Robust Facial Expression Recognition via Sparse Representation and Multiple Gabor filters

Rania Salah El-Sayed
Dept. of Computer science
Faculty of Science,
Al-Azhar University
Cairo. Egypt

Prof.Dr. Ahmed El Kholy
Dept. of Mathematics
Faculty of Science
Al-Azhar University
Cairo. Egypt

Prof.Dr. Mohamed Youssri El-Nahas
Dept. of Systems
Engineering & Computers
Faculty of Engineering
Al-Azhar University, Cairo Egypt.

Abstract—**Facial expressions recognition plays important role in human communication. It has become one of the most challenging tasks in the pattern recognition field. It has many applications such as: human computer interaction, video surveillance, forensic applications, criminal investigations, and in many other fields. In this paper we propose a method for facial expression recognition (FER). This method provides new insights into two issues in FER: feature extraction and robustness. For feature extraction we are using sparse representation approach after applying multiple Gabor filter and then using support vector machine (SVM) as classifier. We conduct extensive experiments on standard facial expressions database to verify the performance of proposed method. And we compare the result with other approach.**

Keywords—Facial expression recognition (FER); L1-minimization; sparse representation; Gabor filters; support vector machine (SVM).

I. INTRODUCTION

Facial expressions are the changes of facial appearance in response to a person's internal emotional states, intentions, or social communications. And facial expression is one of the most powerful, natural, non-verbal, and immediate means for human beings to communicate their emotions and express their intentions [1]. By recognizing facial expressions from facial images, a number of applications in the field of human computer interaction can be facilitated. We can use facial expressions not only in human computer interaction and data-driven animation but also we can use in video cameras which have recently become an integral part of many consumer devices and can be used for capturing facial images for recognition of people and their emotions.

Last two decades, the developments, as well as the prospects in the field of multimedia signal processing have attracted the attention of many computer vision researchers to concentrate in the problems of the facial expression recognition.

The task of automatically recognizing different facial expressions in human-computer environment is significant and challenging. A variety of systems have been developed to perform facial expression recognition and each system consists of three stages: first, face acquisition; second, facial feature extraction then facial expression classification [16].

Face acquisition is a preprocessing stage to detect face regions in the input images or sequences. One of the most widely used face detector is the real-time face detection algorithm developed by Viola and Jones [14].Once a face is detected in the images, the corresponding face regions are usually normalized.

The automatic facial expression recognition literature can be divided into three approaches: geometric features-based systems, appearance features-based systems and both geometric and appearance feature-based systems [6].

In the geometric features-based systems, the shape and locations of major facial components such as mouth, nose, eyes, and brows, are detected in the images. The approaches used in such system are Active Appearance Model (AAM) [5], manually locating a number of facial points. And recently, Piecewise Bezier Volume Deformation tracker (PBVD) was proposed [6].

In the appearance features-based systems, the appearance changes with changes in facial expressions including skin texture changes, wrinkles, bulges, and furrows, are presented. And in this case image filter can be applied to either the whole-face or specific face regions to extract the facial appearance changes. The existing algorithms are Principal Component Analysis (PCA), Independent Component Analysis (ICA) [20], Linear Discriminant Analysis (LDA) [15], Gabor wavelet analysis, Local Binary Patterns (LBP) [19], Local Phase Quantization (LPQ) [15], The Local Directional Pattern (LDP) [11],Fisher Discriminant Analysis (FDA)[21], Local Feature Analysis (LFA), Circular Local Binary Pattern (CLBP) [8], Speeded Up Robust Features (SURF), and Neural Network Architecture-based 2-D Log Polar Gabor Transform.

In geometric and appearance feature-based methods Yang et al use local Gabor Binary Pattern (LGBP) [14]. Histograms with shape and texture parameters of an Active Appearance Model (AAM)are presented in[9].

