

Local Independent Components Analysis-Based Facial Features Detection Method

M. Hassaballah^{1,2}, Tomonori KANAZAWA³, Shinobu IDO³, Shun IDO¹¹Department of Computer Science, Ehime University, 790-8577, Japan²Department of Mathematics, Faculty of Science, South Valley University, Qena, Egypt³eCompute Corporation, 2821-4, Matsuyama, Ehime 791-8042, Japan

E-mail: m.hassaballah@ic.cs.ehime-u.ac.jp, ido@cs.ehime-u.ac.jp

Abstract

目や鼻や口といった顔特徴の検出は、その後の顔画像の解析において重要なステップである。本稿では、グレースケールの画像から目と鼻を検出する方法を紹介する。独立成分分析法(ICA)を顔特徴の外観や形状の学習に用いることで、ICA基底ベクトルに対して高応答の領域を顔特徴として検出した。一層の性能改善のためにICAに加えて目と鼻の局所特徴を適用し、異なるデータベースによる実験において有望な結果を得た。

1 Introduction

Facial feature detection plays an important role in many face-related applications including face recognition, face validation, and facial expression recognition [1]. These applications need to detect the facial features robustly and efficiently. However, this is not an easy machine-vision task at all. The difficulty comes from high inter-personal variation (e.g. gender, race), intra-personal changes (e.g. pose, expression), and from acquisition conditions (e.g. lighting, image resolution).

In spite of these difficulties, several methods have been proposed to detect certain facial features, especially eye as it is the most salient and stable feature of the human face. Ryu and Oh [2] introduced a method based on eigenfeatures derived from the eigenvalues and eigenvectors of the binary edge data set and neural networks to detect eye region. The eigenfeatures extracted from the positive and negative training samples of eye are used to train a multilayer perceptron (MLP) whose output indicates the degree to which a particular

image patch contains an eye within itself. An ensemble network consisting of a multitude of independent MLPs was used to enhance the general performance of a single MLP. It was tested on 180 images without glasses from ORL database and its performance was 91.7% and 85.5% for left and right eye respectively. The advantage is that it does not need a large training set by taking advantage of eigenfeatures and sliding window. Kroon *et al.* [3] address the problem of eye detection for the purpose of face matching in low and standard definition image and video content. They present a probabilistic eye detection method based on well-known multi-scale local binary patterns (LBPs). The detailed survey of the most recent eye detection methods [4] concluded that the development of a general eye detection technique involves addressing many challenges, requires further theoretical developments, and is consequently of interest to many other domains problems in computer vision and beyond.

On the other hand, there are few methods that address nose detection problem, even though it is not less important than eye or other facial features. It does not affect so much by facial expressions and in several cases is the only facial feature which is clearly visible during the head motion. Most of the existing approaches detect nose depending basically on the prior detection of eye center and considering nose can be located within certain pixels below the line connecting the centers of two eyes [5]. Other approaches detect nose depending on the reflection of light on the nose tip or using projection methods [6]. However, all existing projection methods do not consider complex conditions such as illumination

and pose variation, thus they will fail under these imaging conditions. In [7] a method for facial landmarks detection based on extracting oriented edges and constructing edge maps at two resolution levels is proposed. It can achieve average nose detection of 78% on 330 images from Pictures of Facial Affect database. The method was not fully automatic and required manual classification of the located edge regions.

In spite of considerable amount of previous work on the subject, detection of facial features will remain a challenging problem because the shape and texture of facial features vary widely under changing expression, head pose and illumination. Therefore, new robust methods are required. In this paper, we present a method to detect facial features (e.g. eye and nose). The method depends on higher-order statistics and second-order moments of the training data decorrelated by ICA basis vectors as well as local characteristics of facial features. The novelty of the method is to apply ICA on parts of face containing facial features rather than whole face image as done in face recognition field. This new use of ICA might capture the advantages of spatially localized of its basis vectors which will help in highlighting salient regions in the data and provide a better probabilistic model of the training data. It was evaluated on different databases under various imaging conditions.

The remainder of this paper is organized as follows. A brief background about ICA is introduced in Section 2. The outline of the proposed method for detection of facial features is presented in Section 3. Experimental results are reported in Section 4, the conclusions are given in Section 5.

