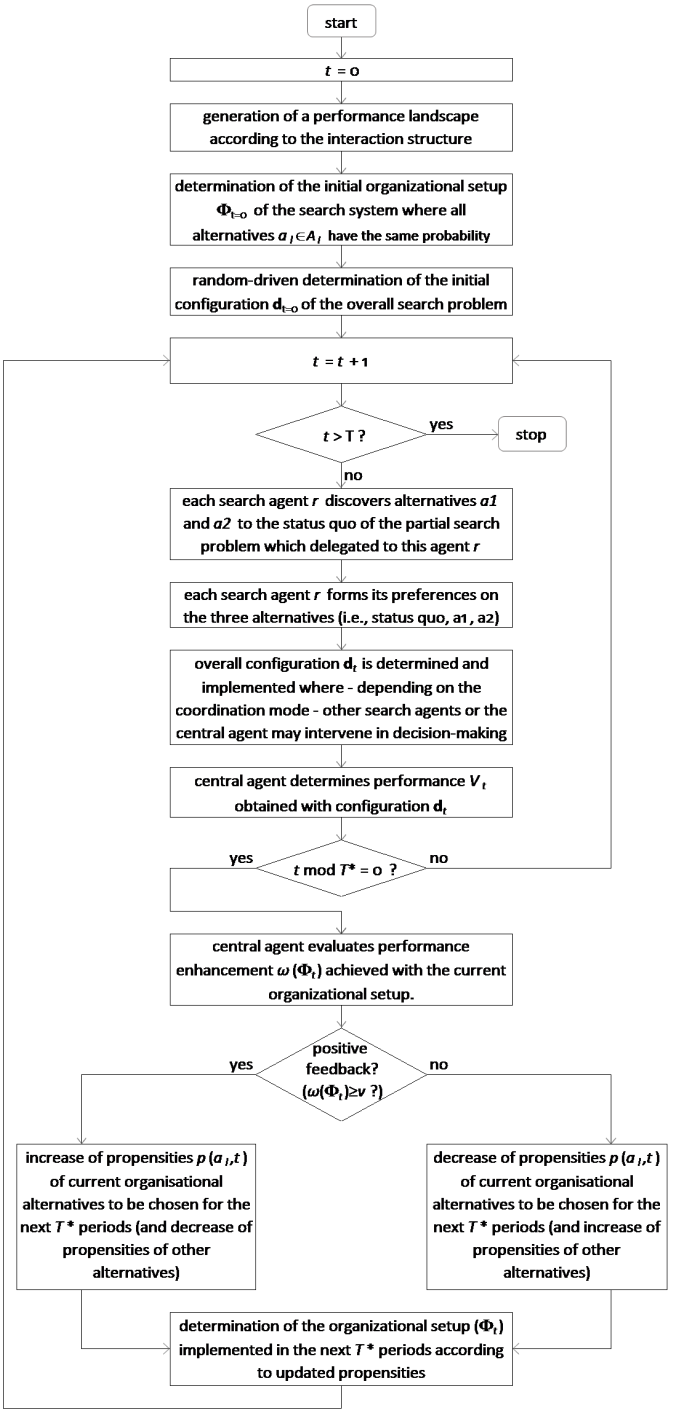



Fig. 1. Schematic representation of one simulation run over $T$ periods including potential changes of the search system's organizational setup.

In order to oppose search systems with learning capabilities (i.e., with $\lambda > 0$) to non-learning systems employing organizational change, simulations for $\lambda = 0$ are conducted. Moreover, search systems which do not alter their organization within the observation time $T$ (i.e., with $T^* > T$) are simulated.

In order to capture the complexity of the underlying search problem, simulations for two interaction structures are performed which, in a way, represent two extreme scenarios [22]: in the *block-diagonal* structure the overall search problem can be segmented into disjoint parts with maximal intense intra-sub-problem interactions but no cross-sub-problem interactions ($K^*$). An example is given in Figure 2.a with $K = 2$ and $K^* = 0$ where each of the four sub-problems is assigned to one search agent. In this setup, one agent's decisions do not affect the performance contributions of the other agents' choices.

The second case is characterized by full interdependence, i.e., all single options $d_i$ affect the performance contributions of all other choices $d_{j\neq i}$ and the search problem's complexity is raised to its maximum, i.e., intensity of interactions $K$ as well as the cross-sub-problem interactions $K^*$ are maximal (see Figure 2.b for an example with $K = 11$ and $K^* = 9$).

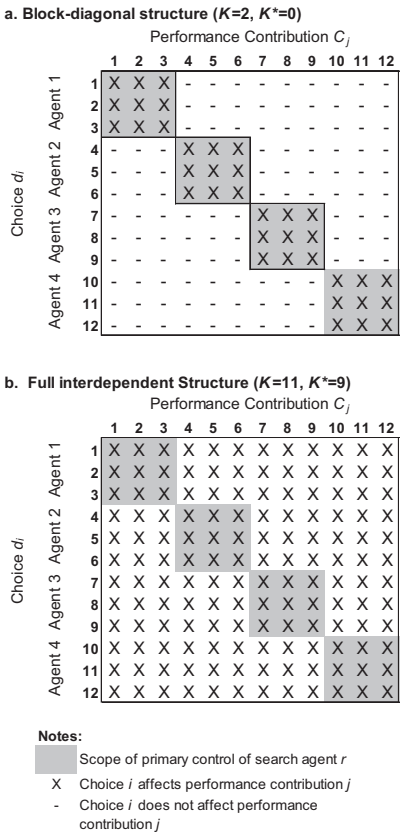



Fig. 2. Interaction structures and assignment of choices to search agents for the a. block-diagonal and b. full interdependent structure.

## IV. Results

The simulation experiments are conducted for two baseline scenarios of complexity (see Figures 2.a and 2.b) and for four modes of organizational adaption: (I) no change, (II) change without learning, (III) learning-based change with low aspiration level and with (IV) high aspiration level. These baseline scenarios are, then, modified in the number of search agents. In the modified scenarios, two or six search agents are employed instead of four.

### A. Baseline Scenarios

Table II displays condensed results of the simulated scenarios. The final performance achieved in the end of the observation period ($V_{t=500}$) and the performance achieved on average in each of the 500 periods ($\overline{V}_{[0;500]}$) may serve as indicators for the effectiveness of the search process. The same applies to the relative frequency of how often the global maximum is found in the final period. The ratio of alterations of d informs about the diversity of the short termed search processes, while the average number of alterations of organizational dimensions informs about the diversity of organizational setups that are implemented during adaptive walks and, eventually, modified within in the (mid-

termed) search for appropriate organizational setups. Figure 3 depicts the averaged adaptive walks for the different modes of change in the block-diagonal structure of interactions, and Figure 4 reports on the full interdependent structure correspondingly. In particular, Figures 3 and 4 show the performance differences of the "change, no learning" mode and the two modes employing learning, against the "no change" mode.

