

Recognition of Facial Gestures using Gabor Filter

Subhashini Ramalingam
VIT University, India

Dr Ilango Paramasivam
VIT University, India

Mangayarkarasi Ramiah
VIT University, India

ABSTRACT

Vision is the task of “seeing”. When human see things, their eyes (sensing device) capture the image, then pass the information to brain (interpreting device). The brain interprets the image, gives us meanings of what human see [35]. Similarly, in computer vision, camera serves as sensing device, and computer acts as interpreting device to interpret the image what the camera captures. Gestures are expressive meaningful body motions i.e., physical movements of the hands, arms, fingers, head, face or other parts of the body with the intent to convey information or interact with the environment[12]. Gestures are used for everything from pointing at a person or an object to change the focus of attention, to conveying information. Gestures, which function independently of speech, are referred to as autonomous gestures. Autonomous gesture can also represent motion commands to use in communication [15] and machine control. Gesture recognition is the process by which gestures made by the user are made known to the intelligence system. The core objective of the proposed work is to detect and recognize various facial gestures that are present in a given image using Gabor filter and use it for automation. The performance of the proposed method is evaluated using Gabor filtering and compared with the other methods namely wavelet and neural networks. Finally, it is concluded that the proposed method shows better performance over the other methods.

Keywords

Image processing, Gabor filters, computer vision.

1. INTRODUCTION

Human machine interaction (HMI) takes an important role in computer vision technology. The HMI takes place by using gestures. Gestures may be hand, face, some other body gestures or combination of all these things. The major steps that are used to identify the gesture of the person on the given image are as follows. Capturing of face image, identification of the face on the image, preprocessing of the image for enhancing the facial features, applying the Gabor filter on the face (processing) [11] at different angles, position of the kernel, extracting the features of the image on the filtered images, forming the feature vector for supervised training on the Bayes Classifier system[32]. Identifying the gesture of the face using the same procedure. The main challenges faced by the system are finding out the correct mechanism of finding out the face images using object detection [21]. Marking of the Region of Interest (ROI) for the face, Extracting out the features of the face using Gabor, Formation of feature vector and correct training of the system with 100's of samples for each gesture of the face. (2500 images are totally used for training and testing).

Gesture Analysis for

- HCI- Human Computer Interaction
- providing Security for safe driving
- Improving Computer Vision
- Context Based working

Facial analysis plays a vital role in Human Computer Interaction. This can be a good step for solving problems which are faced by the today's HCI. If computer can understand human emotions then that will be a great leap in the HCI. This HCI has many sub domains like face gesture, hand gesture, body gesture etc., which the humans can in easily learn but computers cant. In transportation, Safety measures play the important role in avoiding accidents. Through modern technology we can improve the safety by ensuring the driver's consciousness on the driving. [31]Using facial gesture analysis driver's attention in driving can be monitored. This can be performed by ensuring driver's eyes open or not, checking driver is attentive to the road etc. If the systems finds fault then it can give alerts to the driver to concentrate in driving or in advanced cars automatic mode can be carried over and stop the car by the side of the road [3]. Like Human vision computer can't differentiate objects whichever it sees. This problem is the heart of the robotics [29] which is a very hot research area people all over the world. Facial analysis is a sub domain of the main problem in the computer vision i.e. making computer to imitate human eye. This facial analysis [39] can be a good step solves that problem.

1.1 Context Based Working

Suppose if a person shows a sign in a home and the same sign in an office. Meaning of the sign will change entirely from place to place. This can't be easily recognizable by computer vision. This is called as context based image processing. Facial analysis can be used with context based to come out with different new features in HCI which will have lots of commercial applications. Example for this context based gesture analysis[28] is [15].The objective of the proposed work is to detect and recognize various hand gestures and special hand gestures that are present in a given image and use for automation. This Hand Gesture Recognition technique is fast (easy to calculate the orientation histograms), robust to illumination changes and takes less processing time due to little training. If the systems finds fault then it can give alerts to the driver to concentrate in driving or in advanced cars automatic mode can be carried over and stop the car by the side of the road.

