Local Gray Code Pattern (LGCP): A Robust Feature Descriptor for Facial Expression Recognition

Mohammad Shahidul Islam

Atish Dipankar University of Science & Technology, School, Department of Computer Science and Engineering, Dhaka, Bangladesh.

Abstract: ***This paper presents a new local facial feature descriptor, Local Gray Code Pattern (LGCP), for facial expression recognition in contrast to widely adopted Local Binary pattern. Local Gray Code Pattern (LGCP) characterizes both the texture and contrast information of facial components. The LGCP descriptor is obtained using local gray color intensity differences from a local 3x3 pixels area weighted by their corresponding TF (term frequency). I have used extended Cohn-Kanade expression (CK+) dataset and Japanese Female Facial Expression (JAFFE) dataset with a Multiclass Support Vector Machine (LIBSVM) to evaluate proposed method. The proposed method is performed on six and seven basic expression classes in both person dependent and independent environment. According to extensive experimental results with prototypic expressions on static images, proposed method has achieved the highest recognition rate, as compared to other existing appearance-based feature descriptors LPQ, LBP, LBP_{U2}, LBP_{RI}, and LBP_{RIU2}.***

Keywords: Expression Recognition, Image Processing, Local Feature, Pattern Recognition, CK+, JAFFE, Computer Vision

1. Introduction

FER (Facial Expression recognition) has gained substantial importance in the applications of human-computer interactions (C. Shan *et al*., 2009) as this is one of the most effective, natural, and immediate means for human beings to communicate their emotions and intentions (C. Shan *et al*., 2005), (Y. Tian *et al*., 2003). It has attracted much attention from behavioral scientists since the work of Darwin in 1872. Expression recognition with high accuracy remains difficult due to ethnicity and variation of facial expressions (G. Zhao *et al*., 2009), though many works have been done with automatic facial expression analysis. Extracting relevant features from human face images is very important for any successful facial expression recognition system. The extracted features should retain essential information having high discrimination power and stability which minimizes within-class differences of expressions whilst maximizes between-class differences (C. Shan *et al*., 2005).

1.1 Motivation

Facial features can be two types- global feature and local feature. Global feature is a feature, which is extracted from the whole face whereas local feature considers small local region from the whole face. Some global feature extraction methods are PCA (Principal component analysis), LDA (Linear Discriminant Analysis) etc. Even they are popular and widely used but their performance fluctuates with the environment. Therefore, I have chosen local feature methodology for feature extraction as it is robust in uncontrolled environment. Some popular local feature descriptors are Gabor filter, Local binary pattern, Local phase quantization etc. Facial feature representations using Gabor filter is time and memory intensive. S. Lajevardi *et al*. 2012 solved some limitations of Gabor-filter using log-Gabor filter but the dimensionality of resulting feature vector was still high. Local binary pattern is popular but sensitive to non- monotonic illumination variation and shows poor performance in the presence of random noise (T. Jabid *et al*., 2010). LPQ (V. Ojansivu *et al*., 2008) is also very time and memory expensive. Keeping all these sensitive issues in mind, I have proposed LGCP, which overcomes almost all those weaknesses.

1.2 Paper Review

Some surveys of existing research on facial expression analysis can be found in B. Fasel *et al*., 2003 and M. Pantic *et al*., 2000. Three types of facial feature extraction approaches are there: the geometric feature-based system (Y. L. Tian *et al*., 2001), the appearance-based system (Y. Tian *et al*., 2003) and hybrid, which uses both the approaches. Geometric feature vectors represent the shapes and spatial locations of facial parts by encoding the face geometry from the location, distance, angle, and other geometric relations between these parts. A most commonly used facial descriptor is the facial action coding system (FACS), in which, facial muscle movements are encoded by 44 Action Units(AUs) (P. Ekman, 1978). Y.Zhang *et al*. (2005) proposed IR illumination camera for facial feature detection, tracking and recognized the facial expressions using Dynamic Bayesian networks (DBNs). M.Yeasin *et al*. (2006) used discrete hidden Markov models (DHMMs) to recognize the facial expressions. Z. Zhang *et al*. (1998) used 34 fiducial points as facial features to present a facial image. Y.L. Tian *et al*. (2001) proposed a multi-state face component model of AUs and neural network (NN) for classification. I. Cohen *et al*. (2003) employed Naive–Bayes classifiers and hidden Markov models (HMMs) together to recognize human facial expressions from video sequences. M.F. Valstar *et al*. (2005, 2006) used several fiducial points on face and mentioned that geometric approaches are better in feature extraction than appearance-based approaches. Geometric feature-based methods need exact and accurate facial components detection, in many situations, which is not possible (C. Shan *et al*., 2009). Recent psychological research concluded that spatial relations of the facial features from the full face could be a good source of information for facial emotions (M. Meulders *et al*., 2005, M. Zia Uddin *et al*., 2009). Principal Component Analysis (PCA) is a holistic method widely used to extract features from faces (A. Turk *et al*., 1991). PCA is also very useful in reducing feature dimension. Lately, Independent Component Analysis (ICA) (M.S. Bartlett *et al.*, 2005) and Enhanced ICA (EICA) are used for feature extraction. G. L. Donato *et al*. (1999) did a comprehensive analysis of different techniques, including PCA, ICA, Local Feature Analysis (LFA), Gabor-wavelet and local Principal Components (PCs). Since then Gabor-wavelet representation is widely used for feature extraction. However, the size of filter bank needed for Gabor filters to extract useful information from a face makes the Gabor representation time and memory intensive. Afterwards log-Gabor filters proposed by S. Lajevardi *et al*. (2012) overcame some limitations of Gabor-filter but the dimensionality of resulting feature vector was still high. Another popular feature descriptor is Local Binary Pattern (LBP) (T. Ojala *et al*., 2002) and its variants (G. Zhao *et al*., 2009). It is widely used in many research works (C. Shan *et al*., 2005), (C. Shan *et al*., 2009). Originally, LBP was introduced for texture analysis. LBP labels each pixel of an image by thresholding it's P neighbors gray color intensity with the center gray color intensity and derives a binary patter using Equation (1),