TABLE II

CONDENSED RESULTS OF THE BASELINE SCENARIOS

| Scenario of learning and change | Final Perform. $V_{t=500}$ [a] | Average Perform. $\overline{V}_{[0;500]}$ [a] | Frequ. glob. max at $t$=500 | Ratio of altered configs. of **d** | Average no. of altered organiz. dimens. over $T$ |
|---|---|---|---|---|---|
| *Block diagonal interaction structure* ($K$=5) [b] | | | | | |
| I. No change | 0.954 ±0.0042 | 0.9519 ±0.0035 | 36.08% | 20.00% | n/a |
| II. Change, no learning | 0.9599 ±0.0044 | 0.9582 ±0.001 | 45.40% | 22.97% | 38.2 |
| III. Learning low asp.lvl | 0.9696 ±0.0035 | 0.9609 ±0.0025 | 47.40% | 19.85% | 6.9 |
| IV. Learning high asp.lvl | 0.914 ±0.0055 | 0.9296 ±0.0022 | 17.76% | 43.19% | 20.0 |
| *Full interdependent interaction structure* ($K$=11) [b] | | | | | |
| I. No change | 0.8372 ±0.0081 | 0.811 ±0.0075 | 4.52% | 5.63% | n/a |
| II. Change, no learning | 0.8843 ±0.0075 | 0.8421 ±0.0048 | 7.84% | 10.24% | 37.9 |
| III. Learning low asp.lvl | 0.8684 ±0.007 | 0.8317 ±0.0059 | 4.96% | 8.74% | 6.2 |
| IV. Learning high asp.lvl | 0.8669 ±0.0093 | 0.8261 ±0.0058 | 7.76% | 15.98% | 36.2 |

[a] Confidence intervals, at a confidence level of 99.9%, for final performance range between 0.002 and 0.005 and for average performance between 0.001 and 0.003.
[b] Scenarios: "no change": $T^* > 500$; "change, no learning": $\lambda = 0$; $v = 0.01$; "learning, low aspiration level": $\lambda = 0.5$; $v = 0$; "learning, high aspiration level": $\lambda = 0.5$; $v = 0.01$. For further parameter settings see Table I.
Each data row shows the results of 2,500 adaptive walks: 10 walks on 250 distinct landscapes.

In the following, three aspects of the presented results are discussed in detail: (1) performance differences of scenarios in which the organizational setup is changed against scenarios in which the organizational setup is modelled to be constant, (2) the effects of learning-based adaptation compared to purely random adaptions of the organizational setup, (3) the intensity of organizational change (which is captured by the average number of altered dimensions.

Concerning the first aspect, Table II as well as Figures 3 and 4 indicate that – with one exception – performance levels of scenarios employing change persistently go beyond the level of performance achieved without change. This behavior can be observed after approximately 40 periods. These results confirm findings of research which indicate that altering the organizational setup in the course of distributed search processes may be favorable [8], [10], [11]: It has been argued that this is driven by the increased diversity of search which reduces the peril of sticking to local peaks. This is broadly confirmed by the ratio of alterations of configurations d and the frequency of how often the global maximum is found in $t = 500$ (cf. Table II).

However, results also suggest that learning by reinforcement with high aspiration levels is not universally beneficial. Apparently, the complexity of the search problem together with the aspiration level subtly affects the benefits of learning. In case of the block-diagonal structure, employing learning-based change with a high aspiration level leads to performance levels that are remarkably below the performances achieved without change throughout the adaptive walk from about time-step 75 to 500 and the final performance $V_{t=500}$ is about 4 points of percentage below the "no change" case. An explanation why, in case of the block-diagonal interaction structure, a high aspiration level apparently induces such a rather poor performance, may be based in the specific selective effects induced in this scenario:

With increasing aspiration level it becomes more unlikely that a positive stimulus $\tau(t)$ is achieved under the regime of a certain organizational setup – even if the setup had brought some (lower than $v$) performance enhancements in the last $T^*$ periods. Hence, even potentially appropriate organizational setups are likely to receive low probabilities to be re-chosen for the next $T^*$ periods. In the block-diagonal structure with its fairly low level of interactions ($K$ = 2), it is rather likely that the global maximum is found [22]: of course, no further performance enhancement is possible in these cases and the aspiration level is not reached. Whenever the global maximum is found (with aspiration level $v > 0$) the organizational setup is likely to be modified. An altered organizational setup also induces a modified evaluation of the current configuration **d** [11]. As a result, a move away from the global maximum in the performance landscape becomes likely.

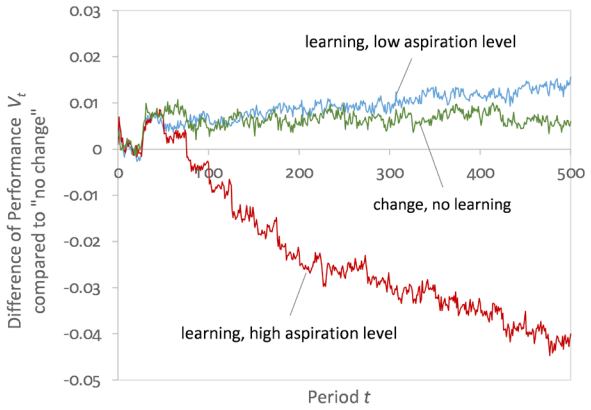



Fig. 3. Performance differences of adaptive search processes employing organizational change against search processes without alterations of the organizational setup in case of the block-diagonal interaction structure. Each curve represents the difference of means of the average of 2,500 adaptive walks, i.e., 250 distinct performance landscapes with 10 adaptive walks on each. For parameter settings see Table I.

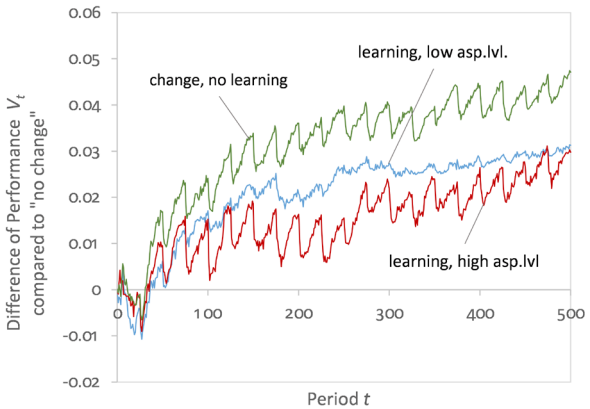



Fig. 4. Performance differences of adaptive search processes employing organizational change against search processes without alterations of the organizational setup in case of the full interdependent interaction structure. Each curve represents the difference of means of the average of 2,500 adaptive walks, i.e., 250 distinct performance landscapes with 10 adaptive walks on each. For parameter settings see Table I.

The second aspect to be discussed in detail is related to the performance effects of learning-based adaptation compared to the purely random-driven alterations. The results suggest that learning-based change is not universally more beneficial than purely random-driven organizational change. Rather, it appears that the aspiration level $v$ is of remarkable relevance: in both interaction structures, learning-based adaptation employing a high aspiration level leads to a level of final performance that is inferior to the performances achieved under purely random-driven change. Employing a low aspiration level performs best in the block-diagonal structure it leads to a medium performance in the case of high complexity.

As argued above, a high aspiration level induces more organizational alterations which leads to more diversity of search, i.e., more alterations of d, as compared to the low aspiration level. For highly complex interaction structures, a particular peril is that the search processes may stick to a local optimum, and, hence, increasing diversity of search "per se" may be beneficial. This might explain the good performance of the "change, no learning" mode. However, a high aspiration level, in a way, "penalizes" particularly those search processes which have reached a good solution from which further improvements are hard to achieve: as argued above, the block-diagonal interaction structure is particularly prone to this effect; however, the rather low performance in the full-interdependent structure (Figure 4) might also be caused by this effect.

