ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 14, May 2016

Facial Expression Recognition Using Features of Active Face Regions

Jansi A
PG student, Dept of CSE
Mepco Schlenk Engineering College, Sivakasi
a.jansi0811@gmail.com

Beaulah Jeyavathana R
Asst Professor, Dept of CSE
Mepco Schlenk Engineering College, Sivakasi
rbeaulah@mepcoeng.ac.in

***Abstract*— Automatic facial expression recognition system is used to detect facial expressions and facial landmarks using localization of salient regions in face. Extraction of discriminative features from salient facial regions plays an important role in expression recognition in face. The accurate detection of facial landmarks improves the localization of the salient regions on face images. This paper proposes a novel framework for expression recognition by using appearance features of selected facial regions. A few prominent facial regions, depending on the position of facial landmarks, are extracted which are active during emotion evocation. These active regions are further processed to obtain the salient regions which contain discriminative features for classification of each pair of expressions, thereby selecting different facial regions as salient for different pair of expression classes. One-against-one classification method is adopted using these features. The proposed method is found to perform well consistently for different resolutions of images. The Cohn-Kanade (CK) database is released for the purpose of promoting research into automatically detecting individual facial expressions. Experiments on JAFFE facial expression databases show the effectiveness of the proposed system.**

***Keywords*— Facial expression analysis, facial landmark detection, feature selection, salient facial regions, CK+ dataset.**

I. Introduction

Automatically detecting facial expressions has become a hot research topic in recent years as they are needed in various fields such as computer vision, machine learning and behavioral sciences and for various applications such as security, human-computer-interaction, driver safety, and health-care. Effective expression analysis hugely depends upon the accurate representation of facial features in face images for providing proper identification of expressions across the face images for which Facial Action Coding System (FACS) are used that represents face by measuring all visually observable facial movements in terms of Action Units (AUs) and associates them with the facial expressions. Accurate detection of AUs depends upon proper identifying and traversing of different facial muscles that are irrespective of pose, face shape, illumination and image resolution across the face images. Permanent facial features like eyebrows, eyes, mouth and nose are extracted using the methods such as SUSAN edge detection operator, facial geometry, edge projection analysis. The detection of all facial fiduciary points is even more challenging than expression recognition across the face images. Therefore, most of the existing algorithms are based on geometric and appearance based features. The geometric models based on geometric features track the shape and size of the face and facial components such as eyes, lip corners, eyebrows etc., and their expressions categorized based on their relative position of these facial components.

The shape models are based on a set of characteristic points on the face to classify the expressions. However, these methods usually requires very accurate and reliable detection as well as tracking of the facial landmarks which are difficult in most of the scenarios practically adding to it the distance between facial landmarks vary across person to person, thereby making the expression recognition system less reliable for various applications. Facial expressions involve change in local texture. In appearance-based methods a bank of filters such as Gabor wavelets, Local Binary Pattern (LBP) etc. there are applied to either the whole face or specific face regions to encode the texture. The performance of appearance based methods to the geometry based methods are studied which shows that the appearance based methods generates high dimensional vector which are further represented in lower dimensional subspace by applying dimensionality reduction techniques such as principal component analysis (PCA), linear discriminate analysis (LDA) . Finally, the classification is performed in learned subspace. Although the time and space costs are higher in appearance based methods, the preservation of discriminative information makes them very popular. Extraction of facial features by dividing the face region into several blocks achieved better accuracy. However, this approach fails due to improper face alignment and occlusions. Some of the earlier works used the extraction system that were employed, recognized these following as AUs: 1 (inner brow raiser), 2 (outer brow raiser), 4 (brow lower), 5 (upper eye lid raiser), 9 (nose wrinkler), 10 (upper lip raiser), 12 (lip corner puller), 14 (dimpler), 15 (lip corner depressor), 17 (chin raiser), 20 (lip stretcher), and 45 (blink), as well as a detector of social Smiles. Once the active regions are selected, the expression recognition becomes easy irrespective of the data. Effective computing aims at effective emotion recognition in low resolution images.

