A Study On Some Aspects Of Human Emotion Recognition Using Facial Expression Analysis

Sanjoy Mitra[1], Priya Saha[2], Mrinal Kanti Bhowmik[2], Debotosh Bhattacharjee[3]
[1]Department of Computer Science and Engineering
Tripura Institute of Technology
Narsingarh - 799009, Tripura, India
E-mail: mail.smitra@gmail.com
[2]Department of Computer Science and Engineering
Tripura University (A Central University)
Suryamaninagar - 799022, Tripura, India
E-mail: mkb_cse@yahoo.co.in
[3]Department of Computer Science and Engineering
Jadavpur University
Kolkata-700032, India
E-mail: debotosh@indiatimes.com

Abstract

In this particular study, we will demonstrate the high demanding trend of emotional intelligence in human-computer interaction paradigms. The most prospective channel for machine perception of emotions is through the medium of facial expressions. A facial expression investigates the emotions of fear, anger, joy, relief and sadness and also analyses other mental activities, social interactions and other associated physiological signals. There are many obstacles in the development of such automated systems which can easily detect human faces and can recognize emotional status by interpreting facial expressions. The problems may be categorized as – i) detection of image segment as a face, ii) extracting the information of the facial expression, and iii) classification of expression/categorizing the variety of expressions. This paper highlights the literature survey on the efforts aimed at solving these problems, challenges, comparisons, database of human expressions and applications of the past as well as of the future.

1. Introduction

The detection of faces and the analysis of facial expressions under variable circumstances is an effortless simplified task for human beings. The identity, probable age, gender and the emotional state can be easily understood from someone's face (9). The impression as revealed from a displayed expression affects our interpretation of the oral statement and also our impression about the speaker himself. The expressions of Humor and Sympathy are typical instances which convey the essential information which are primarily conveyed via facial expressions. Hence, they occupy a high importance in our routine life though we are not very often conscious of the same. For setting up of a computer based system, it is still very difficult to open it up as a important channel of communication. Progresses in this field assure to equip our technical environment by means of more effective with human beings and probably this will happen tactfully (9). Recent advances in the field of image analysis and pattern recognition blossomed the occurrence of automatic detection and classification of emotional and conversational facial signals. Analysis of automated facial expressions may enable facial expressions into man-machine interaction as a new modality

and make the interaction more effective and suitable. Such a system can also classify facial expressions as an accessible tool for research and in behavioral science and medicine. This paper aims to survey the extent of work done in automating facial expression analysis in facial images and image sequences. Here, we are identifying three basic problems related to analysis of facial expressions. These problems are – i) face detection in a facial image or image sequence, ii) extraction of data from a facial expression, and iii) classification of facial expression.

2. Previous Work

The study of analysis of facial expressions dates back to the 19th century when Bell (1) and Darwin (2) studied the anatomical and physiological basis of the facial expressions of man and animal. Right from the mid 1970s, the analysis of automatic facial expressions has grabbed the interest of many computer vision research groups. Research activities performed on facial data extraction before the mid 1990s is based on the usage of facial motion analysis. Samal and Iyengar (34), Pantic and Rothkrantz (10) and more recently Fasel and Luettin (6) presented quality surveys regarding the various approaches that are intended to meet the challenges the come parallel with facial analysis. Dona.to et al. (3) concentrated on systems for automatic FACS-coding and offers an interesting comparative study on the different features of extraction technologies.

Pantic et al. (28) analyzed Facial Expression using MMI face database with 25 subjects with the help of rule based classifier taking only frontal and profile facial points. Later on, they have laid stress upon on temporal rules for tracking a set of 20 facial fiducial points and succeeded to achieve 90% recognition rate with better accuracy than the previous method[29]. Zheng et al.[30] have used KCCA to recognize facial expressions taken from JAFFE and Ekman's database and for feature extraction, 34 landmark points are converted into a Labeled Graph (LG) using Gabor wavelet transform. It has been noticed that Leave one image out (LOIO) cross validation generates 85.79% accuracy and Leave one subject out (LOSO) generates 74.32% in case of JAFFE database whereas Ekman's database gives 81.25% accuracy. Anderson and McOwen[31] suggested the use of motion by averaging the over specified regions of the face to condense the data which is fed to the classifier. It has been seen that both MLP and SVM were giving almost similar performance. Aleksic and Katsaggelos[32] suggested a way to assign the stream weights where the improvement in performance is demonestrated with the use of MS HMM classifier. Cohn Kanade database with 284 recordings of 90 subjects are used for the experiment and PCA is used for dimension reduction. Pantic and Patras [33] have strssed on automatic segmentation of input video into facial expressions. The experiment is based on MMI face database where test is conducted by using the rule based classifier giving 86.6% accuracy on 96 test profile sequences. Sebe et al.[15] evaluated several machine leaning algorithms like Bayesian Nets, SVMs and Decision Trees. They also used voting algorithms like bagging and boosting to improve the classification results.