With the third aspect to be discussed more into detail the intensity of organizational change (right-most column in Table II), and, thus, the *efficiency* of the mode of change and learning comes into play. The average number of organizational dimensions in which alterations occur during the adaptive search may be regarded as an indicator for the effort ("costs"), if any, of organizational dynamics.

Obviously, the context of the search organization is relevant for whether, or not, and, if so, in which shape costs of organizational change occur. For example, in case of a network of unmanned aerial vehicles, collaboratively serving a certain service area, the switch from one coordination mode to another might not cause any costs (apart from activating another of already available coordination mechanisms); however, in case of firm managers, collaboratively searching for better configurations of key performance drivers, reorganizations are rather costly, including, for example, learning costs of new organizational procedures or the adjustment of incentive schemes. Hence, the average number of dimensions changed may be rather critical for the efficiency of inducing organizational dynamics of search.

Results suggest that, in both interaction structures under investigation, learning with a low aspiration level yields good performance and a high level of efficiency as compared to the other scenarios:

In case of the block-diagonal interaction structure the final performance achieved with a low aspiration level exceeds the performance reached via purely random-driven change by about 7 points of percentage while the average number of organizational alterations is remarkably lower (i.e., 6.9 altered dimensions on average in case of learning with low aspiration level versus 38.2 in case of purely randomized change). If the complexity of the search problem is high the performance of the "change, no learning" scenario exceeds the performance of learning-based adaptation with low aspiration levels; however, this comes along with, on average, 37.9 organizational alterations compared to 6.2 alterations in the latter case.

In sum, it appears that learning with low aspiration level may provide rather high performance levels combined with few organizational alterations. Thus, whenever organizational alterations do not come along without any cost, learning with low aspiration level appears to be particularly interesting with respect to the efficiency of search.

## *B. Sensitivity Analysis*

In the sensitivity analysis, the baseline scenarios are modified with respect to the number of search agents: the simulations additionally are conducted for systems with two and with six search agents. In particular, the interactions among decisions remain unchanged, but the *assignment of decisions* is modified. Figures 5.a and 5.b show the assignment for the case of two agents and six agents, respectively, in the block-diagonal structure as compared to Figure 2.a.

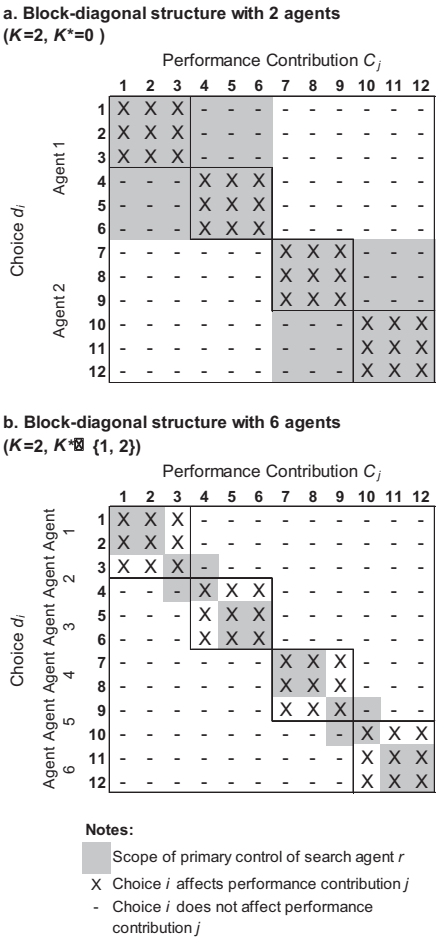



Fig. 5. Assignment of four independent sub-problems to a. two and b. six search agents.

In the simulation model, with increasing (decreasing) the number of search agents the diversity of search is increased (decreased): in each time step, each search agent discovers two alternatives to the status quo of the own partial sub-problem (Section II.A) – one alternative where one bit is flipped and another with two bits flipped. Thus, in case of two search agents, at maximum four bits of the entire configuration d could be flipped in time step $t$; in contrast, with six agents at maximum 12 bits could be flipped. Thus, with increasing number of agents the need for coordination is increased too, and viceversa.

Figures 6 and 7 plot the final performance $V_{t=500}$ achieved in the block-diagonal and the full-interdependent interaction structure, respectively, with two, four and six search agents.

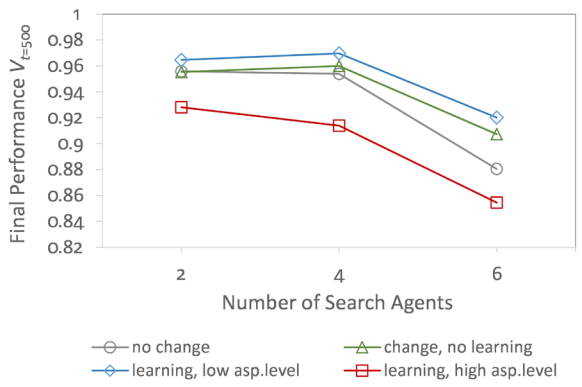



Fig. 6. Sensitivity of final performance to number of search agents in the block-diagonal interaction structure. Each mark represents the average of 2,500 adaptive walks, i.e., 250 distinct performance landscapes with 10 adaptive walks on each. For parameter settings see Table I.

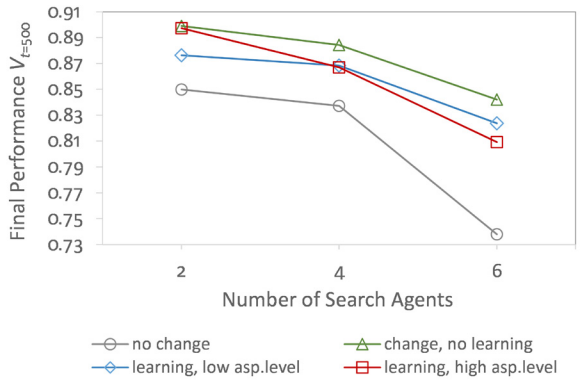



Fig. 7. Sensitivity of final performance to number of search agents in the full interdependent interaction structure. Each mark represents the average of 2,500 adaptive walks, i.e., 250 distinct performance landscapes with 10 adaptive walks on each. For parameter settings see Table I.

The results suggest that the "change, no learning" mode and learning with low aspiration level are least sensitive to the number of agents. In contrast, the final performance obtained by learning with high aspiration level varies considerably with the number of agents. However, the "no change" mode appears most sensitive to an increase in the number $M$ of search agents compared to the modes employing organizational change.

With the transition from two to four agents, the final performance shows rather slight de- or increases – depending on the mode of change and the interaction structure. However, with the transition from four to six agents the final performance obtained decreases remarkably in both interaction structures.

An interesting question is what might cause these effects. A reason might be given by the relation between assignment of decisions to search agents and the interactions among agents' decisions. For example, with six search agents in the block-diagonal structure (Figure 5.b), cross-agent interactions among search agents' choices occur whereas for two and four agents no cross-agent interactions show up (Figure 5.a.). Hence, in this interaction structure the need for coordination among agents' choices ranges from no need at all (i.e., $K^* = 0$ for $M = 2$ and $M = 4$) to some coordination need as captured by $K^* = 1$ or $K^* = 2$ (see Figure 5.b).