1.2 Improving Computer Vision

Like Human vision computer can't differentiate objects whichever it sees. This problem is the heart of the robotics which is a very hot research area people all over the world. Facial analysis is a sub domain of the main problem in the computer vision i.e. making computer to imitate human eye [28]. This facial analysis can be a good step solves that problem.

The way human interacts with computer started evolving from days when people started using computers for their daily lives. Most successful interfaces know to common man is keyboard and mouse, other than that many other interfaces have been tried out by researchers for decades. Some hardware devices like kindest [9] are most successful in the present day scenario, but these kinds of devices are costly as per now. Considering the human interactions with computers with help of gestures is long awaited goal. Now a day's these is becoming reality. In this survey we can see the importance of gestures and different methods which are tried out for facial gestures are discussed.

As indicated by Mehrabian [8] in face-to-face human communication only 7% of the communicative message is due to linguistic language, 38% is due to paralanguage, while 55% of it is transferred by facial expressions. Development of an automatic facial expression analyzer has attracted great attention in these decades. Tina *et al.* developed an Automatic Face Analysis (AFA) system to analyze facial expressions based on both permanent facial features (brows, eyes, mouth) and transient facial features (deepening of facial furrows). The AFA system recognizes fine-grained changes in facial expression and turns them into action units (AUs) of the Facial Action Coding System (FACS) [33], instead of a few prototypic expressions. However, the AFA[14] system requires accurate locations of the facial features, and further efforts are demanded to implement a corresponding model-driven facial expression synthesis system under the framework of AFA. Ekman and Friesen [2] developed the most comprehensive system for synthesizing facial expressions based on what they call Action Units (AU). They defined the facial action coding system (FACS) [15]. FACS consists of 44 action units (AU), which describe

2. FACIAL ACTION CODING SYSTEM

FACS (Facial Action Coding System) gives us the six standard facial expressions namely *joy, anger, fear, disgust, sadness* and *surprise*. There are two main ways to in doing facial analysis they are using static images [12],[13] and another is using motion video facial expression.[1],[9].The main aim of this proposed work is to find these six expressions. Various papers which concentrate on this area mainly deals with two approaches one is the training based and then the mathematical based. The advantage of using training based is result we are getting will be accurate but more sampling is required, this approach uses neural networks etc for training a machine. Second approach is the mathematical approach which doesn't require training but it fails sometime more than the other technique. This proposed work concentrate on the training based method which has the following steps, image capture from camera, preprocessing, feature extraction, classification and recognition. Face detection is the primary step of facial gesture recognition for detecting the face people use different methods one such is

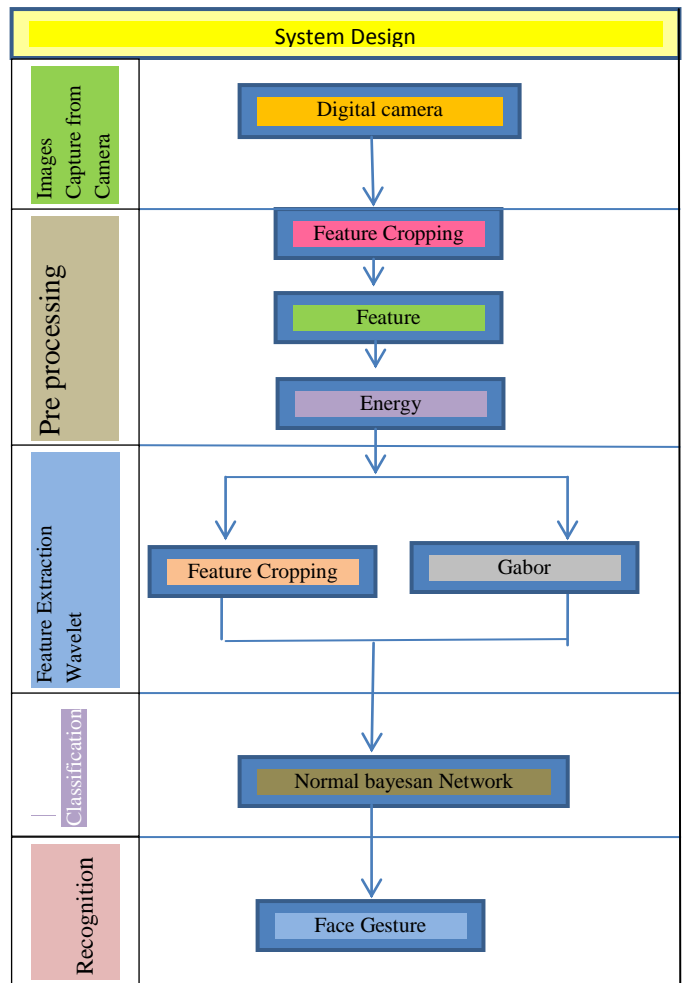
the Viola-Jones object detection framework [10] this object detection framework can be used for simple positioning of the ROI (Region of Interest).

