Active Shape Model of Combining Pca and Ica: Application to Facial Feature Extraction

DENG Lin, RAO Ni-ni, WANG Gang
School of Life Science and Technology, University of Electronic Science and Technology of China Chengdu 610054 China

Abstract Active Shape Model (ASM) is a powerful statistical tool to extract the facial features of a face image under frontal view. It mainly relies on Principle Component Analysis (PCA) to statistically model the variability in the training set of example shapes. Independent Component Analysis (ICA) has been proven to be more efficient to extract face features than PCA. In this paper, we combine the PCA and ICA by the consecutive strategy to form a novel ASM. Firstly, an initial model, which shows the global shape variability in the training set, is generated by the PCA-based ASM. And then, the final shape model, which contains more local characters, is established by the ICA-based ASM. Experimental results verify that the accuracy of facial feature extraction is statistically significantly improved by applying the ICA modes after the PCA modes.

Key words facial feature extraction; Active Shape Model (ASM); Principle Component Analysis (PCA); Independent Component Analysis (ICA)

Facial feature extraction is important for facial image analysis, which is an indispensable processing step between face detection and recognition. Extensive researches[1-7] have been performed over the past 20 years. Active contour model, proposed by Kass et al.[1], is an energy minimization approach for facial contour representation. It is too sensitive to its initialization. Henceforth, the deformable template is introduced to extract the contour of the eyes and mouth[2]. The template makes use of global information and improves the reliability of location the contour. However, it is computationally expensive and the convergence speed is comparatively slow. Active Shape Models (ASM) and Active Appearance Models (AAM), proposed by Cootes et al., are two popular shape and appearance models for face localization and face images interpretation[3-4]. In ASM, the modes of shape variation are generally extracted from the training set of example shapes using Principle Component Analysis (PCA). Whereas, PCA supposes that the training data is from a Gaussian distribution, which often is not the case. Independent Component Analysis (ICA) is an alternative shape decomposition method, which does not assume a normal distribution of the input data[5]. Furthermore, in PCA the objective is to find global shape variations, whereas in ICA the objective is to show localized shape variations.

Recently, ICA is used by Üzümçü et al. to construct statistical shape models from 2-D cardiac datasets[6] and to perform 2-D cardiac segmentation[7].

In this paper, we propose a new method of facial feature extraction, which combines the advantages of both PCA and ICA. The method not only represents global shape variations but also has local characters. We first briefly introduce ASM in Section 1. Next, PCA modes and ICA modes are described in Section 2. Our algorithm is presented in Section 3. Experimental results and conclusions are given finally.

1 Review of ASM

This method includes two processes: statistical shape models and local appearance models.

1.1 Statistical Shape Models

A shape model is described by a landmark points that represent the boundary, internal features, or even external features. Each training shape $x$ is represented as a vector,

$$x = (x_0, y_0, x_1, y_1, \ldots, x_k, y_k, \ldots, x_{n-1}, y_{n-1})^T$$

where $x_i$ represents the $i$th training shape, $(x_k, y_k)$ is the coordinate of the $k$th landmark point and $T$ represents the transpose operation.

The training shapes are all aligned by scaling, rotation and translation for minimizing the weighted
sum of the squared distances between their equivalent landmark points. Then principle component analysis is applied to aligned shape. Therefore, a shape model can be approximated as follows
\[ x \approx \bar{x} + Pb \] (2)
where \( \bar{x} \) is the mean shape, \( P \) is a set of orthogonal models of shape variation and \( b \) is a vector of shape parameters.

1.2 Local Appearance Models
The local appearance models describe local texture features around each landmark. For example, at test image location \( s \), when searching points, the local appearance models find the best candidate in the neighborhood of the search point by minimizing
\[ D_{s}(g_{s}) = D_{g g}(s) - C_{s}^{-1}(g_{s} - \bar{g}_{s}) \] (3)
The ASM search procedure is an iterative process. On each iteration it uses the local appearance models to find a new shape and then updates the model parameters to fit the new search shape best[3].

2 Methods
2.1 PCA Modes
The objective of PCA approach is to find the modes of shape variation that explain maximized amount of the variance in the training set. It is assumed that the data are drawn from multidimensional Gaussian distribution and hence completely determined by its covariance matrix, i.e. the eigenvectors \( p \) and eigenvalues \( \lambda \) of the covariance matrix. That is
\[ [G] = \frac{1}{N-1} \sum_{i=1}^{N} (x_{i} - \bar{x})(x_{i} - \bar{x})^{T} \] (4)
where \( N \) is the number of aligned shapes, \( \bar{x} \) is the mean shape, and \([G]\) is the covariance matrix of \( x_{i} \). \( P_{\text{PCA}} = (p_{1}, p_{2}, \cdots, p_{t}) \) is a matrix consisting of the first \( t \) eigenvectors. \( P_{\text{PCA}} \) can approximately explain maximal amount of the variance in the training set.

2.2 ICA Modes
The objective of ICA approach is to find such modes of shape variation that are statistically independent. It simulates the receptive field functions found in early stage of the human vision system[8]. Meanwhile, it assumes that everything we see is composed of some hidden and independent signals and their correlative basis functions. This concept can be illustrated as follows. Let \( X \) be a shape vector which is the linear combinations of statistically independent components \( S \).
\[ X = AS \] (5)
where \( A \) is an unknown mixing matrix and \( S \) is the source signals. Then, the basic problem of ICA is to calculate the source signals \( S \):
\[ S = WX \] (6)
The weight matrix \( W \) can be found by optimizing a cost function. In ICA, several different cost functions are available, such as FastICA, InfoMax and the JADE algorithm. Among these methods, JADE algorithm does not have a random initialization and also does not have any adjustable parameters, and therefore is more robust. In this work, we have selected this algorithm to perform ICA.

