K-077

TV Rating System Using Facial Expressions

Uwe Kowalik, Terumasa Aoki, Hiroshi Yasuda

Research Center of Advanced Science and Technology,
The University of Tokyo
Yasuda/Aoki Lab.,
4-6-1 Komaba, Meguro-ku,
Tokyo, Japan
{uwe,aoki,yasuda}@mpeg.rcast.u-tokyo.ac.jp

Abstract

Today the television plays the probably most important role in the area of entertainment, education and information services. Since a TV can be found in almost every household, television business has the biggest amount of customers amongst all media services. A lot of money is invested annually in programs, commercials, equipment and people. Therefore marketing research institutions investigating since years the watching behavior of the television audience to provide statistical proven data for increasing the effectiveness of program making. We propose a system that automatically extracts the emotional state of the audience based on facial expressions while watching TV. This paper describes the system components as well as the mapping between facial features and emotion space.

1. Introduction

The current technology provides means to focus on the 'who is watching when which channel' problem. A representative amount of households is provided with special equipment to measure statistical watching behavior by tracking channel switch events synchronously along with the program. A weak point of the current technology is the fact, that the level of attention, that is actually paid to the program by the audience, can not be verified. People might leaving the room or doing something else than watching, while the program is still running. This research focuses on this problem and will provide a solution by facilitating a vision/thermal based detection system for human emotions that could be integrated into a set top box provided to the participant of a statistical enquiry. The system detects and classifies the emotional state of the audience while watching the program. The collected data is automatically edited and the result will be sent back to the program provider. A program provider can use the statistical data to improve the quality of the content which leads to a higher level of satisfaction at the audience. In the following the proposed system and its functionality will be described.

2. Related Work

Humans communicating there emotions by the means of facial expressions[1]. Ekman and Friesen developed the a system of facial action units (AU), containing all possible basic patterns of facial expressions to classify any possible complex facial expression as a superposition of one or more of those AUs. The Facial Action Coding System (FACS) is a comprehensive and widely used method of objectively describing facial activity. At present FACS coding is done manually by experienced FACS coders. In [3] an approach towards an automatic FACS coding system is described.

3. System Overview

The proposed System outline can be seen in Fig. 1. The system uses input from a visual camera as well as input from a thermal camera.

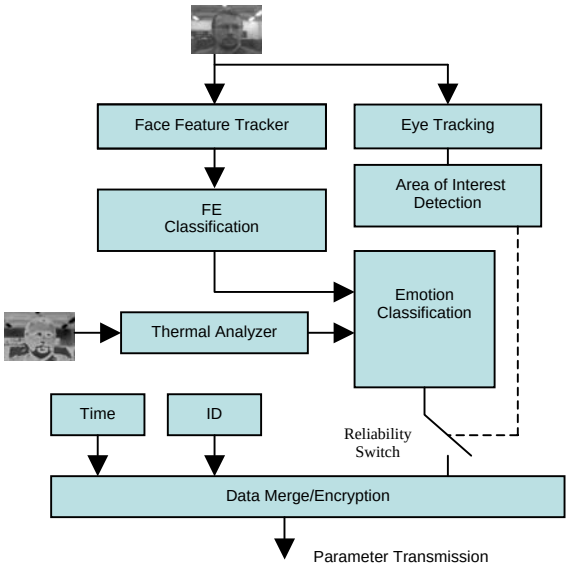

Fig. 1 Proposed system outline

Tracking of 22 significant face features will done by the tracking module providing information about position and movement of eyebrows, pupil, mouth, chin, etc. [2]. A feature vector containing 44 elements, i.e. the x- and y-components of the features, is build and feed into the facial expression

classification block. Here the positions of features and their changes over time are classified into intensities of different AUs as defined in [4].

Fig. 2 The system tracks 22 face features

A neural network will be trained to accomplish this task. Training data are generated manually accordingly to [4]. The output of the network is vector of 44 intensities referring to the 44 AUs that can be classified in the next step by the emotion classifier.

4. Mapping into Emotion Space

In [5] a data base is described, that is currently under construction and will serve for the purpose of mapping facial expressions to psychological interpretations based on the AU code. We propose to exploit the circumflex-model of affect in order to achieve a parameterization of the emotion space. This model can be seen in Fig. 3. It was introduced by Russel (1980) and is widely used in behavioral science to classify emotions. The emotion classification will be performed by a neural network, that produces an output vector consisting of two elements, i.e. x and y component of the two dimensional emotion space. Training data will be generated manually by referring to the FACSAID knowledge base [5].

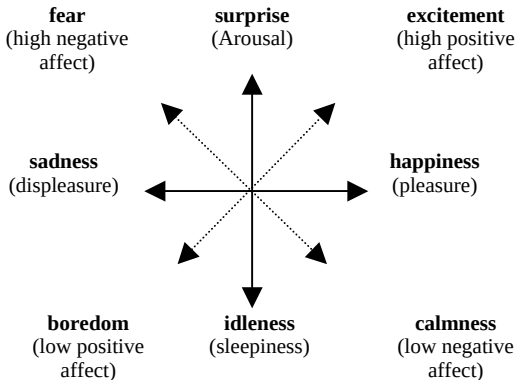

Fig. 3 Circumflex Model of Affect (Russel 1980)

To improve the performance we introduce the skin temperature as an additional parameter, since we believe that the body temperature carries important information of the emotional state as well. Therefore the input vector contains 45 elements (44x AUs and 1x temperature measure). The area of interest detection ensures, that the recorded emotion is related to the program by tracking the customers view direction.

5. Parameter Transmission

For the practical usage of the proposed TV program rating system it is very important to keep the time relationship between the program, the parameterized emotions and the information of focus of attention. Therefore the merge step ensures the synchronous multiplexing and delivery of the extracted statistical data to the program provider. Another important point is to protect the audiences privacy. No personal information will be included into the transmission. Instead an unique identifier will be added to the data stream for later assignment.

6. Conclusion

The emotional component plays an important role in almost every human decision making process. To understand mass media's emotional influence at the audience is necessary and desirable in order to keep and improve the quality of television programs and to provide customized content for better satisfaction. A lot of effort is done to investigate the desires and interests of TV program consumers. Today's technique of program rating does not allow to perceive information about the level of customers satisfaction. We propose a system, that focuses on this issue by introducing a vision/thermal based analysis of audience' emotions while watching TV. It facilitates an automatic emotion classification based on facial expressions. Parameterized emotion data are sent back to the program provider. These statistical data can be used to receive direct user feedback about the quality of provided television content in order to improve or customize the program for better customers satisfaction.

References

[1] Darwin, Charles: The expression of the emotions in man and animals. London, John Murray, 1872

[2] Joseph C. Hager & Paul Ekman: The Inner and Outer Meanings of Facial Expressions. Chapter 10 From: J. T. Cacioppo & R. E. Petty (Eds.) Social Psychophysiology: A Sourcebook. New York: The Guilford Press, 1983.

[3] J.R. Movellan and M.S. Bartlett: The next generation of automatic facial expression measurement. Afterword in P. Ekman (Ed.), What the Face Reveals, 2nd Edition, Oxford University Press

[4] Paul Ekman, Wallace V. Friesen, and Joseph C. Hager: The new (2002) Facial Action Coding System (FACS)

[5] Facial Action Coding System Affect Interpretation Dictionary (FACSAID) http://face-andemotion.com/dataface/facsaid