Trends in Engineering and Technology (NCTET-2K17)
International Journal of Advanced Engineering Research and Science (IJAERS) Special Issue-4
https://dx.doi.org/10.22161/ijaers/nctet.2017.ece.2 ISSN: 2349-6495(P) | 2456-1908(O)

Real-time Facial Expression Recognition System using Raspberry Pi

Priyanka[1], T.R.Revanth kumar[2]
[1]Department of ECE, SVCE, Tirupathi
Email: cpriyanka579@gmail.com
[2]second author details: Assistant professor ,SVEW
Email: revanthkumar.tr@svcolleges.edu.in

***Abstract*—*In present day technology human-machine interaction is growing in demand and machine needs to understand human gestures and emotions. Emotions can understand by text, vocal, verbal and facial expressions. Facial expressions are a rich source of communicative information about human behaviour and emotion. Facial Expression Recognition is challenging problem up till now because of many reasons, moreover, it consists of three sub challenging tasks face detection, facial feature extraction and expression classification. Automatic facial expression analysis is an interesting and challenging problem which impacts important applications in many areas such as human-computer interaction and data driven animation. Deriving effective facial representative features from face images is a vital step towards successful expression recognition. Most of the systems are able to recognize basic prototype emotions like Happy, Sad, Surprise, Anger, Fear and Disgust. These general expressions are detected using certain variations of the facial features like broadening of mouth, closing of eyes, twitching of nose, etc.*

The proposed method achieves a fast and robust facial feature extraction based on consecutively applying filters to the image. The proposed method implements the real time emotion recognition from facial image using three steps face detection, features extraction and classifier for classification of emotions. The proposed method uses raspberry pi for implementing emotions recognition.

***Keywords*—Feature extraction, Active shape Model, Adaboost classifier, Raspberry pi**

I. INTRODUCTION

In present day technology human-machine interaction is growing in demand and machine needs to understand human gestures and emotions. If machine can identify human emotion, it can understand human behaviour better, thus improving the task efficiency. It can serve as a vital measurement tool for behavioural science and socially intelligent software can be developed which can be used for robots. Emotions are the strong feelings which are govern by the surroundings and play a great role in daily task like decision making, learning, attention, motivation, coping, perception, planning, cognition, reasoning and many more, which leads to emotion recognition a big research field. Emotion recognition can be done by text, vocal, verbal and facial expression. FACIAL EXPRESSIONS synchronize the dialogue by means of brow rising and nodding, clarify the content and intent of what is said by means of lip reading and emblems like a wink, signal comprehension, or disagreement, and convey messages about cognitive, psychological, and affective states [1], [2]. Therefore, attaining machine understanding of facial behaviour would be highly beneficial for fields as diverse as computing technology, medicine, and security in applications like ambient interfaces, empathetic tutoring, interactive gaming, research on pain and depression, health support appliances, monitoring of stress and fatigue, and deception detection. Because of this practical importance [3], [4] and the theoretical interest of cognitive and medical scientists [5], [6], machine analysis of facial expressions attracted the interest of many researchers in computer vision and AI.

Two main streams in the current research on automatic analysis of facial expressions consider facial affect (emotion) detection and facial muscle action detection [7]–[10]. These streams stem directly from the two major approaches to facial expression measurement in psychological research [11]: message and sign judgment. The aim of the former is to infer what underlies a displayed facial expression, such as affect or personality, while the aim of the latter is to describe the "surface" of the shown behaviour, such as facial movement or facial component shape.

Thus, a frown can be judged as "anger" in a message-judgment approach and as a facial movement that lowers and pulls the eyebrows closer together in a sign-

https://dx.doi.org/10.22161/ijaers/nctet.2017.ece.2 ISSN: 2349-6495(P) | 2456-1908(O)

judgment approach. While message judgment is all about interpretation, sign judgment is agnostic, independent from any interpretation attempt, leaving the inference about the conveyed message to higher order decision making. Most facial expression analyzers developed so far adhere to the message judgment stream and attempt to recognize a small set of prototypic emotional facial expressions such as the six basic emotions proposed by Ekman [7]–[9], [12]. Even though the automatic recognition of the six basic emotions from face images and image sequences is considered largely solved, reports on novel approaches are published even to date (e.g., [13]–[16]). Exceptions from this overall state of the art in machine analysis of human facial affect include few tentative efforts to detect cognitive and psychological states like interest [17], pain [18], [19], and fatigue [20]. Facial emotion recognition from 2D images is well studied field but lack of real-time method that estimates features even low quality images. Most of the work [22]-[24] are based on frontal view images of the faces. More work need to be done on non-frontal images with different illumination conditions as in real time these global conditions are not uniform.