$$LBP(i,j) = \sum_{p=1}^{P} A(g_c - g_p) \times 2^p; \qquad A(x) = \begin{cases} 1, & x < 0 \\ 0, & x \geq 0 \end{cases} \tag{1}$$

Where g_c and g_p are the gray color intensity of the center pixel (i, j) and p neighboring pixel respectively. A comprehensive study of LBP can be found in T. Ojala *et al.* (1996). Later he observed that binary patterns with less transition from 1 to 0 and 0 to 1 occur more frequently in a facial image. Therefore, patterns having more than two transitions are discarded in LBP_{U2}. He also proposed rotation invariant LBP (LBP_{RI}), in which he made a bit wise right shift, eight times for a 8-bit binary pattern. He counted all eight shifted patterns as a single bin (T. Ojala *et al*., 2002). This method is widely adopted by many researchers, but it is sensitive to non-monotonic illumination variation and shows poor performance in the presence of random noise (T. Jabid 2010). To overcome this problem, T.Jabid *et al*. (2010) proposed a facial descriptor, named as Local Directional Pattern (LDP), which is more robust than LBP. LDP is derived from the edge responses, which are less sensitive to illumination changes and noises. H.Kabir *et al*. (2012) extended LDP to LDPv by applying weight to the feature vector using local variance and found it to be more affective for facial expression recognition. V. Ojansivu *et al*. (2008) proposed LPQ (Local Phase Quantization), a facial feature extraction method that is blur insensitive. J. Li *et al* (2012) extended LPQ to RI-LPQ along with SRC (Sparse Representation-based Classification) classifier and obtained better accuracy than LBP. K. Anderson *et al.* (2006) used the multichannel gradient model (MCGM) to determine facial optical flow. The motion signatures achieved and then classified using Support Vector Machines.

1.3 Contribution

A novel feature extraction technique, LGCP (Local Gray Code Pattern) is proposed, which characterizes both contrast and texture information for more accurate facial expression recognition performance. Figure 1 shows an overall flow of the expression recognition system based on proposed LGCP descriptor coupled with LIBSVM. The performance of proposed method is evaluated using Multiclass Support Vector Machine (LIBSVM) (C.C. Chang *et al*., 2011) with different kernel setup. Proposed method LGCP is more

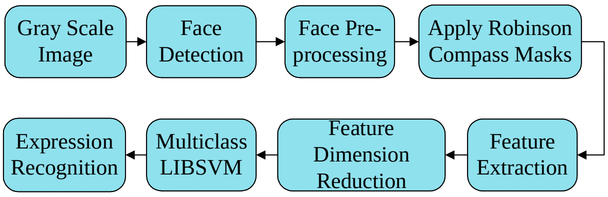

Figure 1: Overview of the facial expression recognition system based on LGCP representation

robust in extracting facial features, and has a better recognition rate, as compared to LBP, Gabor-wavelet features, and other appearance-based methods. LGCP descriptor is stable in presence of noise. The rest of the paper is organized as follows: The proposed LGCP feature is described in section 2. Section 3 presents the experimental setup used for evaluating the effectiveness of the proposed feature representation, Section 4 lists the expression recognition performances of LGCP compared with existing representations. Finally, section 5 concludes the paper.

