Feature Based Method For Human Facial Emotion Detection using Optical Flow Based Analysis

Gurpreet Singh[1], **Baljit Singh**[2]
[1]*A.P, CSE (IET Bhaddal, Ropar),*
[2]*A.P, CSE& IT (BBSBEC, Fatehgarh Sahib),*
gurpreet32@gmail.com, baljitkhehra@rediffmail.com

Abstract: *Computer has been widely deployed to our daily lives, but human computer interaction still lacks intuition. Researchers intend to resolve these shortcomings by augmenting traditional systems with human like interaction mechanism. Today, dedicated hardware often infers the emotional state from human body measures.These have been a considerable amount of research done into the detection and implicit communication channels, including facial expressions. Most studies have extracted facial features for some specific emotions in specific situations. In this paper we uses a feature point tracking technique applied to five facial image regions to capture basic emotions. The used database contains 219 images, 10 Japanese female, six expressions and one neutral. We use grayscale images which are ethically not diverse. We use optical flow based analysis to detect emotions from human facial image data. Our proof of data demonstrates the feasibility of our approach and shows promising for integration into various applications.*

1. Introduction

Artificial recognition of facial expressions has attracted a lot of attention due to its potential commercial value in fields like Lie detection, Surveillance, Criminal Investigation, Security and Forensic applications. Facial expression recognition can be utilized for automated analysis of human emotion.

© 2011 Journal Anu Books

Communication between humans is influenced by emotion. Interpersonal behavior is affected by facial expressions during communication. The study of human facial expressions started with Darwin (1965) in the 19th century and is still being studied. In 1971, Ekman and Frisen classified emotions into six primary categories, all universal across different ethics and cultural groups with each being represented by a unique facial expression. Six emotional categories are: Happiness, Sadness, Surprise, Fear, Anger and Disgust. Recent approaches for facial expression detection are Template Based Method Edward, Cootes and Taylor (1998) and Feature Based Method by Black and Yacoob (1997). The difference between these two methods depends on the use of still images or successive image sequences and whether they are template of feature based (Pantic&Rothkrantx, 2000). The template approach uses the average face for each category of emotion and classifies the individual facial expressions according to the best match of each template. The feature based approach uses a training set of images for different emotional expressions. The features are extracted fro each emotion subset for all facial expressions and then are subsequently tested unseen facial images. Feature based technique involves detecting changes of the features in different facial regions. The selection of these facial regions is based on the Facial Action Coding System (FACS). The Facial Action Coding System is a human observed based system designed to detect changes in facial features.

