

Meta-analysis of the First Facial Expression Recognition Challenge by Using Embedded Systems

* Vutukuri Samba Siva Rao 1, Dr.R.V.Krishnaiah2

1PG Student (M.Tech VLSI&ES), Dept. of ECE, DRK Institute of Science & Technology, Hyderabad, AP, India

2 Professor, Dept. of ECE, DRK Institute of Science & Technology, Hyderabad, AP, India

Abstract : *Facial expressions convey non-verbal cues, which play an important role in interpersonal relations. Automatic recognition of facial expressions can be an important component of natural human-machine interfaces; it may also be used in behavioral science and in clinical practice. Although humans recognize facial expressions virtually without effort or delay, reliable expression recognition by machine is still a challenge. Automatic facial expression recognition has been an active topic in computer science for over two decades, in particular facial action coding system action unit (AU) detection and classification of a number of discrete emotion states from facial expressive imagery. Our system is designed by using ARM 32-bit micro controller which supports different features and algorithms for development of first facial recognition. The webcam combines video sensing, video processing and communication within a single device it captures a video stream like different Expressions of face, computes the information and Transfers the compressed video stream to the ARM Microcontroller. The image it received is processed by using image processing algorithms and processed image is classified by using PCA algorithms and identified expressions are displayed on display unit. Our system is designed by using S3C2440 micro controller developed by Samsung which was called as friendly ARM or mini 2440 board.*

Keywords: ARM, PCA, S3C2440.

1. INTRODUCTION

A facial expression is a visible manifestation of the affective state, cognitive activity, intention,

personality, and psychopathology of a person [6]; it plays a communicative role in interpersonal relations. Facial expressions, and other gestures, convey non-verbal communication cues in face-to-face interactions. These cues may also complement speech by helping the listener to elicit the intended meaning of spoken words. As cited in [14] (p. 1424), Mehrabian reported that facial expressions have a considerable effect on a listening interlocutor; the facial expression of a speaker accounts for about 55 percent of the effect, 38 percent of the latter is conveyed by voice intonation and 7 percent by the spoken words.

Facial expression recognition, in particular facial action coding system (FACS) action unit (AU) detection and classification of facial expression imagery in a number of discrete emotion categories, has been an active topic in computer science for some time now, with arguably the first work on automatic facial expression recognition being published in 1973. Many promising approaches have been reported. The first survey of the field was published in 1992 and has been followed up by several others. However, the question remains as to whether the approaches proposed to date actually deliver what they promise. To help answer that question, we felt that it was time to take stock, in an objective manner, of how far the field has progressed.

2. PREVIOUS WORK

This section describes previous work on FACS and prior work on automatic recognition of AUs from video.

2.1. FACS

The Facial Action Coding System (FACS) [14] is a comprehensive, anatomically-based system for measuring nearly all visually discernible facial movement. FACS coding typically involves frame-by-frame inspection of the video, paying close attention to transient cues such as wrinkles, bulges, and furrows to determine which facial action units have occurred and their intensity. Full labeling requires marking onset, peak and offset and may include annotating changes in intensity as well. Left to right, evolution of an AU 12 (involved in smiling), from onset, peak, to offset. FACS coding typically involves frame-by-frame inspection of the video, paying close attention to transient cues such as wrinkles, bulges, and furrows to determine which facial action units have occurred and their intensity. Full labeling requires marking onset, peak and offset and may include annotating changes in intensity as well. Left to right, evolution of an AU 12 (involved in smiling), from onset, peak, to offset. FACS coding typically involves frame-by-frame inspection of the video, paying close attention to transient cues such as wrinkles, bulges, and furrows to determine which facial action units have occurred and their intensity. Full labeling requires marking onset, peak and offset and may include annotating changes in intensity as well. Left to right, evolution of an AU 12 (involved in smiling), from onset, peak, to offset.

example for AU12. Comprehensive reviews of automatic facial coding may be found in [23, 32, 26].



Figure 2. FACS coding typically involves frame-by-frame inspection of the video, paying close attention to transient cues such as wrinkles, bulges, and furrows to determine which facial action units have occurred and their intensity. Full labeling requires marking onset, peak and offset and may include annotating changes in intensity as well. Left to right, evolution of an AU 12 (involved in smiling), from onset, peak, to offset.