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 14, May 2016

II. RELATED WORK

In [1], Affect detection is, however, a very challenging problem because emotions are constructs with fuzzy boundaries and with substantial individual difference variations in expression and experience. The review of current affect detection systems is organized with respect to individual modalities or channels (e.g., face, voice, and text). The some important progress is being made toward the development of fully automated face-based affect detection. Camera based facial expression detection also enjoys some advantages because it is nonintrusive (in the sense that it does not involve attaching sensors to a user), has reasonable accuracy, and current systems do not require expensive hardware. Affect detection is critical because an affect sensitive interface can never respond to users' affective states if it cannot sense their affective states.

In [2], Feed forward back propagation neural network is used as a classifier for classifying the expressions of supplied face into seven basic categories like surprise, neutral, sad, disgust, fear, happy and angry. The face portion segmentation and localization, morphological image processing operations are used. Permanent facial features like eyebrows, eyes, mouth and nose are extracted using SUSAN edge detection operator, facial geometry, edge projection analysis. The combination of SUSAN edge detector, edge projection analysis and facial geometry distance measure is best combination to locate and extract the facial feature for gray scale images in constrained environments and feed forward back-propagation neural network is used to recognize the facial expression. It consume more time due to extract the whole face image features.

In [3], the concept of Manifold of Facial Expression based on the observation that images of a subject's facial expressions defines a smooth manifold in the high dimensional image space. Such a manifold representation can provide a unified framework for facial expression analysis. To apply Active Wavelet Networks (AWN) on the image sequences for facial feature localization. To learn the structure of the manifold in the feature space derived by AWN, it investigated two types of embeddings from a high dimensional space to a low dimensional space: locally linear embedding (LLE) and Lipchitz embedding. The manifold of facial expression shows promise as a unified framework for facial expression analysis. The expression classification can be performed effectively on the manifold. To learn more properties of expression manifolds with the images of blended expressions.

In [4], a system present for automatic recognition of facial action units (AUs) and their temporal models from long, profile-view face image sequences. It introduces facial-action-dynamics recognition from continuous video input using temporal rules. The algorithm performs both automatic segmentation of an input video into facial expressions pictured and recognition of temporal segments (i.e., onset, apex, offset) of 27 AUs occurring alone or in a combination in the input face-profile video. To automatically segment an arbitrarily long video sequence into the segments that correspond to expressive and expressionless facial displays, it uses a sequential facial expression model. In [5], a multidetector approach to facial feature localization is utilized to spatially sample the profile contour and the contours of the facial components such as the eyes and the mouth. Most approaches to automatic facial expression analysis attempt to recognize a small set of prototypic emotional facial expressions, i.e., fear, sadness, disgust, anger, surprise, and happiness. It achieve automatic AU coding in both frontal-face and face-profile static images. The term direct indicates that as the inference process is executing, it creates the proper chain of reasoning. Detection of the features related to the image intensity and brightness distribution in certain facial areas was not robust since it required highly constrained illumination conditions. The system was not capable of dealing with minor inaccuracies of the utilized detectors. The proposed algorithm cannot handle distractions like occlusions (e.g., by a hand), glasses, and facial hair. Hence, its analysis is limited to non occluded faces without a beard, moustache, and glasses.

In [6], it introduces facial expression recognition from live video input using temporal cues. The existing methods exploit and propose a new architecture of hidden Markov models (HMMs) for automatically segmenting and recognizing human facial expression from video sequences. The architecture performs both segmentation and recognition of the facial expressions automatically using a multi-level architecture composed of an HMM layer and a Markov model layer. An important aspect is that while the 'static' classifiers are easier to train and implement, the dynamic classifiers require more training samples and many more parameters to learn. Naive-Bayes classifiers have a surprisingly very good record in many classification problems, although the independence assumption is usually violated in practice. The difficulty of this model is in estimating the parameters of the Cauchy distribution. The problem with the Naive-Bayes approach is that the independence assumption may be too strong for our application because the facial motion measurements are highly correlated when humans display emotions.

