

Compare between CLA-EC and PSO-Great Deluge Mechanism as Feature Selection Methods in Facial Expression Recognition System

Akhtarhazratibishak^{*1}, Shirinkhezri², SajjadGhatei³

¹Department of Computer Engineering, Ahar Branch, Islamic Azad University, Ahar, Iran

²Department of Computer Engineering, Payame Noor University, PO BOX 19395-3697, Tehran, I.R of IRAN

³Department of Computer Engineering, Payame Noor University, I.R of IRAN

ABSTRACT

In this paper, we propose a new method of facial expression recognition based on local average binary patterns (LABP). In my method, Firstly, the LABP features are extracted from the original facial expression images. Then the best subset from LABP features find by CLA-EC and PSO-Great Deluge methods and each subset represented as a histogram descriptor separately. In features selection methods, the classification accuracy considered as fitness function. We use support vector machines (SVM) classifier for facial expression classification. We evaluate the efficacy of our approach by conducting experiments on benchmark dataset - JAFFE. Our representation yields improved facial expressions recognition rates relative to other methods. Additionally, experiments show that the proposed system using CLA-EC chooses a subset of features which will lead to the best results.

KEYWORDS: Facial expression, Feature selection, Local binary pattern, Cellular Learning Automata, Evolutionary Computing, support vector machines (SVM) classifier.

1. INTRODUCTION

Facial expression analysis plays a significant role in a variety of applications. Due to its potential applications, automatic facial expression recognition has attracted much attention over two decades (Tian et al, 2003; Fasel et al, 2003). Researchers are joined this area and a thorough survey of the existing works can be found in (Fasel et al, 2003; Pantic et al, 2000). From the survey, it is revealed that most of the facial expression recognition systems are based on the Facial Action Coding System (FACS) which is developed by Ekman and Friesen for describing facial expressions by action units (AUs) (Ekman, 1978). In another survey by Fasel and Luetttin (Fasel et al, 2003), the most prominent automatic facial expression analysis methods and systems are introduced and some facial motion and deformation extraction approaches are discussed. It is a system designed for human observers to describe changes in the facial expression in terms of visually observable activations of facial muscles.

Automatic facial recognition involves two vital aspects: facial feature representation and classifier design. Many classifiers have been applied to recognition, such as Nearest Neighbor (Akhshabi et al, 2011), neural network (NN) (Tian, 2004; Zhang et al, 1998), Support Vector Machine (SVM) (Bartlett et al, 2003; Littlewort et al, 2004), Linear Discriminant Analysis (LDA) (Lyons et al, 1999), and Bayesian network (Cohen et al, 2003). Because, a previous successful technique to facial expression classification is Support Vector Machine (SVM) (Bartlett et al, 2003; Bartlett et al, 2005; Valstar et al, 2005; Valstar et al, 2006), so we adopted SVM as alternative classifiers for expression recognition. While regarding feature representation, generally, there are two categories of feature representation: geometric features and appearance feature. Appearance features have been demonstrated to be better than geometric features, because geometric features are very sensitive to noises, especially illumination noise (Daugman, 1997). In recent years, local binary patterns (LBP) (Ojala et al, 1996), originally proposed for texture analysis and a non-parametric method efficiently summarizing the local structures of an image, have received increasing interest for facial expression representation. The most important property of LBP features is their tolerance against illumination changes and their computational simplicity. LBP has been successfully applied as a local feature extraction method in facial expression recognition (Shan et al, 2005; Shan et al, 2009; Moore et al, 2011).

Since the calculations within original LBP are performed between two single pixel values, it is much affected by small changes in the pattern and it is too local to be strength. In order to obtain better feature representation, in this paper we introduced a new method based on Uniform LBP called Local Average Binary Pattern (LABP) that employs a larger number of sample points. The proposed scheme could fully utilize the information of LBPs in multiple scales. Also, in my proposed method, Cellular Learning Algorithm Based Evolutionary Computing (CLA-EC) and PSO-Great Deluge are used to search the space of LABP patterns with the goal of selecting a subset of patterns encoding important information about the target concept of interest. The proposed method used in facial expression recognition systems. Support vector machines (SVM)

***Corresponding Author:** Akhtar Hazrati Bishak, 1Department of Computer Engineering, Ahar Branch, Islamic Azad University, Ahar, Iran, E-mail: a-hazrati@iau-ahar.ac.ir

classifier used for facial expression classification and classification accuracy considered as quality measure in CLA-EC and PSO-Great Deluge algorithms. Results show that the proposed method has much superiority comparing to classical LBP method and CLA-EC method finds a better subset of features than PSO-Great Deluge method.