In the proposed system, a real-time emotion recognition system that recognizes basic emotions like anger, disgust, happiness, surprise and neutral using database [25] consisting 2D images with different illumination and poses. The software system developed using our proposed method is deployed on Raspberry Pi as it can be used with robots as the size of Raspberry Pi is very small, light weighted and very less power supply is needed for it. As a result it can be mounted over any robot very easily and can be used for many applications such as surveillance security, monitoring senior citizen or children at home, monitoring critical patients in ICU, for customer satisfaction and many more. Circuit board of Raspberry Pi is shown in Fig. 1

LBP features were originally proposed for texture analysis [24], and recently have been introduced to represent faces in facial image analysis. The most important properties of LBP features are their tolerance against illumination changes and their computational simplicity. LBP has been used along with linear programming (LP) [21] to recognize facial expression. In [22], LBP features were compared with Gabor-filter features and it was concluded that LBP features perform better for low resolution images. In this paper, a simple but effective localized facial features based on local binary patterns has been proposed. Various machine learning approaches including

Support Vector Machine (SVM) [25] and Adaptive Boosting (AdaBoost) [26, 27] have also been examined for facial expression recognition. One limitation of the existing techniques is that they are slow in extracting the facial features and recognizing the expression, therefore a real-time implementation of the proposed approach has also been presented.

II. LITERATURE SURVEY

Many research works on emotion recognition and analysis have been carried out for a decade due to applications in the field of human-machine interaction. Many studies on facial expression recognition and analysis have been carried out for a long time because facial expressions play an important role in natural human-computer interaction as one of many different types of nonverbal communication. Paul Ekman *et al*. postulated six universal emotions (anger, disgust, fear, happiness, sadness, and surprise), and developed Facial Action Coding System (FACS) for taxonomy of facial expressions. Their significant works have formed the basis of the existing automatic facial expression recognition systems. For automatic facial expression recognition system, a variety of research approaches have been proposed. According to types of facial features that the proposed systems take more interest, they extract geometric features, appearance features, or a hybrid of geometric and appearance features on targeting face. For example, Active Shape Models (ASMs) are one of the most popular statistical models fitted to a new face image which can be successfully used for good geometric features such as measurements among coordinates of landmarks on the face. On the other hand, Gabor wavelets representation and local binary patterns (LBP) are successful appearance features with changes of the facial appearance, e.g. wrinkles and furrows. For hybrid features of shape and appearance, the Active Appearance Model (AAM) is a well-known method of good performance.

In addition, there are different approaches for classifying facial expressions in video sequences depending on spatial and spatio-temporal information. For the frame-by-frame approaches relying on static image or only a frame of video sequences without temporal information, diverse classifiers such as Neural Network (NN), Bayesian Network (BN), rule-based classifiers, Support Vector Machine (SVM) achieve good results for facial expression recognition. On the other hand, spatio-temporal approaches result in better

https://dx.doi.org/10.22161/ijaers/nctet.2017.ece.2 ISSN: 2349-6495(P) | 2456-1908(O)

performance in video sequences, compared to spatial approaches without temporal information. Above all, Hidden Markov Models (HMM) is one of the most popular classifiers among spatio-temporal approaches and works well for facial expression recognition. Although most systems are obviously interested in achieving high performance in terms of accuracy rate of recognition, they have no great concern about mobile platform.

The face has a complex three-dimensional surface structure, and thus for the formation of two-dimensional image, the change is very large, especially for different face pose and facial expression, as well as different lighting conditions, the two-dimensional image difference is very clear, accurate and efficient method is a very challenged task. Feature location is very critical for analysis of related face issues, its accuracy is directly related to the reliability of the subsequent application. It is not only to provides an important geometric information for face image processing and analysis, and but also it plays an important application for face recognition, expression analysis of the face pose, face synthesis, facial animation, model-based face image coding and so on. Compared with the traditional feature locating methods, deformable models based methods have attracted much attention in the areas of features extraction and image matching because of their ability of adapting themselves to fit objects more accurately. Active contour model or Snakes model proposed by Kass et al. can match any arbitrary shape according to the guidance of some energy terms, which it has been used successfully to represent facial contours and locate objects' feature. Then Yuille and etc. used the deformable templates to extract positions of the eyes and mouth in a face image. The templates are specified by a set of parameters, which allows the a priori knowledge of the expected shape to be used in the matching. To improve robustly and accurately for facial feature extraction, we need to consider all the facial features in its representation and searching process. Recently, statistical approaches had been proposed to use for feature extraction.