Jabid et al. [7] developed local facial descriptor based on Local Description Patterns (LDP) codes and obtained better performance than LBP features. In [8], face image is divided

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 14, May 2016

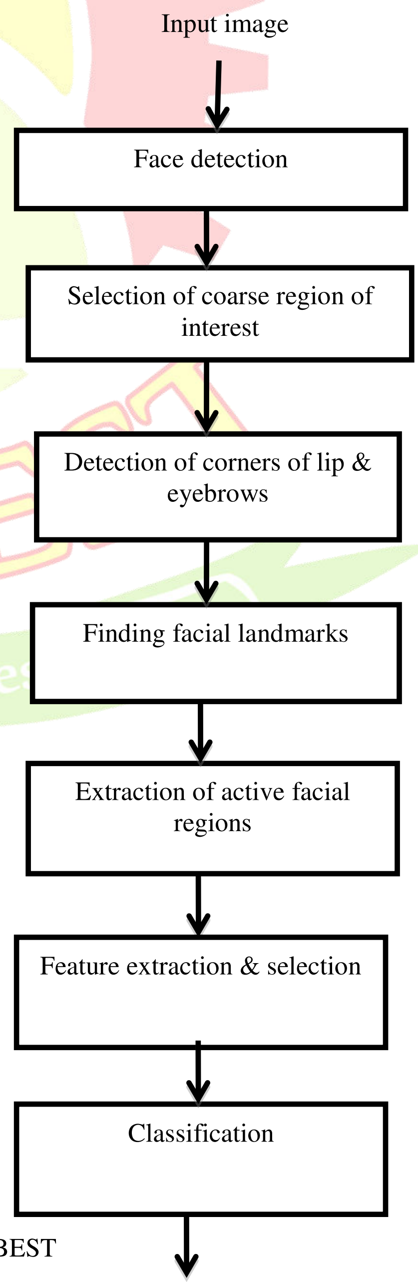
into several sub regions (7 × 6) and local features (7 × 6 × 59 dimensional features) are extracted. Then, the discriminative LBP histogram bins are selected by using Adaboost technique for optimum classification. Similar approaches are reported in [9], [10], and [11].In [12], authors extracted Gabor features of different scales from the face image and trained using Adaboost to select the salient regions for each expression. However, the salient patch size and position is different when trained with different databases. Therefore, a unique criterion cannot be established for recognition of expressions in unknown images. Some issues related to real-time detection of facial landmarks and expression recognition remains unaddressed so far. Most of the researches in this field are carried out on different data sets with suitable performance criteria befitting to the database. For example, selection of prominent facial areas improves the performance. However, in most of the literature, the size and position of these facial regions are reported to be different for different databases. Therefore, our experiments attempt to identify the salient facial areas having generalized discriminative features for expression classification. Selection of salient regions retaining discriminating features between each pair of facial expressions improved the accuracy. The size and location of regions are kept same for different databases for the purpose of generalization. In addition, the proposed framework has the potential to recognize expressions in low resolution images.

In [13], an active Infra-Red illumination along with Kalman filtering is used for accurate tracking facial components. Performance is improved by the use of both geometric and appearance features. Here the initial positions of facial landmarks are figured out using face geometry, given the position of eyes, which is not convenient.

Tian et al. [14] also used relative distance (lip corner, eye, brow etc.) and transient features (wrinkles, furrows etc.) for recognizing AUs present in lower face. However, the use of canny edge detector for extracting appearance features is not flexible in different illumination and determining the presence of furrows using threshold is uncertain.

III. SYSTEM DESIGN

This paper attempts to understand the contribution of different facial areas toward automatic expression recognition. In other words, the paper explores the facial regions which generate discriminative features to separate two expressions effectively. The overview of the proposed method is finding that facial landmark detection and extraction of appearance features from active face regions which improves the performance of expression recognition across the face images. Therefore, the first step is to localize the face followed by detection of the landmarks. A learning-free approach is proposed in which the eyes and nose are detected in the face image and a coarse region of interest (ROI) is marked around them. The lip and eyebrow corners are detected from respective ROIs. Locations of active regions are defined with respect to the location of landmarks. In training stage, all the active facial regions are evaluated, in which features of maximum variation between pairs of expressions are selected. These selected features are further projected into lower dimensional subspace and classified into different expressions using a multi-class classifier. The training phase includes pre-processing, selection of facial features, extraction of appearance features and learning of the multi-class classifiers. In an unseen image, the process first detects the facial landmarks, then extracts the features from the selected salient regions, and finally classifies the expression.



ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 14, May 2016

Recognized Facial Expression

Fig 1: System Design

IV. PROPOSED SYSTEM

A. Pre-processing

A low pass filtering was performed using 3×3 Gaussian mask to remove noise from the facial images followed by face detection or face localization. Viola-Jones face detector is applied for this task. Each image is first converted to gray scale and enhanced with histogram equalization which was done as preprocessing steps before applying the detection. Default scale factor is set to 1.1, minimal neighbor connectivity is set to 4, and minimal size is set to 96×96. After detecting the faces in the image, cropping of the largest face in the image with a border is done. The localized face was extracted and then it was scaled to bring it to a common resolution for further processing. This made the algorithm shift invariant, which is insensitive to the location of the face on image. Histogram equalization was carried out for lighting corrections.

Fig.2 Face Detection

B. Landmark Detection

To extract these regions from face image, we need to locate the facial components first followed by the extraction of the regions around these organs. Using gradient analysis, this method model so that its projection on image will match the facial feature points. In this proposed work, the learning-free approach was adopted for localization of facial landmarks that helps in detecting the facial landmarks in an easy way. The extraction of the active facial regions with respect to the location of eyes, eyebrows, nose, and lip corners using the geometrical statistics of the face was done.

Eye and Nose Detection

To reduce the computational complexity involved during the expression recognition and false detection rate problems to overcome this the coarse region of interest (ROI) for eyes and nose were selected using geometrical positions of face. The regions of both of the eyes were identified seperately using popular Haar classifiers that are trained for each eye. The Haar classifier is shown as the vertices of the rectangular area across the detected eye regions and the eye centers were also computed as the mean of these coordinates. Similarly,the process is carried for nose position which was also detected using Haar cascades.

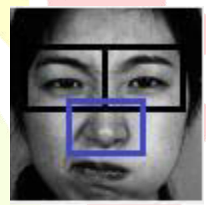


Fig. 3 coarse ROI selection of eyes and nose

Lip and Eyeborw Corner Detection

The ROI for mouth was extracted using the position of nose as reference. The upper lip always produces a distinct edge which can be identified using a horizontal edge detector. In images with different expressions, a lot of edges were obtained which was further thresholded by using Otsu method. In this process, a binary image was obtained containing many connected regions. Using connected component analysis, the spurious components having an area less than the threshold were eliminated. Further, morphological dilation process was carried out on the resulting binary image. Finally, the connected component analysis is done by finding the largest area which were identified just below the nose region was selected as upper lip region. Similarly eyebrows were detected by following the same steps as that of upper lip detection.

The landmark that was obtained from manually labelled facial landmarks which serves the pupose of ground truth during training and testing. The average Euclidian distance error form point to point for each landmark location is used as the distance measure, which is given as:

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 14, May 2016

$$e = \frac{1}{ns}\sum_{i=1}^{n} d_i$$

Here d_is are the Euclidean distance error for the landmarks, *n* is the number of landmarks, and *s* is the distance between the eyes pupils used as the scaling factor.

Fig. 4 Landmark detection

C. Extraction of Active Facial Regions

The expressions that occurs in the face, in which the local regions were extracted corresponding to the positions of active facial muscles. Consider the appearance of facial regions exhibiting considerable variations during one expression. The size of all facial regions was kept equal and was approximately one-ninth of the width of the face. Here onwards, the regions are represented as numbers assigned as R1, R2, R3, and R4 were directly extracted from the positions of lip corners and inner eyebrows respectively. R6 was at the center both the eyes; and R5 was the region above R6. R8 and R9 were located in the midway of eye and nose. R7 and R10 were located just below eyes. R11, R12 and R15 were grouped together and located at one side of nose position. In a similar fashion R13, R14, and R16 were located.