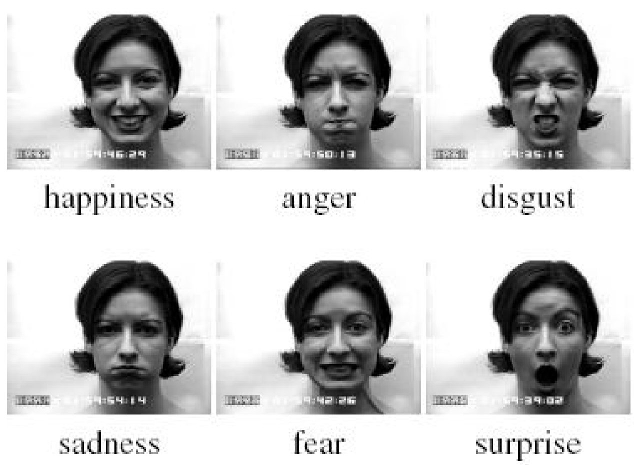

Fig 1. Six universal facial expressions

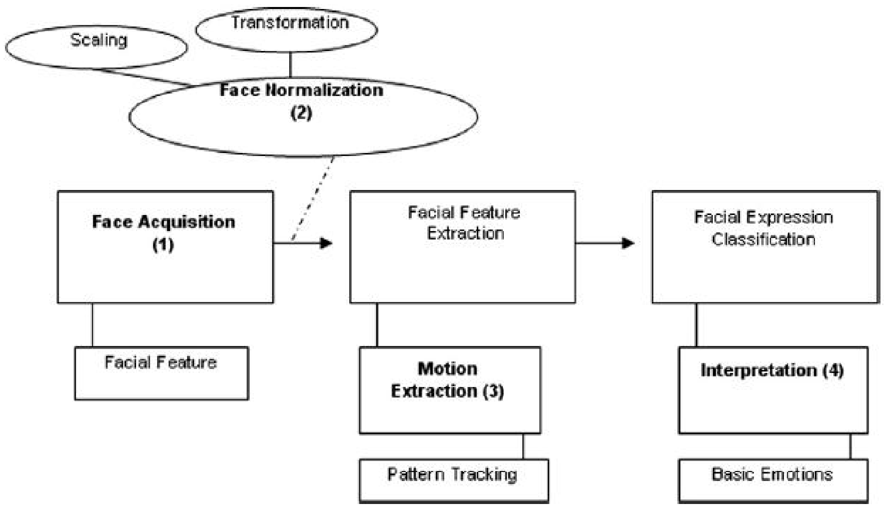

Fig 2.Overview of the emotion extraction and classification Process from captured facial Image

FACS consists of 44 anatomically based action units, which individually or in combination can represent all visible discriminate expressions. The tracking algorithm used in this research was separately applied to the five facial image regions (eyebrows, eyes and mouth) each represented by their feature points. This approach was chosen to accelerate the computation and to classify the images based on the movement of individual facial regions rather than the entire face. This approach was to create independence for each facial feature and thus it was envisaged that the accuracy of the tracking process would be increased and thus achieve a better classification result.

The tracking algorithm generally compared the pixel values of a feature point to that of the surrounding pixels to determine their movement. Misinterpretation of movement increases with changing light conditions faces whit glasses and facial hair. To minimize this, images of faces with glasses and facial hairs were not used in this research. The advantage of tracking feature points of individual facial regions allowed each area to be classified individually. Based on threshold, the feature points of a region were either discarded or used as input for the classifying process.

© 2011 Journal Anu Books

2. Method

This method section will explain the different steps required forthe facial emotion recognition process. The following describeswhich pictures were used and how they were pre-processed. Thiswill be followed by an explanation about the developed emotionextraction and interpretation process. The flow chart in Fig. 2shows the process by which the facial feature extraction process isperformed.

2.1 Image pre-processing

The facial image pre-processing includes: facial feature acquisitionand face normalization (angle and size).Such 'normalization' is necessary if the subjects in the images exhibitout of focal plane motion. To compensate for this motion an'affine transformation' is used (Weisstein, 2000). This transformationensures no changes occur to the normalized face positionand maintains magnification. Colored images of the dataset were then transformed intogreyscale images and were resized into 256 X 256 pixel dimensions.Using FACS as a basis, the five facial regions were croppedfrom the image. The resultant five feature regions (botheyebrows, both eyes and mouth), formed the input data for theextraction process.

2.2. Facial feature extraction

The obtained input data was used for the selection of the keyfeature points. The key feature points were determined for the firstframe of each image sequence. These key points were tracked insubsequent frames by the tracking algorithm.

2.2.1. Pattern tracking

The aim of the tracking process is to trace the displacementof each selected feature point within each of the regions.Each point serves as the centre of a 5 X5 pixel window.Fig. 2c and d demonstrates the upward movement of the centrepoint on the left-eyebrow (marked by _ in Fig. 2c) from a twoframe sequence.This method estimates the optical flowbetween two images. In cropped images, the displacement of eachfeature point (centre of a 5 X 5 pixel window) was calculated bysubtracting its current position from the position of the previousframe. Based on the position of the feature points in each first imageand the position of the feature points (after the tracking process)in the last image, vectors were calculated. A vectorrepresents a line segment connecting the initial with the terminalpoint. The resultant set was 26 feature vectors withangles (degree) and magnitudes (length). All 26 feature vectorswere divided into 8 eyebrow vectors, 8 eye vectors and 10 mouthvectors. These vectors (depicted by the lines) demonstrate the displacementof the selected feature points

© 2011 Journal Anu Books

between the first mouthimage and the last image. The white squaresrepresent the selected feature points in the first image.

