Mindanao Journal of Science and Technology Vol.15 (2017) 35-53

An MSULBP Features Selection based on GA and Majority Voting Mechanism in Facial Expression Recognition

Jafar Sheykhzadeh[1 *], Akhtar Hazrati Bishak[2] and Mohammad Bagher Moradi Gheshlagh[3]

[1]Department of Computer Engineering
Islamic Azad University-Ahar Branch
Ahar, Iran
[*]*j-sheykhzadeh@iau-ahar.ac.ir*

[2]Young Researchers and Elites Club, Science and Research Branch
Islamic Azad University
Tehran, Iran

[3]Department of Computer Engineering
Islamic Azad University-Shabestar Branch
Shabestar, Iran

Date received: February 25, 2017
Revision accepted: June 1, 2017

Abstract

Because of excellent capability of description of local texture, local binary patterns (LBP) have been applied in many areas. Also, to extract individual features the efficacy of the uniform LBP has been validated. In this study, a proposed new facial expression recognition system based on Multi-Scale Uniform LBP (MSULBP) schema can fully utilize LBP information. In the previous works, the selection of the optimal subset of the extracted features has not been considered. But in the proposed algorithm, the MSULBP features were extracted from the original facial expression images. The best subset from MSULBP features was found by genetic algorithm (GA) and majority voting mechanism (MVM) and was represented as a histogram descriptor. Finally, support vector machines (SVM) classifier was used for facial expression classification. The experimental results on the popular Japanese female facial expression (JAFFE) database illustrate that the presented facial expression recognition method based on MSULBP obtains the best recognition accuracy rate. Additionally, experiments show that the MSULBP features are robust to low-resolution images, which is critical in real-world applications.

***Keywords**:* features selection, support vector machines classifier, facial expression, local binary pattern, genetic algorithm

---

1. Introduction

Facial expression analysis plays a significant role in a variety of applications (Tian *et al.*, 2003). Due to its potential applications, automatic facial expression recognition has attracted much attention (Fasel and Luettin, 2003) over two decades ago. More and more researchers have joined this area (Kato *et al.*, 1991; Cottrell and Metcalfe, 1991; Rosenblum *et al.*, 1994) and a thorough survey of the existing works can be found found. According to Pantic and Rothkrantz (2000), most of the facial expression recognition systems were on the basis of facial action coding system (FACS) which was developed by Ekman and Friesen (1978) for describing facial expressions by action units (Aus). In another survey by Fasel and Luettin (2003), the most prominent automatic facial expression analysis methods and systems were introduced and some facial motion and deformation extraction approaches were discussed. This system describes the changes in the facial expression according to visually observable activations of facial muscles.

Recently, the research on facial expression recognition has been performed using dynamic facial image sequences. Pantic and Rothkrantz (2000) surveyed the work done in automating facial expression analysis in facial images and image sequences. According to psychologists (Bassili, 1979), analysis of sequences of images produces more accurate and robust recognition of facial expressions than using only single frames. Kotsia *et al.*, (2007) applied facial wire frame model and a support vector machine (SVM) was used for classification. In the other work which Performed by Zhang *et al.*, (2005), the Dynamic Bayesian networks (DBNs) was applied for facial expression recognition. In the method, R illumination camera was proposed for facial feature detection and tracking. Discrete hidden Markov models (DHMMs) proposed by Yeasin *et al.*, (2007) to recognize the facial expressions. Anderson *et al.*, (2006) used the multichannel gradient model (MCGM) to reveal facial optical flow. The other works that was done in the facial expressions Recognize field, Pantic *et al.*, (2006), achieved 84.9% accuracy rate. Face-profile-contour exploring and rule-based reasoning were used to identify 20 Aus occurring only or in a combination in approximately left-profile-view face image sequences.

Automatic facial expression recognition involves two vital aspects: facial feature representation and classifier design. Many classifiers have been applied to expression recognition such as neural network (NN) (Tian, 2004), SVM (Littlewort, *et al.*, 2004), linear discriminant analysis (LDA), and Bayesian network (Cohen, *et al.*, 2003). Since the previous successful technique to facial expression classification was SVM (Valstar, *et al.*, 2005), the study adopted SVM as an alternative classifiers for expression recognition. Generally, there are two categories under feature representation: geometric features and appearance feature. Appearance features have been demonstrated to be better than geometric features, because the latter are very sensitive to noise, especially illumination noise. Also, texture features have an invariant pattern locally and they are resistant against changes like pose, illumination, occlusion, and expression. Proposed by Ojala *et al.*, (1996), LBP is a powerful means of texture description. After it was extended to Unicode LBP, it was used in many areas because of its high, efficient code way and low, excellent local texture description. Recently, Ahonen *et al.*, (2006) used the LBP operator used to represent the face in static images. In this procedure, the face image was divided into multiple regions and then the LBP operator was applied to extracted feature of each region and the feature of each region was concatenated into an enhanced feature vector. This approach was verified to be a growing success.