## V. Conclusion

The major finding of this study is that employing self-adaptation for the organizational setup of distributed search systems via reinforcement-based learning potentially leads to high levels of performance and this, in particular, with a rather high level of efficiency, as given by the extent of reorganization. These findings are particularly interesting when reorganizing the search system causes marginal costs – may it be due to learning of new organizational procedures on the agents' site or adjustments required in institutional arrangements.

However, the results also suggest that the complexity of the search problem together with the aspiration level considerably shapes the effects of reinforcement learning – which, at worst, may even be harmful if compared to refraining from any organizational alterations. These findings may sensitize the designer of a distributed search system to employing learning by reinforcement as the level of the aspired performance enhancements should not be overstretched in order to avoid "hyper-actively and ineffectively" alternating search systems. Moreover, the sensitivity analysis suggests that learning with high aspiration level is particularly sensitive to the need for coordination among search agents.

These findings emphasize the need for further research efforts. An obvious next step is to test the key idea of inducing learning-based organizational change in more practical settings than the one presented here. Though some preliminary results obtained for learning-based selection of the coordination mode in terms of the job scheduling policy employed by a swarm of unmanned aerial vehicles [23] provide some support for the ideas presented in this paper, further applications are definitely of interest.

Moreover, further studies should perform in-depth analyses of the role of the aspiration level and other parameters like the interval between of organizational alterations or the learning strength which were fixed in the simulation experiments presented in this paper. Furthermore, the basic search problem captured in this study is rather unstructured in terms of randomized performance contributions (apart from the structure of interactions); hence, in further research studies learning-based organizational adjustments of the search system may turn out to be even more beneficial in case of more structured search problems.

## References

[1] C. Yongcan, Y. Wenwu, R. Wei and C. Guanrong, "*An overview of recent progress in the study of distributed multi-agent coordination,*" IEEE Trans. on Industrial Informatics, vol. 9, pp. 427–438, Jan. 2013.

[2] T. Gross and B. Blasius, "*Adaptive coevolutionary networks: a review*," Journal of the Royal Society Interface, vol. 20, pp. 259–271, Mar. 2008.

[3] K. M. Carley and L. Gasser, "*Computational organization theory*," in: *Multiagent systems: A modern approach to distributed artificial intelligence*, G. Weiss, Ed., Cambridge: MIT Press, 1999, pp. 299–330.

[4] A. H. Bond and L. Gasser, "*Chapter 1 – Orientation,"* in *Readings in Distributed Artificial Intelligence,* A. H. Bond and L. Gasser (Ed.), San Mateo: Morgan Kauffmann, 1988, pp. 1–56.

[5] M. Wooldridge, *An Introduction to Multiagent Systems* (2nd ed.), Chichester: Wiley, 2009.

[6] F. von Martial, *Coordinating Plans of Autonomous Agents*, Lecture Notes in Artificial Intelligence, vol. 610. Berlin/Heidelberg/New York: Springer, reprint 2008 (1992).

[7] Y. Brun, G. Di Marzo Serugendo, C. Gacek, H. Giese, H. Kienle, M. Litoiu, H. Müller, M. Pezzè, and M. Shaw, "*Engineering Self-Adaptive Systems through Feedback Loops"*, in *Software Engineering for Self-Adaptive Systems,* BHC Cheng, R. de Lemos, H. Giese, P. Inverardi, and J. Magee, Ed., Berlin/Heidelberg: Springer, 2009, pp 48–70.

[8] O. Baumann, "*Distributed Problem Solving in Modular Systems: the Benefit of Temporary Coordination Neglect*," Systems Research and Behavioral Science, vol. 32, pp. 124–136, Jan./Feb. 2015.

[9] F. Wall, "*Organizational dynamics in adaptive distributed search processes: Effects on performance and the role of complexity*," in Frontiers of Information Technology & Electrical Engineering, vol. 17, pp. 283–295, Apr. 2016.

[10] F. Wall, "*Effects of organizational dynamics in adaptive distributed search*

*processes*," in *Distributed Computing and Artificial Intelligence, 12th Int. Conf.*, S. Omatu, Q.M. Malluhi, S. R. Gonzalez et al., Ed., Advances in Intelligent Systems and Computing vol. 373, Berlin: Springer, 2015, pp. 121–128.

[11] F. Wall, "*Beneficial Effects of Randomized Organizational Change on Performance*," in Advances in Complex Systems, vol. 18, 05n06:1550019, Nov. 2015.

[12] F. Wall, "*Self-adaptive organizations for distributed search: The case of reinforcement learning,*" in *Distributed Computing and Artificial Intelligence, 13th International Conference*, S. Omatu, A. Semalat, G. Bocewicz et al., Ed., Advances in Intelligent Systems and Computing, vol 474, Cham: Springer International Publishing, 2016, pp 23–32.

[13] S. A. Kauffman and S. Levin, "*Towards a general theory of adaptive walks on rugged landscapes*," Journal of Theoretical Biology, vol. 128, pp. 1145, Jan. 1987.

[14] S. A. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford: Oxford University Press, 1993.

[15] R. Li, M. M. Emmerich, J. Eggermont, E. P. Bovenkamp, T. Bäck, J. Dijkstra and J. C. Reiber, "*Mixed-integer NK landscapes*," in *Parallel Problem Solving from Nature IX,* T. Runarsson, H.-G. Beyer, E. Burke, J. Merelo-Guervós, L. D. Whitley and X. Yao, Ed., Lecture Notes in Computer Science, vol. 4193, Berlin: Springer, 2006, pp. 42–51.

[16] B. Levitan and S.A. Kauffman, "*Adaptive walks with noisy fitness measurements*." Molecular Diversity, vol. 1, pp. 53–68, Jan. 1995.

[17] F. Wall, "*The (beneficial) role of informational imperfections in enhancing organisational performance*," in *Progress in Artificial Economics, 6th Artificial Economics,* M. Li Calzi, L. Milone and P. Pellizzari, Ed., Lecture Notes in Economics and Mathematical Systems, vol. 645, Berlin: Springer, 2010, pp. 115–126.

[18] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed., Cambridge (Mass.): MIT Press, 2012.

[19] L. P. Kaelbling, M. L. Littman and A. W. Moore, "*Reinforcement learning: a survey,*" Journal of Artificial Intelligence Research, vol. 4, pp. 237–285, May 1996.

[20] R. R. Bush and F. Mosteller, *Stochastic Models for Learning,* Oxford (Engl.): Wiley, 1955.

[21] T. Brenner, "*Agent learning representation: Advice on modelling economic learning*," in *Handbook of Computational Economics*, vol. 2, L. Tesfatsion and K. L. Judd, Ed., Amsterdam: Elsevier, 2006, pp. 895–947.

[22] R. W. Rivkin and N. Siggelkow, "*Patterned interactions in complex systems: Implications for exploration*," in Management Science, vol. 53, pp. 1068–1085, July 2007.

[23] A. Lembacher, "*The Application of Learning Algorithms in Job Scheduling Problems*" Master Thesis (supervisor F. Wall), Universitaet Klagenfurt 2015 (unpublished).