In classical PCA, the modes can be ordered according to the associated eigenvalues. In ICA however, such ordering can not obtained automatically. We can derive such information from the independent components[6].

3 The Proposed Algorithm
Facial feature extraction is a challenging issue because the human face is a three dimensional (3-D) object, and any changes in illumination and facial expression will affect the accuracy of extracting the facial features. ASM is an effective tool to extract the facial feature. It generally uses PCA to describe the main direction of shape variation in a training set. Recently, ICA has been proven to be more efficient to extract face features than PCA does. The main difference between these two methods is that the objective of PCA is to find the modes of shape variation that represent global variations, whereas in ICA, the obtained vectors describe local variations. Therefore, the combination of PCA and ICA will obtain a global-to-local and coarse-to-fine face represent.

In terms of the different optimization criterion in PCA and ICA, we adopted the consecutive strategy to combine PCA and ICA and established a novel ASM approach for facial feature extraction, that is, the PCA modes are done first, and then, the obtained results are refined using the ICA modes. The combined modes are
implemented as follows:

1) for the PCA-based ASM, the optimal weights \( b_{\text{PCA}} \in [-3, 3, -3] \) are searched and the new shape is computed. (See Eq.2):

\[
X_{\text{PCA}} = \bar{X} + P_{\text{PCA}} b_{\text{PCA}} \quad (7)
\]

2) the obtained shape \( X_{\text{PCA}} \) is used as an initialization, and the optimal weights \( b_{\text{ICA}} \) for the ICA-based ASM are searched. The final shape model is represented as:

\[
X_{\text{FINAL}} = X_{\text{PCA}} + P_{\text{ICA}} b_{\text{ICA}} \quad (8)
\]

4 Experiment Results

We have performed our approach for the facial images from the Olivetti Research Lab (ORL) face image database. The database contains 400 images with different illumination and facial expression, and each image is of size 92×112. In the experiment, both ASM and our algorithm are trained with the same 100 frontal-view face images, and 100 images are used for testing. On each image 53 landmarks are labeled.

For all test images, we calculate the overall average error of PCA-based ASM and our proposed approach, respectively. The overall average error is defined as follows:

\[
E = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{1}{n} \sum_{j=1}^{n} \text{Dis}(P_{ij}, P'_{ij}) \right) \quad (9)
\]

where \( N \) is the total number of test images, \( n \) is the total number of feature points, \( \text{Dis} \) represents Euclidean distance between each search shape and the labeled shape, \( P_{ij} \) represents a search feature point and \( P'_{ij} \) represents a corresponding labeled feature point. \( E \) can measure the results for face feature extraction: the smaller the value of \( E \) is, the better the results are. In Tab.1 below, we can see that ASM of combining PCA and ICA works better than PCA-based ASM.

| Tab.1 The results (the overall average error) for locating feature points |
|-----------------------------|-----------------|-----------------|
| The overall average error   | PCA  | PCA+ICA        |
| \( E \)                     | 3.80 | 2.10           |

The average runtime for the PCA-based ASM is about 0.2 s to 0.5 s with a 1.7 GHz Pentium IV computer. Our proposed algorithm, which requires about 1.7 s to 2.2 s, is computationally expensive. This is because the JADE algorithm is based on the joint diagonalization of the cumulant matrices and all of the matrices are diagonalized at once in the process of calculating ICA.

We also test the novel algorithm on other face images, which have significant variation in illumination and expression. The overall average errors of ASM and our proposed approach are 3.8 and 2.3, respectively. Six search results are shown in Fig. 2 and Fig. 3. These results show that the locations of the eyes and eyebrows achieved by our approach are more accurate than those achieved by ASM. Note that our proposed algorithm is more robust for the perspective variations and expression changes.

5 Conclusions

In this paper, we have presented a new ASM that combines PCA and ICA. The ICA modes effectively represent local texture features, thereby improving the locating accuracy and increasing robustness for low-quality images. Thus, the novel model can improve the extraction accuracy compared with the
conventional ASM. Experimental results demonstrate also the accuracy of facial features location by applying the ICA modes after the PCA modes is significantly improved. Meanwhile, the better face representation can be obtained under different perspective variations and facial expressions than the conventional ASM can. This method can also be applied to other fields where ASM is used. However, the effect of the different cost optimization functions on the independent component analysis is a challenging issue all the time. Because the obtained independent components using different cost optimization functions are highly similar, the locations of the main variation direction in the independent components seem to be equal. Therefore, future work must be done to develop a more efficient standard to select the cost optimization function.

References


Brief Introduction to Author(s)

DENG Lin (邓琳) was born in Jilin, China, in 1978. She received the B.S. degree in biomedical engineering from the University of Electronic Science and Technology of China (UESTC) in 2002. She is now pursuing M.S. degree with School of Life Science and Technology, UESTC. Her research interests include facial feature extraction and facial recognition.

RAO Ni-ni (饶妮妮) was born in Sichuan, China, in 1963. She is now a professor and doctoral advisor with School of Life Science and Technology, UESTC. Her current interests include security in telemedicine, signal processing and Bioinformatics.

WANG Gang (王刚) was born in Liaoning, China, in 1977. He received the B.S. degree in communication and information engineering from UESTC in 2000, and the M.S. degree in biomedical engineering from UESTC in 2005. He is now a Ph.D. candidate in biomedical engineering, UESTC. His research interests include signal processing and image processing.