The last step of an automatic facial expression recognition system is facial expression classification. A classifier is employed to identify different expressions based on the extracted facial features, such as Nearest Neighbor (NN), K-Nearest Neighbor (KNN)[18],Support Vector Machine (SVM)

[17], and Hidden Markov Model (HMM).Facial expression recognition system can be presented as figure1.
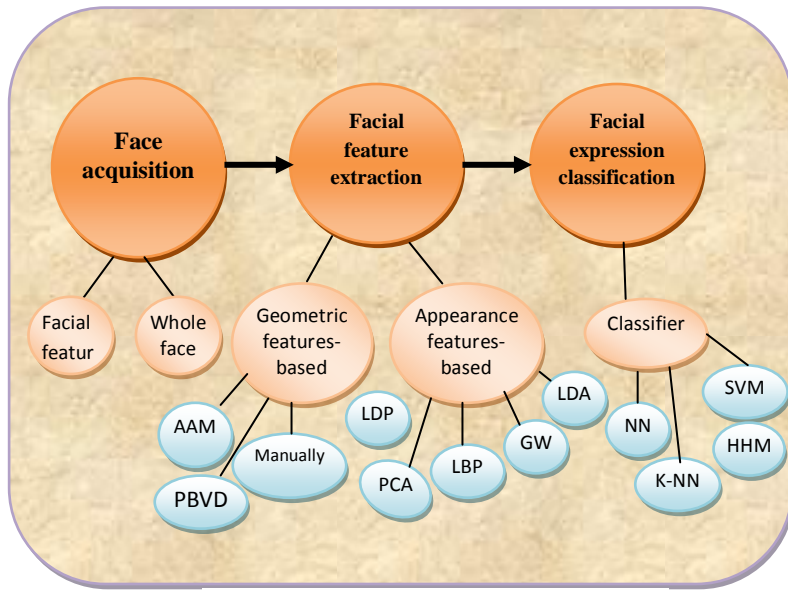

Fig.1. facial expression recognition with associate methods

Sparse representation, also known as compressed sensing [2], has been applied recently to image based facial expression recognition. With sparse representation we can represent each expression by a set of features, which sufficiently characterize each individual expression. With the prior knowledge that expressions of the same individual are similar to each other, a test expression can be considered as being well approximated by linearly combining the number of expressions that have same individual in the training set.

With all the promising advantages, the use of sparse representation for facial expression recognition does not exterminate obstacles posed by several practical issues such as lighting, pose, and robustness to facial expression .In this research we propose a framework which can achieve robustness to facial expressions variations, by using sparse representation of multiple Gabor wavelet filters and SVM as classifier.

This paper is organized as follows. Section 1 is this introduction. Section 2 discusses relevant related works from the literature. Then we introduce Gabor filter's principle, property, and the feature characterization in details in Section 3.In section 4, discussion of sparse representation for facial expression matching is provided. In section 5, details of proposed algorithm are presented. Section 6, provides the details of the conducted experiments along with results .The paper concludes with a brief summary of results and proposal of future research directions in section 7.

II. RELATED WORK

Most facial expression recognition systems use the six principle emotions of Ekman [10]. Ekman considers six basic emotions: happiness, surprise, fear, anger, disgust, sadness; and categorizes facial expressions with these six basic emotions.

A variety of systems have been developed to perform facial expression recognition. Some systems concentrate in two stages: feature extraction and expression classification. Luiz S. Oliveira, Alessandro L. Koerich, and Marcelo Mansano [8] applied feature selection method on the feature matrix extracted by the 2DPCA method, then using KNN and SVM as classifier.

Little wort et al. [22] select the best set of Gabor filters using Gentle Boost and train support vector machines (SVMs) to classify Action Unit (AU) activation.

Tingfan Wu et al. [4], facial expression recognition systems are modeled asa serial pipeline. The first layer detects and normalizes (registers) face images. Second layer applies Gabor energy filters (GEFs) and local Binary Patterns (LBPs) respectively. The final layer consists of a SVM classifier.

Hansung Lee et al. [11], introduce expression recognition systems consist of firstly; Gabor-LBP histogram for face image representation. Secondly, Sparse Representation Classification (SRC) for the face image retrieval.

Praseeda lekshmi.V et al.[12], present a method to analyze facial expression from images by applying Gabor wavelet filter banks on face images. Then theyuse the support vector machine(SVM) for classification.

Zhen Wang et al. [8], expression systems consist of firstly using Gabor filters with different angles, then 2DPCA was used to extract face features and finally Nearest Neighbor (NN) classifier.

In this paper we introduce new hybrid system based on sparse representation, Gabor wavelet filter for feature extraction and support vector machine as classifier. Experimental results show how these techniques reduce error rate of facial expression recognition.