**Friederike Wall** earned her Diploma for Business Economics in 1988 and her Doctoral degree in 1991, both at the Georg-August-Univesität Göttingen, Germany. In 1996 she received the "venia legendi" (Habilitation for Business Economics) from the Universität Hamburg, Germany. After being a scientific Project Referent for Accounting in the context of the implementation of SAP R/3 at Max-Planck Society, Munich, Germany she became Full Professor of Business Administration, esp. Controlling and Information Management at the Universität Witten/Herdecke, Germany. Since 2009 she is Full Professor and Head of the Department of Controlling and Strategic Management at the Alpen-Adria-Universität Klagenfurt, Austria. Prof. Wall's scientific work is focused on distributed decision-making, management control systems and the quality of the information provided by these systems. Her main research approach is defined by agent-based simulation methods and agent-based technologies.

# Electronic Health Record in Bolivia and ICT: A Perspective for Latin America

PhD Eugenio Gil López[1], PhD Karina Medinaceli Díaz[2]

[1]ESIT, International University of La Rioja, Logroño, Spain
[2]Mayor de San Andrés University, La Paz, Bolivia

***Abstract* — The emergence of new technologies in society through its application to many areas and very diverse realities is a clear element in the time in which we live. The health sector has been unable to escape this reality and has been renovated many of its traditional structures with new options brought by the application of information technology and communication (ICT) in areas such as management and hospital administration. This paper focuses on analyzing from the point of view of medical diagnosis the importance of electronic medical records as a unifying element of the information essential for this type of diagnosis, and the use of artificial intelligence techniques in this field. To this end the current situation of electronic medical records is analyzed in a country like Bolivia exhaustively analyzing three of the most important health centers. Is used for this unstructured interview experts on the subject reflect the current status of electronic medical records from the point of view of protection of the right to privacy of individuals and will serve as a model for development, not only in Bolivia but also in other Latin American countries.**

***Keywords* — Electronic Health Record, Artificial Intelligence, ICT, Medical Software, Health.**

## I. Introduction

Usually, when we combine the concepts of computer science and medicine, identification leads us to think about the use of information and computer systems for hospital management and patient management, as well as of the possibilities offered by telemedicine. These elements, while remaining important, are not really more than instrumental realities constantly evolving to meet the needs of healthcare centers and patients. Today, we must find the value of what some have described as Biomedical Informatics not in this kind of realities, but in the new fields created by Information Technologies, such as information management and data protection, and through them the possibility to assist the medical professional in conducting diagnostic assessments and decision-making, primarily through artificial intelligence techniques.

Information thus becomes the central element of medical practice. Each case must be individually analyzed, taking into account the different parameters that can influence each patient and only in that way can we achieve a minimally correct diagnosis. What we must do is systematize all that information and allow access to the various agents of the healthcare chain in order to obtain a complete medical history of the patient and to make decision-making as effective as possible.

All this information should be collected in an instrument that goes beyond mere medical history on paper, and should take advantage of all the benefits of the digital format. This brings us to the Electronic Health Record (hereinafter EHR), which means incorporating Information and Communication Technologies (ICTs) in the core of the healthcare activity, and results in the medical record ceasing to be a record of the information generated in the relationship between a patient and a medical professional or healthcare center, to become part of an integrated healthcare information system.

The analysis performed in EHRs is done from the point of view of protecting the fundamental right to privacy of individuals by establishing the measures that have been implemented in order to protect this right in the field of security. To do this we proceeded to use the unstructured interview to qualified personnel from different fields, which has come to reflect the current status of electronic health records. The various interviews with experts on the protection of personal data, information technology and communications (ICT) and bioethics, seek to gather their views, their consensus, their differences and predictions on the matter.

## II. The Electronic Health Record

To reinforce the progressive development of these systems, the Institute of Medicine (IOM) published a new report focusing on nine core features that an EHR should have in order to improve patient safety, achieve effective service delivery, facilitate the management of chronic diseases and improve efficiency [1]. These features were determined as follows:

- **Health information management.** An EHR should contain Health information management. An EHR should contain complete information about the patient's treatment history, any current diagnosis problems and their background, medications, allergies and any contacts with the healthcare provider. This includes clinical developments of patients' diagnoses in narrative text −entered by physicians, nurses, technicians, or through structured templates.
- **Results management**. It refers to the representation of laboratory and other complementary test results, such as images, pathological anatomy tests and others. Quick access to information on additional tests by all clinicians involved in treating a patient saves time and money, avoids redundancy and improves coordination of healthcare.
- **Medical order management**. The entry of orders, whether laboratory test results or other ancillary services requested, or the entry of prescriptions through computerized entry systems, is the first link in the healthcare chain, where the EHR is no longer a passive system but plays an active role in the patient's health. The system should contain a knowledge base that allows for a more efficient management of information and one which interacts with professionals to collaborate with their decisions.
- **Decision Support Systems**. Initially, support systems were in direct relation to order handling systems, supporting diagnoses, treatment and care through the use of alerts or reminders about potential interactions or problems. Its usage has expanded and now it covers a wide variety of functions.



DOI: 10.9781/ijimai.2017.4412

- **Electronic communication and connectivity systems**. In order to receive information from external auxiliary equipment and other systems, EHRs should allow communication via a standard messaging system as well as a common terminology.
- **Patient support**. Most EHRs provide output means for sending information to patients on health conditions, diagnostic tests or treatments. This information improves the doctor-patient relationship as well as the education of the latter.
- **Administrative processes**. Depending on the level of care provided, an EHR may be closely linked to administrative processes through the electronic scheduling of visits, the electronic submission of benefits, a verification of eligibility, the sending of automated messages of prescription renewal, the automatic registration of patients for research and artificial intelligence purposes.
- **Reporting and public healthcare systems**. New EHRs allow for automatic reporting to national databases. Other systems may allow for the enrollment of patients in clinical trials, providing the patient with information on how to follow a protocol.
- **Issuing of medical reports, medical discharge and consultation reports**. In addition to providing support for the management of medical orders and results, it should allow for multiple ways to display information and add data to various medical reports [2].

But in order to accomplish all these features an HER must comply with certain basic requirements, such as being complete, accessible, flexible, confidential and interoperable. In our view such an interoperability becomes a key factor in an EHR, as it must serve as a reliable tool for the exchange and use of the information entered, and able to assist the medical professional in the field of diagnosis. The EHR should not be understood therefore as an isolated entity. It requires information from other systems –in or outside the organization– so it needs to be developed taking into account the possibility of electronic data interchange. There are different levels of interoperability and ideally it is achieved through the use of standards [3].

From the point of view of health informatics, the *Institute of Medicine of the National Academies* (IOM) defines interoperability as "the ability of systems to work together, in general through the adoption of standards. Interoperability refers not only to the ability to exchange health information, but also to the need to understand the information that has been exchanged" [4].

One of the fundamental requirements for the implementation of e-Health systems is interoperability between systems, understood as the capacity of various systems or components to exchange information, understand such data and use them. Thus, information is shared and is accessible from any point of the health care network in which consultation is required and the consistency and quality of data across the system are guaranteed, with the consequent benefit for patient safety and the continuity of care. The centerpiece in the interoperability of systems is the use of «standards» that define the methods to carry out such information exchanges.

For the development of interoperability it is essential to consider the use of technology standards. In the development and implementation of EHRs there are many standards that can be used, among which we could cite those oriented to the exchange of data and electronic messaging, those oriented towards terminology, documents, the conceptual ones, and finally those that are architecture-oriented [5].