3. Basic Definition of Facial Expressions

Facial Expressions as defined by Fasel and Luettin [6] is temporally deformed facial features such as eye lids, eye brows, nose, lips and skin texture generated by contractions of facial muscles. The typical changes noticed by them is for the movement of muscles, which

in brief , "lasting for a few seconds, but rarely more than five seconds or less than 250 ms."[6]. They also highlighted on the important fact that felt emotions are only one source of facial expressions besides verbal and non-verbal communication or physiological activities. Though facial expressions obviously are not equally compatible with emotions, in the computer vision community, the term "facial expression recognition" often refers to the classification of facial features in one of the six so called basic emotions: happiness, sadness, fear, disgust, surprise and anger, as introduced by Ekman in 1971 [4], This interpretation is based on the assumption that the appearances of emotions are universal in all human beings of varied ethnics and cultures.

4. Facial Expression Analysis

In general, the problems for facial expression analysis [10] can be differentiated in three main steps. First, before analyzing a facial expression, the face must be detected in a scene. Next is to employ devise mechanisms for extracting the information of the facial expression from the observed facial image or image sequence. In the case of static images, the process of extracting the facial expression information is referred to as localizing the face and its features in the scene. In the case of facial image sequences, the process is referred to as tracking the face and its features in the scene. At this point, a clear distinction should be made between two terms, namely, facial features and face model features. The facial features are the prominent features of the face eyebrows, eyes, nose, mouth, and chin. The face model features are the features used to represent (model) the face. The face can be represented in various ways, e.g., as a whole unit (holistic representation), as a set of features (analytic representation) or as a combination of these (hybrid approach). The applied face representation and the kind of input images determine the choice of mechanisms for automatic extraction of facial expression information. The concluding step to define some set of categories, which we want to use for facial expression classification and/or facial expression interpretation, and to devise the mechanism of categorization

4.1 Face Detection, Tracking and Feature Extraction

In maximum of the research works conducted in the field of automatic facial expression analysis, the conditions for obtaining a facial image or image sequence are well controlled. Usually, the image has the face in frontal view. Hence, the presence of a face in the scene is well ensured and some global location of the face in the scene is known a priori. However, determination of the exact location of the face in a digitized facial image is much more complex. Firstly, it is because of the fact that the scale and the orientation of the face may vary from image to image. If the mugshots are taken with a fixed camera, faces can occur in images at various sizes and angles due to the movements of the observed person. Thus, it is complex to search for a fixed pattern (template) in the image. The presence of noise and occlusion makes the problem even more critical [10].

The process of locating the face within an image is termed as face detection or face localization and the procedure of locating the face and tracking it across the different frames

of a video sequence is known as face tracking. Research in the fields of face detection and tracking has been very popular and there is wide range of literature available on the same topic [11].

Tao and Huang developed Piecewise Bezier Volume Deformation (PBVD) tracker which is used extensively by the face expression recognition researchers. The tracker uses a generic 3D wire frame model of the face which is associated with 16 Bezier volumes. It may be noted here that the same PBVD model can also be used for the analysis of facial motion and also in the field of the computer animation of faces.

Apart from the mentioned facial expression trackers, others which are also beneficial in the same context are the ones based on Kalman Filters, extended Kalman Filters, Mean Shifting and Particle Filtering. Out of these methods, particle filters (and adaptive particle filters) are more extensively used since they can deal successfully with noise, occlusion, clutter and a certain amount of uncertainty. In recent years, several improvements in particle filters have been proposed, some of which are mentioned here: Motion prediction has been used in adaptive particle filters to follow fast movements and deal with occlusions [12], mean shift has been incorporated into particle filters to deal with the degeneracy problem [7] and AdaBoost is used with particle filters to allow for detection & tracking multiple targets.