Because of face images with composite consequences of multiple factors, such as illumination, pose, and expression and so on, all factors will increase a few challenges and difficulties to extract facial features. Currently, many researchers have proposed a number of extraction methods, but deformable models applied for locating the unknown face image are very efficiently to overcome multiple influences. The statistics based parameterized model, especially ASM (Active Shape models) and AAM (Active Appearance models) [14-15] proposed by Cootes, the two classical algorithms have obtained more and more applications especial for features extraction. The two models' advantages include: for one thing, the deformable models can extract object's features precisely. On the other hand, these features can be utilized for face related areas. Its merit is very similar to the Active contour models (ACMs) [7] proposed by Kass. ASM has inherited these advantages of the variable models from ACM, moreover, it has extended the power to features extraction. ASM consists of two important models: the deformable shape model and the local profile model. ASM research an unknown image depending on the training set, and the training set can ensure the accuracy of matching. ASM locates images with a variable model deriving from prior knowledge of the training set [16-17]. ASM allows a little shape changing, but it can ensure the changed shape to represent objects' structures.

III. TECHNIQUES FOR EXTRACTING THE FACIAL FEATURES

a) Filtering and Feature Extraction Stage

Among the various stages, filtering and feature extraction is the core for the development of an emotion recognition system. As there is a large amount of raw data (material) on facial images, it is necessary to analyse and synthesize it into a small and concrete set of information which is called a 'feature space'. The performance analysis is directly dependent on the feature space. The more the relevance and preciseness of the feature space, it is not only better but also easier for the task of performance analysis. A number of researches have been conducted using different approaches on different areas of analysis for feature extraction. For instance, some use information dependent on geometric features in 2-D and 3-D facial images while others use static image information obtained by filtering the image. Several filter image that have been of use include- Principal Component Analysis (PCA), Independent Component Analysis (ICA), Discrete Cosine Transform (DCT), Gabor Filters, Fast Fourier Transform (FFT), Singular Value Decomposition (SVD), Harr Wavelet transforms, etc. Sometimes a combination of multiple techniques and filtering are used for better performance analysis.

b) Principal Component Analysis (PCA)

Also known as Karhunen-Loeve Transform (KLT) or the Hotelling Transform. The basic strength of ACP lies in the fact that it can extract the most essential abstract facial features by the Eigen face/ Fisher face calculations.

c) Gabor Wavelets

International Journal of Advanced Engineering Research and Science (IJAERS) Special Issue-4
https://dx.doi.org/10.22161/ijaers/nctet.2017.ece.2 ISSN: 2349-6495(P) | 2456-1908(O)

Gabor acts as a filtering device. It transforms the facial image into small wavelets that helps in easier recognition of the desired feature. Gabor is a very effective tool because the Gabor filtered images stand strong and unaffected to the variations or changes made in illumination and facial expressions. Further Gabor wavelet representation has higher degree of correlation with human semantic ratings.

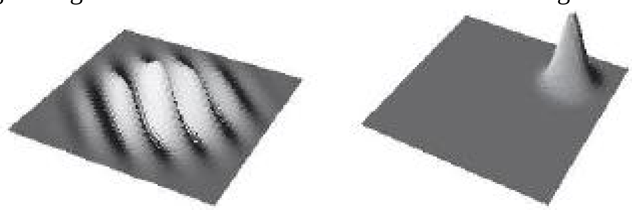
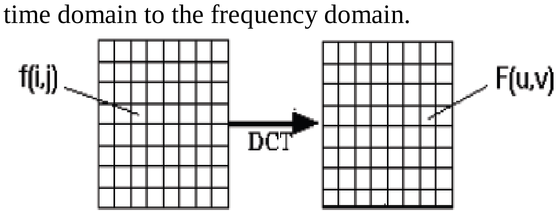


Fig. 1 Gabor Filter in space (left) and frequency (right)

Gabor wavelets are represented by a 2-D plane waves in the spatial domain. One characteristic of wavelets is that they can be located somewhere between the space and the frequency domain. In the frequency domain, as shown in Fig. above, the Gabor wavelet filters can be represented as Gaussian windows.

d) Discrete Cosine Transform (DCT)