D. Feature Extraction

Local binary patterns were originally used for texture classification. A LBP operator computes a value for each pixel in the image based on its relationship with neighborhood pixels so that it can capture the local texture of the image. To compute a LBP of one pixel, the values of the neighborhood pixels are found, and then the values are binarized via thresholding with the value of the current pixel. Finally, the binary values are concatenated together as the LBP of the current pixel. LBPs are usually aggregated into histogram for further processing.

In order to make LBP distinguishable, a concept called uniform pattern is used. A LBP is called uniform iff the binary pattern contains atmost two bitwise transitions. Consider the following example, the pattern 01000000 contains 2 transitions and so it is said to be uniform; while the pattern 10001001 contains 4 transitions, is said to be non-uniform pattern. Uniform patterns accounts for most of the patterns in face images and thereby they are more important for face recognition. There are totally 58 different uniform patterns. When aggregating LBPs into histogram, all non-uniform patterns are assigned into the same bin, so the final dimension of the uniform LBP histogram is 59. To reduce the feature dimension use the Principle Component Analysis (PCA), and employ cosine similarity to compute the similarity between two expressions.

Table 1: Pixel value of sample block (10×10)

120	163	185	195	189	178	157	133	128	108
109	140	174	195	194	172	147	134	132	112
136	143	163	182	190	188	167	156	152	135
149	158	162	175	175	177	160	146	138	135
133	163	169	170	164	163	162	156	144	140
117	149	167	178	179	185	181	161	149	139
121	152	166	180	185	194	183	157	142	139
124	151	166	178	186	193	175	154	148	128
132	154	169	181	181	187	178	155	135	121
116	154	169	181	181	187	178	155	135	121

Table 2: LBP value of sample block (10×10)

0	1	0	0	0	2	0	0	2	1
0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	0	0	0
0	0	1	0	0	1	0	4	1	0
0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0
0	0	0	0	1	0	0	3	0	0
0	0	0	0	0	0	0	0	0	0

E. Feature selection

The facial regions responsible for recognition of each expression can be used separately to recognize that particular expression. Based on this hypothesis, the performance was evaluated for each of facial region for recognition of different expressions. The LBP histogram features in lower resolution images are sparse in nature because of the smaller area. LDA was applied for projecting these features to the discriminating dimensions and to choose the salient regions according to their discriminating performance. LDA finds the hyper-plane that minimizes the intra-class scatter (Sw), while maximizing the inter-class scatter (Sb). It is also used as a tool for interpretation of importance of the features. Hence it can be considered as a transformation into a lower dimensional space which can be used as optimal discrimination between

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 14, May 2016

classes.The intra-class scatter (S_w) and inter-class scatter (S_b) are given by

$$S_b = \sum_{i=1}^{n} N_i \left(\overline{x_i} - \overline{x}\right)\left(\overline{x_i} - \overline{x}\right)^T$$

$$S_w = \frac{1}{N_{i-1}} \sum_{i=1}^{n} \sum_{j=1}^{N_i} \left(x_{i,j} - \overline{x_i}\right)\left(x_{i,j} - \overline{x_i}\right)^T$$

Where $x_{i,j}$ is the j^{th} feature vector in i^{th} class. Here n number of classes, having N_i number of images in i^{th} class, have $\overline{x}$, the mean vector of all the training data, and $\overline{x_i}$, the mean of i^{th} class. LDA aims at maximization of Sb while minimizing $S_{w.}$

The saliency of a region represents the ability of the features from the region to accurately classify a pair of expressions. We applied PCA to reduce the dimensionality o the feature vector. Thus, by projecting the feature vectors from salient regions to the optimal sub-space obtained by above method, we can find the lower dimensional vector with maximum discrimination for different classes.

F. Multi-class classification

SVM is a binary classifier and it is also used for classification of extracted feature into different expression categories.SVM has three methods are One-against-all, One-against-one, DAGSVM method. The one-against-one method constructs k(k-1)/2 classifiers where each one is trained on data from two class. There are different methods for doing the feature testing after all k(k-1)/2 classifier are constructed. After some test, it will decide to use the following voting strategy suggested.

V. EXPERIMENTAL RESULT

In our experiments on JAFFE database, we used 36 images for extract the features from each region: anger (6), disgust (6), fear (6), happiness (6), sadness (6), and surprise (6). An overall accuracy was better compared to the existing approach DRMF method based CLM model.