4.2 Facial Expression Classification & Related Classifiers:-

The last stage of the entire procedure of any face expression recognition system is the classification module (after the face detection and feature extraction modules). Over the years many kind of classifier / feature combinations were tested in facial expression recognition. Facial features and classifier cannot be compared independently as they work parallel. For instance, Hidden Markov Models (HMM) advantageously classifies signals over time but require much more training data than Support Vector Machines (SVM). A lot of recent researches have been carried on the study and evaluation of the different classifiers. Cohen et al. have studied static classifiers like the Naïve Bayes (NB), Tree Augmented Naïve Bayes (TAN), Stochastic Structure Search (SSS) and dynamic classifiers like Single Hidden Markov Models (HMM) and Multi Level Hidden Markov Models (ML – HMM][13][14].

Cohen et al. experimented with the use of Gaussian distribution and Cauchy distribution as the model distribution of the NB classifier while studying NB classifiers. They discovered that a model with better results [14] can be achieved by using the Cauchy distribution. The independence assumption of NB classifiers is very much well known to us. Although this independence assumption is not applicable in many of the real world scenarios, NB classifiers are known to work surprisingly well. For example, if the word 'Bill' appears in an email, the probability of 'Clinton' or 'Gates' appearing become higher. This is a clear violation of the independence assumption. However NB classifiers have been successfully used in classifying email as spam and non spam. Cohen et al. suggest that problems exist in similar independence assumption when it comes to face expression recognition. This is because of the fact that there is a high degree of correlation between the display of emotions and facial motion [14]. Then they recommended that TAN classifier is better than NB. As a

generalized thumb rule, they suggest the use of a NB classifier when data is insufficient since the TAN's learnt structure becomes unreliable and the use of TAN when sufficient data is available [12]. Cohen et al. have also suggested the scenarios where static classifiers can be used and the scenarios where dynamic classifiers can be used [12]. They reported that dynamic classifiers are sensitive to changes in the temporal patterns and the appearance of expressions. So they suggested the use of dynamic classifiers while performing the person dependent tests and the use of static classifiers when performing the person independent tests. But there are other remarkable differences too: static classifiers are easy to implement and train when compared to dynamic classifiers. But on the flip side, static classifiers give poor performance when given expressive faces that are not at their apex [14].

4.3 Facial Expression Databases:-

Facial expression analysis using databases can be divided into two groups: One is fixed according to the FAC System [5] and another one for which the main focus lies on person identification but also contains samples with different expressions. Sebe and colleagues [15] did an initial and important study in the direction of creating a spontaneous expression database. They listed out the major problems that are associated with the capturing of spontaneous expressions. Different subjects reveals the same emotions at different heights, if the subject becomes aware that he or she is being photographed, their expression loses the genuine, even if the subject is not aware of the recording, the laboratory conditions may not encourage the subject to display his spontaneous expressions. We have studied only those databases that have mostly been exhausted in the past few years.

The BioID Face Database [18]

The collection of images used by researchers to work in the field of face detection is known as The BioID Face Database. The interesting feature of this BioID Face Database is that face images are recorded in "real world" conditions i.e. test set employs a large variety of illumination, background and face size. Some typical sample images are shown below.BioID database consists of 1521 images (384 × 288 pixel, gray scale) of 23 different persons.For comparing reasons, the set also contains manually set eye positions. The images are labeled "BioID_xxxx.pgm" where the characters xxxx are replaced by the index of the current image (with leading zeros). Similar to this, the files "BioID_xxxx.eye" contain the eye positions for the corresponding images. The images are stored in single files using the portable gray map (pgm) data format.

Frontal face dataset (Caltech Faces) [19]

The Frontal Face Image Database of Caltech taken from 27 unique people under different lighting, expressions and backgrounds consists of 450 face images (896 x 592 pixels). The images are stored in Jpeg format.

CBSR NIR Face Dataset [20]

The CBSR NIR face database was taken by an NIR camera with active NIR lighting and contains 3940 images pertaining to 197 persons; the images sizes being 480x640 pixels. Images are divided into a gallery set and a probe set. In the probe set, 12 images per person. In the gallery set, there are 8 images per person.