2.3. Facial expression classification

After resizing and normalization, the feature points weretracked and the displacement of the points calculated (facial featuredata). The next step of the classification process identifiesthe facial expressions (emotion) of each sequence and classifieseach sequence into the four categories: happiness, sadness, angerand null category. The classification process compared the magnitudeand angle of the feature vectors of an image sequence fromthe templates which were calculated for each of the three emotionalcategories: happiness, sadness and anger

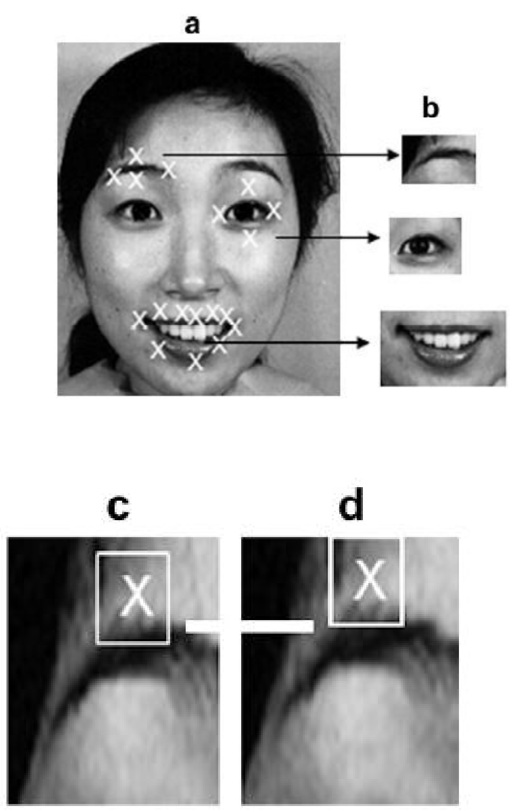

Fig.3. (a) Feature points _ and (b) cropped regions, (c, d) the displacement of a feature point.

© 2011 Journal Anu Books

3. Facial expression databases

There have been some attempts to createcomprehensive test-beds for comparative studies offacial expression analysis. The most famous andcommonly used of these are given below. Chosen here as representative examples, rather thanan exhaustive survey. We provide the details of thesedatabases in Table 1. To date, the Cohn-KanadeAUCodedFacial Expression Database is the mostcommonly used database in research on automated facial expression analysis.

Table 1. Examples of Existing Facial Expression Databases

Name	JAFFE [17]	Cohn-Kanade [11]	MMI [18]
Number of data	219 images	2105 videos	740 images and 848 videos
Subjects	10 Japanese female	100 subjects; 65% female, 15%t African-American, and 3% Asian/Latino, 18 to 30 year old	19 subjects, 3 ethnicities, 19-62 year old
Expressions	6 basic facial expressions + 1 neutral	facial expressions and combination of facial action units	6 basic facial expressions single action units combination of action units
features	Each image rated on 6 emotion adjectives by 60 Japanese subjects.	in-plane and limited out-of-plane motion beginning with a neutral face FACS coding for each expression , codes refer only to the final frame	Data from frontal and profile view For 169 samples analysis of AU temporal activation patterns supported
Limitations	only gray scale images ethnically not diverse	limited image data from the frontal camera available for distribution data available for public use is gray scale mostly single action units are rarely included	

4. Algorithm

Microsoft Visual Studio 2008 is used as the software development environment to normalize and resize the images and extract the facial data. This algorithm can be categorized into two main parts: 1) The face normalization process and 2) The feature extraction and classifying process. The first step converts the color image to grayscale and normalizes them for all of the 210 uploaded facial image sequences. Images exhibiting out of plane motion were subjected to the affine transformation process.