III. GABOR FILTERS

In recent years, Gabor filters have been successfully used in many applications, because they possess the optimal localization properties in both spatial and frequency domain. And it is obtained by modulating a sinusoid with a Gaussian[12].

The Gabor wavelet representation of images allows description of spatial frequency structure in the image while preserving information about spatial relations.

Let r(x,y) be the original image f(x,y) convolved with a 2D Gabor function g(x,y)[14]

r(x,y) = f(x,y) * g (x,y)

Where * denotes the convolution operator and with g(x,y) being a 2D Gabor filter defined as [16]

$$g_{\lambda,\theta,\varphi,\sigma,\gamma}(x,y) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda} + \varphi\right) \quad (1)$$

$$x' = x\cos\theta + y\sin\theta$$

$$y' = -x \sin \theta + y \cos \theta$$

The parameters used in Eq.(1)are defined as follows. Wavelength (λ); this is the wavelength of the cosine factor of the Gabor filter kernel, Orientation(s) (θ); it specifies the orientation of the normal to the parallel stripes of a Gabor function. Its value is specified in degrees. Valid values are real numbers between 0 and 360.The Phase offset(s) (φ) in the argument of the cosine factor of the Gabor function is specified in degrees. Valid values are real numbers between -180 and 180.The Aspect ratio (γ) called more precisely the spatial aspect ratio, specifies the ellipticity of the support of the Gabor function, and we can calculate wavelength of Gabor filter using Bandwidth (b) in eq. 2.

$$\frac{\sigma}{\lambda} = \frac{1}{\pi}\sqrt{\frac{ln2}{2}}\cdot\frac{2^b + 1}{2^b - 1} \quad , \quad b = log_2 \frac{\frac{\sigma}{\lambda}\pi + \sqrt{\frac{ln2}{2}}}{\frac{\sigma}{\lambda}\pi - \sqrt{\frac{ln2}{2}}} \quad (2)$$

Figure 1, show Gabor filter kernels with values of the orientation parameter of 0, 45 and 90degrees, from left to right, respectively. The values of the other parameters are as follows: wavelength 10, phase offset 0, aspect ratio 0.5, and bandwidth 1.
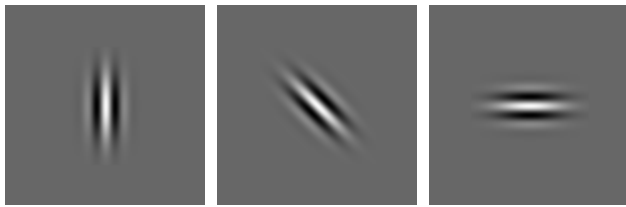

Fig.2. Gabor filter with three orientations in the spatial and frequency domain

IV. SPARSE REPRESENTATION

Sparse representation aims at describing a high-dimensional signal as a linear combination of a few generating elements.. Specifically, a feature vector in a given class of dimension M should be represented by a small subset of a collection of all elements N. when M=N, the elements are independent, the collection is termed a basis and the representation is unique [13].

Let the training samples be denoted as A = [A₁ A₂ ... A_r], where each column A_i denotes a facial image and r the total number of training samples.

For a test sample $\mathbf{y} \in R$, it can be approximated by linear combination of selected training samples. The test sample can be expressed as [3]

$$\mathbf{y} = x_1 A_1 + x_2 A_2 + \dots + x_r A_r = A\mathbf{x} \qquad (3)$$

Where $\mathbf{x} = [x_1, x_2, \dots, x_r]$ is coefficient vector whose entries are zero except some important training samples. The problem of sparse representation is to find the sparse solution $\mathbf{x}$. This sparse solution $\mathbf{x}$ explains how to combine all training samples to obtain a global representation of test sample. This is an L1 optimization problem formally defined as

$$\mathbf{x}_1 = \min \| \mathbf{x} \|_1 \text{ subject to } \mathbf{y} = A\mathbf{x} \qquad (4)$$

V. SPARSE REPRESENTATION CLASSIFIER

Sparse representation classifier (SRC) chooses the training samples which can best represent each test sample. For i th class, let δ_i be the function that selects the sparse coefficients associated with the i the class. In $\delta_i(x)$, only the elements that are associated with ith class are nonzero [11]. That is