Cohn-Kanade AU-Coded Facial Expression Database [21]

The automatic facial image analysis is studied with the aid of The Cohn-Kanade AU-Coded Facial Expression Database and is available for use by the research community. Version 1 includes 486 sequences from 97 posers (the original or initial release (Kanade, Cohn, & Tian, 2000)). Each sequence begins with a neutral expression and proceeds to a peak expression. Version 2, referred to as CK+, includes both posed and non-posed (spontaneous) expressions and additional types of metadata. For posed expressions, the number of sequences is increased from the initial release by 22% and the number of subjects by 27%.

Computer Vision Science Research Projects [22]

The images are stored in 24 bit RGB, JPEG format and records 7900 images of 395 individuals and each individual contains 20 images. The data is held in four directories (faces94 , faces95 , faces96 ,grimace) in order of increasing difficulty. Faces96 and grimace are the most difficult, though for two different reasons (variation of background and scale, versus extreme variation of expressions).

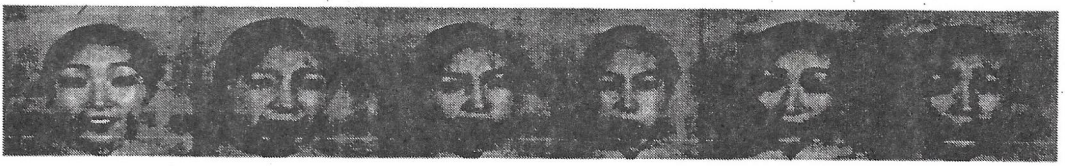
Face94

Face95

Face96

Japanese Female Facial Expression (JAFFE) Database [23]

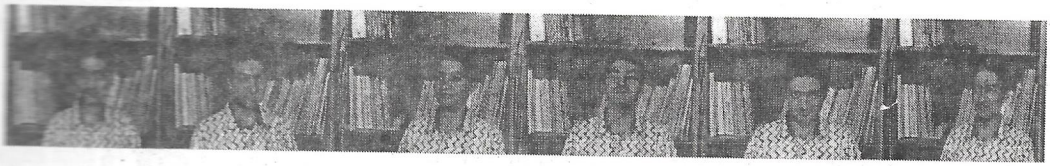
The database contains 213 images of 7 facial expressions (6 basic facial expressions + 1 neutral) posed by 10 Japanese female models. Each image has been rated on 6 emotion adjectives by 60 Japanese subjects.



Georgia Tech face database [24]

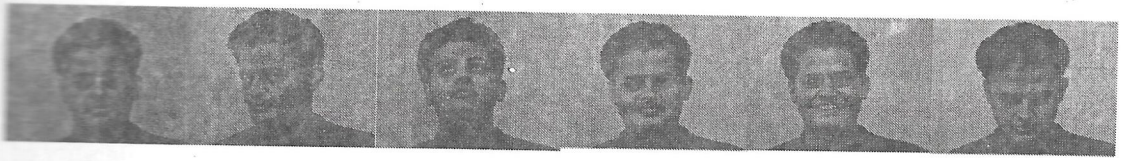
The pictures show frontal and/or tilted faces with different facial expressions, lighting conditions and scale. Each image is manually labeled to determine the position of the face in

the image. All people in the database are represented by 15 color JPEG images with cluttered background taken at resolution 640x480 pixels. The average size of the faces in these images is 150x150 pixels.



Indian Database [25]

The files are in JPEG format. The size of each image is 640x480 pixels, with 256 grey levels per pixel. For each individual, we have included the following pose for the face: looking front, looking left, looking right, looking up, and looking up towards left, and looking up towards right, looking down. In addition to the variation in pose, images with four emotions - neutral, smile, laughter, sad/disgust - are also included for every individual. This database contains images of 40 distinct subjects with eleven different poses for each individual. All the images have a bright homogeneous background and the subjects are in an upright, frontal position.



ORL Face Database [26]

The files are in PGM format, and can conveniently be viewed on UNIX (TM) systems using the 'xv' program. The size of each image is 92x112 pixels, with 256 grey levels per pixel. There are ten different images of each of 40 distinct subjects. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses / no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement).

IRIS Thermal/Visible Face Database [27]

The actual size of the images is 320 x 240 pixels (for both visual and thermal). 176-250 images per person, 11 images per rotation (poses for each expression and each illumination).In this database; all the thermal and visible unregistered face images are taken under variable illuminations, expressions, and poses. Total 30 classes are present in that database and the size of the database is 1.83 GB.