The remainder of the paper was organized as follows: in section II we briefly reviewed local binary patterns and Local Average Binary Pattern (LABP) operators. In Section 3, the feature selection systems based CLA-EC and PSO-Great Deluge algorithms are described. To demonstrate the effectiveness of my proposed method, Experiments on benchmark dataset - JAFFE are presented In Section 4. Finally, the conclusions are given in Section 5.

2. FEATURE EXTRACTION

Ojala et al in 1996 (Ojala et al, 1996), introduced a powerful means of texture description namely Local Binary Pattern (LBP) operator as 3×3 square operator. This method works as follows: the 8 neighborhood of operator is compared with central pixel, if each of the eight neighboring pixels has values greater than or equal to central pixel value, so, takes value 1 otherwise take value zero. Finally, the central pixel replaces by the weighted binary sum of neighboring pixels value and the window with size of 3×3 is moved to the next pixel. The histogram of these values is used as a descriptor of the image texture. Figure (1) shows the local binary pattern Operator.

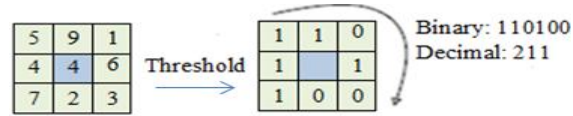


Figure 1 Local Binary Pattern operator

Two extensions of the original operator were made in (Ojala et al, 2002). 1- The LBP with neighborhoods of different sizes, that making it possible to handle textures at different scales. 2- The *uniform* LBP: an LBP that it contains at most one 0-1 and one 1-0 transition when viewed as a circular bit string. A good source of references for various extensions and reformations of the original LBP can be found in http://www.ee.oulu.fi/mvg/page/lbp_bibliography.

Since computation in original LBP, are performed between two single pixel values, so, it is too local to be powerful and any small changes in the pattern will affect edits. In order to obtain better feature representation, in (Hazratiet al, 2011) modified LBP that employs a larger number of sample points is offered which name is LABP. Calculation in this modified operator, are performed on average gray-values of pixels that is located in the window with size $n \times n$, instead of single pixel values. Also in LABP, the value of fixed threshold θ is computed based on the standard deviation value of P-neighbor values of each pixel in image (σ) by using (1).

$$\theta = \sigma \times \alpha \quad 0 < \alpha \leq 1 \quad (1)$$

Where: α is a scaling factor. In this paper, we use of LABP feature extraction method.

3. Feature selection

Feature selection is an important stage in pattern recognition systems. There are a number of advantages of feature selections: (1) dimension reduction to reduce the computational cost; (2) reduction of noises to improve the classification accuracy; (3) more interpretable features or characteristics that can help classify the expression more accurately. In this paper we use a two new methods based on Cellular Learning Automata-Computing Evolutionary (CLA-EC) and PSO-Great Deluge. The CLA-EC algorithm is an Evolutionary algorithm that is obtained combining from Cellular Learning Automata (CLA) and Computing Evolutionary concept (CA). The PSO-Great Deluge approach is obtained from combine particle swarm optimization (PSO) with great deluge algorithm. In this two methods classification accuracy considered as fitness function.

3.1 CLA-EC method

The CLA-EC model is Cellular Learning Automata (CLA) (Masoodiet al, 2007; Masoodifaret al, 2006) based on evolutionary computing (Rastegar et al, 2004), and like most other evolutionary computing algorithms, code the parameters of the search space in the form of genomes. Each genome has two major components: the genome's model and genome's string. The genome's model is a set of learning automata. Second component of the genome is the set of actions selected by the automaton. According to a local law, are in for cement signal vector is constructed for each cell, and then this vector based on a learning algorithm case the internal structure of the automaton is updated. Each cell in cellular learning automata based on evolutionary computing model produce a genome's string and calculates the fitness of genome. If the new genome has a better fitness than current cell's genome, so, new genome is replaced instead of current genome of the cell. This process continues until convergence to the desired solution. In this paper to simplify the algorithm, we assume that sight search space is a binary finite search space.

