A Robust Rotation Invariant Multiview Face Detection in Erratic Illumination Condition

G.Nirmala Priya
Associate Professor, Department of ECE
Sona College of Technology
Salem

R.S.D.Wahida Banu, PhD.
Principal, Govt. College of Engineering
Salem

ABSTRACT

The major challenge of face detection techniques lies in handling varying poses, i.e., detection of faces in arbitrary in-depth rotations. The face image differences caused by rotations are often larger than the inter-person differences used in Rotation the research toward pose-invariant face recognition in recent years and many prominent approaches have been proposed. However, several issues in face recognition across pose still remain open. The aim of *Rotation* invariant multiview face detection (MVFD) is to detect faces with arbitrary rotation-in-plane (RIP) and rotation off-plane (ROP) angles in still images. MVFD is crucial as the first step in automatic face processing for general applications since face images are usually upright and frontal unless they are taken in cooperation with the person. This paper, proposes a innovative methods to construct a high-performance rotation invariant multiview face detector. This multiview face detector reduces the computational complexity and has broad detection scope. The detection accuracy is high on the testing set of images. The existing techniques are discussed in detail and are compared with the proposed method. A new pose invariant face recognition system based on MBWM histogram matching is proposed. The classification is performed by using the Multiclass support vector machine of a test face and training faces in the database. The proposed system gives 98.80% recognition rate on the HP database of 15 face subjects.

KEYWORDS

Face detection, Pose variation, LBPH, SLBMH, MBWMH, SVM

1. INTRODUCTION

For many applications computer vision systems have been used for Human-Computer-Interaction (HCI) in the recent past. One of the most important biometric techniques used is Face recognition and this has benefited as it is natural and has no direct inflation over the applications. Other bio- metric technique such as fingerprint recognition and iris recognition requires cooperation of subjects. In general, there are two methods for face detection techniques, Knowledge-based method and learning-based method. The knowledge-based method needs the prior knowledge of the face pattern which may be the intensity of faces, elliptic face contour or equilateral triangle relation between eyes and mouth. A face detection system should identify even an uncooperative face in uncontrolled environment and an arbitrary situation without the notice of the subject. This generalization of environment and situations, added serious challenges to face recognition techniques, e.g., the appearances of a face as viewed through a camera. Many face recognition techniques proved satisfactory performances, with the constraint on environment. In real world application this will be major criteria. Pose variation was identified as one of the prominent unsolved problems in the research of face detection and it has gained great interest in the computer vision and pattern recognition research community. A few methods have been proposed to solve this problem of recognizing faces in arbitrary poses.

Tied factor analysis (TFA) [4], 3D Morphable model (3DMM) [1], Eigen light-field (ELF) [3], illumination cone model (ICM) [2], etc are some of the techniques for face detection with varying pose. But, all the methods have their own limitations and are not able to fully solve pose problem in face detection. The techniques of face recognition across pose may be broadly classified into three categories, i.e., general algorithms, 2D techniques, and 3D approaches. General algorithms, means that the algorithms do not include any specific strategy to handle pose variations. They were designed for general purpose of face recognition equally handling all image variations (e.g., illumination variations, expression variations, age variations, and pose variations, etc.). The general algorithm includes the holistic approach of PCA, FDA etc. The 2D techniques includes Active appearance model (AAM), linear shape model, Correlation filters etc. The 3D techniques include Automatic texture synthesis, multi-level quadratic variation minimization.

Head pose estimation is the ability to infer the orientation of a person's head relative to the view of a camera. Still more, head pose estimation is the ability to infer the orientation of a head relative to a global coordinate system, but this subtle difference requires knowledge of the intrinsic camera parameters to undo the perceptual bias from perspective distortion. The range of head motion for an average adult male encompasses a sagittal flexion and extension (i.e., forward to backward movement of the neck) from 60.4° to 69.6°, a frontal lateral bending (i.e., right to left bending of the neck) from 40:9° to 36.3°, and a horizontal axial rotation (i.e., right to left rotation of the head) from 79.8° to 75.3° [5]. The combination of muscular rotation and relative orientation is an often overlooked ambiguity (e.g., a profile view of a head does not look exactly the same when a camera is viewing from the side as compared to when the camera is viewing from the front and the head is turned sideways). Despite this problem, it is often assumed that the human head can be modeled as a disembodied rigid object. Under this assumption, the human head is limited to three Degree of freedom (DOF) in pose, which can be characterized by pitch, roll, and yaw angles as shown in Fig. 1.