Uniform patterns play a significant role in texture classification. Their priority has been shown in face recognition (Ojala, *et al.*, 2002). However, as the percentage of non-uniform patterns in a large-scale ones increases, the use of the uniform patterns causes much information loss. Some works of He *et al.*, (2005) were suggested to address this issue. This paper has proposed a multi-scale uniform local binary patterns (MSULBP) scheme for face recognition. The proposed scheme could fully utilize the information of non-uniform LBPs in big scale. The study used GA (Montazeri-Gh, and Mahmoodi-K, 2016) and MVM to select the optimum features of MSULBP for classification. The proposed method used in facial expression recognition systems on MSULBP features and classification accuracy and number of unselected feature (zeros) was a quality measure. The ensemble of supervised SVM classifier was trained using best subset of MSULBP features and its variants which give a robust system which was capable of better generalization accuracy.

2. Methodology

2.1 Feature Extraction

The feature extraction phase represents a key component of any pattern recognition system. A feature extraction algorithm encodes image properties into a feature vector and a similarity function computes the similarity between two images as a function of the distance between their feature vectors. In recent years, LBP has received increasing attention for facial description.

2.1.1 Local Binary Pattern

Introduced by Ojala *et al.* (1996), all the original LBP operator is a powerful means of texture description. The operator labels the pixels of an image by thresholding the 3x3-neighbourhood of each pixel with the center value and converts the result into a binary number by using (1).

$$LBP_{P,R(x,y)} = \sum_{p=0}^{P-1} s(g_p - g_c)2^P$$

$$s(x) = \begin{cases} 1 & if \quad x \geq 0 \\ 0 & if \quad x < 0 \end{cases} \tag{1}$$

Where g_c is an intensity of central pixel and g_p is a gray level intensity of neighborhood pixel and 2^p is a relevant factor for any neighborhood. S (.) is a sign function. The process is demonstrated in Figure 1.

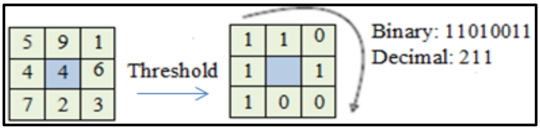

Figure 1. The basic LBP operator

Two extensions of the original operator were made by Ojala, *et al.* (2002). The first is defined LBPs for neighborhoods of different sizes, thus making it possible to handle textures at different scales. The use of circular neighborhoods and bilinearly interpolating the pixel values allows any radius and number of pixels in the neighborhood. In this extension, P sampling points on a circle of radius of R, are shown to form a (P, R).

The second is defined as the so-called uniform patterns: an LBP is uniform if it contains at most one 0-1 and one 1-0 transition when viewed as a circular bit string. For example, 00000000, 00011110 and 10000011 are uniform patterns. Uniformity was important because it characterizes the pieces that include primitive structural information such as edges and corners. According to Ojala *et al.* (1996), less than 90 % of all patterns are uniform patterns when using the (8, 1) neighborhood and approximately 70 % in the (16, 2) neighborhood. There are various extensions and reformations of the original LBP following its first introduction by Ojala *et al.* (2006).

2.1.2 Multi-scale Uniform LBP Schema

A single LBP operator has a restricted efficiency. Further image features could be displayed by multi-scale or multi-resolution under various settings (Chan, 2008). The framework of multi-scale approximation, called multi-resolution analysis, has been developed by the computer vision, image analysis and signal processing communities with complementary motivations regarding physics and biological aspects. The motivation for having a multi-scale representation of the face image comes from the basic observation that real-world objects are composed of different structures at different scales. Nevertheless, the effectiveness of the uniform LBP has been confirmed to extract representative features (Ahonen *et al.*, 2006; Zhang *et al.*, 2004), but in small scale operators of uniform LBP, the beneficial information of non-uniform patterns will be ignored. In this paper, to overcome the limitations of LBP, and apply it to face recognition; a simple but powerful texture representation, called multi-scale local binary pattern, was proposed for face recognition. This multi-resolution, representation-based uniform LBP was obtained using a window with three scales. The principal advantage of the proposed multi-scale LBP operator is that it can fully take advantage of LBP information without requiring any training phase (Notice that training phase may be sensitive to training samples). (See Figure 2).