2. LGCP (Local Gray Code Pattern)

I have followed three steps to compute LGCP code from a gray scale facial image.

$$E = \overset{m4}{\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}} \quad NE = \overset{m3}{\begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix}} \quad N = \overset{m2}{\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}} \quad NW = \overset{m1}{\begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -2 \end{bmatrix}}$$

$$W = \underset{m8}{\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}} \quad SW = \underset{m7}{\begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}} \quad S = \underset{m6}{\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}} \quad SE = \underset{m5}{\begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}}$$

Figure 2 : Robinson Edge response masks in eight directions

Step 1: I have used Robinson Compass Mask (Robinson 1977) to maximize the edge value, which makes LGCP descriptor stable in the presence of noise. It is a single mask in eight major compass orientations: E, NE, N, NW, W, SW, S, and SE as shown in Figure 2. The edge magnitude = the maximum value found by the convolution of each mask with the image. The mask that produces the maximum magnitude defines the edge direction. The Robinson compass mask computes the edge response values in all eight directions at each pixel position and generates a code from the relative strength magnitude. The obtained 3x3 matrix describes the local curves, corners, and junctions, more stably and retains more information. Given a central pixel in the image, the eight directional edge response values {Ri}, i= 0, 1... 7 are computed by

$$R_i = \sum_{i=0}^{7} a * m_i \tag{2}$$

Where 'a' is a local 3x3 pixels region and m_i is the Robinson masks in eight different orientations centered on its position. Figure 3 shows an example of obtaining edge response using equation (2).

4	6	8
7	4	5
3	2	1

(a) Local 3x3 pixels region

R1	R2	R3
R8	C	R4
R7	R6	R5

(b) Eight directional edge response position.

==

12	16	12
2	C	-2
-12	-16	-12

(c) Obtained matrix with edge response from (a)

R1=(a).*m1=7x1+4x2+6x1+3x0+4x0+8x0+2x(-1)+1x(-2)+5x(-1)

R2=(a).*m2=4x1+6x2+8x1+7x0+4x0+5x0+3x(-1)+2x(-2)+1x(-1)

| | | | | | | | | |

R8=(a).*m8=4x1+7x2+3x1+6x0+4x0+2x0+8x(-1)+5x(-2)+1x(-1)

Figure 3: Detailed example of obtaining directional edge response using Robinson Compass Mask

Step 2: The matrix with 8-edge response is then used to get 8-bit gray code pattern as shown in Figure 5. Gray code is a reflected binary code where only one bit changes at a time. For an example, in a three bit binary code 001(1) becomes 010(2) by changing two bits, but in case of gray code only one bit changes at a time, so for gray code '1'=001 and '2'=011. The advantage of gray code is that it makes very stable position digitizers, because only one bit changes at a time, resulting in uncertainty of only one bit.

Step 3: in this step, I have computed the TF (Term Frequency) of the gray code image using the following formula, (Equation (3))

$$\mathrm{TF}(\mathrm{bin_i}, \mathrm{I}) = \frac{f(bin_i, I)}{\max(bin\ value)} \quad ; \qquad (3)$$

Where, 'I' is the image and i= 1,2,... ,n (number of bins). For an example, if I have an image with four possible bins, bin(1-4) and the bin counting value or histogram is 20, 10, 30, 25, then the TF for bin1 is 20/max(20,10,30,25)=2/3. 8-bit gray code pattern can produces up-to 256 combinations. Therefore, I have histogram of 256 bins. I have observed the histograms of all the images and found only 15 gray code patterns are mostly responsible for expression classification. Rests of the patterns are arbitrary. The most prominent patterns are

0,2,10,22,30,31,64,160,191,224,225,233,245,253 and 254 in decimal value.

For the experiments, I have reduced the bin number from 256 to 15 by discarding the arbitrary patterns. Therefore, the length of the TF is also reduced to 15 and the value range from 1 to 0. After obtaining the term frequency for all images, I have used it as a weight while building histogram for each block using Equation (4).

$$H(block_i, y) = TF(w, I) \times H(block_i, x) \qquad (4)$$

Where i=1, 2... 81., x is the non-weighted bin value and y is the weighted bin value.

LGCP gives more stable patterns in the presence of noise as I have used Robinson compass masks to intensify the edges. Figure 4 shows local 3x3 pixels region from an original gray scale image and the corresponding image after adding Gaussian white noise.

start ➡

end

4	6	8
7	4	5
3	2	1

LBP = 11110001
LGCP= 10111111

(a)

3	5	7
7	5	6
4	3	0

LBP = 01110001
LGCP= 10111111

(b)

Figure 4: Noise sensitivity of LGCP, (a) 3x3 pixels region from original Image, (b) same region with noise

It is clear from the figure that LBP pattern changes with noise but LGCP remains stable.