The grey scaling was applied to minimize the problem of lighting variation within the image sequences.

The next step in the process was to crop the facial image region of interest from the image sequences. Eight key feature points were manually selected for both eyebrows and both eyes respectively and a further 10 key feature points for the mouth. All categorized image sequences were used as the input data forthe tracking algorithm. The final position of the tracked featurepoints of each facial image region for each sequence formed theoutput of the tracking process. Based on the location of the initialbaseline and the final feature point positions for each image sequence,the feature vectors for each facial region (8 vectors fromboth eyebrows, 8 for both the eyes and 10 for the mouth) were calculatedfor the three emotional categories.Using the calculated feature vectors of each emotional category,the mean of the angle and the magnitude of the feature vectors ofeach facial image region was determined. The outcome were 8template feature vectors from both eyebrows, 8 template featurevectors for both the eyes and 10 template feature vectors for themouth for each emotional category. In total each category contained26 template feature vectors.

5. Results

A total of 83.33% of the image sequences were classified in either of the three emotion categories, while 16.77% were not classified at all (no category folder). A total of 63.33% of the 83.33% were classified correctly (right category). Those 20% that were not classified correctly formed the false category happy image sequences have been detected with an accuracy of 80% the angry and sad images with an accuracy of 50% and 60%, respectively.

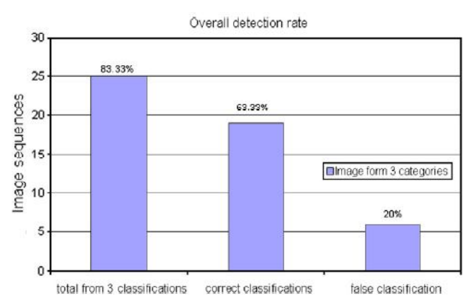

Fig The overall, the correct and false detection rate in either of the three emotion categories (happy, angry and sad).

© 2011 Journal Anu Books

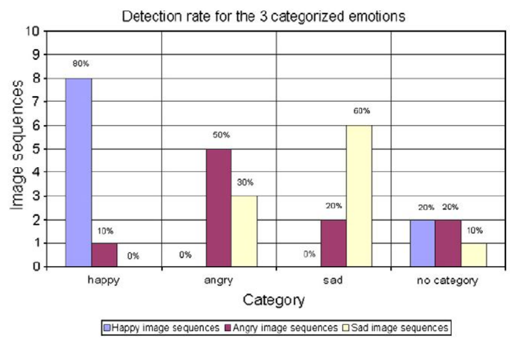

Fig. The detection accuracy for the happy, angry and sad image sequences.

Happy image sequences were classified in either the happy category or in the null category, this demonstrates the capability of the algorithm to distinguish happy from angry and sad categories.

6. Conclusion

This study has investigated the possibility to detect the threeemotions happy, angry and sad in video sequences by applying atracking algorithm. Therefore, the 10 images per each sequencewere pre-processed and the facial movement tracked. Some significantdifferences regarding the detection rate was detected betweenhappy image sequences and the similar angry, sad imagesequences.Based on that knowledge the future work will be to increase thedetection accuracy of the system using images sequences of similaremotions. A solution to achieve this better detection accuracycould be use of images of better quality and especially the use ofhigher amounts of images per sequence.The current work used 10 images per sequence. A sequencecontaining more images contains a higher amount of information.The more information would provide a more accurate descriptionof an emotion and thus result in higher detection accuracy. The researchshows, that the tracking algorithm separately applied to thefive facial image regions show promising results.

References

[1] Facial Emotion Recognition using Multi-Model information by Liyanage C. DE SILVA, **I** Tsutomu MIYASATO , Ryohei NAKATSU in International conference

on "Information, Communications and Signal Processing" ICICS' 97 , Singapore 9-12 September, 1997.