2.1 Image Processing

Image processing is a rapidly growing area of computer science. Technological advances in digital imaging, computer processors and mass storage devices have fueled its growth. Important examples are medicine, film and video production, photography, remote sensing, and security monitoring [22]. These and other sources produce huge volumes of digital image data every day, more than could ever be examined manually.

3. PROPOSED ARCHITECTURE

Main design of this proposed work can be given by the following



3.1 Data Capture from the Facial Expression Database

Database used for this proposed work is the Cohn-Canada AU-Coded Facial Expression Database. This database consist of the database includes approximately 2000 image sequences from over 200 subjects. They ranged in age from 18 to 30 years. Sixty-five percent were female, 15 percent were African-

American, and three percent were Asian or Latino. Note: Selected images from the database are given for training and testing of the data [23].

3.2 Pre-Processing

In the pre-processing step, converting the color images into gray scale images happens followed by smoothing or enhancement technique to make the dark images into intelligible / clear images.

3.3 Region of Interest

Before processing the image region of interest is selected from the whole image so that processing is done only on the face not the areas around it. This step is very crucial in the system. A wrong selection of region of interest will lead to wrong recognition of gesture. For this region detection, face detection algorithm can be used in this module. Haar cascade face detection [10] is used for detecting the face on the image which follows the principle of Viola-Jones object detection framework [27].

3.4 Processing

After the Preprocessing Step is completed then image's Energy is found and then the image is Energy normalized. Energy Value of the image can be got from the following

$$F(m, n) = \sum_1^m \sum_1^n f^2$$

After Finding the Energy of the image then

We have to normalize the whole image with that energy by the following way.

$$F(m, n) = f(m, n) / (\text{energy})^{1/2}, m, n = 1, 2, 3, \dots, N$$

Energy Normalized image is sent for Wavelet decomposition text. For two addresses, use two centered tabs, and so on [37]. For three authors, you may have to improvise.

3.5 Wavelet Decomposition Using Db2

Wavelet decomposition consists of series of steps such as convolution and de-sampling the image in different levels. Three types of details we can obtain from each level of the decomposition. They are the Horizontal [18], Vertical and Diagonal Details. Then further image is decomposed. This below image represents the sub band labeling scheme for 2-level 2D Wavelet.

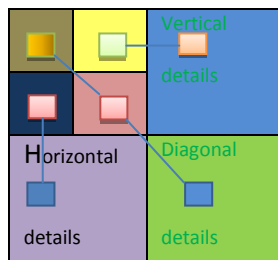


Figure 1 Wavelet representation

The three images which we get from the one level of decomposition can be expressed as LH-Low High Band, HL-High Low Band, and HH- High-High Band [22]. The three bands of images are calculated using the wavelet analysis. In the wavelet analysis series of (shift multiple and sum) called convolution.

3.6 Gabor Filter

Features based on Gabor filters have been used in image processing due to their powerful properties. Gabor kernels are characterized as localized, orientation selective, and frequency selective [41]. A family of Gabor kernel is the product of a Gaussian envelope and a plane wave. A 2D Gabor filter is expressed as a Gaussian modulated sinusoid in the spatial domain and as shifted Gaussian in the frequency domain. The Gabor wavelet [21] representation of images allows description of spatial frequency structure in the image while preserving information about spatial relations. The Equation of Gabor is customized as

$$G(x, y; \varphi, \sigma, \psi) = \exp(-(x^2 + y^2) / 2 \sigma) * \cos(w * y * \cos(\varphi) + w * x * \sin(\varphi) + \psi)$$

Filter of Gabor can be drawn as follows. We can vary φ (i.e. angle); ψ (i.e. psi) of the filter to derive at these different filters as below.