3. Experimental Setup

There are several techniques for expression classification. Among them support vector machine is popular due to its input output data simplicity. In (C. Shan *et al.*, 2009), the author conducted experiments using few machine learning techniques, namely Template matching, Linear Discriminant Analysis, Linear programming, and Support Vector Machine. He found SVM to be the best. Therefore, I have decided to use SVM as the classifier in all through the experiments. I have used Cohn-Kanade (CK+) (P. Lucey *et al.*, 2010) and Japanese Female Facial Expression (JAFFE)

B1	B2	B3	B4	B5	B6	B7	B8

(8-bit Gray Code pattern)

$$B1 = \begin{cases} 0, & R3 - R1 < 0 \\ 1, & R3 - R1 \geq 0 \end{cases} \quad B2 = \begin{cases} 0, & R4 - R8 < 0 \\ 1, & R4 - R8 \geq 0 \end{cases} \quad B3 = \begin{cases} 0, & R5 - R7 < 0 \\ 1, & R5 - R7 \geq 0 \end{cases} \quad B4 = \begin{cases} 0, & R1 - R7 < 0 \\ 1, & R1 - R7 \geq 0 \end{cases}$$

$$B5 = \begin{cases} 0, & R2 - R6 < 0 \\ 1, & R2 - R6 \geq 0 \end{cases} \quad B6 = \begin{cases} 0, & R3 - R5 < 0 \\ 1, & R3 - R5 \geq 0 \end{cases} \quad B7 = \begin{cases} 0, & R1 - R5 < 0 \\ 1, & R1 - R5 \geq 0 \end{cases} \quad B8 = \begin{cases} 0, & R3 - R7 < 0 \\ 1, & R3 - R7 \geq 0 \end{cases}$$

12	16	12
2	C	-2
-12	-16	-12

(a) Obtained matrix with edge response from Figure 3(a)

R1	R2	R3
R8	C	R4
R7	R6	R5

(b) Eight directional edge response position.

10111111

(c) Binary patterrn for the matrix (b)

191

(d) Corresponding Decimal value

	213	

(e) Corresponding Gray Code

Figure 5: Example of obtaining Gray code pattern value from 3x3 matrixes with edge response obtained from Figure 3

(J. Michael 1997) as the facial image datasets. I have used two datasets to clarify the outcome of proposed method. There are 326 posed images in the CK+. No person has multiple instances for same expression in CK+. There are 7 prototype expression classes in it. I have not used neutral expression class for the experiment for this dataset. However, JAFFE has multiple instances of same expression from the same subject. It does not have contempt expression class like CK+. Neutral class replaces the contempt class. This dataset has total 213 images from 10 different women with seven expression classes. Figure 6 shows the feature vector building steps starting from raw image.

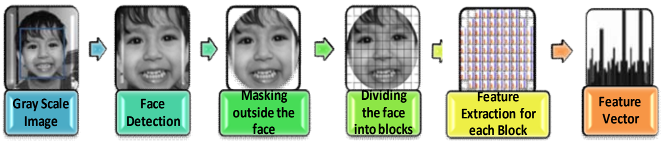

Figure 6: Facial Feature Extraction

I have first converted all the images to gray scale if they are in different format. Then I have used fdlibmex face detector for detecting face area, which is free to use in MATLAB. There is no publication for this detector. However, for the experiments I have found it better than manual cropping. After successfully detecting the face, I have re-dimensioned detected face areas to 180x180 pixels for CK+ dataset and 99x99 pixels for for JAFFE dataset. I have conducted several experiments with different combination of dimension and found the above dimension as optimum one. Then I have used an elliptical masking to remove some unnecessary areas from the head corners and neck corners. The masked face is then divided into 9x9=81 blocks. I have done this to preserve the spatial information from the facial area. I have used LGCP to extract feature from each block to build histogram and concatenated them to obtain the final feature vector. Therefore the length of the feature vector is, 81 x Number of bins e.g. 15.

I have followed the above procedure for both training and testing images. As a classifier, I have used LIBSVM (C.C. Chang *et al.*, 2011) with 10-fold cross validation and the folds are not overlapping in a person dependent and independent environment. The whole dataset is divided into 10 equal folds and each fold is validated against rest 90% of the images in each fold validation. LIBSVM is a multiclass classifier with variety of options that can be set depending on input data types. I have set the kernel parameters for the classifier to: s=0 for SVM type C-Svc, t=0/1/2 for linear/polynomial/RBF kernel function respectively, c=100 is the cost of SVM, g=0.000625 is the value of 1/ (length of feature vector), b=1 for probability estimation. This setting for LIBSVM is found to be suitable for CK+ dataset as well as JAFFE dataset with both six seven classes of data.

