Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
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

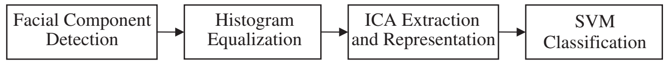

Fig. 1. The overall system.

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

Table 1
Recognition rates for the upper part of faces

Patterns	Images	Number of samples	Recognition rate (%)
AU4		12	$19/20 = 95$
AU6		48	$50/50 = 100$
AU1 + AU2		12	$19/20 = 95$
AU1 + AU4		9	$10/10 = 100$
AU4 + AU7		12	$17/20 = 85$
AU1 + AU2 + AU5		48	$50/50 = 100$
Total		141	$165/170 = 97.06$

Table 2
Recognition rates for the lower part of faces

Patterns	Images	Number of samples	Recognition rate (%)
AU17		12	$20/20 = 100$
AU9 + AU17		9	$10/10 = 100$
AU12 + AU25		39	$40/40 = 100$
AU15 + AU17		9	$10/10 = 100$
AU20 + AU25		12	$16/20 = 80$
AU25 + AU27		45	$40/40 = 100$
Total		126	$136/140 = 97.13$

Table 3
Recognition rates for the whole face

Patterns	Images	Number of samples	Recognition rate (%)
Neutral		48	100
AU4 + AU17		15	100
AU6 + AU12 + AU25		39	100
AU1 + AU2 + AU5 + AU25 + AU27		33	100
Total		135	100

1 + AU 2 + AU 5. The second experiment is constructed to classify six individuals and their combinations of the lower face AUs containing AU 17, AU 9 + AU 17, AU 12 + AU 25, AU 15 + AU 17, AU 20 + AU 25, and AU 25 + AU 27. The third experiment is designed to categorize four combinations of AUs on the whole face, which include neutral, AU 1 + AU 2 + AU 5 + AU 25 + AU 27, AU 6 + AU 12 + AU 25,

Table 4
Recognition rates between genders on action units

	Male (%)	Female (%)
Action units on the upper face	100	95.38
Action units on the lower face	100	93.33
Action units on the whole face	100	99.29

Table 5
Performance comparison of different systems

Systems	Database	AUs to be recognized	Correct rate (%)
Tian et al. [2]	Cohn–Kanade	AU 9, 10, 12, 15, 17, 20, 25, 26, 27, 23 + 24	95.6
Donato et al. [3]	Ekman–Hager	AU 1, 2, 4, 5, 6, 7	96.9
Bazzo and Lamar [4]	Cohn–Kanade	Upper AU 0, 6, 1 + 2, 4 + 7, 1 + 2 + 5, 4 + 6 + 7 + 9, 4 + 7 + 9	86.55
		Lower AU 0, 25, 26, 27, 12 + 25, 15 + 17, 20 + 25	81.63
Proposed system	Cohn–Kanade	Upper AU 4, 6, 1 + 2, 1 + 4, 4 + 7, 1 + 2 + 5	97.06
		Lower AU 17, 9 + 17, 12 + 25, 15 + 17, 20 + 25, 25 + 27	97.13
		Whole face AU neutral, 4 + 17, 6 + 12 + 25, 1 + 2 + 5 + 25 + 27	100

and AU 4 + AU 17. The fourth experiment is used to measure the effect of gender factor on the aforementioned three experiments.

There are 27 subjects randomly selected from the Cohn–Kanade AU-Coded Face Expression Image Database in terms of AUs. Among them, 20 are female and 7 are male. In each image sequence of 27 subjects, we select the first three and the last three images as the neutral and their corresponding AUs, respectively. Totally, there are 141 images for the upper part of AUs, 126 for the lower part, and 135 for the whole face. Since the number of images in the data set is limited, we randomly select 90% as the training set and the remaining as the test set. We repeat the same procedure 10 times and average the recognition rates of 10 independent tests for each experiment. Tables 1–3 list the recognition rates for different AUs. Table 4 compares the recognition rates between genders on AUs.

From Table 1, we observe that most misclassified patterns are derived from the combination of AU4 + AU7. This category at times is misclassified as AU6 because the combination of AU4 + AU7 makes the eyes narrow as close as to AU6. From Table 2, most misclassification stems from the combination of AU20+AU25. This group is sometimes recognized as the group of AU12 + AU25 since the difference between these two groups is simply the motion of lip corners. From Tables 1 and 2, there is no significant difference for our system to recognize the upper and the lower parts of faces. Their recognition rates are, respectively, 97.06% and 97.13%.

From Table 3, we observe that the recognition rate of 100% is achieved for classifying the AUs on the whole face. The combination of facial feature components can produce better results than individual components in facial expression recognition. We also explore the gender effect on the classification of AUs. From Table 4, we deduce that males are capable of expressing the AUs more accurately than females.

We compare our method with Tian et al. [2], Donato et al. [3], and Bazzo and Lamar [4], and the results are summarized in Table 5. Our proposed system receives the highest recognition rates: 97.06% on the upper part of faces, 97.13% on the lower part of faces, and 100% on the whole faces. Our system is implemented in Matlab on a Pentium IV with 2.80 GHz PC under window XP. It takes only 1.8 ms for classifying a test image.

5. Conclusions

In this paper, we present the recognition of facial action units, i.e., the subtle change of facial expressions. We deal with the comprehensive and heterogeneous subject database: Cohn–Kanade facial expression database. Our proposed system is demonstrated to be superior to three other systems. We deduce two arguments from experimental results. First, the combination of facial features is better than individual features in facial expression recognition. Second, males are able to express the action units more accurately than females.

References

[1] T. Kanade, J. Cohn, Y. Tian, Comprehensive database for facial expression analysis, Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, Grenoble, France, March 2000, pp. 46–53.
[2] Y.-L. Tian, T. Kanade, J.F. Cohn, Recognizing action units for facial expression analysis, IEEE Trans. Pattern Anal. Mach. Intell. 23 (2) (2001) 97–115.
[3] G. Donato, M.S. Bartlett, J.C. Hager, P. Ekman, T.J. Sejnowski, Classifying facial actions, IEEE Trans. Pattern Anal. Mach. Intell. 21 (10) (1999) 974–985.
[4] J.J. Bazzo, M.V. Lamar, Recognizing facial actions using Gabor wavelets with neutral face average difference, Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, Seoul, Korea, May 2004, pp. 505–510.
[5] P. Ekman, W.V. Friesen, Facial Action Coding System: A Technique for the Measurement of Facial Movement, Consulting Psychologists Press, San Francisco, 1978.
[6] F.Y. Shih, C.-F. Chuang, Automatic extraction of head and face boundaries and facial features, Inform. Sci. 158 (1) (2004) 117–130.