2.2. Automatic FACS recognition from video

Two main streams in the current research on automatic analysis of facial expressions consider emotion-specified expressions (e.g., happy or sad) and anatomically based facial actions (e.g., FACS). The pioneering work of Black and Yacoob [5] recognizes facial expressions by fitting local parametric motion models to regions of the face and then feeding the resulting parameters to a nearest neighbor classifier for expression recognition. De la Torre et al. [13] use condensation and appearance models to simultaneously track and recognize facial expression. Chang et al. [8] use a low dimensional Lipschitz embedding to build a manifold of shape variation across several people and then use I-condensation to simultaneously track and recognize expressions. Lee and Elgammal [17] use multi-linear models to construct a non-linear manifold that factorizes identity from expression. Recently there has been an emergence of 2 Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on January 28, 2010 at 11:13 from IEEE Xplore. Restrictions apply. efforts toward explicit automatic analysis of facial expressions into elementary AUs [29, 21]

as they are very suitable to be used as mid-level parameters in automatic facial behavior analysis [9]. Several promising prototype systems were reported that can recognize deliberately produced AU in either near frontal view face images (Bartlett et al., [2]; Tian et al., [26]; Pantic & Rothkrantz, [22]) or profile view face images (Pantic & Patras, [21]). These systems employ different machine learning methods and different image representations as they are the key stages for automatic AU recognition. Most work in automatic analysis of facial expressions differs in choice of features and/or classifiers. Bartlett et al. [3] investigate machine learning techniques including SVMs, Linear Discriminant Analysis, and AdaBoost, concluding that the best recognition performance is obtained through SVM classification on a set of Gabor wavelet coefficients selected by AdaBoost. However, the computational complexity of Gabor and SVMs are considerable. To develop and evaluate facial action detector, large collections of training and test data are necessary. Although high scores have been achieved on posed facial action data [28, 31, 25], only a small number of studies being conducted on nonposed spontaneous data [7, 3, 19]. The latter are preferable to posed as they are representative of real world facial actions. In our paper, we focus on a problem common to almost all approaches to facial expression analysis; that is, how best to exploit the training data to improve classification performance. We evaluate our approach by detecting FACS action units (AU) in a relatively large data set of nonposed, spontaneous facial behavior.

3. GEMEP-FERA DATA SET

To be suitable to base a challenge on, a data set needs to satisfy two criteria. First, it must have the correct labeling, which in our case means frame-by-frame AU labels and event coding of discrete emotions. Second, the

database cannot be publicly available at the time of the challenge. The GEMEP database is one of the few databases that meet both conditions and was therefore chosen for this challenge.

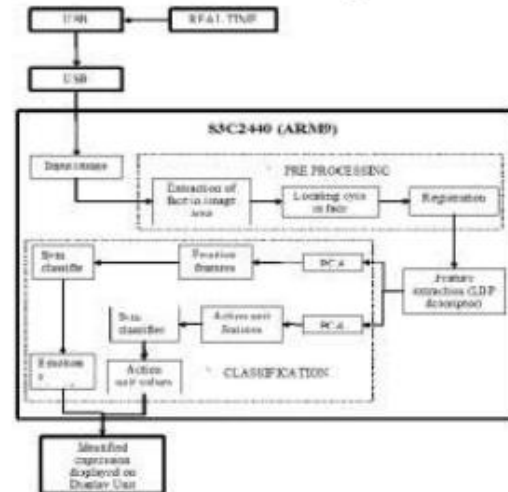


Figure 2 Block Diagram of The System

By no means does the GEMEP-FERA data set constitute the entire GEMEP corpus. In selecting videos from the GEMEP corpus to include in the GEMEP-FERA data set, the main criterion was the availability of a sufficient number of examples per unit of detection for training and testing. It was important that the examples selected for the training set were different from the examples selected for the test set.