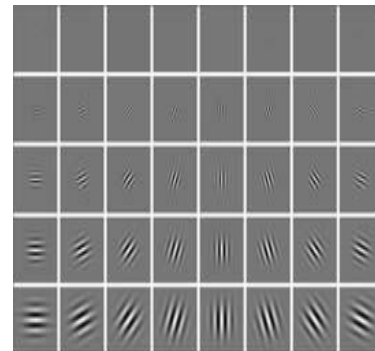


Figure 2 Forty bank filter of Gabor

Using these filters images we can convolve it with the face image to produce the Gabor filtered images. From these images we can extract the features of the face like eye brows, eyes, mouth, etc., and a feature vector is generated and then given to the classifier. Different people use different ways to form the feature vector for training the classifier[34]. Some of them even use whole image as a feature vector and perform classification which needs high computation. So here feature vector is taken from important values of the image from each filter Energy, mean and standard deviation forming a 32 value feature vector for every image.

3.7 Classification

After the feature vector is formed then it is given to the classifier to train the feature under supervised learning and this process is repeated for many images until training data becomes sufficient. For classification of gestures [37] we can

you Normal Bays Classifier to classify the different gestures of the face. A **Bays classifier** is a simple probabilistic classifier based on applying Bays' theorem (from Bayesian statistics [16]) with strong (naive) independence assumptions. A more descriptive term for the underlying probability model would be "independent feature model).Input image either from camera or from database is given into the system as input [20]. ROI of the image is selected and then Gabor features are taken from the image and then the values are stored in the data file for training. Description of training and testing of datasets are as follows.

3.8 Training

Training of the system is done with 100 images of each gesture. Each image input is given and the system is taught with the gesture type of it, this process is the supervised learning. After these processes are done, feature vector values which are stored in a text file is taken and classification is done with help of **Bays Classifier** [38].Bays Classifier gives out the model file which is saved in a XML format. This can be used for testing. For testing 30 images of each gesture is taken and tested with the created model of the system.

Details of Training and testing

Minimum Number of training images per Gesture – 100

Minimum Number of testing images per Gesture – 30

After getting the feature for training images all the data is stored in data file is taken for training. In the data file feature vector are stored as follows

Gesture type value1, value2, value3.....

When given to the training system for classification it will produce the XML file. That is taken as the model for the system.

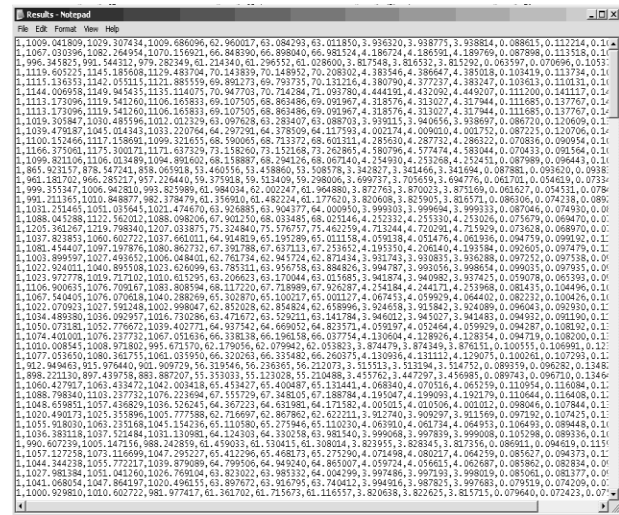


Figure 3. Feature Vectors in a text file

For evaluating a system we can compare the system's performance with a previous existing system. By this method the proposed method of facial gesture recognition with Gabor filters and Normal Bays network [19] can be compared with the existing system of wavelet and neural network ensemble. This comparison can bring out the system performance in real time. System's output can also be given in a matrix format called confusion matrix [35]. System can recognize six different gestures namely Surprise, Smile, Sad, Disgust, Angry and Fear. Failure of recognition can happen by identifying one gesture as one of the other five. So percentage of this can be tabulated in a table.