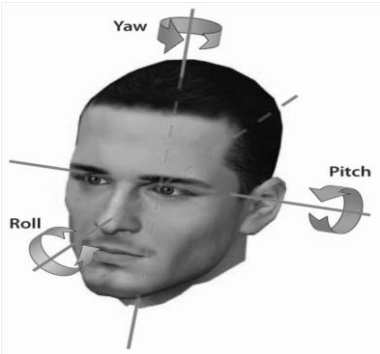

Figure 1. The three degrees of freedom of a human head described by the egocentric rotation angles pitch, roll, and yaw

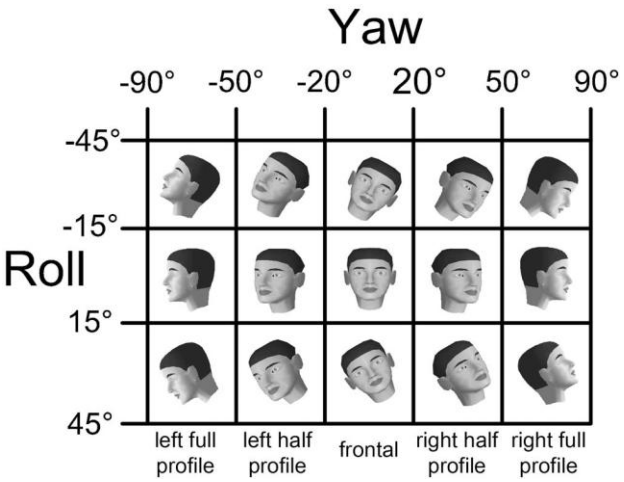

Figure. 2 Views as defined in the upright multiview face detector.

In general, there are two trends in developing face recognition techniques (1) improving the capability of general face recognition algorithms so that the variation in the image is made tolerable (2) Specific algorithm that can eliminate or at least compensate the difficulties brought by image variations (e.g., pose variations). The details of the view as based on upright multiview face detector is shown in figure 2.The problem of face recognition across pose is discussed in Section 2.In Section 3 related works done in this field so far is discussed. In Sections 4, existing and proposed methods of feature extraction and classification are discussed. In section 5 results of the various methods with the robustness of pose variation is discussed. Finally, in section 6 the conclusion and future work to be done are arrived at..

2. PROBLEM FORMULATION

Face recognition across pose refers to recognizing face images that vary in their poses by computers. This is of great interest in many face recognition applications, mostly in application that handle indifferent or un-cooperative subjects, such as surveillance systems. Face recognition is used mainly in airport security system to recognize terrorists and arrest them or stop them from boarding plane. The faces of terrorists are collected and stored in the database. The scanned passengers faces are compared against the database created. The face of passengers moving through a security checkpoint will be scanned. Whenever match is found, cameras will be turned on to survey people with a live video feed, and then the authorities will verify the match and decide whether to stop the individual whose face matches one in the database. The most natural solution for this task might be to collect multiple gallery images in all possible poses to cover the pose variations in the captured images, which requires a fairly easy face recognition algorithm. In many real situations, it is tedious and difficult to collect the multiple gallery images in different poses and therefore the ability of face recognition algorithm to tolerate pose variations is desirable. For instance, if only a passport photo per person was stored in the database, a good face recognition algorithm should still be able to perform the above airport surveillance task. In such sense, face recognition across pose refers to recognizing face images whose poses are different from the gallery (known) images. If face recognition does not have a good pose tolerance, given a frontal passport photo, the system appears to require cooperative subjects who look directly at the camera [6] and face recognition is no longer passive and non-intrusive. Therefore, pose invariance or tolerance is a key ability for face recognition to achieve its advantages of being non-intrusive over other biometric techniques requiring cooperative subjects such as fingerprint recognition and iris recognition. Due to the complex 3D structures and various surface reflectivities of human faces, however, pose variations bring serious challenges to current face recognition systems. The image variations of human faces under 3D transformations are larger than that conventional face recognition can tolerate. Specifically, innate characteristics of the faces, which distinguish one face from another, do not vary greatly from individual to individual, while magnitudes of image variations caused by pose variations are often larger than magnitudes of the variations of the innate characteristics. The challenging task faced by pose-invariant face recognition algorithms is to extract the innate characteristics free from pose variations. Generally, if more gallery images in different poses are available, the performance of recognizing a face image in an unseen pose will be better.