Table 3: Confusion Matrix of proposed method on JAFFE Database

Anger	Fear	Disgust	Happiness	Sadness	Surprise
90.2	0	0	0	0	0
0	83.4	0	0	0	7.3
7.54	6.91	82.76	0	0	0
0	0	0	89.7	0	4.32
9.87	6.32	0	7.81	78.72	0
0	4.56	0	0	0	97.5

VI. CONCLUSION

The facial expression recognition using features of active face regions has been proposed. The face expression recognition using features of active face regions is proposed to recognize the expression for a given query image in a JAFEE dataset. The image is extracted based on the features of landmark position instead of extracting features from regions which will be efficient to recognize the expressions.

ACKNOWLEDGMENT

The author wish to thank Dr. S. Arivazhagan M.E. Ph.D., Principal and the Management of Mepco Schlenk Engineering College, Sivakasi for extending generous help. The author also very much thankful to Dr. K Muneeswaran M.E., Ph.D., HOD/CSE for his precious guidance and supportive suggestions. The authors like to convey sincere thanks to guide Mrs.R.Beaulah Jeyavathana M.E., Assistant Professor/CSE Department for her valuable suggestions and complete support in carrying out this project work.

REFERENCES

[1] R. A. Calvo and S. D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications," IEEE Trans. Affective Comput., vol. 1, no. 1, pp. 18–37, Jan. 2010.

[2] J. Whitehill, M. S. Bartlett, and J. Movellan, "Automatic facial expression recognition," in Social Emotions in Nature and Artifact.London, U.K.: Oxford Univ. Press, 2013.

[3] Y. Chang, C. Hu, and M. Turk, "Manifold of facial expression," in Proc. IEEE Int. Workshop Anal. Modeling Faces Gestures, 2003,pp. 28–35.

[4] M. Pantic and I. Patras, "Dynamics of facial expression: Recognition of facial actions and their temporal segments from face profile image sequences," IEEE Trans. Syst., Man, Cybern., vol. 36, no. 2,pp. 433–449, Apr. 2006.

[5] M. Pantic and L. Rothkrantz, "Facial action recognition for facial expression analysis from static face images," IEEE Trans. Syst.,Man, Cybern., vol. 34, no. 3, pp. 1449–1461, Jun. 2004.

[6] I. Cohen, N. Sebe, A. Garg, L. Chen, and T. Huang, "Facial expression recognition from video sequences: Temporal and static modeling," Comput. Vis. Image Understanding, vol. 91, pp. 160–187, 2003.

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 14, May 2016

[7] T. Jabid, M. Kabir, and O. Chae, "Robust facial expression recognition based on local directional pattern," ETRI J., vol. 32, pp. 784–794, 2010.
[8] C. Shan and R. Braspenning, "Recognizing facial expressions automatically from video," in Handbook of Ambient Intelligence and Smart Environments, New York, NY, USA: Springer-Verlag, 2010,pp. 479–509.
[9] C. Shan, S. Gong, and P. W. McOwan, "Robust facial expression recognition using local binary patterns," in Proc. IEEE Int. Conf. Image Process., 2005, pp. II–370.
[10] C. Shan and T. Gritti, "Learning discriminative LBP-histogram bins for facial expression recognition," in Proc. Brit. Mach. Vis.Conf., 2008, pp. 1–10.
[11] L. Zhang and D. Tjondronegoro, "Facial expression recognition using facial movement features," IEEE Trans. Affective Comput., vol. 2, no. 4, pp. 219–229, Oct. 2011.
[12] Y. Zhang and Q. Ji, "Active and dynamic information fusion for facial expression understanding from image sequences," IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 5, pp. 699–714, May 2005.
[13] Y. Tian, T. Kanade, and J. F. Cohn, "Recognizing lower face action units for facial expression analysis," in Proc. IEEE Int. Conf. Autom.Face Gesture Recog., 2000, pp. 484–490.
[14] S. Moore and R. Bowden, "Local binary patterns for multi-view facial expression recognition," Comput. Vis. Image Understanding, vol. 115, no. 4, pp. 541–558, 2011.