EG/OG	SURPRISE	SMILE	SAD	DISGUST	FEAR
SURPRISE	90.2	5.3	0	3	0
SMILE	5	92	0	0	3
SAD	0	0	88.8	7.1	0
DISGUST	0	0	6	89	3
FEAR	0	0	3	7.2	82

EG – Expected Gesture OG – Output Gesture

This above table clearly explains the missing recognition of the system and its confusion with the rest of the gestures. By this way of evaluation it will be useful for future enhancements of the system for better performance. Results of the system compared to the previous existing system are given below in table format.

EG/OG	AVERAGE RECOGNITION RATE	
	WAVELET + NETWORK	GABOR
SURPRISE	81.8	90.2
SMILE	77.7	92
SAD	70.2	88.8
DISGUST	74.2	89
FEAR	75.2	82
AVERAGE	75.9	87.6

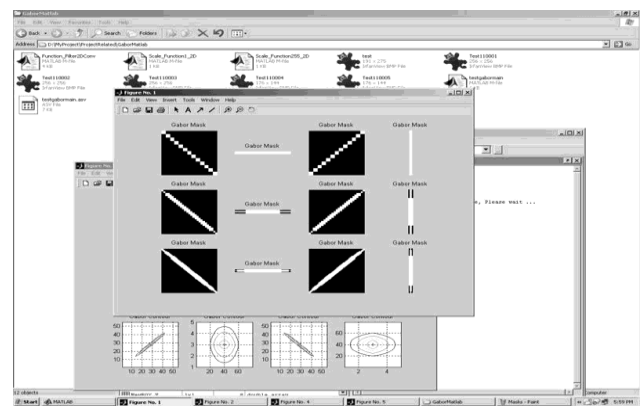


Figure 5.2 Mat lab output of Gabor filters

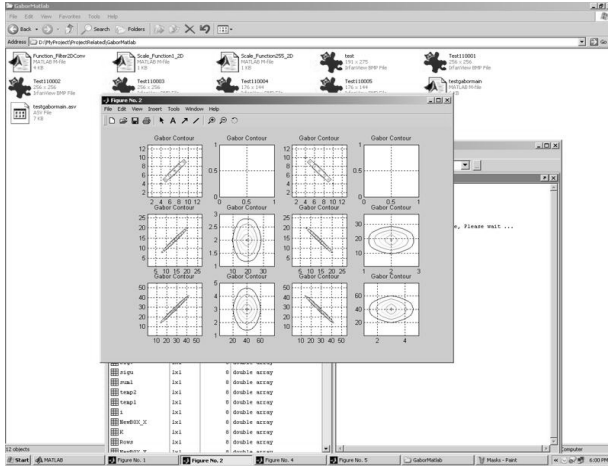


Figure 5.3 Mat lab output of filters in contour format.

4. CONCLUSION AND FUTURE WORK

A novel approach having the following is different face gestures like Surprise, Smile, Sad, Disgust, Angry and Fear. Use of Gabor filters with the proposed way of getting feature extraction using energy, mean and standard deviation, uses less computation power with a pretty decent result compare to the complex way of using mathematical approaches. This system can be enhanced by using proper face model inclusion in the system. In future this proposal can be combined with hand gesture or some other modal of input creating a multi-modal human computer interface which can be used in the area of affective computing and other fields. Connecting with some other input device such as kindest will give us input in 3D and it can be used alone with the proposed algorithm to give out better result and can be used for wide area of applications and other fields as mentioned in the motivation of this proposal.