In general, an EHR system is a complex structure. EHR systems or services incorporate many elements of information, and today there are five main approaches that are competing to become the dominant platform for interoperability of an EHR:

- OSI (*Open Systems Interconnection*).
- CORBA (*Common Object Request Broker Architecture*).
- GEHR (*Good European Health Record*).
- HL7-CDA (*Clinical Document Architecture*).
- openEHR and the generic approach XML / Ontology.

Along with these requirements it may be necessary to overcome a critical obstacle in the idiosyncrasies of the medical community. Traditionally, physicians have considered ICTs unimportant both from a scientific and clinical point of view. Computerization programmes of medical records should take the peculiarities of clinical practice into account and facilitate the work of healthcare professionals without introducing new activities that are not essential; they should facilitate work, not complicate it; for example, a common mistake is to guide IT solutions towards data mining solutions aimed at managers rather than considering it a common instrument of clinical practice. Sometimes major projects with a large budget have failed due to not taking these issues into account [6].

On the basis of these considerations we can say that the process of decision making in the medical field is being closely scrutinized and solutions are emerging through the analysis and massive processing of data, or big data, and the application of artificial intelligence techniques of the Watson or IBM type. We are talking about an alternative utility to support decision making, which, when combined with proactive monitoring of the environment, ensures a better understanding of the problem context, which will increase the quality of decisions [7], in our case medical diagnoses. IBM's Watson cognitive system is more than mere big data. Its databases incorporate worldwide scientific literature, handbooks and physicians' desk references, and particular cases of patients included in their databases. These are cases where, if treatments have worked, they could serve as a precedent for similar instances, just as jurisprudence is applied in the judicial process; the system could provide oncologists with the assistance they need to make more informed decisions. The system understands the natural language used in the medical history and physicians' annotations. It will record all the variables included in the reports, and also has the ability to learn, and physicians may incorporate new variables.

From there, when an oncologist queries Watson for the most appropriate treatment, the system will offer several options recommended and others less recommended, so that the doctor will have the last word. Beside each treatment option, professionals can see why the system recommends such treatment[1].

In this connection, we could go a step further and record the knowledge and emotions that arise in the communication between patient and physician.

## III. Application to the Bolivian Case

Let us now analyze the situation of certain healthcare centers in Bolivia from the viewpoint of the implementation of EHR systems. We will focus on the most advanced companies in this field, such as the Health Insurance Fund of the Private Banking (*Caja de Salud de la Banca Privada, C.S.B.P*), the Military Social Insurance Corporation (*Corporación de Salud Militar*), and *Arco Iris* Hospital.

It should be noted that a general deficiency that occurs in Bolivia is the difficulty of access to information since it is not available from official websites and some institutions in the public or private sector do not have annual institutional memories to access information. In order to know the state-of-the-art of managing and archiving Medical Records on paper, the projects integrating Electronic Health Records and the implementation of Information and Communication Technologies

1 ConSalud.es. "Inteligencia Artificial que aprende como un humano para asesorar a los oncólogos". Ed. Grupo Mediforum. Online. http://consalud.es/tecnologia/inteligencia-artificial-que-aprende-como-un-humano-para-asesorar-a-los-oncologos-19108. [Accessed on 14/09/2016]

in the medical field, we proceeded to perform the open unstructured interviews cited above. In addition to the experts interviewed in the above-mentioned health centers and in order to provide a much more accurate picture of the situation, among the current ICT government actors we conducted interviews with the Agency for the Development of the Information Society in Bolivia (*Agencia para el Desarrollo de la Sociedad de la Información en Bolivia*, *ADSIB*) in order to collect information on the certificate and digital signature service as well as the activities being carried out for the implementation of the Electronic Government, the National Telecommunications Company (*Empresa Nacional de Telecomunicaciones*, ENTEL) for the implementation of Telecentres and service utilization of the Bolivian satellite Tupac Katari and the Directorate General of Electronic Government of the Ministry of Development Planning in charge of the socialization of the Electronic Government Plan. Among the ICT actors in the private sector, we conducted an interview with the Association of Private Banks of Bolivia (ASOBAN) in order to know about the digital certification service being provided since 2002.

The most interesting issues arising from these interviews on the matter at hand are reflected in the following lines. It is worth mentioning that not all of them have been included here due to lack of space, but the total number of participants who provided information has nevertheless been incorporated and these are shown in the table below, along with the issue dealt with and the dates thereof, in order to get an overview of the healthcare industry in Bolivia, in general, and in the city of La Paz, in particular.

### *A. Caja de Salud de la Banca Privada (CSBP)*

With over 16 years of experience −since 1988−, and presence throughout Bolivia −except for the Department of Pando− , it is widely acknowledged as the best company in providing medical and welfare assistance in the country, intelligently developing products and services that allow us to view the future with optimism, and with good growth projections in the region, managing to maintain their privileged position as the institution with the highest level of excellence in the provision of medical services in Bolivia.

It maintains a national central bureau located in the city of La Paz, with the main functions of planning, regulating, assessing, operating and monitoring policies, strategies, plans and healthcare programs, which are then to be implemented and managed from Regional offices.

The CSBP has computerized medical records of eight regional departments in Bolivia, except for the department of Pando.

Electronic health records in the CSBP are implemented vertically, that is, from top to bottom, and training plans are made. Physicians are trained first, but they are the professionals that spend less time training for their job. At the beginning there are complaints by the patient because the doctor is more concerned with filling the medical record than listening to the patient, so very often there is also a computer professional next to the doctor to solve questions of the physician about software management. Subsequently nurses and administrative staff are trained.

The implementation of electronic health records has been positive for the CSBP, it forces the physician to fill all the boxes that require information before closing the medical record. In the medical history on paper we often see problems such as being unsigned, unstamped, badly presented, the physician's handwriting is not understood, among others.

With electronic health records we have all the information of the insured population – i.e. Name, National Identity card, Organization, Rights enforcement at national level, and medical history of the patient.

The cost of *Software 9000* was US $ 150,000; it was acquired in 2002 from a Peruvian development company specialized in the healthcare area. From 2002 to 2004 we worked remotely with the Peruvian company; basically we had an outsourcing contract in place, which eventually proved costly for the CSBP due to its training, maintenance, new product development, implementation, and validation costs of approximately US $ 40,000 per year.

In 2005, the CSBP management assumed the development of a functional prototype, based on the experience of *Software 9000* acquired from the Peruvian company. This new prototype is tailored to the particular needs of the CSBP; to this effect, Alfa testing was performed in order to validate the information, and finally Beta testing was conducted and the product was fully operational and running smoothly with all the information of the insured [8].

### *B. Arco Iris Hospital (HAI)*

The first plans to build a clinic were born in 1997 at Arco Iris Foundation[2] as a result of the search for a system that would meet the healthcare needs of Street Children - NDLC. The first healthcare project was set up in a school in the central area of La Paz, where children living on the streets were treated on an outpatient basis.

In August 1998, the *Papstiiches Missionswerk fur Zinder* (PMK – Papal Missionary Work for Children) of Germany made a request for aid from the European Union for a five-year project. With this aid the Arco Iris Foundation planned the construction of a modern, secondary-care hospital, to be located in the city of La Paz. The construction of the hospital was completed in September 2001, the opening ceremony was held on September 27, and the opening of Arco Iris Hospital took place on October 23, 2001.