5. Applications

The study of Automatic facial expression recognition systems is applicable in several interesting areas. With the latest advances in robotics sciences, especially humanoid robots, the urgency in the requirement of a robust expression recognition system is evident. As robots are constructed to interact more and more with humans and start becoming a part of our living, it is important to note that Darwin's work is based on the work of Sir Charles Bell, one of his early contemporaries.

It will be unwise to compare and confuse Facial Expression recognition with human Emotion Recognition. As Fasel and Luettin pointed out, "Facial Expression recognition deals with the classification of facial motion and facial feature deformation into classes that are purely based on visual information" whereas "Emotion Recognition is an interpretation attempt and often demands understanding of a given situation, together with the availability of full contextual information.

The system of Expression recognition creates intelligent visual interface between the man and the machine. Human beings effectively communicate and are responsive to the emotional states of each other. Computers must also gain this ability. Human Computer Interaction research community is precisely focusing on it: namely, Affective Computing. Expression recognition plays a significant role in recognizing one's affect and in turn helps in building meaningful and responsive HCI interfaces. The interested reader can refer to Zeng et al.'s[30] comprehensive survey to get a complete picture on the latest developments in Affect Recognition and its applications to HCI.

HCI, expression recognition systems find uses in a host of other domains like Telecommunications, Behavioral Science, Video Games, Animations, Psychiatry, Automobile Safety, Affect sensitive music juke boxes and televisions, Educational Software, etc apart from the two main applications, namely robotics and affect sensitive.

Bartlett et al. have successfully used their face expression recognition system to develop an animated character that mirrors the expressions of the user (called the CU Animate) [35]. They have also been successfully deployed the recognition system on Sony's Aibo Robot and ATR's RoboVie. Another interesting application has been demonstrated by Anderson and McOwen [31], called the 'EmotiChat'. It consists of a chat room application where users can log in and start chatting. The face expression recognition system is connected

to this chat application and it automatically inserts emoticons based on the user's facial expressions. As expression recognition systems become more real times and robust, we will see many other innovative.

6. Conclusion:

The face expression research community is shifting its focus to the recognition of spontaneous expressions. As discussed earlier, the major challenge that the researchers face is the non availability of spontaneous expression data. Capturing spontaneous expressions on images and video is one of the biggest challenges ahead.

Apart from the six prototypic expressions there are a host of other expressions that can be recognized. But capturing and recognizing spontaneous non basic expressions is even more challenging than capturing and recognizing spontaneous basic expressions. This is still an open topic and no work seems to have been done on the

Face expression recognition systems have improved a lot over the past decade. The focus has definitely shifted from posed expression recognition to spontaneous expression recognition. The next decade will be interesting since robust spontaneous expression recognizers will be developed and deployed in real time systems and used in building emotion sensitive HCI interfaces. This is going to have an impact on our day to day life by enhancing the way we interact with computers or in general, our surrounding living and work spaces.

References:

[1] Sir C. Bell. "Essays on the Anatomy of Expression in Painting." Longman, Reese, Hurst & Orme, London, third edition, 1896.

[2] C. Darwin" The Expression of the Emotions in Man and Animal" J. Murray, London, 1872.

[3] G. Donato, M.S. Barlett, J. Hager, P. Ekman, and T. Sejnowski "Classifying facial actions" IEEE Transactions on Pattern Analysis and Machine Intelligence, 21(10) ,pp:974–989, 1999.

[4] P. Ekman and W. Friesen "Constants across cultures in the face and emotion" Journal of Personality and Social Psychology, 17(2),pp:124–129, 1971.

[5] P. Ekman and W. Friesen. "The Facial Action Coding System" Consulting Psychologists Press Inc., Palo Alto, Calif., 1978.

[6] B. Fasel and J. Luettin" Automatic facial expression analysis: A survey" Pattern Recognition, 36(1),pp:259–275, 2003.

[7] F. Xu, J. Cheng and C. Wang, "Real time face tracking using particle filtering and mean shift," Proc. IEEE Int. Conf. Automation and Logistics (ICAL), pp. 2252–2255, Sept. 2008.

[8] H. Tao and T.S. Huang, "A Piecewise Bezier Volume Deformation Model and Its Applications in Facial Motion Capture," in Advances in Image Processing and Understanding: A Festschrift for Thomas S. Huang, edited by Alan C. Bovik, Chang Wen Chen, and Dmitry B. Goldgof, 2002.

[9] R. Fischer "Automatic facial expression analysis and emotional Classification" MIT, October, 2004.