5. REFERENCES

- [1] Maria E. Jabot, Jeremy N. Balinese, Emmanuel Pontiacs, Leila Takayama, and Clifford Nass "Facial Expression Analysis for Predicting Unsafe Driving Behavior" IEEE Pervasive Computing 2010.
- [2] Gwen L, Marian S B, Ian F, et al. "Dynamics of facial expression extracted automatically from video", Image and Vision Computing, 2006 (24), pp. 615-625.
- [3] L.Ma, K.Khorasani "Facial Expression Recognition Using Constructive feed forward Neural Networks" IEEE Transactions on Systems, Man and Cybernetics. 2004 pg-1588-1595
- [4] B. V. Kumar, M. Savvies, K. Venkataramani and X. Xie "Spatial frequency domain image processing for biometric recognition", Proc. IEEE. Intl. Conf. Image Process., vol. 1, pp.53 -56 2002
- [5] M. Savvies, B. Kumar and P. Kholo "Corefaces Robust shift invariant PCA based correlation filter for illumination tolerant face recognition", Proc. IEEE, Compute. Vis. Pattern Recognition., vol. 2, pp.834 -841 2004
- [6] Qingshan Zhang, Zicheng Liu, Baining Guo, Demetri Terzopoulos, Heung-Yeung Shum "Geometry-Driven

Photorealistic Facial Expression Synthesis" IEEE Transactions on visualization and computer graphics 2006 pg 48 – 60

- [7] Douglas Fidaleo, Mohan Trivedi "Manifold analysis of facial gestures for face recognition", WBMA'03 ACM 2003 pg 65-69
- [8] Chuang C F, Shih F Y. "Recognizing facial action units using independent component analysis and support vector machine", Pattern Recognition, 2006 (39), pp.1795-1798
Viola, Jones "Robust real time object detection", International journal of computer vision. 2001.
- [9] P. J. Phillips , H. Moon , S. A. Rizvi and P. J. Rauss "The FERET evaluation methodology for face recognition algorithms", IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 10, pp.1090 -1104 2000
D. M. Blackburn , M. Bone and P. J. Phillips Facial recognition vendor test 2000 evaluation report., 2000 [online] Available:
- [10] P. Phillips , P. Grother , R. Micheals , D. Blackburn , E. Tabassi and M. Bone Face recognition vendor test 2002: evaluation report., 2003
- [11] K. Messer , J. Kittler , M. Sadeghi , M. Hamouz and A. Kostin "Face authentication test on the BANCA database", Proc. Int. Conf. Pattern Recognit., vol. 4, pp.523 -532 2004 .
- [12] P. J. Phillips , P. J. Flynn , T. Scruggs , K. Bowyer , J. Chang , K. Hoffman , J. Marques , J. Min and W. Worek "Overview of the face recognition grand challenge", Proc. IEEE. Comput. Vis. Pattern Recognit., vol. 1, pp.947 -954 2005
- [13] P. N. Belhumeur and D. J. Kriegman "What is the set of images of an object under all possible lighting conditions?", Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp.270 -277 1996
- [14] R. Ramamoorthi and P. Hanrahan "On the relationship between radiance and irradiance: Determining the illumination from images of a convex Lambertian object", J. Opt. Soc. Amer., vol. 18, no. 10, pp.2448 -2459 2001
- [15] A. Shashua and T. Riklin-Raviv "The quotient image: Class-based re-rendering and recognition with varying illuminations", IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 2, pp.129 -139 2001 .
- [16] H. Wang , S. Li and Y. Wang "Generalized quotient image", Proc. IEEE. Comput. Vis. Pattern Recognit., vol. 2, pp.498 -505 2004 .
- [17] Q. Li , W. Yin and Z. Deng "Image-based face illumination transferring using logarithmic total variation models", Int. J. Comput. Graph., vol. 26, no. 1, pp.41 -49 2009 .
- [18] E. H. Land "The Retinex theory of color vision", Sci. Amer., vol. 237, no. 6, pp.108 -128 1977.
- [19] D. J. Jobson , Z. Rahman and G. A. Woodell "Properties and performance of a center/surround Retinex", IEEE Trans. Image Process., vol. 6, no. 3, pp.451 -462 1997 .
- [20] R. Gross and V. Brajovie "An image preprocessing algorithm for illumination invariant face recognition", Proc.