3. RELATED WORKS

In [7], Eigen faces and self organizing map and convolution network (SOM+CN) approaches is used. The recognition rate

Figure .3 HP face database

for Eigen faces is from 61.4% to 89.5% and the performance of self-organizing map with convolution network is in the range of 70.0% to 96.2%. The increase in performance is due to the capability for face recognition algorithms of tolerating small pose variations. When the number of gallery images increases, the probability that the test image lies closely to one of the training image increases. The recognition rate decreases in case of real world images. When multiple training images is used the recognition rate increases. In, multi-level quadratic variation minimization (MQVM) [10] two training images of frontal view and side view is used. In feature- based reconstruction of 3D human faces for recognition, the inclusion of additional side view training image provides more depth information of human face structures, and this results in better reconstructed models than those using single training images. The inclusion of multiple gallery images puts restricted requirements for data collections, because many existing face database might only contain a limited number of (even single) gallery images such as a passport photo database (single gallery images) or police mug-shot database (one front view image and one side view image per face).Therefore, the requirement of multiple gallery images (in different poses)limits the applicability of face recognition algorithms and the most generic scenario is to recognize a probe image in an arbitrary pose from only a single. Training image in another pose, which is also more challenging than multiple gallery view scenario. In [11], a single gallery preface face recognition from a single gallery image per person is discussed. In [12] modular PCA (MPCA) is used in which the pose angle of the input image can be estimated before recognition.. Head pose can be estimated either simultaneously in the process of recognition (asdonein3Dmorphablemodel [1] and cylindrical 3D pose recovery [13]or separately in an independent process. Alternative methods have been reviewed in [14]. Recently, many pose-invariant face recognition approaches have been proposed. Also, a number of face image database have been established for the purpose to comparing the performances of different face recognition algorithms across pose. The widely used databases for face recognition across pose are FERET database [15-16] and CMU-PIE database. CMU- PIE database contains larger pose variations and vertical in-depth rotations Compared to FERET database, but with fewer faces.

4 PROBLEM SOLUTION

4.1 Histogram Based Face Recognition

Image histogram is one method that describes an image in lower dimension. Histogram of an image can be considered as feature vector representing of the image. In general, histogram of an image is a statistical description of the distribution in terms of occurrence of pixel intensities. The size of the image histogram depends on the number of quantization levels of the pixel intensities. In case of monochrome image, 8- bit representation is used for 256 gray levels. An image histogram is simply a mapping η_i that counts the number of pixel intensity levels that fall into various disjoint intervals, known as bins. The bin size determines the size of the histogram vector. In this paper the bin size is assumed to be 256 and the size of the histogram vector is 256. Histogram of

a monochrome image, η_i, is the sum of all the 255 bins. Where N is the number of pixels in an image. Then, histogram feature vector, H, is the total histogram of the image.The similarity between two images can be measured using the multiclass support vector machine. If H1,H2,….,HM be a set of training face images with different poses and M be the number of image samples, then for a given test face image, the histogram of the test image H_t can be used to compare the test image histogram with the training image histogram.

Head Pose face database, contains 15 subjects with 10 selected different poses. The face dataset is divided into training set and test set. The images used in the test set are not included in the training set. The correct recognition rates in percent are included in Table 1. The results of the proposed system are outstanding, because even a single image in the training set provides a correct recognition rate as high as 94.89%. This rate is down to 68.89% in the PCA based face recognition systems respectively. The proposed method shows slight improvement as the number of training set images is increased.

4.2 LBPH Based Face Recognition

LBP operator was originally designed for texture description and was proved a powerful means of texture description. [21-22]The operator labels the pixels of an image by thresholding a 3 X3 neighborhood of each pixel with the centre value and considering the results as a binary number Figure.4 shows the, extraction of Local Binary Pattern (LBPs). It has shown a good performance in the area of facial image description. The derived binary numbers (called Local Binary Patterns or LBP codes) codify local primitives including different types of curved edges, spots, flat areas, etc .so each LBP code can be regarded as a micro-texton. The basic LBP operator labels the pixels of an image (Ip) by thresholding each 3X3 pixel neighborhood of the input image with the centre pixel value (Ic), multiplying the threshold values by weight (powers of two) and summing them. Thresholding is done using the centre pixel as in Eq. 1:

$$f(I_P - I_C) = \begin{cases} 1 & I_P \geq I_C \\ 0 & Otherwise \end{cases} \quad (1)$$

