Pattern Recognition Letters 38 (2014) 26–33

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

Feature selection for improved 3D facial expression recognition

Kamil Yurtkan [a,*], Hasan Demirel [b]

[a] Faculty of Engineering, Department of Computer Engineering, Cyprus International University, Nicosia, Northern Cyprus, Mersin 10, Turkey
[b] Electrical and Electronic Engineering Department, Eastern Mediterranean University, Gazimağusa, Mersin 10, Turkey

ARTICLE INFO

Article history:
Received 22 October 2012
Available online 8 November 2013

Communicated by M.S. Nixon

Keywords:
Facial expression analysis
Facial expression recognition
Facial feature selection
Faces biometrics

ABSTRACT

Automatic recognition of facial movements and expressions with high recognition rates is essential for human computer interaction. In this paper, we propose a feature selection procedure for improved facial expression recognition utilizing 3-Dimensional (3D) geometrical facial feature point positions. The proposed method classifies expressions in six basic emotional categories which are anger, disgust, fear, happiness, sadness and surprise. The most discriminative features are selected by the proposed method based on entropy changes during expression deformations of the face. Developed system uses Support Vector Machine (SVM) classifier organized in two levels. The system performance is evaluated on 3D facial expression database, BU-3DFE. The experimental results on classification performance are superior or comparable with the results of the recent methods available in the literature.

1. Introduction

Facial image processing has attracted a lot of interest with the recent improvements in computer graphics and digital image processing. Human face contains most of the information about the feelings of a human; hence human–computer interaction highly depends on accurate facial analysis. Therefore, the desire to automatically extract this information has been continuously increasing.

The earlier studies of facial expressions were pioneered by Ekman and Friesen (1976) in 1970s. Ekman's studies were about the classification of the human facial expressions into seven basic classes which are anger, disgust, fear, happiness, sadness, surprise and neutral. Later, Ekman and Friesen have proposed the Facial Action Coding System (FACS) to code the facial expressions using the facial movements called action units (AUs) Ekman and Friesen, 1978. In this study, we have analyzed the facial expressions for the six basic classes proposed by Ekman. The improvements in the field of computer graphics and imaging in late 1990s have let facial expression analysis and synthesis to be considered as an important topic in multimedia systems. In 1999, MPEG-4 standard introduced a face model with 83 feature points defined, Facial Definition Parameters (FDPs), describing a face in its neutral state. MPEG-4 standard also defined 68 Facial Animation Parameters (FAPs) which are used to animate the face by the movements of the feature points. FAPs can be used to animate the faces and to synthesize basic facial expressions (Yurtkan and Demirel, 2010). Besides, FAPs can be used for facial expression representation on a generic face model. MPEG-4 FAPs are widely used in most of the research labs for facial expression synthesis and analysis studies (Abrantes and Pereira, 1999).

Recent research activities in facial expression recognition are focused on automatic facial expression recognition and achieved acceptable recognition performances under controlled environments. Unfortunately, most of the systems developed have several limitations. Classification accuracy of an automatic facial expression recognition system is limited by the pose, lighting conditions, resolution and orientation of the face image. These limitations make automatic recognition of facial expressions an important research topic. Additionally, most of the systems in the literature recognize some expressions with high rates, and some others with low rates resulting in an acceptable average recognition rate. Our study also focuses on this problem and the proposed method safeguards high recognition rates for all expressions recognized.

The researchers working in automatic facial expression recognition have been working to improve the recognition rates. There has been developments in feature extraction, feature selection and classifier design. Chen and Huang proposed a clustering-based approach for facial expression recognition (Chen and Huang, 2003). Wang et al. (2006) employed LDA based classifier system and achieved 83.6% overall recognition rate on the BU-3DFE database. Lyons et al. (1999) achieved 80% average recognition rate using 2D appearance feature based Gabor-Wavelet (GW) approach. Jiangang Yu and Bhanu improved evolutionary feature synthesis for facial expression recognition (Yu and Bhanu, 2006). In late 2011, Zhang and Tjondronegoro proposed a facial expression recognition system using features extracted from facial movements (Zhang and

* Faculty of Engineering, Department of Computer Engineering, Cyprus International University, Nicosia, Northern Cyprus, Mersin 10, Turkey. Tel.: +0392 6711111 2434.

E-mail address: kyurtkan@ciu.edu.tr (K. Yurtkan).

0167-8655/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.patrec.2013.10.026

Tjondronegoro, 2011). Soyel et al. (2011) proposed NSGA-II based feature selection algorithm and achieved 88.3% overall expression recognition rate using 3D feature distances on BU-3DFE database.

In this paper, we focus on person independent facial expression recognition. The developed system considers 83 3-Dimensional (3D) geometrical feature point positions on the face provided in BU-3DFE database. Instead of using all of the 83 facial feature points, a novel feature selection procedure is proposed which selects the most discriminative feature points. The feature selection process employs information content metric, which is the entropy. Higher entropy refers to the feature point which is affected significantly from the expression deformation. Hence, high entropy feature points contain most of the information on the expression. Therefore, features are selected with the assumption that the level of entropy is correlated with the level of information that respective feature point carries.