PARTITIONING

For the AU detection subchallenge, we used a subset of the GEMEP corpus annotated with the FACS. The 12 most commonly observed AUs in the GEMEP corpus were selected (see Table I). To be able to objectively measure the performance of the competing facial expression recognition systems, we split the data set into a training set and a test set. A total of 158 portrayals (87 for training and 71 for testing) were selected for the AU subchallenge. All portrayals are recordings of actors speaking one of the two

pseudolinguistic phoneme sequences. Consequently, AU detection is to be performed during speech. The training set included seven actors (three men), and the test set included six actors (three men), half of which were not present in the training set. Even though some actors were present in both training and test sets, the actual portrayals made by these actors were different in both sets. For the emotion sub-challenge, portrayals of five emotional states were retained: anger, fear, joy, sadness, and relief. Four of these five categories are part of what Ekman called basic emotions as they are believed to be expressed universally by specific patterns of facial expression. The fifth emotion, relief, was added to provide a balance between positive and negative emotions but also to add an emotion that is not typically included in previous studies on automatic emotion recognition. Emotion recognition systems are usually modeled on the basic emotions; hence, adding “relief” made the task more challenging.

4. CONCLUSION

The paper “Meta-Analysis of the First Facial Expression Recognition Challenge” has been successfully designed and tested. It has been developed by integrating features of all the hardware components and software used. Presence of every module has been reasoned out and placed carefully thus contributing to the best working of the unit. Secondly, using highly advanced ARM9 board and with the help of growing technology the project has been successfully implemented.

Another issue that arose during the challenge is the choice of performance measure. It is well known that, in a heavily unbalanced data, such as that of the AU detection subchallenge, the classification rate is not a suitable measure. A naive classifier based on the prior probability of the classes will give an overoptimistic

representation of the problem and is very likely to outperform systems that try to detect both classes with equal priority. The detection of AUs, however, is still far from solved, and this should definitely remain a focus in future events. In the future, it would be desirable to have a data set that will allow a competition on detection of all 31 AUs, plus possibly a number of FACS action descriptors. Aside from addressing the detection of the activation of AUs, it would be a good thing to move toward the detection of the intensities and temporal segments of AUs, as it is these characteristics that prove to be crucial in many higher level behavior understanding problems.

5. RESULT

Present system is designed for identifying facial expressions of a human. It identifies first facial expression provided by the person, based on expressions generated by person corresponding images are just displayed on screen.

In future we can use same method of recognizing facial expressions based on facial expression generated by a person we can play music. Suppose if person keeps sad face expression camera captures first facial expression and by using internal algorithms it identifies type of expression based on that expression corresponding sad song is played.

6. REFERENCES

- [1] T. Ahonen, A. Hadid, and M. Pietikäinene, “Face description with local binary patterns: Application to face recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.
- [2] N. Ambady and R. Rosenthal, “Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis,” *Psychol. Bull.*, vol. 11, no. 2, pp. 256–274, Mar. 1992.

[3] A. B. Ashraf, S. Lucey, J. F. Cohn, T. Chen, Z. Ambadar, K. M. Prkachin, and P. E. Solomon, "The painful face—Pain expression recognition using active appearance models," *Image Vis. Comput.*, vol. 27, no. 12, pp. 1788–1796, Nov. 2009.

[4] A. Asthana, J. Saragih, M. Wagner, and R. Goecke, "Evaluating landmark fitting methods for facial expression recognition," in *Proc. Int. Conf. Affective Comput. Intell. Interact.*, 2009, pp. 1–8.

[5] T. Baltrusaitis, D. McDuff, N. Banda, M. Mahmoud, R. El Kaliouby, P. Robinson, and R. Picard, "Real-time inference of mental states from facial expressions and upper body gestures," in *Proc. IEEE Int. Conf. Autom. Face Gesture Anal.*, 2011, pp. 909–914.

BIOGRAPHY

Vutukuri Samba Siva Rao is a PG Student. He is doing Masters in Electronics and Communication Engineering from DRK Institute of Science&Technology, Hyderabad with specialization in VLSI&ES.

Dr.R.V.Krishnaiah received his B.Tech degree in ECE from Bapatla Engineering College. He received his M.Tech degree in Computer Science Engineering from JNTU and also M.Tech-(EIE) from NIT,Warangal. He did Ph.D (MIE, MIETE, MISTE) from JNTU Ananthapur.

IJERT