ARTICLE IN PRESS

Pattern Recognition ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Emotion recognition from geometric facial features using self-organizing map

Anima Majumder*, Laxmidhar Behera, Venkatesh K. Subramanian

Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur 208016, India

ARTICLE INFO

Article history:
Received 4 February 2013
Received in revised form
17 July 2013
Accepted 6 October 2013

Keywords:
Facial expression analysis
Geometric facial features
Self-organizing map
Features extraction
System identification
Radial basis function
Multi-layer perceptron
Support vector machine

ABSTRACT

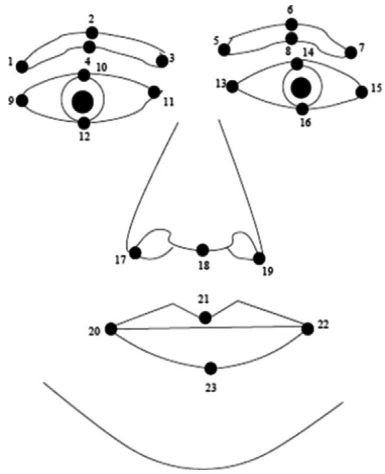
This paper presents a novel emotion recognition model using the system identification approach. A comprehensive data driven model using an extended Kohonen self-organizing map (KSOM) has been developed whose input is a 26 dimensional facial geometric feature vector comprising eye, lip and eyebrow feature points. The analytical face model using this 26 dimensional geometric feature vector has been effectively used to describe the facial changes due to different expressions. This paper thus includes an automated generation scheme of this geometric facial feature vector. The proposed non-heuristic model has been developed using training data from MMI facial expression database. The emotion recognition accuracy of the proposed scheme has been compared with radial basis function network, multi-layered perceptron model and support vector machine based recognition schemes. The experimental results show that the proposed model is very efficient in recognizing six basic emotions while ensuring significant increase in average classification accuracy over radial basis function and multi-layered perceptron. It also shows that the average recognition rate of the proposed method is comparatively better than multi-class support vector machine.

1. Introduction

In our day-to-day life, communication plays a very important role. With the growing interest in human–computer interaction, automation of emotion recognition became an increasingly crucial area to work on. Facial expressions are a kind of nonverbal communication. They are considered to be one of the most powerful and immediate means of recognizing one's emotion, intentions and opinion about each other. Mehrabian [16] found that when people are communicating feelings and attitudes, 55% of the message is conveyed through facial expression alone, vocal cues provide 38% and the remaining 7% is via verbal cues. Ekman and Friesen [3] did a rigorous study on facial expression and came to conclusion that facial expressions are universal and innate. They also stated that there are six basic expressions, these include happiness, sadness, disgust, anger, surprise and fear. Much efforts have gone towards the study of facial expression and emotion recognition, initially by cognitive scientists and later by computer vision researchers [27]. The Facial Action Coding System (FACS) [3] is a human observer based system, developed to detect the changes in facial features or facial muscles movements using 44 anatomically based action units. Determining FACS from images is a very laborious work, and thus, during the last few decades a lot of attention is given towards automating it. Automatic analysis of facial features requires extraction of relevant facial features from either static images or video sequences, which can either be further classified into different action units (AUs) or can be applied directly to the classifiers to give the respective emotion. Efficient extraction of facial features from faces of different persons is a crucial step towards accurate facial expression recognition. Generally, two common types of features are used for facial expression recognition: geometric features data and appearance features data. Geometric features give clues about shape and position of the feature, whereas appearance based features contain information about the wrinkles, bulges, furrows, etc. Appearance features contain micro-patterns which provide important information about the facial expressions. But one major drawback with appearance based methods is that it is difficult to generalize appearance features across different persons. Although geometric based features are sensitive to noise and the tracking of those features is rather difficult, geometric features alone can provide sufficient information to have accurate facial expression recognition [28]. We humans have a very extraordinary ability to recognize expressions. Even if we are given a cartoon image having only some contours, we can easily recognize the expression [5]. In many cases, it is observed that features obtained from facial

* Corresponding author. Tel.: +91 512 259 7854; fax: +91 512 259 0063.
E-mail addresses: animam@iitk.ac.in (A. Majumder), lbehera@iitk.ac.in (L. Behera), venkats@iitk.ac.in (V.K. Subramanian).