4. Results and Analysis

Proposed method is tested against some popular methods e.g., LBP (Local Binary Pattern) (T. Ojala *et al.*, 1996), LBP_{RI} (Rotation Invariant LBP) (T. Ojala *et al.*, 2002), LBP_{U2} (Uniform LBP) (T. Ojala *et al.*, 2002), LBP_{RIU2} (Rotation invariant and Uniform LBP) (T. Ojala *et al.*, 2002) and LPQ (V. Ojansivu *et al.*, 2008), (J. Li *et al.*, 2012). To make the comparison more effective, I have used the same experimental setup for all above methods as well as the setup used by the authors.

Table 1 : Confusion matrices for facial expression recognition systems using LGCP on CK+ and JAFFE

LGCP (CK+) = 10-fold cross vali.
Feature Extraction time for single Image = 0.034 Seconds
Average Classification Accuracy = 91.9%
LIBSVM Kernel parameter: = (-s 0 -t 1 -c 100 -g 0.0008 -b 1)
Confusion Matrix:

C: prediction \ Actual	Angry	Contempt	Disgust	Fear	Happy	Sad	Surprise
Angry	**80.0**	4.4	6.7	2.2	0.0	6.7	0.0
Contempt	5.6	**83.3**	0.0	0.0	0.0	5.6	5.6
Disgust	3.4	0.0	**93.2**	1.7	1.7	0.0	0.0
Fear	8.0	4.0	4.0	**76.0**	4.0	0.0	4.0
Happy	0.0	0.0	0.0	0.0	**100.0**	0.0	0.0
Sad	14.3	0.0	0.0	0.0	0.0	**78.6**	7.1
Surprise	0.0	0.0	0.0	0.0	0.0	0.0	**100.0**

LGCP (JAFFE) = 10-fold cross vall.
Feature Extraction time for single Image = 0.021 Second
Average Classification Accuracy = 93.3%
LIBSVM Kernel parameter: = (-s 0 -t 1 -c 100 -g 0.0008 -b 1)
Confusion Matrix:

C: prediction \ Actual	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
Angry	**100.0**	0.0	0.0	0.0	0.0	0.0	0.0
Disgust	3.4	**93.1**	3.4	0.0	0.0	0.0	0.0
Fear	0.0	3.1	**81.3**	0.0	6.3	6.3	3.1
Happy	0.0	0.0	0.0	**96.8**	0.0	3.2	0.0
Neutral	0.0	0.0	0.0	0.0	**100.0**	0.0	0.0
Sad	0.0	0.0	3.2	3.2	3.2	**90.3**	0.0
Surprise	0.0	0.0	3.3	3.3	0.0	0.0	**93.3**

Table 1 shows Confusion matrices for facial expression recognition systems using LGCP on CK+ and JAFFE dataset.

Table 2 : Results (Classification Accuracy) comparison using different methods on CK+ and JAFFE Dataset

Method	*CK+ (%)*	*JAFFE (%)*
LGCP	91.9	93.3
LBP	91.6	88.7
LBP(RI)	90.7	85.4
LBP(U2)	90.1	84.4
LBP(RIU2)	89.8	83.8
LPQ	80.2	79.6

Table 2 shows results (Classification Accuracy) comparison using different methods on CK+ and JAFFE Dataset. All the results of Table 1 and Table 2 are obtained using LIBSVM with polynomial kernel. Experiments are also performed with linear and RBF (radial basis function) kernel. The results are shown in Table 3 on CK+ dataset:

Table 3: Results of proposed method and some other popular methods in same experimental setup but different SVM kernel setup on CK+ dataset (person dependent)

Methods	*Linear Kernel*	*Polynomial kernel*	*RBF kernel*
LGCP (Proposed method)	**91.89%**	**91.89%**	**92.04%**
LBP	91.59%	91.59%	91.88%
LBP_{RI}	90.68%	90.68%	90.95%
LBP_{U2}	90.12%	90.12%	90.65%
LBP_{RIU2}	89.83%	89.83%	90.03%
LPQ	80.21%	80.21%	80.43%

Table 4 shows the results comparison on JAFFE dataset with different SVM kernel setup:

Table 4: Results of proposed method and some other popular methods in same experimental setup but different SVM kernel setup on JAFFE dataset (person dependent)

Methods	*Linear Kernel*	*Polynomial kernel*	*RBF kernel*
LGCP	93.31%	93.31%	93.52%
LBP	88.72%	88.72%	89.31%
LBP_{RI}	85.37%	85.37%	86.21%
LBP_{U2}	84.43%	84.43%	85.06%
LBP_{RIU2}	83.82%	83.82%	84.12%
LPQ	79.56%	79.56%	80.14%

I have also experimented on 6-class expressions by removing 'contempt' from CK+ dataset and 'natural' from JAFFE dataset. This is because most of the previous works are done on 6 expressions class. Therefore, to compare with those results, I have removed one class from both the datasets. Table 5 shows classification accuracy comparison on 6-class expressions on Cohn-kanade dataset with different SVM kernel setup. I have removed contempt class from CK+ dataset for proposed LGCP method.