$\mathbf{x} = [0, \dots, 0, \dots, x_{i,1}, \dots, x_{i,ni}, \dots, 0, \dots, 0]^T$. The test sample $\mathbf{y}$ can be approximated by training samples from each class as $\hat{\mathbf{y}}_i = A \delta_i(\hat{\mathbf{x}})$. Define the residual between test sample y and approximation $\hat{\mathbf{y}}_i$ from each class sample as $r_i = \|y - \hat{y}\|_2$. The test facial expression are assigned to the class that minimizes the residual between $\mathbf{y}$ and $\hat{y}$

$$\text{identity}(\mathbf{y}) = \arg \min r_i(\mathbf{y}) \qquad (5)$$

$r_i(y)$ represent the classification result of the i th component of test sample. The final result is determined by voting methods. Let the output of the SRC classifier be defined as [7]

$$r = [r_1, r_2, \dots, r_r] , r_i \in \{c_1, c_2, \dots, c_r\}$$

Then we define the binary classification function as

$$f(c_r) = \begin{cases} 1 & if \ r = c \\ 0 & if \ r \neq c \end{cases} \qquad (6)$$

The test sample is assigned to the class label which is most observed at the classifier outputs.

For reference we include the algorithm1 given in [23] as follows:

A. Algorithm 1

- **Input:** n training samples divided into r classes, A₁, A₂,...,A_r and a test image y.
- Set $A = [A_1 A_2 \dots A_r]$.
- Normalize columns of matrix A
- Solve the L1 minimization problem :
  $\min \| x \|_1$ subject to $y = Ax$
- **For** each subject i ,
  Compute the residuals: $r_i = \| y - Ax_i \|_2$ where x_i is obtained by setting the coefficients in x, corresponding to training samples not in class i, to zero.
- **End**
- **Output**: the class with the smallest residual
  identity($\mathbf{y}$) = arg min $r_i(\mathbf{y})$.

The system architecture for facial expression recognition using sparse representation is presented as shown in Figure 2
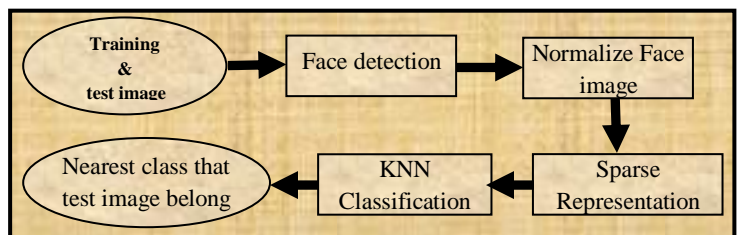

Fig.3. diagram of sparse representation approach

VI. PROPOSED FRAMEWORK

In this paper we propose a new and robust framework for facial expressions recognition. Robustness to variations in illumination conditions and pose variations is achieved by using both Multiple Gabor filters and sparse representation. Then using a SVM for classification should increase system robustness due to generalization property of SVM. The system architecture is shown in Figure3. Face detection is implemented by Viola-Jones method. We focus in this paper on the procedures for feature extraction, sparse representation and SVM classification.
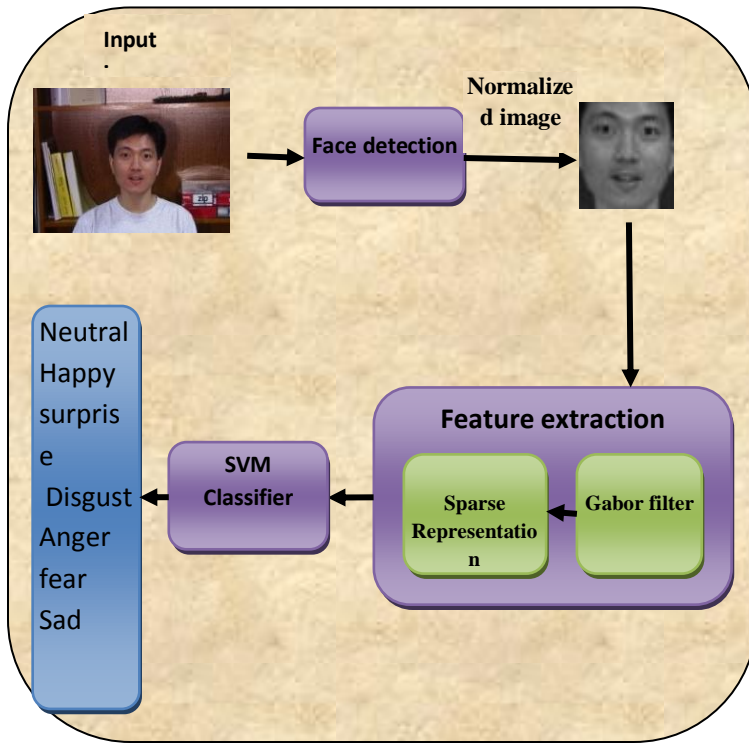