48	25	26	27	28	29	30
47	24	9	10	11	12	31
46	23	8	1	2	13	32
45	22	7	**c**	3	14	33
44	21	6	5	4	15	34
43	20	19	18	17	16	35
42	41	40	39	38	37	36

Figure 2. Proposed multi-scale LBP operator

The study defined multi-scale uniform LBP operator, $MLBP_{P,R}^{u2}$ as:

$$MLBP_{P,R}^{u2} = \begin{cases} \sum_{P=1}^{8} \frac{S(g_p - g_c)2^p}{1} & \text{if} \quad u(NLBP) = 3 \\ \sum_{P=9}^{24} \frac{S(g_p - g_c)2^p}{2^9} & \text{if} \quad u(NLBP) = 2 \\ \sum_{P=25}^{49} \frac{S(g_p - g_c)2^p}{2^{25}} & \text{if} \quad u(NLBP) = 1 \\ (P-1)P + 2 & \text{otherwise} \end{cases} \tag{2}$$

Where

Int u (NLBP) **{**If $\left(\sum_{t=25}^{47} \begin{matrix} |S(g_t - g_c) - S(g_{t+1} - g_c)| + \\ |s(g_{49} - g_c) - s(g_{25} - g_c)| \end{matrix} \right) \leq 2$

Then Return u (NLBP) =1

(3)

Else If $\left(\sum_{t=9}^{23} \begin{matrix} |S(g_t - g_c) - S(g_{t+1} - g_c)| + \\ |s(g_{24} - g_c) - s(g_9 - g_c)| \end{matrix} \right) \leq 2$

Else If $\left(\sum_{t=1}^{8} \begin{matrix} |S(g_t - g_c) - S(g_{t+1} - g_c)| + \\ |s(g_7 - g_c) - s(g_1 - g_c)| \end{matrix} \right) \leq 2$

Then Return u (NLBP) =3**}**

In the new operator, the pixels of an image were labeled based on the equations two (2) and three (3). Thresholding the 3x3, 5x5 and 7x7 neighborhoods of each pixel with the center value converts the result into a binary number, if the result is a uniform pattern. In equation three (3), the first row checks the uniformity of the 7x7 neighborhood. If the resulting LBP pattern of this Neighborhood was a uniform pattern, the function results in (NLBP) =1. Otherwise, the function checks the uniformity of the 5x5 and 5x5 neighborhood, respectively. Then, the resulting uniform pattern converts into a binary number and leads to MLBP in equation two (2).

Notice that in equation three (3), gc is an intensity of central pixel and gp is a gray level intensity of neighborhood pixel and 2p is a relevant factor for any neighborhood. S (.) is a sign function and "p" is the neighborhood number in equation two (2).

Texture descriptor for this operator is obtained from the histogram of the labels. A histogram of the labeled image fl(x, y) can be defined as:

$$H(k) = \sum_{i=1}^{N} \sum_{j=1}^{M} f\big(LBP_{P,R}(i,j), k\big) \qquad k = [0, k]$$

$$f(x,y) = \begin{cases} 1 & x = y \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

2.2 MSULBP Features Selection

Feature selection is an important stage in pattern recognition systems. Instead of using all features in the data, one selectively chooses a subset of features to be used in the discriminant system. 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; and (3) more interpretable features or characteristics that can help classify the expression more accurately. In this paper GA was used for feature selection. In this method classification accuracy and number of unselected feature (zeros), considered as fitness function. The goal was to search the MSULBP space using GAs to select a subset of patterns encoding important information about the target concept of interest.

2.2.1 Brief Review of GAs

GAs is a general purpose optimization algorithm based on the mechanics of natural selection and genetics. In the past, they have been used to solve various problems including: target recognition (Katz and Thrift, 1994), object recognition (Bebis *et al.*, 2002; Swets *et al.*, 1995), face recognition (Liu and Wechsler, 2000), face detection (Bebis *et al.*, 2001). Associated with the characteristics of exploitation and exploration search, GA can deal with large search spaces efficiently, and hence, there a little chance to get local optimal solution than other algorithms. This section presents a brief introduction about genetic algorithms. A more detailed introduction can be found in (Mitchell, 1996).