The origin of the entropy concept arises from the studies of Ludwig Boltzmann in 1877. Entropy has been given a probabilistic interpretation in information theory by Claude Shannon in 1948. In the 1960s, Henri Theil developed applications of information theory in economics collected in Economics and Information Theory (1967) and Statistical Decomposition Analysis (1972) based on entropy (Frenken, 2007). In our study, entropy is used as a metric of information content for feature selection procedure.

The main novelty of the paper is the proposed feature selection procedure that detects the most discriminative facial feature points in 3D space contributing to facial expressions. This goal is achieved by the use of entropy on 3D geometrical feature positions that extracts the most discriminative features in 3D. Another contribution of the paper is the implementation of the classifier in two levels to perform a coarse-to-fine classification. Feature selection algorithm is applied to each level separately, and the system selects different feature points for the different modules of the classification process.

The paper is organized as follows. The face representation is described in Section 2. Section 3 explains the details of the classifier system. Then, Section 4 describes the proposed novel feature selection procedure applied to 3D geometrical facial feature point positions for improved facial expression classification. The performance analysis of the proposed system and the discussions are reported on BU-3DFE database in Section 5.

2. Face representation

Representing a face from the face image is of utmost importance for accurate facial expression recognition. On the other hand, selection of the most discriminative facial features is a vital step for correct classification. There are appearance based or facial geometry based representations in the literature. The conventional methods for facial expression recognition focus on the extraction of data needed to describe the changes on the face. A number of techniques were successfully developed using 2D static images (Shan et al., 2009). In these studies, the face is considered as a 2D pattern with certain textures so that expression variations can be observed. However, facial features that affect changes on the face are mostly in 3D space rather than 2D surface. Also, many expressions include skin wrinkles, for example, forehead deformations. Due to the limitations in describing facial surface deformations in 2D, there is a need for 3D space features in order to represent 3D motions of the face successfully (Yin et al., 2006). In this context, we employed 3D geometrical feature point data from BU-3DFE database in our study.

The BU-3DFE database consists of 100 individuals with 6 basic prototypic expressions and the neutral expression. All expressions contain 4 different intensities, 1 being the lowest and 4 being the highest intensity for the corresponding expression. The aim is to model spontaneous facial expressions. The database includes facial shape models, frontal view textures and 83 3D geometrical feature point positions for each subject. The 3D pose of the face affects geometrical feature extraction process, so obtained facial models inherently contain varying poses. The feature detection algorithm used in the creation of BU-3DFE database already incorporates some of the corruptions that can be introduced by possible movements of the head, including rotations. Model projections with respect to the frontal projection plane are open to corruptions on some of the 3D feature point positions, and those corruptions are already embedded in the available data. Therefore, we employed 83 3D feature point positions from BU-3DFE database that they reflect facial behavior of real life application and can represent a face with high accuracy in 3D.

Consider a 3D facial feature point consisting of three vertices as given in Eq. (1). By using facial feature positions, each face is represented by a face vector, FV. This face vector is obtained from the ordered arrangement of 3D feature point vertices (x, y and z for each point) and is created for each expression of the subject. Eq. (2) shows how a face is represented as a vector of 3D feature positions. In total, excluding the neutral expression, we have 100 individuals with 6 expressions and 4 intensities. It makes 2400 face vectors. Face vectors are then combined into a matrix, face matrix, FM, and recognition tests are performed using this matrix, shown in Eq. (3), where n denotes number of face vectors which is equal to 600. Training and test sets for the classifier are derived from the subdivisions of FM into two parts. Detailed explanations about training and test phases are given in Sections 3 and 4.

$$V_i = [\,V_{i_x} \quad V_{i_y} \quad V_{i_z}\,] \tag{1}$$

$$FV_j = [\,V_1 \quad V_2 \quad V_3 \ldots V_k\,] \tag{2}$$

$$FM = \begin{bmatrix} FV_1 \\ FV_2 \\ \ldots \\ FV_n \end{bmatrix} \tag{3}$$

3. Support Vector Machine classifier

After representing a face with facial feature positions, facial expression recognition can be considered as a vector classification problem. The system employs Support Vector Machine (SVM) which is a well known classifier for the vector classification problems (Shan et al., 2009).

Classifier system includes 15 2-class SVM classifiers in total. Classifiers are designed in two classes including all the combinations of six expression classes with two, forming a classifier module as shown in Fig. 1 Level 1. The classifiers with expression couples are anger–disgust, anger–fear, anger–happiness, anger–sadness, anger–surprise, disgust–fear, disgust–happiness, disgust–sadness, disgust–surprise, fear–happiness, fear–sadness, fear–surprise, happiness–sadness, happiness–surprise and sadness–surprise. Each SVM classifier employs a linear kernel function (dot product) that maps the training data into kernel space. Penalty coefficient C used for each SVM classifier is 1.

Our initial experiments were performed on BU-3DFE database with the classifier module depicted in Fig. 1 Level 1, using all 83 available feature point data. Table 2 shows the recognition results for our initial experiments. 80% average recognition rate was achieved with this setup after applying 8-fold cross validation.

In order to create a discriminative expression space, we adapt SVM classifier into two-level classification process. This proposal is motivated by our preliminary results showing confusions