Fig.4. System architecture for proposed facial expressions recognition

A. The Proposed Procedure

- **Input**: n training samples partitioned into r classes, $S_1, S_2,...,S_r$ and a test sample z.
- **For** each subject i ,
    Apply Gabor filter for N times; and
    generate A_i
- **End**
- Apply Gabor filter on test sample z for N times; and generate **y**.
    - Set matrix A based on Gabor filter of images
        $A = [A_1 A_2 ... A_r]$.
- Normalize columns of matrix A
- Solve the L1 minimization problem
    $\min \|x\|_1$ subject to $y = Ax$
- **Compute** yEstimated = Ax_i,
    where x_i is obtained by setting the coefficients in x,

corresponding to training samples not in class i, to zero
- **Classify** yEstimated with **SVM** classifier
- **Output:** class to which yEstimated belongs.

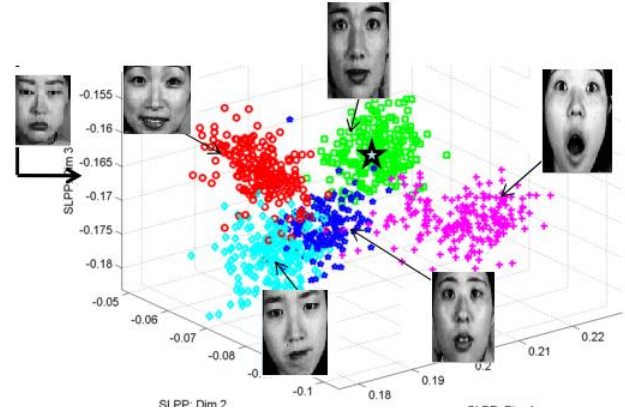Figure 4schematize show SVM classifies the facial expressions.

Fig.5. Concept of SVM classifier for facial expressions.

VII. EXPERIMENTS

In this section, we quantitatively verify the performance of different facial expression recognition systems. Such as algorithm 1 with SVM classifier, proposed algorithm, proposed algorithm with K-NN classifier, and Gabor filter with SVM classifier. Using public facial expression databases namely, JAFFE Database which contains 213 images of female facial expressions. Each image has a resolution of 256×256 pixels. The number of images corresponding to each of the seven categories of expressions (anger, joy, sadness, neutral, surprise, disgust and fear) is the same. Figure 5 shows example of facial expression images from JAFFE Database [2].
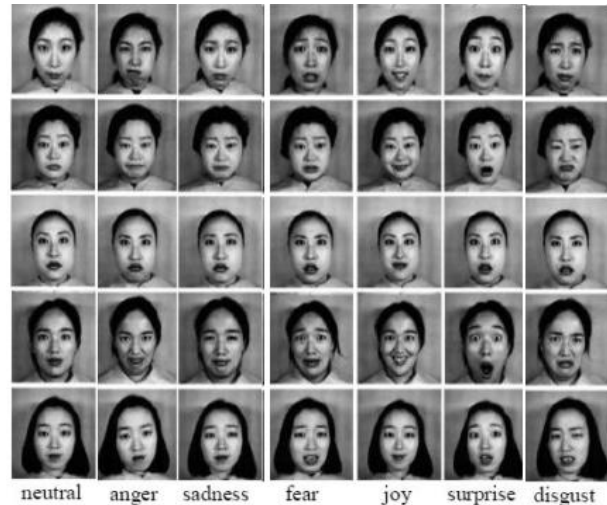

Fig.6. example facial expression images from JAFFE database

In this experimental work, we evaluate five approaches for facial expression recognition and compare the recognition rate of the previously published work with recognition rate of the

proposed approach. First approach [1] used sparse representation to extract features and support vector machine (SVM) for classification, the second applies multiple Gabor filter and uses NN as classifier [25][26], the third approach as done in [2] uses sparse representation classifier after application of Gabor filter. Fourth approach is similar to third one but this time SVM is used as classifier [24], the last is our proposed approach which uses multiple Gabor filters and sparse representation to extract features and then apply SVM as classifier.