The 100-bed hospital shows a high standard of infrastructure and medical equipment and is considered one of the best hospitals in La Paz, with a workforce of nearly 270 people. It has 26 specialties and serves about 80,000 people annually, of which about 4,000 belong to street children from the main town. Beginning in 2008 the management implemented an Integrated Financial Management System – SIAF, along with the System for Statistical and Clinical Information - SICE, both developed by the Non- Governmental Organization *Medicus Mundi* and recognized by the Bolivian Ministry of Health and Sports through Ministerial Resolution No. 0853 dated November 18, 2005.

There are access profiles available for the use of SICE. The procedure for obtaining a medical history consists of retrieving a record from the system, printing it, and handling it over to the Archiving unit which seeks for the patient's medical record on paper and takes it to the Nursing Unit, who provide the different doctors' offices with the medical records. The process of recording entries and exits in the medical history is manual.

### *C. Military Social Insurance Corporation (COSSMIL)*

This organization has had bad experiences with the implementation of health information systems; in fact, in 2004 a Bolivian company developed the Hospital Management System (SIGEH), including software, equipment, and structured cabling; the system included twenty one modules, but it encountered problems, therefore the staff of

2 Arco Iris Foundation is a non-governmental organization, based on the principles of the Catholic Church, which since 1994 has been combating discrimination, marginalization and lack of opportunities suffered by thousands of children and youth: orphans whose parents are in prison, victims of domestic violence, abuse, rape, those living or working on the streets of La Paz - Bolivia. To fulfill its main purpose it runs a large number of projects of comprehensive support and conducts awareness campaigns among people with a spirit of solidarity and generosity. The Foundation relies almost entirely on the support of individual people, with no major funding from international or governmental cooperation. It seeks to be a sign of friendship, solidarity and cooperation in favor of the poor, the needy and vulnerable. Arco Iris Foundation (2015): "Hospital Arco Iris" [online]: http://www.arcoirisbolivia.org/mision.html [Retrieved on: 01/10/2014].

TABLE I. Interviews with members of sectors studied

| | INSTITUTION | THEME | NAME | DATE |
|---|---|---|---|---|
| **Public Subsector** | Ministry of Health | General Direction of Planning | Dr. Ronal Machaca. Head of Systems Unit | 21/07/2015<br>04/08/2015 |
| | National Health Information System (SNIS) | SNIS-VE | Rocco Abruzzese. Responsible Information and Production Services Level I | 17/06/2015<br>24/06/2015 |
| | | SNIS | Ing. Jorge Bailey e Ing. Gabriel Jiménez | 17/06/2015 |
| | | Primary Health Care Software (SOAPS) | Ing. Mauricio Bustillos | 23/07/2015 |
| | | Statistical Clinical Information System (SICE) | Ing. Gabriel Jiménez | 03/08/2015 |
| | | Statistical Clinical Information System (SIAF) | Ing. Gabriel Jiménez | 27/08/2015 |
| | Program VIH-AIDS | Departmental Monitoring and Reference Center (CDVIR) | Dr. David Segurondo Responsible CDVIR La Paz | 20/07/2015<br>27/07/2015 |
| **Social insurance subsector** | University Social Insurance - SSU | Health Record | Jaime Riveros<br>Biostatistics Responsible | 10/09/2015 |
| | | Use of Health Record | Lic. Sonia Apaza Head of Nurses | 14/10/2015 |
| | | Archive of Health Record | Lic. Elizabeth Saravia. Head of the Admission, Archive and Transfer Unit | 14/10/2015 |
| | Private Banking Health Bank - CSBP | Electronic Health Record | Dr. Gonzalo Maldonado<br>Hospital Regional La Paz Director<br>Peditrician | 04/03/2015<br>30/07/2015 |
| | | Electronic Health Record | Dr. David Martínez Médico Traumatologist | 22/07/2015<br>24/07/2015 |
| | | Medical Software and Medical Administrative System - SAMI | Dra. Tania Cherro<br>Responsible Medical Software | 25/08/2015 |
| | Military Social Security Corporation (COSSMIL) | Archive of Health Record | Lic. Katia A. de Auza<br>Head of Clinical Archive | 25/09/2015<br>30/09/2015 |
| | | Integrated Hospital Follow-up System for Patients - SISHAP | Coronel Grover Quiroga<br>Director Nacional de Sistemas | 04/11/2015 |
| **Private subsector** | Arco Iris Hospital | Electronic Health Record | Ing. Julio Alarcón<br>Head of Systems Unit | 04/08/2015<br>07/08/2015 |
| | | Electronic Health Record<br>– OPEN HAI | Dr. Igor Salvatierra | 01/09/2015 |
| | | Archive of Health Record | Mr. Rubén Heredia | 08/09/2015 |
| **Member ICT** | Agency For The Development Of The Information Society In Bolivia (ADSIB) | Digital Signature | Lic. Kantuta Muruchi Head of Planning and Projects | 14/08/2015 |
| | | Digital Signature | Ing. Sylvain Head of Innovation and Development Unit | 02/09/2015 |
| | Entel | Telecentres | Ing. Rolando Álvarez Head of Telecentres Unit | 28/07/2015 |
| | | | Ing. Wilson Cuellar<br>Rural Development Professional | 28/07/2015 |
| | Ministry of Development Planning | Directorate General of Electronic Government | Ing. Rodrigo Siles<br>General Director of Electronic Government | 29/09/2015 |
| | Association of Private Banks of Bolivia | Compensation and Compensation Chambers Settlement – ACCL S.A. | Ing. Ricardo Primintela<br>Systems administrator | 21/09/2015 |

the Informatics Directorate developed a new project design.

Since 2014 they have developed the Integrated Information System Control and Monitoring Hospital System (SISHAP), *Sistema de Información Integrado de Control y Seguimiento Hospitalario;* it currently has twelve modules, including the Outpatient module, which allows physicians access to the various electronic health records; physicians access their profile through user name and password; a menu option allows them to filter some diseases such as HIV-AIDS; for safety reasons the physician cannot change the information in the electronic health record until the scheduled appointment time has completed; if any changes are needed, these should be included in the comments field. The SISHAP has also been implemented in the cities of Sucre and Puerto Suarez.

With regard to information security, they have a CIO; a weakness is observed in regional healthcare centers, as some non-commissioned officers seen in the office are not computer professionals; there is also great staff mobility; they have redundant servers; backups are performed on a weekly basis; they use proprietary software, they pay for software licenses and antiviruses; the development of SISHAP modules is done in the form of consultancy per product; training in the use of SISHAP modules is performed by the IT Department and is provided both generally and specifically to doctors.

As a result of the interviews conducted little coordination is perceived among the IT Department, the Unit of Clinical Archiving, the Medical Audit and other COSSMIL units for the development of SISHAP modules, particularly with regard to electronic health records. The Clinical Archiving and Medical Audit units consider that for the moment the SISHAP does not preserve the confidentiality, privacy and security of health records on paper; they expect a change in the Technical Standard for the management of Medical History to recognize the probative value of electronic health records and electronic signatures.