Table 5: Results of proposed method and some other popular methods in different SVM kernel setup on 6-class expression (person dependent)

Methods	*Linear Kernel (%)*	*Polynomial Kernel (%)*	*RBF Kernel (%)*
LGCP (Proposed Method)	*94.9*	*94.9*	*95.9*
Gabor Feature(M.S. Bartlett *et al.*, 2005)	89.4	89.4	89.8
LBP(C. Shan *et al.*, 2009)	91.5	91.5	92.6
LDP(T. Jabid *et al.*, 2010)	92.8	92.8	94.5

Table 6 shows the accuracy comparison of proposed method with some other methods on CK+ dataset.

Table 6: Classification Accuracy Comparison on CK+ dataset in a person dependent environment

Author	*Classifier*	*Classification accuracy (%)*
LGCP (Proposed Method)	*Multi Class SVM(Poly)*	*92%*
(S.W.Chew *et al.*, 2011)	SVM(RBF)	75%
(G.Littlewort *et al.*, 2011)	SVM(Linear)	90%
(L.A. Jeni *et al.*, 2012)	SVM	87%
(S. Naika *et al.*, 2012)	Multi Class SVM(RBF)	82%

Table 7 shows the accuracy comparison of proposed method with some other methods on JAFFE dataset. I have done this in person independent expression recognition environment. Nine from ten subjects are chosen as the training sample and the remaining one is chosen as test samples. Therefore, training and testing subjects are different from each other. I have repeated this process for all 10 subjects .The 10 results are then averaged to get the final facial expression recognition rate.

Table 7: Classification accuracy comparison in person dependent environment; (SRC: Sparse Representation-based Classification, GP: Gaussian Process classifier)

Author	*Classifier*	*Classification accuracy (%)*
LGCP (Proposed Method)	*LIBSVM*	*63.12%*
(Y. Zilu *et al.*, 2008)	SVM	58.20%
(J. Li *et al.*, 2012)	SRC	62.38%
(F. Cheng *et al.*, 2010)	GP	55.00%
(J. Lyons *et al.*, 1998)	SVM	56.80%

Table 8 shows results comparison in a person dependent environment on JAFFE dataset. I have followed a 10-fold cross validation with non-overlapping folds. I have divided the dataset into 10 random non-overlapping folds with nearly equal number of instances e.g. 21/22 each. I have used one fold for testing and rest 9 folds for training support vector machine. I have repeated this for 10 times with each independent fold and averaged the results to get the final recognition rate.

Table 8: Comparison of classification accuracy of proposed system with some other systems on JAFFE dataset in a person dependent experimental environment (NN: Neural Network, LDA: Local discriminant analysis)

Author	*Classifier*	*Classification accuracy (%)*
LGCP (Proposed Method)	*Multi Class SVM(Poly)*	*93.3%*
(K. Subramanian *et al.*, 2012)	SVM	88.09%
(J. B. M. J. Lyons *et al.*, 1999)*	LDA-based classification	92.00%
(Z. Zhang *et al.*, 1998)	NN	90.10%
(G. Guo *et al.*, 2003)	Linear Programming	91.00%

*used a subset of the dataset

In all cases, proposed system outperforms all the existing systems. 6-class expression gives more accuracy than 7-class expression. Reason for misclassifying is lack of proper face registration and variety of face shapes.

5. Conclusion

In this paper, I have proposed a novel method for facial feature extraction, which is insensitive to noise, and non-monotonous illumination changes. The method encodes spatial structure along with local contrast information for facial expression recognition. Extensive experiments prove that the proposed method is effective and efficient for expression recognition. I are planning to incorporate a effective face registration method in the preprocessing phase and boosting technique with support vector machine e.g. adaboost in the recognition phase in future.

References

[1] A. Turk, P. Pentland. "Face Recognition Using Eigenfaces." *in Proceedings of Computer Vision and Pattern Recognition*, 1991: 586-591.

[2] B. Fasel, J. Luettin. "Automatic Facial Expression Analysis: A Survey." *Pattrn Recognition* 36, no. 1 (2003): 259-275.