The recognition rate for facial expressions for the five approaches is summarized as show in Table 1.

TABLE I. CONFUSION MATRIX FOR 7-CLASSES OF FACIAL EXPRESSIONS

methods	happy	sadness	fear	disgust	surprise	anger	natural	average
SR+SVM	75%	80%	75%	90%	95%	90%	95%	85.71%
Gabor filter NN [26][25]	85%	75%	75%	90%	95%	90%	95%	86.42%
Gabor filter SRC[2]	85%	85%	75%	90%	95%	90%	95%	87.85%
Gabor filter SVM[24]	90%	75%	75%	90%	85%	90%	90%	85%
Proposed (Multiple Gabor filter SR+SVM)	90%	80%	80%	95%	100%	90%	90%	89.28%

TABLE II.
CONFUSION MATRIX WITH PROPOSED ALGORITHM ON THE JAFFE DATABASE

Expression	Anger %	Disgust%	Joy %	Fear%	Sad %	Surprise%	Neutral%
Anger	90	0	0	5	5	0	0
Disgust	0	95	0	5	0	0	0
oy	2	0	90	2	0	1	5
Fear	0	12.5	0	80	7.5	0	0
Sad	0	0	5	5	80	0	10
Surprise	0	0	0	0	0	100	0
Neutral	0	0	0	0	5	5	90

Table 1, shows that the proposed approach achieves the highest performance in facial expression recognition (89.28%) because we used the properties of Gabor filter and sparse representation in the stage of extract feature and then using the properties of support vector machine (SVM) in classification. Compared with the previously reported work [1,2,24,25,26] in which the experimental settings are similar to ours, In [24], on

7-class facial expression recognition tasks they used Gabor filter to extract features, SVM for classification and reported an accuracy of (85%). Additionally, they also make the same test [25] but using NN classifier and reported an accuracy of (86.42%).

In [2], they used Gabor filter to extract features and sparse representation classification and reported an accuracy of (87.85%).

Table 2, shows the confusion matrix of 7-class expressions when using multiple Gabor multiple filters and sparse representation for feature extraction and SVM as classifier. From Table 2, it can be seen that 5-class facial expressions are identified very well with an accuracy of over 90%.

VIII. CONCLUSION

We have presented a hybrid framework for facial expressions recognition based on sparse representation and multiple Gabor filter. Designing a good filter and classifier is a crucial step for any successful facial expression recognition system. An average recognition rate of (89.28%) is achieved under facial expressions variations. This means that our approach achieves a high recognition rate compared to other approaches in published literature. Using multiple Gabor filters rendered the method robust to facial expression variations because each filter has specific property to extract. In addition using generalization property of SVM classifier increased the recognition rate in presence of facial class variations.

We believe that facial expressions recognition under varying conditions is still an interesting area of research, and we anticipate that there will be many further advances in this area.