DCT is basically a technique for image compression. It compresses the image by removing the information which is not of use. The DCT mechanism transfers an image from the time domain to the frequency domain.



$$F(u,v) = \frac{\Lambda(u)\Lambda(v)}{4}\sum_{i=0}^{7}\sum_{j=0}^{7}\cos\frac{(2i+1)\cdot u\pi}{16}\cdot\cos\frac{(2j+1)\cdot v\pi}{16}\cdot f(i,j)$$

$$\Lambda(\xi) = \begin{cases}\frac{1}{\sqrt{2}} & for\ \xi = 0\\ 1 & otherwise\end{cases}$$

e) Harr Wavelet Transforms

It is another frequently used image filtering method. It is based on the mechanism of filtering the image by separating the frequency bands into two groups- low and high. Wavelet functions for 2-D DWT can be obtained by

Multiplying two wavelet functions or wavelet and scaling function for one-dimensional analysis.

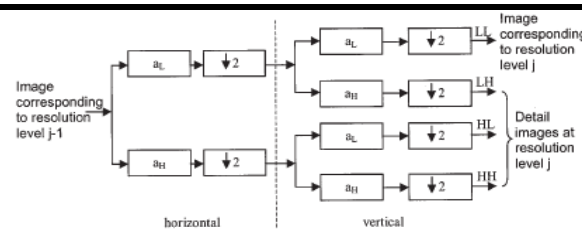


Fig.2 One filter stage in 2-D DWT

IV. FILTERED FEATURE CLASSIFICATION

FFC is the next very important and sensitive stage in the emotion recognition system. It is sensitive in the sense that even the slightest changes in the movements of the muscles in a facial expression may alter the emotion and that is exactly what is needed to be captured and differentiated during this stage. For example, some facial expressions are a blend of more than one emotion as, say a combination that occurs for the case of fear, sorrow and disgust. To overcome the above problem, very recently, some newer approaches have been used. There are two main categories of Feature Classification approach:

1) Statistical non-machine learning approach such as Euclidean and linear discrimination analysis.

2) Machine learning approach such as Feed forward Neural Network, Hidden Markov Model, Multilayer Perceptron,Radial Basis Function Network, etc.

V. DEFICIENCY IN PREVIOUS WORK

i) Recognition rate is not achieved up to the mark due to lack of training in system.

ii) Expression recognition is subject’s age dependent.

iii) Data representation for some expression is identical.

iv) Previous facial expression analysis can’t deal with:

a) Spontaneously occurring facial behavior

b) Feature of wrinkle extraction

c) Feature of spontaneous change in skin color

v) Lack of facial point localization & tracking

vi) Some papers, have shown that their results and performance are database dependent. e.g. JAFFE, CMU, Cohn Kanade, U-Maryland

Issue 1: Feature and good parameter selection process is unyielding

Among others improvements, optimization of the feature selection step and choice of good parameter sets, along with the robustness issue, should be taken up for further study. Also, some other similarity measurements and combination with (unsupervised) clustering approaches might be considered, as that uses the combination of Gabor wavelet neural networks (GWNN), Q-learning, and integrated adaptive fuzzy clustering (IAFC) is used to render some capabilities of adaptation and a long-term learning aspect.

Issue 2: Shortcomings of the CMU database

One thing that is to be kept in mind is the deficiency of existing popular facial expression databases, e.g., the CMU database.

VI. PROPOSED WORK

In light of the deficiencies explored, the proposed system work is to overcome the following deficiencies:

1) Developing the image processing techniques highly relevant for identifying the facial features under 'uneven lighting'.

2) Interpreting the face emotion through the processed facial features.

3) Improving the recognition rate.

First step in process of emotion recognition is face detection in given image. Haar feature selection, creating an integral image, Adaboost training and Cascading classifiers. After face detection depending on facial feature extraction three types of approaches which geometric approaches, appearance based approach and hybrid approach combination of geometric and appearance can be used. Active Shape Model is popular geometric based approach in which detected image is iteratively deformed to fit shape model and extract facial points after comparison with shape model. After extraction of features, different classifiers are used for the classification of emotions. Least mean square method, Support Vector Machine (SVM), Neural Networks (NN), Hidden Markov Model and Adaboost are different types of classifiers used for classification. In classification process first training has to be done to train the software later testing is done using test subject. For training many database are available which are Cohn-Kanade, FEEDTUM, JAFFE and CMU MultiPIE. Later the software developed can be deployed on system development kit or on mobile phones [3] for further use.