Genetic algorithms start their search with a set of randomly generated solutions called initial population. Each candidate solution was appropriately encoded as a string of symbols (e.g., binary). Based on the Darwinian principle of survival of the fittest, the GA obtains the optimal solution after a series of iterative computations. It was assumed that the quality of each candidate solution was evaluated using a fitness function.

GA uses some forms of fitness-dependent probabilistic selection of individuals from the current population to produce individuals for the next generation. The selected individuals were submitted to the action of genetic operators to obtain new individuals that constitute the next generation. Mutation and crossover are two of the most commonly used operators that are used with genetic algorithms representing individuals as binary strings. The crossover operator generated new offspring from parents by interchanging the genes between the parents at the crossover point. Crossovers can be single point, two points or homologue. Mutation alters the genes at random points to generate new offspring. Other genetic representations require the use of appropriate genetic operators. The basic algorithm of GA is shown in Figure 3, where P (t) is the population of strings at generation t. The study could add new operator to GA namely majority voting.

```
Procedure:
begin
   t = 0
   initialize P(t)
   evaluate P(t)
   while (not termination condition)
   begin
      select P(t) from p(t - 1)
      * Majority Voting
      crossover P(t)
      mutate P(t)
      evaluate P(t)
      t = t+1
   end
end
```

Figure 3. Basic algorithm of GA

The process of fitness-dependent selection and application of genetic operators to generate successive generations of individuals was repeated many times until a satisfactory solution was found.

2.2.2 Feature Selection Encoding

The study applied a simple encoding scheme. In the schemes used here, chromosome is like a string whose length is equal to the total number of uniform patterns. Each bit in the string, corresponding to a one pattern that computed by MSULBP. If the i^{th} bit is 1, then the i^{th} pattern was selected, otherwise, that pattern was ignored. Therefore each chromosome indicates a distinct subset of patterns.

2.2.3 Feature Subset Fitness Evaluation

The study has used a fewer features to gain the identical or better efficiency as the aim of feature subset selection. Hence, the fitness evaluation comprises two conditions: number of features that were used, and classification precision.

In the proposed method, between precision and feature subset size, precision was the major concern. But among the two feature subsets containing different number of features, the subset with less features will be preferred if two subsets gain the identical efficiency. Notice that each feature subset include some number of uniform patterns. Hence, by using the two conditions mentioned above, the fitness function is obtained as follows:

$$\text{Fitness}=10^4 \text{ precision} + 0.5 \text{ Zeros} \quad (5)$$

In which, classification precision of one specific feature subset is shown by precision term, and the number of zeros in the chromosome is shown by *Zeros term* (i.e., zeros *term* indicate the number of patterns that is not chosen). The precision ranges approximately from 0.5 to 0.99 and the number of zeros changes between 0 to *n-1* (*n* is the length of the chromosome). It should be mentioned that the number of features in each individual does not matter, since individuals with higher precision will be more significant than individuals with lower precision. Overall, higher precision implies higher fitness. Also, fewer features used imply a greater number of zeros, and as a conclusion, the fitness growth.

2.2.4 Initial Population

The GA starts from a population of randomly selected chromosomes (a coded feature string). Each chromosome consists of a number of genes (features) and corresponds to a possible solution to the problem. In the initial method to explore subsets of different numbers of features, the number of 1's for each individual was generated randomly. Then, the 1's were randomly scattered in the chromosome.

2.2.5 Selection

After evaluation, the population for the next generation was selected from the combined set of parents and offspring. The selection mechanism was responsible for selecting the parent chromosome from the population and forming the mating pool. The selection mechanism emulates the survival-of-the-fittest mechanism in nature. It is expected that a fitter chromosome receives a higher number of offsprings and thus has a higher chance of surviving on the subsequent evolution while the weaker chromosomes will eventually die.

The roulette wheel selection used in this work (Davis 1991) was utilized, which is one of the most common and easy-to-implement selection mechanism. Basically it works as follows: each chromosome in the population is associated with a sector in a virtual wheel. According to the fitness value of the chromosome, the sector have a larger area when the corresponding chromosome has a better fitness value.

2.2.6 Majority Voting

Figure 4 shows the steps of majority voting operator. In majority voting operator, first, one offspring was generated using voting among general population. After which, the general population was divided to several groups with random numbers of individuals (there were more than two individuals in each group) and in each group new offspring was generated using voting between parents that exist in group. Note that in the groups which number of individual is even, one of the individuals that has more different chromosomes compared to other individuals has been removed from group, and then one remaining individual is selected. Finally, if the produced individuals selected do not exist among the population, they are added to population.