3.2 PSO method

Particle swarm optimization is a stochastic population based optimization method, introduced for the first time by Kennedy and Eberhart in 1995 (Kennedy et al, 1995; Kennedy et al, 1995). PSO method works as

each member of the population is called a ‘‘particle’’, and each particle flies with a constantly updated velocity in the multidimensional search space. The particle’s own experience and the experience of particle’s neighbors or the experience of the whole swarm are exploited in the velocity updates. This has been applied in several contexts, such as artificial neural network training, optimizations, pattern classification and fuzzy system control. In comparison to other similar methods, it is noticeable that PSO is rapidly converging towards an optimum as well as convenient computations and implementation as opposed to those of genetic algorithm (e.g., coding/decoding, crossover and mutation). However, PSO have shown some disadvantages. Sometimes it is easy to be trapped in local optima, and the convergence rate decreases considerably in the later period of evolution. Closing a solution near to optimal solution, the algorithm stops optimizing and this limits the accuracy the algorithm may achieve (Yang et al, 2007).

3.3 Great Delugemethod

Dueck proposed Great deluge algorithm in 1993 (Dueck, 1993). The method is a comprehensive approach for solving optimization problems. Like other local search methods, this method also replaces common solution (*New_Config*) with best results (*Best_Config*) that have been found by then. The process continues until stop conditions is met. In this algorithm, new candidates are selected from neighbors. Selection strategy is different from other approaches. Great deluge algorithm selects results among values equal or better than (for optimization problems) the value of *Water Level (WL)*. Value of *WL* also rises at a steady pace at each step. Increasing of *WL* continues until value of *WL* reach to the best result achieved by then. In this step, the algorithm is repeated for some times and if not a better answer comes up, it ends. The initial value of *WL* is equal with the primary solutions (*f(s)*).

3.4 PSO-Great Delugemethod

A new hybrid model of the particle swarm optimization algorithm and great deluge algorithm is presented in (Ghatei et al, 2007) referred as MPSO. In normal PSO, achieving a new answer, the new answer is compared with the best answer found so far and in case of being better; the new answer will be accepted. Whereas in the new algorithm first the new found answer is compared by the best found solution so far then if it is better, the comparison continues with another parameter called ‘‘*Water Level*’’ or *WL*. If it is better than the both, it is accepted as new solution. In fact, inside of the PSO there is a level of acceptance for new answers and this procedure gives a second chance to trapped particles to be able to get rid of local optima’s. Depending to the nature of the problem in the sense of being minimum or maximum, value of this level of acceptance decreases or increases over the time. This algorithm differs from the normal PSO, since MPSO tries to exploit the basic functionality of great deluge local search in the PSO.

3.5 Feature selection with CLA-EC and PSO-Great Deluge method

In high space of features that each dimension represents the one type of feature, a subset of feature is a point in the search space. If n is a number of all feature then there are 2ⁿ subset of feature, that each of them are different in size and content of features. Best points are subsets that have high classification accuracy.

3.5.1 Feature selection encoding using CLA-EC and PSO-Great Deluge method

In CLA-EC: In this paper like (Hazratiet al, 2012) the binary form of CLA-EC is used to solve feature selection problem. In this encoding scheme, the genome is a bit string whose length (n) is determined by the number of features. Each feature is associated with one bit in the string. If the *ith* bit is 1, then the *ith* feature is selected, otherwise, that feature is ignored. Thus each chromosome represents a different subset of features.

In BPSO: Binary form of PSO (BPSO) was presented by Kennedy and Eberhart in 1997 (Kennedy et al, 1997). Unlike to the standard PSO, it is able to optimize the discrete spaces. In BPSO, each particle considers as a binary string(Figure3).Each feature is associated with one bit in the string. In this string, if the *ith* bit is 1, then the *ith* feature is selected, otherwise, that feature is ignored. Equation to update the particle's position in standard PSO is corrected like EQ. (2):

$$S(V_{id}^{k+1}) = \frac{1}{1 + \exp(-V_{id}^{k+1})}$$

If $s(V_{id}^{k+1}) > rand$ then $X_{id}^{k+1} = 1$ else $X_{id}^{k+1} = 0$ (2)

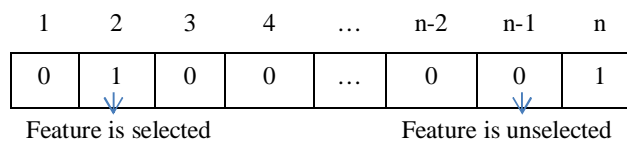


Figure (3): A binary string for each particle

Where: rand is a random number that distributed uniformly in the interval[0, 1].To avoid saturating of sigmoid function, have recommended that the speed is limited in the interval[- 4, 4] (Franken et al, 2005). Proposed system using BPSO chooses a subset of features which will lead to the best results.