0031-3203/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2013.10.010



Geometric features	Number of features	Description
Eyebrow	4x2	Two extreme corners, upper and lower mid points
Eyes	4x2	Two extreme corners, upper and lower mid points
Nose	3	Two nostrils and nose tip
Lip	4	Two extreme corners, upper and lower mid points

Fig. 1. Facial points of the frontal image.

contours alone can convey adequate information to recognize various expressions on the face.

The goal of this work is to introduce a completely automatic method of facial expression recognition using geometric facial features alone. The features extracted from the region of the eyes, eyebrows, lips, etc. play a significant role in providing sufficient information to recognize the presence of any of those six basic expressions. To remove planar head motion effects and scaling issues in subsequent image frames, all the feature parameters are calculated as the ratio of current values to those of the reference frame. This includes methodologies for detection of different facial features, such as eyebrow contours, state of eyes, lip contour and key points detection for each of the features. We also introduce methodologies to make the features rotation and illumination invariant. In order to come up with very accurate facial expression recognition results, a good classifier is extremely desirable. We propose a classification method using Kohonen Self-Organizing Map (KSOM) [31,7] to classify the features data into six basic facial expressions. KSOM has an extra-ordinary ability to arrange the data in an order that maintains the topology of the input data. The features data are first clustered using KSOM, then the cluster centers are used to train the data for recognition of the basic different emotions. To evaluate the performance of the proposed classification method, we compare the proposed approach with three widely used classifiers: radial basis function network (RBFN), 3 layered multilayer perceptron (MLP3) and support vector machine (SVM).

The rest of the paper is organized as follows. Section 3 presents segmentation and key features extraction techniques of the most important geometric features. Section 4 describes the architecture of SOM and the methodologies involved in applying 26 dimensional data to the SOM network for clustering the features data into basic six emotion zones. The section is followed by system identification using self-organizing map that creates a model by solving least square error of a supervised training system. Experimental results are given in Section 5 and finally in Section 6, conclusions are drawn.

2. Related works

Facial expression analysis approaches can be broadly classified into three basic stages: face detection, facial features extraction, facial expression classification. For decades, researchers are working on human facial expression analysis and features extraction. Substantial efforts were made during this period [26,27,30,23]. Major challenge was the automatic detection of facial features. Representation of visual information in order to reveal the subtle movement of facial muscles due to changes in expression is one of the vital issues. Several attempts were made to represent the visual informations accurately. Some of them are: optical flow analysis [14], local binary patterns (LBPs) [22], level set [23], active appearance model (AAM) [14], geometric analysis of facial features [32]. The major drawback with model based methods like AAMs and ASM is that they need prior information about the shape features. Generally, during the training phase of AAM and ASM, the shape features are marked manually [11]. Moore et al. found appearance based features by dividing the face image into sub-blocks. They used LBPs and variations of LBPs as texture descriptors [17]. Gu et al. [5] used contours of the face and its components with a radial encoding strategy to recognize facial expansions. They applied self-organizing map to check the homogeneity of the encoded contours. Kobayashi and Hara [8] modeled local facial features using geometric facial points. Zang et al. [32] used geometric components of facial points along with multi-scale and multi-orientation Gabor wavelet coefficients computed from every pixel of facial images.

Many techniques have been proposed for classification of facial expressions, such as multilayer perceptron (MLP) [33], radial basis function network (RBFN) [21,13], support vector machine (SVM) [1] and rule based classifiers [27].

3. Automatic facial features extraction techniques

The first and most crucial aspect of automatic facial expression recognition is the accurate detection of the face and prominent facial features, such as eyes, nose, eyebrows and lips. We present an analytical model shown in Fig. 1, consisting of 23 facial points which can describe all six basic facial expressions in frontal face images. The details of the 23 facial points are given in Fig. 1. We extract 26 dimensional geometric facial features using the concept of the analytical face model. The 26 dimensional geometric features are consisting of displacement of 8 eyebrow points, 4 lip points along x- and y-direction and projection ratios of two eyes. The displacement or movement of facial features is calculated using the neutral expression as reference where nose tip also plays the role in calculating the features displacement. Explanation of this part is given in Section 3.6.

Please cite this article as: A. Majumder, et al., Emotion recognition from geometric facial features using self-organizing map, Pattern Recognition (2013), http://dx.doi.org/10.1016/j.patcog.2013.10.010