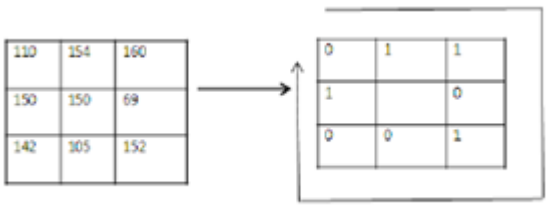

Figure.4 LBP operator

Weights are then assigned and the LBP values are obtained using Eq. 2 and the values are summed to obtain the LBP values for the 3□3 matrix. The LBP feature thus obtained are considered for classification purpose:

$$LBP = \sum_{p=0}^{7} f(I_P - I_C)2^P \quad (2)$$

LBPs have been very effective for image representation as it is being applied to visual inspection, motion detection and outdoor scene analysis. The most important properties of LBP features are their tolerance against monotonic illumination changes and their computational simplicity. The LBP operator mainly detects many texture primitives as spot, line end, edge and corner typically accumulated into a histogram over a region to capture local texture information.

4.3 SLBMH Based Face Recognition

The simplified local binary mean overcomes the disadvantage of the LBP algorithm by using the mean of the 9 pixels for thresholding. This method involves three steps which include subdividing, thresholding and weighing. First, a 3x3 sub image is cropped[8].The pixel values are represented as Ip. Thresholding is done using the mean of the 9 elements of the 3x3 sub image (Im). Thresholding is done based on the rule given in Eq. 3:

$$f(I_P - I_m) = \begin{cases} 1, & I_P \geq I_m \\ 0, & Otherwise \end{cases} \quad (3)$$

Weights are then assigned and summed to obtain the SLMB values for the 3□3 matrix using Eq. 4. The

$$SLBM = \sum_{p=0}^{7} f(I_P - I_m)2^P \quad (4)$$

SLMB feature thus obtained are considered for classification purpose:

The SLMB features are thus calculated. Many images of different types can have similar histograms, because, histograms provide only a coarse characterization of an image. This is the main disadvantage of using histograms. So, the statistical features such as mean and standard deviation of the SLMB features are calculated.

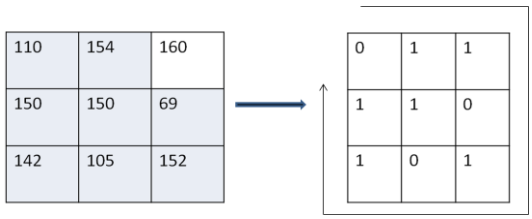

Figure 5. SLBM operator

4.4 MBWM Based Face Recognition

In SLBM, thresholding is exactly at the value of the central pixel ic. This makes it to be sensitive to noise, especially in near-uniform image regions. Many facial regions are relatively uniform; it is potentially useful to improve the robustness of the underlying descriptors in these areas. So, the SLBM is extended to Mean Based Weight Matrix (MBWM) [9].The mean based weight matrix involves three steps which include subdividing, thresholding and weighing. The 3x3 pixels of the image are replaced by a 3-valued function as given in Eq. 5

$$f(I_P - I_m) = \begin{cases} 2, I_P > I_m \\ 1, I_P = I_m \\ 0, Otherwise \end{cases} \quad (5)$$

:

Weights are then assigned and summed to obtain the

MBWM values as in Eq. 6 for the 3□3 matrixes:

$$MBWM = \sum_{P=0}^{7} f(I_P - I_m)2^P \quad (6)$$

The MBWM features are thus calculated. The first order statistical features such as mean and standard deviation of the MBWM features are calculated. These features are used for classification.

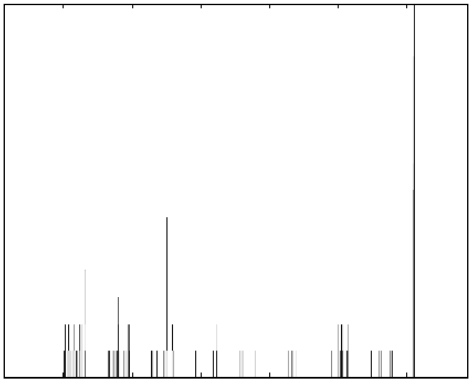

Figure. 6 Histogram of the MBWM feature

4.5 Support Vector Machine (SVM)

As a powerful machine learning technique for data classification, SVM performs an implicit mapping of data into a higher (maybe infinite) dimensional feature space and then finds a linear separating hyperplane with the maximal margin to separate data in this higher dimensional space (Chang and Lin, 2012). Given a training set of labeled examples {(xi, yi), i = 1, 2 ……l} where xi Є Rn and yi Є {-1, 1} a new test example x is classified by the function as in Eq. 7:

$$f(\chi) = sgn\left(\sum_{i=1}^{i} \alpha_i y_i K(x_i, x) + b\right) \qquad (7)$$

where, □i is the Lagrange multiplier of a dual optimization problem that describes the separating hyperplane K (xi, x) is a kernel function and b is the threshold parameter of the hyperplane. The training sample xi with □i > 0 is called support vectors and SVM finds the hyperplane that maximizes the distance between the support vectors and the hyperplane. Given a non-linear mapping Φ that embeds the input data into the high dimensional space, kernels have the form of K(xi, xj) = (Φ (xi), Φ (xj). SVM allows domain-specific selection of the kernel function. Though new kernels are being proposed, the most frequently used kernel functions are the linear, polynomial and Radial Basis Function (RBF) kernels. SVM makes binary decisions. With regard to the parameter selection of SVM, the mean and standard deviation are chosen. These parameters provided the best accuracy. The generalization performances achieved using the two different kernels is discussed.

5. ROBUSTNESS ON POSE VARIANCE

Pose differences (planer rotations) on the images remains the same ie. their general distribution of the grey level (PDFs) is the same. The shape of the PDF is more or less remains the same, only the amplitudes are changing as the angle of rotation of the image is changing. This means, the shape of the histogram of a face image is its signature in representing the image. Hence, if the general distribution is preserved, then the histogram of the image with high and low resolutions will be obtained. Fig.6 shows the performance of the histogram based face recognition system for the changing poses. The performance of the proposed face recognition system is always higher compared to the performance of the histogram, LBPH and SLBMH. The comparison of the histogram method with the proposed LBPH,SLBMH and MBWMH methods for the different training set images are illustrated in table 1.The recognition rate for each of the training set differs for each of the method. The mean recognition rate for each of the method is arrived at. The recognition rate using MBWMH method gives a better recognition rate of 98.8% compared to the histogram method which has a recognition rate of 94.6% . Recognition rate has been improved by 4.2% .

Table 1.Comparision of Proposed LBPH, SLBMH , MBWMH method with existing histogram method.

Training Set	Methods			
	Histogram	LBPH	SLBMH	MBWMH
01	94.89	95.78	96.88	97.78
02	94.62	95.14	94.78	95.89
03	98.14	98.78	98.89	98.78
04	97.33	96.75	98.75	97.15
05	98.80	94.88	97.88	98.88
06	96.45	97.56	98.14	100
07	97.42	98.89	100	98.89
Mean	**94.66**	**96.68**	**97.76**	**98.19**

A graphical representation of recognition rate of the four methods is presented in figure 7. This shows performance of MBWMH method is better for the training set 5, 6 and 7 whereas performance has reached 100% for the training image 7 in SLBMH method.

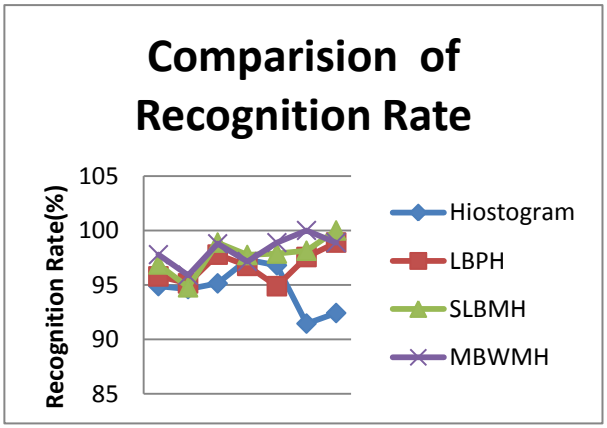

Figure.7 Comparision of recognition rate od the different methods

6. CONCLUSION AND FUTURE WORK

In this paper, a novel faces recognition system based on Mean based weight matrix histogram (MBWMH) is proposed. 98.80% recognition rate on the HP database of 15 face subjects was obtained by using the proposed method while this rate was down to 94.66% in the case of conventional histogram method. It has been shown that due to the high

correlation between the histograms of the faces at different poses, the proposed system is robust for pose changes. The HP database is an apt database for comparison of various methods of face estimation with varying poses. The results show that the proposed method can impressively improve the performance of face recognition across poses.

Further the work may be extended by comparing the results of these proposed methods with the other face databases with varying poses. This may also be extended to the real world face images which shall be the best work in application point of view.