[2] Multi-Model Emotion Recognition Using Canonical Correlations and Acoustic features by RokGajsek, VitomirStruc, France Mihelic in International Conference on Pattern Recognition, 2011

[3] Tailoring Model-based Techniques to Facial Expression Interpretation by Matthias Wimmer,

Christoph Mayer, Sylvia Pietzsch, and Bernd Radig in First International Conference on Advances in Computer-Human Interaction 0-7695-30876-9/08 IEEE DOI 10.1109/ACHI.2008.7

[4] Towards Improving Visual-Facial Emotion Recognition through Use of Complementary Keyboard-Stroke Pattern Information by George A. Tsihrintzis, Maria Virvou, EfthymiosAlepis, and Ioanna-OuraniaStathopoulou in 978-0-7695-3099-4/08, 2008 IEEE DOI 10.1109/ITNG.2008.152

[5] A Multi-Modal Emotion-Diagnosis System to Support e-Learning by KiyhoshiNosu and TomoyaKurokawa in Proceedings of the First International Conference on Innovative Computing, Information and Control (ICICIC'06)0-7695-2616-0/0, 2006

[6] A Bimodal Face and Body Gesture Database for Automatic Analysis of Human Nonverbal Affective Behavior by HaticeGunes and Massimo Piccardi in Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06), 2006

[7] Phased processing of facial emotion: An ERP study by Nugraha P. Utama, Atsushi TakemotoYasuharu Koike, Katsuki Nakamura in Neuroscience Research 64 (2009) 30–40. Journal homepage: www.elsevier.com/locate/neures

[8] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Vot- sis, S. Kollias,W. Fellenz, and J. Taylor. Emotion recog- nition in human-computer interaction. *IEEE Signal Processing Magazine*, 18 (1)(1):32 – 80, January 2001.

[9] F. Eyben, M. WIlmer, and B. Schuller. openear - introducing the munich open-source emotion and affect recognition toolkit. In *Proc. of ACII 2009*, Amsterdam, pages 576–581., 2009.

© 2011 Journal Anu Books

[10] T. Kim, J. Kittler, and R. Cipolla. Discriminative learning and recognition of image set classes using canonicalcorrelations. *TPAMI*, 29(6):1005–1018, June 2007.

[11] M. Mansoorizadeh and N. M. Charkari. Multimodal information fusion application to human emotion recognition from face and speech.Multi.Tools and App, 2009.

[12] O. Martin, I. Kotsia, B. Macq, and I. Pitas. The enter-face'05 audio-visual emotion database.In *ICDEW '06*,Washington, DC, USA, 2006.

[13] M. Paleari, R. Benmokhtar, and B. Huet. Evidencetheory-based multimodal emotion recognition. In*MMM'09*, pages 435–446, Berlin, 2008.

[14] J. Pittermann, A. Pittermann, and W. Minker. *Handling Emotions in Human-Comp. Dialog*.Springer,Dordrecht (The Netherlands), 2009.

[15] B. Schuller. Speaker, noise, and acoustic space adaptation for emotion recognition in the automotive environment.In *Proc. 8th ITG conf. on Speech Comm.*, 2008.

[16] B. Schuller, S. Steidl, and A. Batliner. The interspeech2009 emotion challenge.In *Proc. Interspeech 2009*.

[17] N. Sebe, I. Cohen, and T. G. T. Huang. Multimodalapproaches for emotion recognition : A survey. In *Proc.of SPIE*, volume 5670, pages 56–67, January 2005.

[18] P. Viola and M. Jones. Robust real-time face detection.*Int. J. of Comp. Vision*, 57(2):137 – 154, 2004.

[19] O. Yamaguchi, K. Fukui, and K. Maeda. Face recognition using temporal image sequence. In *Proc. of AFGR*,pages 318–323, 1998.

© 2011 Journal Anu Books