[10] M. Pantic and Leon J.M.Rothkrantz "Automatic Analysis of Facial Expressions: The State of the Art" IEEE transactions on Pattern Anlysis and Machine Intelligence, Vol 22, No 12, December, 2000

[11] V. K. Bettadapura" Face Expression Recognition and Analysis: The state of the Art", Columbia University, 2010

[12] S. Choi and D. Kim, "Robust Face Tracking Using Motion Prediction in Adaptive Particle Filters," Image Analysis and Recognition, LNCS, vol. 4633, pp. 548- 557, 2007.

[13]I. Cohen, N. Sebe, F. Cozman, M. Cirelo, and T. Huang, "Learning Bayesian Network Classifiers for Facial Expression Recognition Using Both Labeled and Unlabeled Data," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. I-595-I-604, 2003.

[14]I. Cohen, N. Sebe, A. Garg, L.S. Chen, and T.S. Huang, "Facial Expression Recognition From Video Sequences: Temporal and Static Modeling", Computer Vision and Image Understanding, vol. 91, pp. 160-187, 2003.

[15] N. Sebe, M.S. Lew, Y. Sun, I. Cohen, T. Gevers, and T.S. Huang, "Authentic Facial Expression Analysis," Image and Vision Computing, vol. 25, pp. 1856- 1863, 2007.

[16] T. Kanade, J. Cohn, and Y. Tian, "Comprehensive Database for Facial Expression Analysis," Proc. IEEE Int'l Conf. Face and Gesture Recognition (AFGR '00), pp. 46-53, 2000.

[17] http://vasc.ri.cmu.edu//idb/html/face/facial_expression/index.html

[18] http://www.bioid.com/download-center/software/bioid-face-database.html

[19] http://www.vision.caltech.edu/archive.html

[20] http://www.cse.ohio-state.edu/otcbvs-bench/

[21] http://vasc.ri.cmu.edu/idb/html/face/facial_expression/

[22] http://cswww.essex.ac.uk/mv/allfaces/index.html

[23] http://www.kasrl.org/jaffe.html , http://www.kasrl.org/jaffe_info.txt

[24] http://www.anefian.com/research/face_reco.htm

[25] http://vis-www.cs.umass.edu/~vidit/IndianFaceDatabase/

[26] http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html

[27] http://www.cse.ohio-state.edu/otcbvs-bench/

[28] M. Pantic and J.M. Rothkrantz, "Facial Action Recognition for Facial Expression Analysis from Static Face Images," IEEE Trans. Systems, Man arid Cybernetics Part B, vol. 34, no. 3, pp. 1449-1461, 2004.

[29] M. Pantic, I. Patras, "Detecting facial actions and their temporal segments in nearly frontal-view face image sequences," Proc. IEEE conf. Systems, Man and Cybernetics, vol. 4, pp. 3358-3363, Oct. 2005.

[30] W. Zheng, X. Zhou, C. Zou, and L. Zhao, "Facial Expression Recognition Using Kernel Canonical Correlation Analysis (KCCA)," IEEE Trans. Neural Networks, vol. 17, no. 1, pp. 233–238, Jan. 2006.

[31] K. Anderson and P.W. McOwan, "A Real-Time Automated System for Recognition of Human Facial Expressions," IEEE Trans. Systems, Man, and Cybernetics Part B, vol. 36, no. 1, pp. 96-105, 2006.

[32] P.S. Aleksic, A.K. Katsaggelos, "Automatic Facial Expression Recognition Using Facial Animation Parameters and MultiStream HMMs," IEEE Trans.Information Forensics and Security, vol. 1, no. 1, pp. 3-11, 2006.

[33] M. Pantic and I. Patras, "Dynamics of Facial Expression: Recognition of Facial Actions and Their Temporal Segments Form Face Profile Image Sequences," IEEE Trans. Systems, Man, and Cybernetics Part B, vol. 36, no. 2, pp. 433-449, 2006.

[34] A. Samal and P.Iyengar. Automatic recognition and analysis of human faces and facial expressions: A survey. Pattern Recognition, 25(1):65–77, 1992.

[35] M.S. Bartlett, G. Littlewort, I. Fasel, and R. Movellan, "Real Time Face Detection and Facial Expression Recognition: Development and Application to Human Computer Interaction," Proc. CVPR Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction, vol. 5, 2003.