3.5.2 Feature subset Fitness evaluation

Acquired features from CLA-EC and PSO-Great Deluge algorithms were classified by Support Vector Machine (SVM) classifiers (Vapnik, 1998) and classification rate is considered as evaluation measure. It should be noted that individuals with higher classification rate will outweigh individuals with lower classification rate. Overall, higher classification rate implies higher fitness.

4. EXPERIMENTS AND RESULTS

The proposed system designed and tested with Japanese Female Facial Expression (JAFPE) Database (Lyons et al, 1998). The database contains 213 images in which ten persons are expressing three or four times the seven basic expressions (happiness, sadness, surprise, anger, disgust, fear, neutral). Sample images from the database are shown in Fig. 3.



Figure 3 Samples from the Japanese Female Facial Expression Database

The CSU Face Identification Evaluation System provides standard face recognition algorithms and standard statistical methods for comparing face recognition algorithms. In our experiments image pre-processing is conducted by the pre-processing subsystem of the CSU Face Identification Evaluation System (Bolme et al, 2003). The images are registered using eye coordinates and cropped with an elliptical mask to exclude no face area from the image. As a result, the size of each preprocessed image is 150 × 128(see Fig. 4).

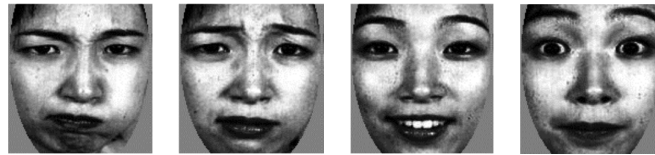


Figure 4 Samples from the preprocessed images

The goal of feature subset selection is to use fewer features to achieve the same or better classification performance. Classification accuracy mostly measured by Receiver Operator Characteristic (ROC) curve. The ROC is a metric used to check the quality of classifiers based on number of features. For each class of a classifier, roc applies threshold values across the interval [0, 1] to outputs. For each threshold, two values are calculated, the True Positive Ratio (the number of outputs greater or equal to the threshold, divided by the number of one targets), and the False Positive Ratio (the number of outputs less than the threshold, divided by the number of zero targets).The ROC chart based on number of LABP features selected by CLA-EC algorithm, PSO-Great Deluge algorithm and based on whole features of LABP operator is shown in figure 5.As is evident from the curves obtained, features selection using CLA-EC algorithm has a better recognition rate than use of PSO-Great Deluge algorithm.

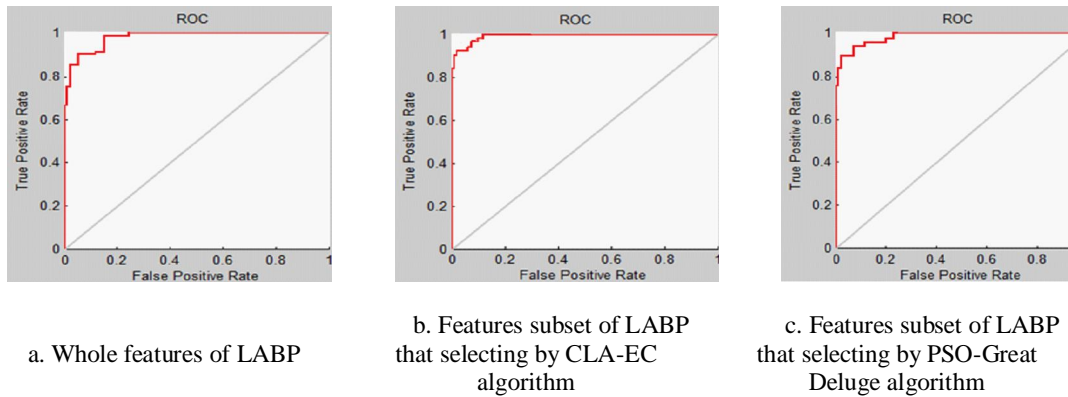


FIGURE5 Classification accuracy measured by ROC curve using a. Whole features of LABP and Features subset of LABP that selecting by b. CLA-EC algorithm& c. PSO-Great Deluge algorithm