In the proposed method, the objective is to develop real-time emotion recognition from facial images to recognize basic emotions like anger, disgust, happiness, surprise and neutral. The CMU MultiPIE database, which is a collection of images from 337 subjects with a variety of different facial expressions including neutral, happiness, surprise, disgust and anger. The subjects include 235 males and 102 females with different level of illuminations and poses. Viola-jones face detection method for face detection, Active shape Model (ASM) for extracting facial points and AdaBoost classifier have been used for developing the emotion recognition software.

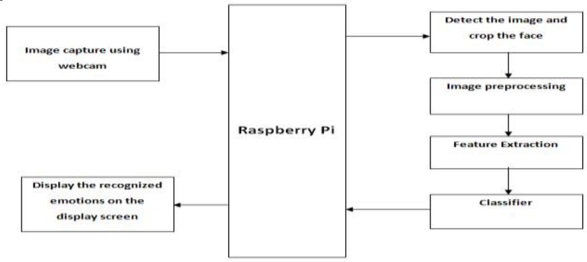


Fig.2 Overview of Real Time Emotion Recognition System

The architecture of proposed system is shown in the above Fig. And explained as follows: The input image in real time is captured through webcam and fed to emotion recognition software as input. Emotion recognition software is deployed in the Raspberry Pi, which gives classified emotion as output. The recognized emotion is displayed in the monitor. The algorithm for real time implementation of emotion recognition using Raspberry Pi II is explained as follows:

Step 1: Input image is captured through webcam.

Step 2: Viola-Jones face detection technique is used to detect the facial image. Viola-Jones used Haar wavelet concept to develop integral image to detect face. Haar features consider the different intensity of values of adjacent rectangular region as different area of face has different value of intensity from other region. After detection, facial image is saved for further processing and non-face area is removed.

Step 3: In image pre-processing, image is cropped according to required size and converted in gray image. This cropped image is used as input to Sobel filter for smoothing to remove the noise.

Step 4: Feature extraction is based on geometric approach for which Active Shape Model (ASM) is used. ASM automatic fiducial point location algorithm is applied first to a facial expression image, and then Euclidean distances between center gravity coordinate and the annotated fiducial points coordinates of the face image are calculated.

In order to extract the discriminate deformable geometric information, the system extracts the geometric deformation difference features between a person's neutral expression and the other basic expressions. In ASM input face shape is iteratively deformed to get the shape model. After comparison with shape model feature point of input facial image is extracted.

Step 5: Classification is done by adaptive boosting classifier (AdaBoost). AdaBoost is a powerful learning concept that provides a solution to supervised classification

https://dx.doi.org/10.22161/ijaers/nctet.2017.ece.2 ISSN: 2349-6495(P) | 2456-1908(O)

learning task. It combines the performance of many weak classifiers to produce a powerful committee as shown in equation. AdaBoost is a flexible classifier which can be combined with any learning algorithm. It is very simple and easy to perform in which only one parameter i.e., number of iteration is varied to get good accuracy.

Step 6: Hardware implementation using Raspberry Pi II:

The software developed for real time implementation is tested and deployed in Raspberry Pi II in linux environment. The proposed system design using Raspberry Pi with external webcam, keyboard and display monitor. Monitor and keyboard are connected to Raspberry Pi as it does not have display and input unit. Laptop can also be used as remote desktop for display and keyboard for input by using Virtual Network Connection (VNC) and putty software. In real time, when a person look into the webcam, his/her image will be taken and given to Raspberry Pi Emotion recognition software that is already deployed will recognize emotions and displays the recognized emotion into the display monitor.

VI. RESULTS AND ANALYSIS

The results of the classification for 5 basic expressions for frontal poses are recorded. The proposed system in real-time, tested using 25 subjects who performed 5 expressions by looking into the webcam connected to Raspberry Pi and achieved better accuracy with less processing time when compared to other methods. Implementation of real time emotion recognition in Raspberry Pi II is a novel method and it is can be used in a variety of applications as it is very small, light weighted and very less power supply is needed. It can be mounted over small size of robot and used for many applications.

VII. CONCLUSION & Future Work

The proposed system is highly useful to the society for different applications where emotion recognition plays a major role. In future work, different algorithm can be implemented to improve recognition accuracy. Robots can also be made to recognize emotion by neurological inspiration. Other modality like speech can be combined along with image for emotion recognition.

REFERENCES

[1] P. Ekman and E. L. Rosenberg, *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System*. Oxford, U.K.: Oxford Univ. Press, 2005.

