Available online at www.sciencedirect.com
SCIENCE DIRECT
Pattern Recognition 39 (2006) 1795 – 1798
PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/patcog

Rapid and Brief Communication

Recognizing facial action units using independent component analysis and support vector machine

Chao-Fa Chuang, Frank Y. Shih*

College of Computing Sciences, New Jersey Institute of Technology, Newark, NJ 07102, USA

Received 31 October 2005; accepted 21 March 2006

Abstract

Facial expression provides a crucial behavioral measure for studies of human emotion, cognitive processes, and social interaction. In this paper, we focus on recognizing facial action units (AUs), which represent the subtle change of facial expressions. We adopt ICA (independent component analysis) as the feature extraction and representation method and SVM (support vector machine) as the pattern classifier. By comparing with three existing systems, such as Tian, Donato, and Bazzo, our proposed system can achieve the highest recognition rates. Furthermore, the proposed system is fast since it takes only 1.8 ms for classifying a test image.

Keywords: Facial expression recognition; Action unit; Independent component analysis; Support vector machine

1. Introduction

Facial expression plays a principal role in human interaction and communication since it contains critical information regarding emotion analysis. Its applications include human–computer interface, human emotion analysis, and medical care and cure. The task of automatically recognizing different facial expressions in human–computer environment is significant and challenging. In order to facilitate this research, Kanade et al. [1] established a comprehensive, heterogeneous database, named Cohn–Kanade expression database for classifying the upper or lower face action units (AUs).

Tian et al. [2] developed an automatic face analysis (AFA) system to analyze individual AUs based on both permanent and transient facial features in frontal face image sequences. Their recognition rate is 95.6% on Cohn–Kanade expression database. Donato et al. [3] used different techniques for classifying six upper and lower facial AUs on Ekman–Hager facial action exemplars. They found that the best performance is achieved by adopting Gabor wavelet decomposition. Their recognition rate is 96.9%. Bazzo and Lamar [4] invented a pre-processing step based on the neutral face average difference. Their system used a neural-network-based classifier combined with Gabor wavelet and received the recognition rates of 86.55% and 81.63%, respectively, for the upper and the lower faces.

2. Facial action coding system and expression database

In 1978, Ekman and Friesen [5] designed the facial action coding system (FACS) for characterizing facial expressions by AUs. This system is a human observed system developed to explain the subtle changes of facial expressions. There are totally 44 AUs. Among them, 30 are related to facial muscle contraction including 12 for upper faces and 18 for lower faces. For example, AU 1 is related to frontalis and pars medialis describing the inner corner of eyebrow raised, and AU 27 is related to pterygoids and digastric depicting mouth stretched open. The remainder of AUs is attributed to miscellaneous actions. For example, AU 21 portrays the status of neck tighten.

The AUs can exist individually or in combinations, which have additive or non-additive effects. Additive combination

* Corresponding author.
E-mail address: shih@cis.njit.edu (F.Y. Shih).

0031-3203/$30.00 © 2006 Published by Elsevier Ltd on behalf of Pattern Recognition Society.
doi:10.1016/j.patcog.2006.03.017

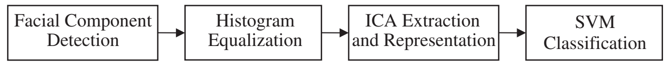

Fig. 1. The overall system.

means the combination does not alter the appearance of comprised AUs. An example is AU 12 + AU 25 indicating smile with mouth opened. Non-additive combination means the appearance of comprised AUs is modified. It represents difficulty and complication for the recognition task. An example is AU 12 + AU 15 indicating that the lip corner of AU 12 is changed by the downward motion of AU 15.

The facial expression image database used in our experiment is the Cohn–Kanade AU-Coded Face Expression Image Database [1]. This database is a representative, comprehensive and robust test-bed for comparative studies of facial expression. It contains image sequences of 210 adult subjects ranging from ages of 18–50. For gender classification, there are 69% females and 31% males. For racial classification, there are 81% Euro-American, 13% Afro-American, and 6% other groups. Lighting conditions and context are relatively uniform. The image sequences also include in-plane and out-of-plane head motion from small to mild. The image resolution is 640 × 480 pixels for 8-bit grayscale and 640 × 490 pixels for 24-bit color images.

3. The proposed system

In our proposed system, we utilize histogram equalization for lighting normalization, independent component analysis (ICA) for feature extraction and representation, and support vector machine (SVM) for classification measure.

3.1. Independent component analysis

Independent component analysis (ICA) is a statistical and computational technique for finding the hidden factors that are representative and favorable for separating different sets of images, sounds, telecommunication channels, or signals. ICA was originally designed to process the cocktail-party problem and belongs to a class of *blind source separation* (BSS) methods for separating data into underlying representative components. ICA is a general-purpose statistical and unsupervised technique where the observed random vectors are linearly transformed into components that are minimally dependent upon each other. The concept of ICA is an extension from the principal component analysis (PCA), which can only impose independence up to the second order and consequently define the directions that are orthogonal.

3.2. Support vector machine

Support vector machines (SVMs), introduced by Vapnik, are learning systems that separate sets of input pattern vectors into two classes using an optimal hyperplane. The set of vectors is said to be optimally separated by the hyperplane if it is separated without an error and the distance between the closest vector and the hyperplane is maximal. SVMs produce a pattern classifier by (1) applying a variety of kernel functions (e.g., linear, polynomial, and radial basis function (RBF)) as the possible sets of approximating functions, (2) optimizing the dual quadratic programming problem, and (3) using structural risk minimization as the inductive principle, as opposed to classical statistical algorithms that maximize the absolute value of an error or its square. Different types of SVM classifiers are used according to the type of input patterns. A linear maximal margin classifier is used for linearly separable classes, a linear soft margin classifier is used for linearly non-separable classes, and a nonlinear classifier is used for overlapping classes.

3.3. Facial expression processing and analysis

Our automatic facial expression processing and analysis system includes face detection, facial component extraction and representation, and facial expression recognition. We apply our previously developed algorithm [6] to automatically detect face regions in still images. Facial component extraction and representation are targeted to extract the most representative information derived from facial expression changes to represent the original detected faces. The advantages are to reduce the dimensionality of the detected faces from the previous stage and to speed up the computation in the next stage.

Facial expression recognition is intended to identify different facial expressions accurately and promptly. The facial expressions to be recognized can be categorized into two types. The first type is emotion-specified expressions, such as happy, angry, and surprise; the second type is facial action. In this paper, we focus on the recognition of facial AUs.

The proposed system is outlined in Fig. 1. The first step is to divide the detected face into upper and lower parts. Histogram equalization is then applied to normalize lighting effect. The ICA is used to extract and represent the subtle changes of facial expressions, and the linear SVM is adopted to recognize the individual AUs and their combinations.

4. Experimental results

There are four experiments conducted in this paper. The first experiment is intended to recognize six individuals and their combinations of the upper face AUs including AU 4, AU 6, AU 1 + AU 2, AU 1 + AU 4, AU 4 + AU 7, and AU