- 4th Int. Conf. Audio Video Based Biometric Person Authentication, vol. 2688/2003, pp.10-18,2003.
- [21] J. Malik and P. Perona "Scale-space and edge detection using anisotropic diffusion", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp.629 -639 1990 .
- [22] M. A. Turk and A. P. Pentland "Eigenfaces for recognition", *J. Cogn. Neurosci.*, vol. 3, no. 1, pp.71 -86 1991 .
- [23] P. S. Penev and J. J. Atick "Local feature analysis: A general statistical theory for object representation", *Network: Comput. Neural Syst.*, vol. 7, no. 3, pp.477 -500 1996 .
- [24] M. Lades , J. C. Vorbruggen , J. Buhmann , J. Lange , C. von der Malsburg, R. P. Wurtz and W. Konen "Distortion invariant object recognition in the dynamic link architecture", *IEEE Trans. Comput.*, vol. 42, no. 3, pp.300 -311 1993.
- [25] L. Wiskott , J. M. Fellous , N. Kruger and C. von der Malsburg "Face recognition by elastic bunch graph matching", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp.775 -779 1997
- [26] J. Chang , M. Kirby , H. Kley , C. Peterson , B. Draper and R. Beveridge "Recognition of digital images of the human face at ultra low resolution via illumination spaces", *Proc. Asian Conf. Comput. Vis.*, vol. 4844/2007, pp.733 -743 2007
- [27] W. Hwang and S. Kee "International standardization on face recognition technology", *Adv. Biomet. Person Authentic.*, vol. 3338/2005, pp.349 -357 2004
- [28] J. Yang and C. Liu "Color image discriminate models and algorithms for face recognition", *IEEE Trans. Neural Newt.*, vol. 19, no. 12, pp.2088 -2098 2008
- [29] Z. Liu and C. Liu "Robust face recognition using color information", *Adv. Biomet.*, vol. 5558/2009, pp.122 -131 2009
- [30] X. Tan and B. Triggs "Fusing gabor and lbp feature set for kernel-based face recognition", *Proc. IEEE Int. Workshop Anal. Model. Face Gestures*, pp.235 -249 2007
- [31] Y. Su , S. Shan , X. Chen and W. Gao "Hierarchical ensemble of global and local classifiers for face recognition", *IEEE Trans. Image Process.*, vol. 18, no. 8, pp.1885 -1896 2009
- [32] X. Wang and X. Tang "Random sampling LDA for face recognition", *Proc. IEEE. Comput. Vis. Pattern Recognit.*, vol. 2, 2004
- [33] B. Heisele , P. Ho , J. Wu and T. Poggio "Face recognition: Component-based versus global approaches", *Comput. Vis. Image Understand.*, vol. 91, no. 1/2, pp.6 -21 2003
- [34] T. Kim , H. Kim , W. Hwang and J. Kittler "Component-based LDA face description for image retrieval and MPEG-7 tandardisation", *Image Vis. Comput.*, vol. 23, no. 7, pp.631 -642 2005
- [35] P. Sinha , B. J. Balas , Y. Ostrovsky and R. Russell "Face recognition by humans: 19 results all computer vision researchers should know about", *Proc. IEEE*, vol. 94, no. 11, pp.1948 -1962
- [36] A. Jain , K. Nandakumar and A. Ross "Score normalization in multimodal biometric systems", *Pattern Recognit.*, vol. 38, no. 12, pp.2270 -2285 2005
- [37] L. Wiskott , J. M. Fellous , N. Kruger and C. von der Malsburg "Face recognition and gender determination", *Proc. Int. Workshop Face Gesture Recognit.*, pp.92 -97 1995
- [38] M. Savvides , B. V. Kumar and P. Khosla pp.810 -813 2004 W. Hwang , G. Park , J. Lee and S. Kee "Multiple face model of hybrid fourier feature for large face image set", *Proc. IEEE Comput. Vis. Pattern Recognit.*, vol. 2, pp.1574 -1581
- [39] R. Chellappa , C. L. Wilson and S. Sirohey "Human and machine recognition of faces: A survey", *Proc. IEEE*, vol. 83, no. 5, pp.705 -741 1995.