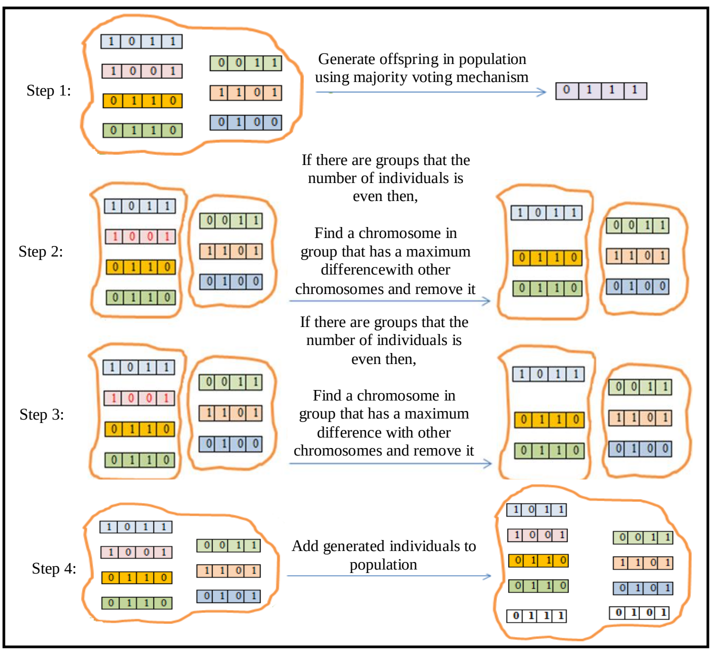

Figure 4. The steps in majority voting operator

2.2.7 Crossover

Crossover is the main genetic operator. It operates on two individuals at a time, and generates an offspring by combining schemata from both parents. There are three basic types of crossovers: one-point crossover, two-point crossover, and uniform crossover. A straight approach to perform crossover would be to select a random cut point, and generate the offspring by combining the left side of cut point in one parent with the right side of the cut point in another parent. For two-point crossover, the chromosomes are thought of as rings with the first and last gene connected (i.e., wrap-around structure). In this case, the rings are split at two common points chosen randomly and the resulting substrings are swapped. Uniform crossover was different from the above two schemes. In this case, each gene of the offspring was selected randomly from the corresponding genes of the parents. Since one-point crossover works well with the bit string representation, so this crossover was used here. The crossover probability used in all of the experiments was 0.65.

2.2.8 Mutation

New individuals was created by mutation operator. Mutation was performed after crossover by randomly choosing a chromosome in the new generation to mutate. A point along the string was selected at random and the character at that point was randomly changed, the alternate individual was then copied into the next generation of the population. Many types of mutation operators exist. The bit flip method was used which was only utilized with a binary chromosome representation, that changes a particular point on the chromosome to its opposite. The mutation probability used throughout current experiments was 0.05.

2.3 Support Vector Machine

SVM is a popular supervised learning algorithm for classification. SVM was used to separate the data that was inseparable with linear rules in the input space (Vapnik, 1998). To perform the separation, first SVM mapped the data into a higher dimensional feature space, then found a linear separating hyper plane with the maximal margin to separate the data in feature space:

Suppose that $T = \{(xi, yi), i = 1...l\}$ is a training set of labeled examples, such that $xi \in Rn$ and $yi \in \{1, -1\}$, A function below was used to classified the new test data x.

$$f(x) = sgn\left(\sum_{i=1}^{l} \alpha_i y_i k(x_i, x) + b\right) \quad (6)$$

In this function: α_i parameters indicate the Lagrange multipliers of a dual optimization problem, b is the parameter of the optimum hyper plane and K (x_i, x) is a kernel function. Kernels have the form of K (x_i, x_j) = (Φ (x_i) ^{T}. Φ (x_j)), with Assuming the Φ is a nonlinear mapping function that embeds input data into feature space, Although, several new kernels were being proposed up to now, the most frequently used kernel function is the linear, polynomial, and RBF kernels.

3. Results and Discussion

The proposed system designed and tested with JAFFE database. The database contained 213 images in which ten persons were expressing three or four times the seven basic expressions (happiness, sadness, surprise, anger, disgust, fear, neutral). Sample images from the database are shown in Figure 5.