## IV. Conclusions

In this study, we proceeded to evaluate the state of the art of electronic health records within the healthcare facilities of the National Health System in Bolivia and, in particular, in the city of La Paz. To that end, twenty unstructured interviews were conducted with key informants from the public sector, i.e. officials from the Ministry of Health (MoH), the National Health Information System (SNIS); the social security subsector, i.e. the Health Fund of Private Banking (CSBP), and the Military Social Insurance Corporation (COSSMIL); as well as with informants from the private non-profit subsector, i.e. Arco Iris Hospital. Since it is beyond our scope at this stage to propose solutions for the problems identified, we only aim at highlighting them, drawing attention to them, and finding new ways to develop and better manage electronic health records through the use of ICTs.

The Health Fund of Private Banking, *Caja de Salud de la Banca Privada (CSBP)* has been serving its insured patient population for 17 years. Since 2005, the CSBP professionals have developed their own Medical Software and Medical Management System (SAMI), which is currently being applied and used by outpatient services in the Policlinic, and also for Hospitalization nationwide. The SAMI has a configurable system of users and passwords with different access levels for each CSBP professional – i.e. physician, nurse, office clerk, pharmacy, laboratory, etc.–, so confidentiality, privacy and information security are ensured for each single patient; patient data cannot be modified or disclosed to any third parties whatsoever. This medical system easily integrates the areas of medical agendas and electronic health records.

An Electronic Health Record, i.e. Historia Clínica Electrónica, of the SAMI contains all of a patient's medical history organized in chronological order, with the date and time of care given, and it allows to sort data according to Insurance Scheme and Healthcare Programme; it can issue laboratory requests, auxiliary tests, medical prescriptions, referral orders, sick leaves, prenatal certificates, among many other documents. It can also get medical reports, comprehensive evaluation reports from medical teams, comparative studies among a wide range of variables available at the SAMI, and it can monitor the quality of the information recorded. Furthermore, SAMI can perform a variety of statistical analyses, medical research tasks, it can detect common risk factors and prevalent diseases among the insured population, and issue all kinds of reports which can be used to develop plans of action for the promotion of health and prevention of disease from the perspective of a Healthcare system that is much more efficient as long as it prevents the occurrence of disease rather than having a healing and remedial approach.

With regard to information security, the CSBP has a National Medical Officer and a National Director of Software Telematics, as well as redundant servers −i.e. mirrors− that perform backup copies; regional agencies submit their daily reports to the National Office, which centralizes all the information; They have an Intranet, and doctors, nurses and administrative staff they all have access to the Internet (except for some specific social networks, such as Facebook, Twitter, etc.); they use proprietary software, they pay for their antivirus software license; they do not perform any computer audits; and the Security Plan is pending final approval.

The *Arco Iris Hospital (HAI)* is a non-profit, private hospital, founded fourteen years ago, which implemented the so-called *openHAI* system two years ago, containing electronic health records (EHRs) of hospitalization, intensive care unit (ICU), emergency and outpatient (since September 2015); this system is a proprietary development by the HAI that had the collaboration of an Italian consultant.

Among the advantages offered by the *openHAI* we find the auto recording capability when the doctor fills in an EHR; being able to access the EHR from any device with Internet access –i.e. smartphone, tablet, laptop or computer SmartTV; It also features a Radiological Imaging Storage System (RIS) and the PACS, two systems where resonances, CT scans, X-rays and other radiological images are stored, and the ease of printing of the EHR since it is configured to use a printer server.

Among the disadvantages we could mention the backing-up of information from the EHR, which includes sensitive data, on external servers to the HAI Data Center, such as any cloud-based service, notably if it is a free service where the actual location of the EHR data is unknown; thus, accessing EHR data from any device with Internet access can go against the confidentiality of patient data contained in the EHR, as these can be easily seen by any third party other than the treating physician or medical personnel of the HAI. Finally, another disadvantage is the printing of electronic health records, where the physician must place his or her name, signature and stamp, because digital signatures have not yet been implemented on the *openHAI*.

## References


[1] Tang, P.C. y otros (2003). "Key Capabilities of an Electronic Health Record System: Letter Report" Washington, DC: Institute of Medicine. National Academy Press [en línea]: http://www.nap.edu/read/10781/chapter/1.

[2] González B., F. y Luna, D. (2012). "La historia electrónica" en Manual de Salud Electrónica para directivos de servicios y sistemas de salud. Publicación de las naciones Unidas enero de 2012 [en línea]: http://repositorio.cepal.org/bitstream/handle/11362/3023/S2012060_es.pdf?sequence=1.

[3] Institute of Medicine (2004). "Patient Safety: Achieving a news Standard for Care". National Academy Press [en línea]: http://www.nap.edu/read/10863/chapter/1#xvii.

[4] González B., F. y Luna, D. (2012). "La historia electrónica" en Manual de Salud Electrónica para directivos de servicios y sistemas

de salud. Publicación de las naciones Unidas enero de 2012 [en línea]: http://repositorio.cepal.org/bitstream/handle/11362/3023/S2012060_es.pdf?sequence=1.

[5] Kim, k. (2005). Clinical Data Standards in Health Care: Five Case Stuides. Oakland, CA: California Health Care Foundation [en línea]: http://www.chcf.org/publications/2005/07/clinical-data-standards-in-health-care-five-case-studies.

[6] Carnicero G., J., Rojas de la Escalera, D. y Blanco R., O. (2014). "La gestión de la función TIC en los servicios de salud: algunos errores frecuentes de los equipos de dirección" en Manual de Salud Electrónica para directivos de servicios y sistemas de salud. Volumen II Aplicaciones de las TIC a la atención primaria de salud" Publicación de las Naciones Unidas [en línea]: http://www.seis.es/documentos/X%20Informe%20SEIS%20-%20COMPLETO.pdf.

[7] Skyrius R, Kazakevičienė G, Bujauskas V. "From Management Information Systems to Business Intelligence: The Development of Management Information Needs". International Journal of Interactive Multimedia and Artificial Intelligence. 2013;2 (Special Issue on Improvements in Information Systems and Technologies), 31-37.

[8] Asawa K, Manchanda P. Recognition of Emotions using Energy Based Bimodal Information Fusion and Correlation. International Journal of Interactive Multimedia and Artificial Intelligence. 2014;2 (Special Issue on Multisensor User Tracking and Analytics to Improve Education and other Application Fields), 17-21.

[9] Caja de Salud de la Banca Privada (2015): "Memoria Anual 2014" [en línea]: http://portal.csbp.com.bo/inicio/attachments/article/1561/MEMORIA%20CSBP%202014.pdf.

**Eugenio Gil** was born in Galicia (Spain) in 1972. He has a degree in Law from Deusto University in Spain. The doctoral degree was obtained in 2014 at the Pontifical University of Salamanca. He is researcher in International University of La Rioja (UNIR) in the field of Information Technologies Law. He has been the Academic Secretary at the School of Engineering and Architecture in Pontifical University of Salamanca, and actually he is In Company Training Director in International University of La Rioja (UNIR).

**Karina Medinaceli Díaz** was born in Tarija (Bolivia) in 1973. She received her law degree from the Universidad Católica Boliviana in 2001, a master's degree in computer science and law from the Universidad Complutense de Madrid in 2002 (Spain), obtained her Ph.D. 2015 by the Universidad Pontificia de Salamanca - UPSA (Spain). From 2006 to the present he is Titular Lecturer of the subject of Computer Law in the Faculty of Law and Political Sciences of the Greater University of San Andrés (La Paz - Bolivia), currently also works in the Defensoría del Pueblo de Bolivia as Departmental Defensorial Delegate Tarija.