To compare the recognition performance to other published methods, we divide the JAFFE database into ten roughly equally sized sets: nine sets are used for training and the remaining one for testing. The above process is repeated so that each of the ten roughly equally sized sets is used once as the test set. The average result over all ten cycles is considered as the recognition rate of one trial. The scaling factor (α) in LABP operator is set to 0.1 in the all experiment. Performance of classic LBP operator for facial expression recognition using SVM classifier with different kernels is shown in Table 1, and Performance of my proposed method based on two feature selection methods (CLA-EC and PSO-Great Deluge) showed in Table 2. The comparative analysis of our result confirms the robustness of the proposed approach. In my proposed method using both feature selection methods we have achieved higher recognition accuracy than classic LBP operator in facial expression system.

Table 2 facial expression recognition rate using classic LBP operator by SVM with different kernels

SVM with different kernels	6-Class recognition (%)	7-Class recognition (%)
	LBP[31]	LBP[31]
SVM (linear)	91.45	84.30
SVM (polynomial)	91.73	84.36
SVM (RBF)	92.86	85.67

Table 2 facial expression recognition rate using my proposed method based on two feature selection methods (CLA-EC and PSO-Great Deluge) using SVM with different kernels

SVM with different kernels	6-Class recognition (%)		7-Class recognition (%)	
	LABP based PSO-Great Deluge	LABP based CLA-EC	LABP based PSO-Great Deluge	LABP based CLA-EC
SVM (linear)	92.52	93.23	86.43	87.74
SVM (polynomial)	92.73	93.81	87.62	88.25
SVM (RBF)	94.23	94.63	91.60	91.07

To further explore the recognition accuracy per expression when my proposed method based on two feature selection methods (CLA-EC and PSO-Great Deluge) performs best, Figure 6 shows the 6-class and 7-class facial expression recognition results obtained by my proposed method. We can see that except two expressions, *i.e.*, sadness and fear, other expressions are classified well with an accuracy of more than 90%. In particular, sadness is recognized with the lowest accuracy than other expressions since sadness is highly confused to neutral and fear. Further in most cases my proposed method based on CLA-EC feature selection method lead to the best results than PSO-Great Deluge method.

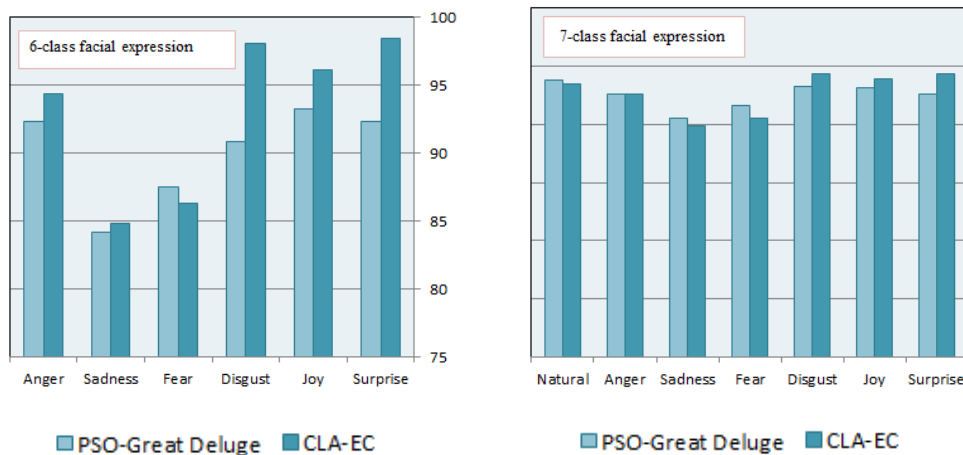





FIGURE6 recognition accuracy of my proposed method based on two feature selection algorithms (CLA-EC and PSO-Great Deluge) using SVM (RBF) on 6 and 7-class facial expressions

In many real-world applications, the resolution of face images is usually low due to various factors such as surrounding environment and imaging equipment. Therefore, it is important to evaluate the recognition performance of the proposed approach also under low-resolution conditions. We further evaluated the LBP-based algorithm over a range of image resolutions, investigating its performance against low-resolution images. The recognition rates are summarized in Table 3. We observed that the proposed LBP-based method is more effective for face images with low resolutions than classic LBP methods.