Figure 5. Samples from the Japanese female facial expression database

In the experiments, image pre-processing was conducted by the pre-processing subsystem of the CSU Face Identification Evaluation System (Bolme, *et al.*, 2003). The images were registered using eye coordinates and cropped with an elliptical mask to exclude nonface area from the image. As a result, the size of each preprocessed image was 150 × 128 (see Figure 6).

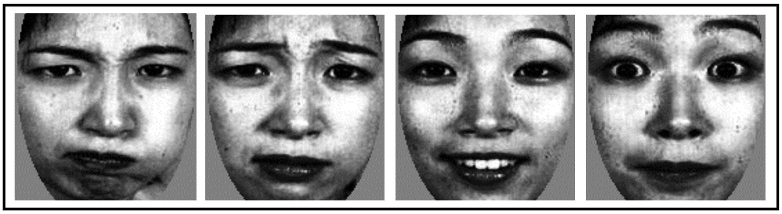

Figure 6. Samples from the preprocessed images

There are mainly two ways to divide the JAFFE database. The first way is to divide the database into several segments, but each segment corresponds to one expresser. The second way is to divide the database randomly into 10 roughly equal-sized segments, of which nine segments are used for training and the last one for testing. The study carried out both experiments way on facial expression recognition to test the validity of the method. Notice that in MSULBP operator, the scaling factor α was set to 0.1 and in GA, a population size of 2000 and 250 generations was used in all of the experiments. In most cases, the GA converged in less than 250 generations.

3.1 Recognition Performance

To compare the recognition performance with other published methods, a set of 193 expression images posed by nine expressers was used. These images were partitioned into nine segments, each corresponding to one expresser. Eight of the nine segments were used for training and the ninth was used for testing. The above process was repeated so that each of the nine partitions was used once as the test set. The average of recognizing the expression of novel expressers is 88.5%. Recognition results of each trail are shown in Table 1.

Table 1. Recognition accuracy of each trail in the method

Trial	Accuracy/Total	% Correct
1	21/22	95.4%
2	17/21	80.9%
3	16/20	80.0%
4	18/21	85.7%
5	20/22	90.9%
6	21/22	95.4%
7	19/20	95.0%
8	19/22	86.3%
9	20/23	86.9%

The recognition performance was compared to other published methods using the same database. In a result of 75% using Linear Discriminant Analysis (LDA) was reported with 193 images (Lyons, *et al.*, 1999). In an average recognition result of 30% was reported with 213 images (Gargesha and Kuchi, 2002). It should be pointed out that, 34 fiducial points have to be selected manually. Other reports of Zhang, 1999; Zhang *et al.*, 1998; Guo and Dyer, 2003; Buciu *et* al., 2003; Fasel, 2002 on the same database did not give the recognition rate for novel expresser's expression. It should be pointed out that in 34 fiducial points have to be selected manually (Lysons *et al.*¸1999). In the method, it only needed the position of two pupils for face normalization and other procedures were completely automatic.

3.4 Performance of the Proposed Method

The database was divided into ten roughly equally sized sets – nine sets were used for training and the remaining one were used for testing. The above process was repeated so that each of the ten roughly equally sized sets was used once as the test set. The average result over all ten cycles was considered as the recognition rate of one trial. Facial expression recognition performance of the proposed method for facial expression recognition using SVMs with different kernels is shown in Table 2. Notice that in 7-Class recognition case, seven basic expressions were considered (happiness, sadness, surprise, anger, disgust, fear and neutral) and in 6-Class recognition case, six basic expressions were included (happiness, sadness, surprise, anger, disgust and fear).

Table 2. Comparisons between LBP features with the proposed method for facial expression recognition using SVMs

SVM with different kernels	6-Class recognition (%)		7-Class recognition (%)	
	LBP (Zhang *et al.*, 2004)	MULBP based GA and Majority Voting	LBP (Zhang *et al.*, 2004)	MULBP based GA and Majority Voting
SVM (linear)	91.45	92.73	84.30	86.84
SVM (polynomial)	91.73	92.80	84.36	86.95
SVM (RBF)	92.86	93.89	85.67	89.76

*Notice that average result over all ten cycles is shown.

Performances of the proposed method compared with other facial expression recognition algorithms are given in the Table 3. Experimental results indicated the stability of the proposed approach, it acquired higher recognition precision for train set 100% and for test set as 95.86%.