[3] C. Shan, S. Gong & P.W. McOwan,. "Robust Facial Expression Recognition Using Local Binary Patterns,." *Proc. IEEE Int'l Conf. Image Procession.* 2005. 370-373.

[4] C. Shan, S. Gong, P.W. McOwan,. "Facial expression recognition based on Local Binary Patterns: A comprehensive study." *Image and Vision Computing* 27, no. 6 (2009): 803-816.

[5] C.C.Chang and C.J.Lin."LIBSVM:a library for support vector machines "*ACM Transactions on Intelligent Systems and Technology* (2011).

[6] F. Cheng, J. Yu, and H. Xiong,. "Facial expression recognition in JAFEE dataset based on Gaussian Process classification." *IEEE Transactions on Neural Networks* 21, no. 10 (2010): 1685–1690.

[7] G. Guo, C. R. Dyer,. "Simultaneous Feature Selection and Classifier Training via Linear Programming: A Case Study for Face Expression Recognition." *Edited by in Proc. of IEEE Conference on Computer Vision and Pattern Recognition*. 2003. 346-352.

[8] G. Zhao, M Pietkainen. "Boosted Multi-Resolution Spatiotemporal Descriptors for Facial Expression Recognition." *Pettern recognition Letters* 30, no. 12 (2009): 1117-1127.

[9] G.L. Donato, M.S. Bartlett, J.C. Hager, P. Ekman, T.J. Sejnowski. "Classifying Facial Actions." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21, no. 10 (1999): 974-989.

[10] G.Littlewort, J. Whitehill, W. Tingfan,I. Fasel, M. Frank, J. Movellan & M. Bartlett,. "The computer expression recognition toolbox (CERT)" *IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG 2011).* 2011. 298-305.

[11] H. Kabir, T. Jabid, and O. Chae."Local Directional Pattern Variance (LDPv):A Robust Feature Descriptor for Facial Expression Recognition." *The International Arab Journal of Information Technology* 9(4) (2012).

[12] I. Cohen, N. Sebe, S. Garg, L. S. Chen and T. S. Huanga.Facial expression recognition from video sequences: temporal and static modelling. *Computer Vision and Image Understanding*, 91(2003), 160-187.

[13] J. Li, Z. Ying. "Facial Expression Recognition Based on Rotation Invariant Local Phase Quantization and Sparse Representation." *Edited by Computer, Communication and Control IEEE Second International Conference on Instrumentation & Measurement*. 2012. 1313-1317.

[14] J. Michael, Lyons, Miyuki Kamachi, Jiro Gyoba. "Japanese Female Facial Expressions (JAFFE),." (*Database of digital images*) 1997.

[15] K. Anderson and Peter W. McOwan. A Real-Time Automated System for the Recognition of Human Facial Expressions. *IEEE Transactions on Systems, Man, and Cybernetics*, 36, 1 (2006), 96-105.

[16] K. Subramanian, S. Suresh and R. V. Babu. "Meta-Cognitive Neuro-Fuzzy Inference System for Human Emotion Recognition." *Edited by IEEE World Congress on Computational Intelligence*. 2012. 10-15.

[17] L A. Jeni, András Lőrincz, Tamás Nagy, Zsolt Palotai, Judit Sebők, Zoltán Szabó & Dániel Takács. 3D shape estimation in video sequences provides high precision evaluation of facial expressions. *Image and Vision Computing*, 30, 10 (October 2012), 785-795.

[18] M. J. Lyons, J. Budynek, and S. Akamatsu. "Automatic classification of single facial images." *IEEE Trans. Pattern Analysis and Machine Intelligence* 21, no. 12 (1999): 1357–1362.

[19] M. J. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba. "Coding Facial Expressions with Gabor Wavelets." *Edited by Proc. of IEEE Int. Conference on Automatic Face and Gesture Recognition*. 1998. 200-205.

[20] M. Meulders, D. Boeck, V. Mechelen, A. Gelman. "Probabilistic Feature Analysis of Facial Perception of Emotions." *Applied Statistics* 54, no. 4 (2005): 788-793.

[21] M. Pantic and L. J. M. Rothkrantz. Automatic analysis of facial expressions: the state of the art. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22, 12 (2000), 1424–1445.

[22] M. Yeasin, B. Bullot and R. Sharma. Recognition of Facial Expressions and Measurement of Levels of Interest From Video. *IEEE Trans. Multimedia*, 8, 3 (2006), 500-508.

[23] M. Zia Uddin, J.J. Lee, T.S. Kim. "An Enhanced

Independent Component-based Human Facial Expression Recognition from Videos." *IEEE Trans. Consumer Electronics* 55, no. 4 (2009): 2216-2225.