REFERENCES

[1] Xia H. "Robust Facial Expression Recognition via Sparse Representation Over Overcomplete Dictionaries". *Journal Of Computational Information Systems*. 2012;1(January):425-433.
[2] Zhang S, Zhao X, Lei B. "Facial Expression Recognition Using Sparse Representation". *Wseas Transactions On Systems*. 2012;11(8):440-452.
[3] Zhang S, Zhao X, Lei B. "Robust Facial Expression Recognition via Compressive Sensing". *Sensors (Peterborough, NH)*. 2012:3747-3761.
[4] Wu T, Butko NJ, Ruvolo P, et al. Multilayer Architectures for Facial Action Unit Recognition. 2012;42(4):1027-1038.
[5] Senechal T, Rapp V, Salam H, et al. "Facial Action Recognition Combining Heterogeneous Features via Multikernel Learning". IEEE. 2012;42(4):993-1005.
[6] Valstar MF, Mehu M, Jiang B, Pantic M, Scherer K. "Meta-Analysis of the First Facial Expression Recognition Challenge". *IEEE*. 2012;42(4):966-979.
[7] Zhi R, Ruan Q. "Facial expression recognition with facial parts based sparse representation classifier". *Pattern Recognition*. 2009;7496:1-7.
[8] Yang J, Oliveira LS, Koerich AL, Mansano M. "2D Principal Component Analysis for Face and Facial-Expression Recognition". *Technology*. 2011:9-13.
[9] Valstar MF, Mehu M, Jiang B, Pantic M, Scherer K. "Meta-Analysis of the First Facial Expression Recognition Challenge". *Emotion*. 2012;42(4):966-979.
[10] [Donato, G.; Bartlett, M.; Hager, J.; Ekman, P.; Sejnowski, T. "Classifying facial actions". IEEE Trans. Pattern Anal. Mach. Intell. 1999, 21, 974–989.
[11] Lee H, Chung Y, Kim J, Park D. "Face Image Retrieval Using Sparse Representation Classifier with Gabor-LBP Histogram ". 2011:273-280.
[12] Paper P. "Analysis of Facial Expression using Gabor and SVM". *International Journal*. 2009;1(2).

[13] Shin Y-suk. "Facial Expression Recognition Based on Dimension Model of Emotion with Autonomously Extracted Sparse Representations". *Emotion*.2004; (Dlm):81-87.

[14] Yang M, Zhang L. "Gabor Feature based Sparse Representation for Face Recognition with Gabor OcclusionDictionary". Springer-Verlag Berlin Heidelberg 2010:pp. 448–461.

[15] Zhao X, Zhang S. "Facial expression recognition using local binary patterns and discriminant kernel locally linear embedding". *EURASIP Journal on Advances in Signal Processing*. 2012(1):20.

[16] Bashyal S, Ã GKV. "Recognition of facial expressions using Gabor wavelets and learning vector quantization". Engineering Applications of Artificial Intelligence. 2008.

[17] I Kotsia, and I Pitas, "Facial expression recognition in image sequences using geometric deformation features and support vector machines", IEEE Transactions on Image Processing, Vol. 16, No. 1, 2007, pp. 172-187.

[18] Yousefi, S.; Nguyen, M.P.; Kehtarnavaz, N.; Cao, Y. Facial expression recognition based on diffeomorphic matching. In Proceedings of 17th IEEE International Conference on Image Processing (ICIP), Hong Kong, China, 26–29 September 2010; pp. 4549–4552.

[19] S. Moore and R. Bowden, "Local binary patterns for multi-view facial expression recognition," Compute. Vis. Image Understand., vol. 115, no. 4, pp. 541–558, Apr. 2011.

[20] M. S. Bartlett, J. R. Moverllan, T. J. Sejnowski, "Face recognition by independent component analysis," IEEE Transactions on Neural Network, 13 (6), 1450--1464 (2002).

[21] Simona E. Grigorescu, Nicolai Petkov, and Peter Kruizing, "Comparison of Texture Features Based on Gabor Filters". IEEE, VOL. 11, NO. 10, october2002

[22] G. Littlewort, J. Whitehill, T. Wu, I. Fasel, M. Frank, J. Movellan, and M. Bartlett, "The computer expression recognition toolbox (cert)," in Proc. IEEE Int. Conf. Autom. Face Gesture Recog., 2011, pp. 298–305.

[23] Wright J, Member S, Yang AY, et al. Robust Face Recognition via Sparse Representation. IEEE. 2009;31(2):210-227

[24] R. Ramanathan, K.P. Soman, Arun. S. Nair, V. VidhyaSagar, and N. Sriram. A support vector machines approach for efficient facial expression recognition. Advances in Recent Technologies in Communication and Computing, International Conference on, 0:850{854, 2009.

[25]] Jia-Jun Wong and Siu-Yeung Cho. " Facial emotion recognition by adaptive processing of tree structures ". In SAC '06: Proceedings of the 2006 ACM symposium on Appliedcomputing, pages 23{30, New York, NY, USA, 2006. ACM.8

[26] curryguinn, thomasjanicki," active appearance models for affect recognition using facial expressions".2010