Table 3. Classification accuracy of proposed method compared with other algorithms

Algorithm	Classification accuracy
Volume Local Binary patterns approach (VLBP) (Zhao and Pietikainen, 2007)	95.19%
Local Binary Pattern on the Three Orthogonal Planes approach (LBP-TOP) (Zhao and Pietikainen, 2007)	94.38%
MSULBP Features Selection based on GA and Majority Voting Mechanism (Proposed algorithm)	95.86 %

To further explore the recognition accuracy per expression where the proposed method performs best, Tables 4 and 5 show the confusion matrix of 6-class and 7-class facial expression recognition results.

Table 4. Confusion matrix of 6-class facial expression recognition using SVM (RBF) on the JAFFE database

	Surprise (%)	Joy (%)	Disgust (%)	Fear (%)	Sadness (%)	Anger (%)
Surprise	97.9	0.5	0	2.1	0	0
Joy	0.2	95.9	0.8	1.0	2.1	0
Disgust	0	0.6	98.3	1.2	0	0
Fear	0	7.6	4.0	85.0	3.4	0
Sadness	5.6	1.5	0	1.8	82.8	8.3
Anger	0	0	2.2	0	5.6	92.2

Table 5. Confusion matrix of 7-class facial expression recognition using SVM (RBF) on the JAFFE database

	Surprise (%)	Joy (%)	Disgust (%)	Fear (%)	Sadness (%)	Anger (%)	Neutral (%)
Surprise	97.4	0.5	0	1.3	0	0	0.8
Joy	0.3	94.8	0	0.7	0.9	0	3.3
Disgust	0	0	97.6	1.6	0.8	0	0
Fear	0	14.5	1.5	79.0	1.0	0	4.0
Sadness	1.3	0	0	1.2	75.4	6.7	15.4
Anger	0	0	2.5	0.2	4.2	87.5	5.6
Neutral	0.3	1.2	0.4	0	4.0	1.3	92.8

Considering Tables 4 and 5, it can be seen that two expressions, i.e., fear and sadness are discriminated with relatively low accuracy (less than 90%). In particular, sadness is recognized with the lowest accuracy since sadness is highly confused with neutral and anger.

It should also be noted that in the JAFFE database, some expressions had been labeled incorrectly or expressed inaccurately. Whether these expressional images are used for training or testing, the recognition result was influenced. Figure 7 shows a few examples with the labeled expression and the recognition results.

Figure 7. Examples of disagreement. From right, the labeled expressions are surprise, sadness, sadness, disgust and the recognition results are happiness, neutral, happiness and angry, respectively

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. The LBP-based algorithm was further evaluated over a range of image resolutions, investigating its performance against low-resolution images. The recognition rates are summarized in Table 6. It was observed that the proposed LBP-based method was more effective for face images with low resolutions than classic LBP methods. Therefore, the proposed method has much superiority comparing to classical LBP method.

Table 6. Comparisons on image resolutions between theclassic LBP algorithm the LBP-based algorithm

			
	48 X 48	32 X 32	16 X 16
LBP	83.09	79.22	69.72
proposed method	89.87	87.42	82.92

4. Conclusions

In this study, a novel and efficient facial expression recognition system is proposed. This proposed system is the MSULBP-based GA and MVM, and ensemble of supervised SVM classifiers for facial expression recognition. The LBP has been proven to be effective for image representation. Also, it has been proven that the uniform LBP draw out the representative features effectively. Fully extracted, useful feature from an image in the proposed method enhances the classic LBP method that it was able to extract beneficial information of non-uniform patterns. For feature selection, GA and MVM were used. In this method, classification accuracy and number of unselected feature (zeros) were considered as fitness functions.

The method was evaluated on JAFFE database. The results showed that the MSULBP method produced the largest improvement in the classification accuracy and in the discrimination between different facial expressions.

5. Acknowledgments

The authors would like to appreciate Ahar Branch, Islamic Azad University for the financial support of this research, which is based on a research project entitled: An MSULBP Features Selection based on GA and Majority Voting Mechanism in Facial Expression Recognition.

6. References

Ahonen, T., Hadid, A., and Pietikainen, M., (2006). Face Description with Local Binary Patterns: Application to Face Recognition, IEEE Trans. Pattern Analysis and Machine Intelligence, 28, 2037-2041.

Bassili, J.N., (1979). Emotion Recognition: The Role of Facial Movement and the Relative Importance of Upper and Lower Area of the Face, J. Personality and Social Psychology, 37, 2049-2059.