Table 3 Comparisons on image resolutions. The first row: classic LBP algorithm; The second row: our LBP-based algorithm

Different Resolutions			
	48 × 48	32 × 32	16 × 16
LBP[31]	83.09	79.22	69.72
My proposed method	90.43	88.12	84.32

5. Conclusion

In this paper, a novel and efficient facial expression recognition system is proposed. The Local Binary Pattern (LBP) has been proved to be effective for image representation, but it is too local to be robust. We enhance the classic LBP method base on facial expression recognition that gives our descriptor more comprehensively: In my method, the calculations is perform based on average gray-values of pixels values within windows with scale S , and we use of standard deviation values of these pixels, Instead of employing a fixed threshold. For feature selection we used two methods; Cellular Learning Automata- Evolutionary Computing (CLA-EC) and PSO-Great Deluge algorithm. The CLA-EC algorithm is an Evolutionary algorithm that is obtained combining from Cellular Learning Automata (CLA) and Computing Evolutionary concept (CA). The PSO-Great Deluge approach is obtained from combine particle swarm optimization (PSO) with great deluge algorithm. In this two methods classification accuracy considered as fitness function.

We evaluated our method on JAFFE database, demonstrating its robustness background clutter and good categorization accuracy even without exploiting geometric information. The results showed that the LBP method producing the largest improvement in the classification accuracy and in the discrimination between different facial expressions. Further experiment show that proposed system using CLA-EC chooses a subset of features which will lead to the best results.

REFERENCES

1. Tian, Y., Tian, L. Brown, A. Hampapur, S. Pankanti, A. Senior, and R. Bolle, (March 2003), "Real world real-time automatic recognition of facial expression", in *IEEE PETS*, Australia.
2. Fasel, B., and J. Luetttin, (2003), "Automatic facial expression analysis: a survey", In *Pattern Recognition*, vol.36, pp.259- 275.
3. Pantic, M., and L. Rothkrantz, (2000), "Automatic Analysis of Facial Expressions: The State of the Art", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1424-1445, Dec.
4. Ekman, P., and W.V. Friesen, (1978), "The facial action coding system: a technique for the measurement of facial movement", In *Consulting Psychologists Press, Inc., San Francisco, CA*.
5. Akhshabi, M., Akhshabi, M., and J. Khalatbari, (2011), "A New Method for Face Recognition with Nearest Neighbor in 3D Images", in *J. Basic. Appl. Sci. Res.*, Vol. 10, No. 1, pp. 1411-1414.
6. Tian, Y.,(2004), "Evaluation of face resolution for expression analysis", in *IEEE Workshop on Face Processing in Video*.
7. Zhang, Z., M. J. Lyons, M. Schuster, and S. Akamatsu, (1998), "Comparison between geometry-based and gabor-wavelets-based facial expression recognition using multi-layer perceptron", In *IEEE FG*, April.
8. Bartlett, M.S., G. Littlewort, I. Fasel, and R. Movellan, (2003), "Real time face detection and facial expression recognition: Development and application to human computer interaction", In *CVPR Workshop on CVPR for HCI*.
9. Littlewort, G., M. Bartlett, I. Fasel, J. Susskind, and J. Movellan, (1999), "Dynamics of facial expression extracted automatically from video", in *IEEE Workshop on Face Processing in Video*, 2004.
10. Lyons, M. J., J. Budynek, and S. Akamatsu, "Automatic classification of single facial images", *IEEE PAMI*.
11. Cohen, I, N. Sebe, S. Garg, L. S. Chen, and T. S. Huang, (2003), "Facial expression recognition from video sequences: temporal and static modelling", *Comput. Vis. Image Understand.*, vol. 91, pp. 160-187.
12. Bartlett, M.S., G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan, (2005), "Recognizing facial expression: machine learning and application to spontaneous behavior", in: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
13. Valstar, M., I. Patras, M. Pantic, (2005), "Facial action unit detection using probabilistic actively learned support vector machines on tracked facial point data", in: *IEEE Conference on Computer Vision and Pattern Recognition Workshop*, vol. 3, pp. 76-84.
14. Valstar, M., M. Pantic, (2006), "Fully automatic facial action unit detection and temporal analysis", in:

- IEEE Conference on Computer Vision and Pattern Recognition Workshop, p. 149.
15. Daugman, J., (1997), "Face and gesture recognition: overview", IEEE Trans, Pattern Anal. Mach. Intel, Vol. 19, No. 7, pp. 675–676.
 16. Ojala, T., M. Pietikäinen, D. Harwood, (1996), "A comparative study of texture measures with classification based on feature distributions", Pattern Recognition, vol. 29, pp. 51–59.
 17. Shan, C.; Gong, S., McOwan, P., (11–14 September 2005), "Robust facial expression recognition using local binary patterns", In *Proceedings of IEEE International Conference on Image Processing (ICIP)*, Genoa, Italy, pp. 370-373.
 18. Shan, C., Gong, S., McOwan, P., (2009), "Facial expression recognition based on local binary patterns: A comprehensive study", *Image Vis. Comput*, 27, 803-816.
 19. Moore, S., Bowden, R., (2011), "Local binary patterns for multi-view facial expression recognition", *Comput. Vis. Image. Und*, 115, 541-558
 20. Ojala, T., M. Pietikäinen, and T. Mäenpää, (2002), "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971-987.
 21. http://www.ee.oulu.fi/mvg/page/lbp_bibliography
 22. Masoodi, B., M. R. Meybodi and M. Hashemi, (2007), "Cooperative CLA-EC", Proceedings of 12th Annual CSI Computer Conference of Iran, ShahidBeheshti University, Tehran, Iran, pp. 558-559, Feb. 20-22.
 23. Masoodifar, B., M. R. Meybodi, and R. Rastegar, (2006), "Asynchronous CLA-EC", Proceedings of 11th Annual CSI Computer Conference of Iran, Fundamental Science Research Center (IPM), Computer Science Research Lab., Tehran, Iran, pp. 447-458, Jan. 24-26.
 24. Rastegar, R., and M. R. Meybodi, (2004), "A New Evolutionary Computing Model based on Cellular Learning Automata", to appear in proceedings of IEEE International Conference on Cybernetics and Intelligent Systems, Singapore.
 25. Kennedy, J., R.C. Eberhart, (1995) "Particle Swarm optimization", in: Proc. of IEEE Int. Conf. on Neural Networks, Perth, Australia, pp. 1942–1948.
 26. Kennedy, J., R.C. Eberhart, (1995) "A new optimizer using particle swarm theory", in: Proc. of the Sixth Int. Symp. on Micro Machine and Human Science (MHS'95), Nagoya, Japan, pp. 39–43.
 27. Yang, X., J., Yuan and H., Mao, (2007), "A modified particle swarm optimizer with dynamic adaptation", Applied Mathematics and Computation, 189: 1205–1213.
 28. Dueck, G., (1993), "New optimization heuristics: The great deluge algorithm and the record-to-record travel", Journal of Computational Physics, vol. 104, pp. 86–92.
 29. Ghatei, S., R.P. Khajei, M. SmimiMaman and M.R. Meybodi, (2012), "A Modified PSO Using Great Deluge Algorithm for Optimization", J. Basic. Appl. Sci. Res., 2(2), PP. 1362-1367.
 30. Hazrati, A., Faez, K., Hazrati, M., (2012), "A New Feature Selection Method based on CLA-EC in Face Recognition System", The Journal of Mathematics and Computer Science Vol. 4 No.2, pp.159 –171.
 31. Kennedy J, Eberhart R., "A discrete binary version of the particle swarm algorithm", Proc. Conf. Systems, Man, and Cybernetics 1997:4104-9.
 32. Franken N, Engelbrecht A., "Investigating binary PSO parameter influence on the knights cover problem", IEEE Cong. Evolutionary Computations 2005; 1: 282-9.
 33. Vapnik, V. N., (1998), "Statistical Learning Theory", Wiley, New York.
 34. Lyons, M., S. Akamastu, M. Kamachi, and J. Gyoba, (1998), "Coding facial expressions with Gabor wavelets", In: Third IEEE Conference on Face and Gesture Recognition, pp. 200-205.
 35. Bolme, D., M. Teixeira, J. Beveridge, B. Draper, (2003), "The CSU face identification evaluation system user's guide: its purpose, feature and structure," In: Third International Conference on Computer Vision Systems, pp. 304 –31.
- HazratiBishak, A., Z. Ghandriz, T. Taheri, "Face Recognition using Co-occurrence Matrix of Local Average Binary Pattern (CMLABP)", Cyber Journals: Multidisciplinary Journals in Science and Technology, Journal of Selected Areas in Telecommunications (JSAT), October Edition, 2011.