[24] M.F. Valstar, I. Patras , M. Pantic. "'Facial Action Unit Detection using Probabilistic Actively Learned Support Vector Machines on Tracked Facial Point Data." *IEEE Int'l Conf. Computer Vision and Pattern Recognition* 3 (2005): 76-84.

[25] M.F. Valstar, M. Pantic. "'Fully automatic facial action unit detection and temporal analysis'." *Proc. IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR'06),* 2006: 149-156.

[26] M.S. Bartlett, G. Littlewort, I. Fasel & R. Movellan. Real Time Face Detection and Facial Expression Recognition: Development and Application to Human Computer Interaction. *In Proc. CVPR Workshop Computer Vision and Pattern Recognition for HumanComputer Interaction* (2003).

[27] P. Ekman and W. Friesen. Facial Action Coding System: A Technique for the Measurement of Facial Movement. *Consulting Psychologists Press, Palo Alto*, 1978.

[28] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar & I. Matthews. The Extended Cohn-Kande Dataset (CK+): A complete facial expression dataset for action unit and emotion-specified expression. *Paper presented at the Third IEEE Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010)* (2010).

[29] Robinson, G. "Edge detection by compass gradient masks." *Computer graphics and image processing*, 6 (1977): 492-501.

[30] S. Lajevardi, Z. Hussain. "Feature extraction for facial expression recognition based on hybrid face regions." *Advances in Electrical and Computer Engineering* 9, no. 3 (2012): 63-67.

[31] S. Naika C.L., S. Shekhar Jha, P. K. Das & S. B. Nair,. "Automatic Facial Expression Recognition Using Extended AR-LBP." *Wireless Networks And Computational Intelligence* 292, no. 3 (2012): 244-252.

[32] S.W.Chew, P.Lucey, S. Lucey, J. Saragih, J.F. Cohn & S. Sridharan,. "Person-independent facial expression detection using Constrained Local Models." *2011 IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG 2011),*. 2011. 915-920.

[33] T. Jabid, H. Kabir, and O. Chae. "Local Directional Pattern for Face Recognition." *Edited by IEEE International Conference on Consumer Electronics.* 2010. 329-330.

[34] T. Ojala, M. Pietikäinen and D. Harwood. A Comparative Study of Texture Measures with Classification Based on Feature Distributions. *Pattern Recognition, 19, 3* (1996), 51-59.

[35] T. Ojala, M. Pietikäinen & T. Mäenpää. Multiresolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24/7(2002), 971-987.

[36] V. Ojansivu and J. Heikkilä. Blur Insensitive Texture Classification Using Local Phase Quantization *in Proc. ICISP, Berlin, Germany* (2008), 236-243.

[37] Y. L. Tian, T. Kanade and J. F. Cohn. Recognizing action units for facial expression analysis. *IEEE Trans. Pattern Anal. Mach. Intell*, 23, 2 (2001), 97-115.

[38] Y. Tian, L. Brown, A. Hampapur, S. Pankanti, A. Senior, R. Bolle. "Real-World real-Time Automatic Recognition of Facial expressions." *Edited by in proceedings of IEEE Workshop on Performance Evaluation of Tracking and Surveillance*. 2003. 9-16.

[39] Y. Tian, T. Kanade, J. Cohn. "Facial Expression Analysis." *Handbook of Face Recognition, Springer*, 2003.

[40] Y. Zhang and Q. Ji. Active and dynamic information fusion for facial expression understanding from image sequences. *IEEE Trans. Pattern Anal. Mach. Intel.*, 27, 5 (2005), 699–714.

[41] Y. Zilu, L. Jingwen and Z. Youwei. "Facial Expression Recognition Based on Two Dimensional Feature Extraction,." *Edited by IEEE International Conference on Signal Processing.* 2008. 1440-1444.

[42] Z. Zhang, M. Lyons, M. Schuster, S. Akamatsu. "Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron." *Third IEEE International Conference on Automatic Face and Gesture Recognition, Proceedings.* 1998. 454-459.

Author Profile

Mohammad Shahidul Islam received his B. Tech. degree in Computer Science and Technology from Indian Institute of Technology-Roorkee (I.I.TR), Uttar Pradesh, INDIA in 2002, M.Sc. degree in Computer Science from American World University, London Campus, U.K in 2005 and M.Sc. in Mobile Computing and Communication from University of Greenwich, London, U. K in 2008. He has successfully finished his Ph.D. degree in Computer Science & Information Systems from National Institute of Development Administration (NIDA), Bangkok, Thailand in 2013. Currently he is serving as an associate professor in the department of computer science and Engineering at Atish Dipankar University of Science & Technology, Dhaka, Bangladesh. His field of research interest includes Image Processing, Pattern Recognition, wireless and mobile communication, Satellite Commutation and Computer Networking.