Bolme, D., Teixeira, M., Beveridge, J., Draper, B., (2003). The CSU face identification evaluation system user's guide: its purpose, feature and structure. In: Third International Conference on Computer Vision Systems. 304 –331

Chan, C.H., (2008). Multi-scale local binary pattern histogram for face recognition, Centre for Vision, Speech and Signal Processing School of Electronics and Physical Sciences University of Surrey Guildford, Surrey, U.K, Submitted for the Degree of Doctor of Philosophy from the University of Surrey, 56-59.

Cohen, I., Sebe, N., Garg, S., Chen, L. S., and Huanga, T. S., (2003). Facial expression recognition from video sequences: temporal and static modelling, Comput. Vis. Image Understand., 91, 160–187.

Cottrell, G.W., and Metcalfe, J., (1991). Face, Emotion, Gender Recognition Using Holos, Advances in NIPS, R.P. Lippman, ed., 564-71.

Ekman, P., and Friesen, W.V., (1978). The facial action coding system: a technique for the measurement of facial movement, In Consulting Psychologists Press, Inc., SanFrancisco, CA, 1978.

Fasel, B., Luettin, J., (2003). Automatic facial expression analysis: a survey, In Pattern Recognition, 36, 259- 275.

Gargesha, M., Kuchi, P., (2002). Facial expression recognition using artificial neural networks, IEEE 511: Artificial Neural Computation Systems, 124-128 .

He, X., Yan, X., Hu, Y., Niyogi, P., and Zhang, H., (2005). Face recognition using Laplacianfaces, IEEE Trans. Pattern Analysis & Machine Intelligence, 27, 328–340.

Kato, M., So, I., Hishnuma, Y., Nakamura, O., and Minami, T., (1991). Description and Synthesis of Facial Expressions Based onIsodensity Maps, Visual Computing, T. Kunii, ed., 39-56.

Kotsia I., Ioannis P., (2007). Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines, IEEE Trans. on Image Processing, 16, 172–187.

Littlewort, G., Bartlett, M., Fasel, I., Susskind, J., and Movellan, J., (2004). Dynamics of facial expression extracted automatically from video, in IEEE Workshop on Face Processing in Video, 12-17.

Lyons, M.J., Budynek, J., and Akamatsu, S., (1999). Automatic classification of single facial images, IEEE PAMI, 23-28.

Mitchell, M., (1996). An introduction to genetic algorithms, MIT Press, Cambridge – MA, 26-35.

Montazeri-Gh, M., Mahmoodi-K, M., (2016). Optimized predictive energy management of plug-in hybrid electric vehicle based on traffic condition, Journal of Cleaner Production, 139, 15 935–948.

Ojala, T., Pietik¨ainen, M., Harwood, D., (1996). A comparative study of texture measures with classification based on feature distributions. Pattern Recognition, 29, 51–59.

Ojala, T., Pietikäinen, M., and Mäenpää, T., (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Transactions on Pattern Analysis and Machine Intelligence, 24, 971-987.

Pantic, M., and Ioannis, P., (2006). Dynamics of Facial Expression: Recognition of Facial Actions and Their Temporal Segments From Face Profile Image Sequences, IEEE Trans. Syst., Man, Cybern. B, Cybern., 36, 433–449.

Pantic, M., and Rothkrantz, L., (2000). Automatic Analysis of Facial Expressions: The State of the Art, IEEE Trans. Pattern Analysis and Machine Intelligence, 22, 1424-1445.

Rosenblum, M., Yacoob, Y., and Davis, L., (1994). Human Emotion Recognition from Motion Using a Radial Basis Function Network Architecture, Proc. IEEE Workshop Motion of Non-Rigid and Articulated Objects, 43-49.

Tian, Y., Brown, L., Hampapur, A., Pankanti, S., Senior, A., and Bolle, R., (2003). Real world real-time automatic recognition of facial expression, in IEEE PETS, Australia, 34-39.

Tian, Y., (2004). Evaluation of face resolution for expression analysis, in IEEE Workshop on Face Processing in Video, 12-18.

Valstar, M., Patras, I., Pantic, M., (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, 3, 76–84.

Vapnik V.N., (1998). Statistical Learning Theory, Wiley, New York, 23-34.

Zhang, G., Huang, Z., and Li, S. Z., (2004). Boosting local binary pattern (LBP)-based face recognition, SinoBiometrics, 179- 186

Zhang, Y., Ji, Q., (2005). Active and dynamic information fusion for facial expression understanding from image sequences, IEEE Trans. Pattern Anal, Mach. Intell., 27, 699–714.

Zhao, G., and Pietikainen, M., (2007). Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions, IEEE Trans. Syst., Man, Cybern. B, Cybern., 29, 915–928.