[2] J. Russell and J. Fernandez-Dols, *The Psychology of Facial Expression*. New York: Cambridge Univ. Press, 1997.

[3] B. Golomb and T. Sejnowski, "Benefits of machine understanding of facial expressions," in *NSF Report—Facial Expression Understanding*, P. Ekman, T. Huang, T. Sejnowski, and J. Hager, Eds. Salt Lake City, UT, 1997, pp. 55–71.

[4] M. Pantic, "Face for ambient interface," in *Ambient Intelligence in Everyday Life*, vol. 3864, *Lecture Notes on Artificial Intelligence*. Berlin,Germany: Springer-Verlag, 2006, pp. 32–66.

[5] M. Cohen, *Perspectives on the Face*. Oxford, U.K.: Oxford Univ. Press, 2006.

[6] A. Young, *Face and Mind*. Oxford, U.K.: Oxford Univ. Press, 1998.

[7] M. Pantic and L. J. M. Rothkrantz, "Toward an affect-sensitive multimodal human–computer interaction," *Proc. IEEE*, vol. 91, no. 9, pp. 1370–1390, Sep. 2003.

[8] Z. Zeng, M. Pantic, G. Roisman, and T. Huang, "A survey of affect recognition methods: Audio, visual and spontaneous expressions," *IEEETrans. Pattern Anal. Mach. Intell.*, vol. 31, no. 1, pp. 39–58, Jan. 2009.

[9] Y. L. Tian, T. Kanade, and J. F. Cohn, *Handbook of Face Recognition*. New York: Springer-Verlag, 2005.

[10] M. Pantic and M. Bartlett, "Machine analysis of facial expressions," in *Face Recognition*, K. Delac and M. Grgic, Eds. Vienna, Austria: I-Tech Educ. Publishing, 2007, pp. 377–416.

[11] J. F. Cohn, "Foundations of human computing: Facial expression and emotion," in *Proc. ACM Int. Conf. Multimodal Interfaces*, 2006, vol. 1, pp. 610–616.

[12] P. Ekman, *Face of Man: Universal Expression in a New Guinea Village*. New York: Garland, 1982.

[13] W.-F. Liu, J.-L. Lu, Z.-F. Wang, and H.-J. Song, "An expression space model for facial expression analysis," in *Proc. CISP*, May 2008, vol. 2, pp. 680–684.

[14] K. Anderson and P. McOwan, "A real-time automated system for recognition of human facial expressions," *IEEE Trans. Syst., Man, Cybern. B,Cybern.*, vol. 36, no. 1, pp. 96–105, Feb. 2006.

[15] S. Park, J. Shin, and D. Kim, "Facial expression analysis with facial expression deformation," in *Proc. 19th ICPR*, Dec. 2008, pp. 1–4.

[16] I. Kotsia and I. Pitas, "Facial expression recognition in image sequences using geometric deformation features and support vector machines," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 172–187, Jan. 2007.

[17] R. E. Kaliouby and P. Robinson, "Real-time inference of complex mental states from facial expressions and head

https://dx.doi.org/10.22161/ijaers/nctet.2017.ece.2 ISSN: 2349-6495(P) | 2456-1908(O)

gestures," in *Proc. IEEE Conf.Comput. Vis. Pattern Recog.*, 2004, vol. 3, p. 154.

[18] A. Ashraf, S. Lucey, T. Chen, J. Cohn, and K. Prkachin, "The painful face—Pain expression recognition using active appearance models," in *Proc. Int. Conf. Multimedia Interfaces*, 2007, pp. 9–14.

[19] G. Littlewort, M. Bartlett, and K. Lee, "Faces of pain: Automated measurement of spontaneous facial expressions of genuine and posed pain," in *Proc. Int. Conf. Multimodal Interfaces*, 2007, pp. 15–21.

[20] H. Gu and Q. Ji, "Information extraction from image sequences of real world facial expressions," *Mach. Vis. Appl.*, vol. 16, no. 2, pp. 105–115, Feb. 2005.

[21] Mehrabian, A. Communication without words, Psychology Today, volume 2, pp 52-55, 1968.

[22] Peng Zhao-yi, Zhu Yan-hui, Zhou Yu, Real-time Facial Expression Recognition Based on Adaptive Canny Operator Edge Detection, IEEE, Multimedia and Information Technology (MMIT), SecondInternational Conference on (Volume:2), pp 154 – 157, 2010.