7. REFERENCES

[1] V. Blanz, T. Vetter, Face recognition based on fitting a 3D morphable model, IEEE Trans. Pattern Anal. Mach. Intell. 25 (9) (2003) 1063–1074.

[2] A.S. Georghiades, P.N. Belhumeur, D.J. Kriegman, From few to many: illumination cone models for face recognition under variable lighting and pose, IEEE Trans. Pattern Anal. Mach. Intell. 23 (6) (2001) 643–660.

[3] R. Gross, I. Matthews, S. Baker, Appearance-based face recognition and light- fields, IEEE Trans. Pattern Anal. Mach. Intell. 26 (4) (2004) 449–465..

[4] S.J.D. Prince, J. Warrell, J.H. Elder, F.M. Felisberti, Tied factor analysis for face recognition across large pose differences, IEEE Trans. Pattern Anal. Mach. Intell. 30 (6) (2008) 970–984

[5] Y. Gao, M.K.H. Leung, W. Wang, S.C. Hui, Fast face identification under varying pose from a single 2-D model view, IEE Proc. Vision Image Signal Process. 148 (4) (2001) 248–253.

[6] S. Cass, M. Riezenman, Improving security, preserving privacy, IEEE Spectrum 39 (1) (2002) 44–49.

[7] S. Lawrence, C.L. Giles, A.C. Tsoi, A.D. Back, Face recognition: a convolutional neural-network approach, IEEE Trans. Neural Network 8 (1) (1997) 98–113.

[8] Priya, G.N. and R.S.D.W. Banu, 2012. A Simplified Local Binary Mean (SLBM) based human gender classification. Eur. J. Sci. Res., 71: 435-442.

[9] Priya, G.N. and R.S.D.W. Banu, 2012. Detection of Occluded Face Image using Mean Based Weight Matrix and Support Vector Machine, J. Computer Sci., 8 (7): 1184-1190

[10] X. Zhang, Y. Gao, M.K.H. Leung, Recognizing rotated faces from frontal and side views: an approach towards effective use of mugshot databases, IEEE Trans. Inf. Forensic Secur. 3 (4) (2008) 684–697

[11] S. Chen, X. Tan, Z.-H. Zhou, F. Zhang, Face recognition from a single image per person: a survey, Pattern Recognition 39 (9) (2006) 1725–1745.

[12] A. Pentland, B. Moghaddam, T. Starner, View-based and modular eigenspaces for face recognition, in: Proceedings of the IEEE Conference on CVPR, 1994, pp. 84–91.

[13] Y. Gao, M.K.H. Leung, W. Wang, S.C. Hui, Fast face identification under varying pose from a single 2-D model view, IEE Proc. Vision Image Signal Process. 148 (4) (2001) 248–253.

[14] E. Murphy-Chutorian, M.M. Trivedi, Head pose estimation in computer vision: a survey, IEEE Trans. Pattern Anal. Mach. Intell. 31 (4) (2009) 607–626..

[15] A. Pentland, B. Moghaddam, T. Starner, View-based and modular eigenspaces for face recognition, in: Proceedings of the IEEE Conference on CVPR, 1994, pp. 84–91.

[16] [62] P.J. Phillips, H. Wechsler, J. Huang, P. Rauss, The FERET database and evaluation procedure for face recognition algorithms, Image Vision Comput. 16 (5) (1998) 295–306.

[17] Z. Zhang, M.J. Lyons, M. Schuster, S. Akamatsu, Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron, in: IEEE International Conference on Automatic Face & Gesture Recognition (FG), 1998.

[18] M. Pantic, I. Patras, Dynamics of facial expression: recognition of facial actions and their temporal segments from face profile image sequences, IEEE Transactions on Systems, Man, and Cybernetics 36 (2) (2006) 433–449.

[19] Y. Zhang, Q. Ji, Active and dynamic information fusion for facial expression understanding from image sequences, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (5) (2005) 1–16.

[20] M.S. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, J. Movellan, Recognizing facial expression: machine learning and application to spontaneous behavior, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2005.

[21] C. Shan, S. Gong, P.W. McOwan, Robust facial expression recognition using local binary patterns, in: IEEE International Conference on Image Processing(ICIP), Genoa, vol. 2, 2005, pp. 370–373.

[22] S. Liao, W. Fan, C.S. Chung, D.-Y. Yeung, Facial expression recognition using advanced local binary patterns, tsallis entropies and global appearance features, in: IEEE International Conference on Image Processing (ICIP), 2006,pp. 665–668.