2 Independent component analysis

Independent component analysis can be seen as an unsupervised learning method based on higher order statistics for extracting independent sources from an observed mixture, where neither the mixing matrix nor the distribution of the sources are known. Formally, the classical model of ICA is given as

$$X = AS \quad (1)$$

where X denotes the observation matrix containing the

‘linear mixtures’ in its rows $X = [X_1, X_2, \dots, X_m]^T$, A is a nonsingular mixing matrix, and S denotes the source matrix containing statistically independent source vectors in its rows $S = [S_1, S_2, \dots, S_m]^T$. Under the assumptions that the sources are statistically independent and non-Gaussian (at most one of them can have Gaussian distribution), an un-mixing matrix W can be found by maximizing some measure of independence. In other words, the estimated separation matrix, W under ideal conditions, is the inverse of the mixing matrix A .

$$Y = WX, \quad W = A^{-1}, \quad \text{and} \quad Y \approx S \quad (2)$$

Generally, ICA aims to estimate the un-mixing matrix W and thus to recover each hidden source using $S_k = W_k X$, where W_k is the k th row of W without any prior knowledge. There are a number of algorithms to estimate the un-mixing matrix W in the ICA model (2). The FastICA algorithm [8] is used in this work to estimate the ICA basis vectors.

The FastICA is basically based on a fixed-point iteration scheme for finding a maximum of the nongaussianity of $w^T X$, where the nongaussianity is measured by the maximization of the so-called *negentropy*. According to [8] and after a series of mathematical calculations, the ICs can be estimated by

$$w_i \leftarrow E\{z g(w_i^T z)\} - E\{g'(w_i^T z)\} w_i \quad (3)$$

where w_i^T is a row vector in W , z is the preprocessing whitening data and the nonlinearity g can be any smooth function that gives robust approximations of negentropy such as,

$$g(y) = \tanh(\alpha y), \quad 1 \leq \alpha \leq 2 \quad (4)$$

The estimation process in (3) is repeated for a chosen n (number of basis vectors) with orthogonalize the matrix $W = (w_1, w_2, \dots, w_n)^T$ using $W \leftarrow (WW^T)^{-1/2} W$, where

$$(WW^T)^{-1/2} = E \text{dig}(d_1^{-1/2}, \dots, d_n^{-1/2}) E^T \quad (5)$$

The convergence will occur when the absolute value of the dot-product of old and new W is (almost) equal 1, i.e. $\langle W_t, W_{t+1} \rangle = 1$. More information and details about ICA and FastICA can be found in [8].

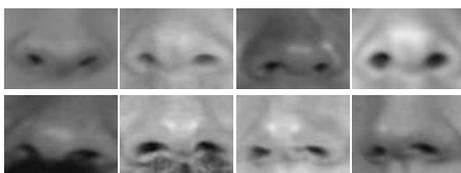
3 Method Overview

3.1 Computing ICA features subspaces

A set of 200 images for eye and nose is used in the training step. In order to minimize the impact of noise or variations in illumination each image is preprocessed with norm normalization. The width w and height h of the training images are determined experimentally with respect to the face width W_F . Examples of eye and nose training images are shown in Fig. (1). Then the FastICA algorithm is implemented to estimate the ICA basis vectors which will form the feature subspace. The number of basis vectors for each subspace is optimized experimentally. Hence, 20 vectors are found sufficient for eye subspace, while 30 vectors are sufficient for nose subspace. The eye and nose subspaces are shown in Fig. (2) and (3) respectively. It is clear that the computed basis vectors of each subspace are spatially localized to the salient regions of eye and nose such as iris, eyelids, nostrils, textures, and edges. This localization might robust the performance to local noise and distortion.



(a) Eye training images 60 x 30 pixels



(b) Nose training images 70 x 50 pixels

Fig. 1: Facial Features training images

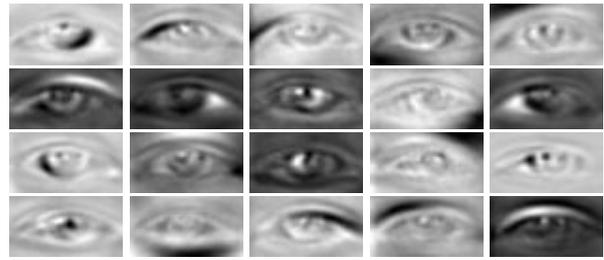


Fig. 2: Eye subspace; the highest 20 ICA basis images

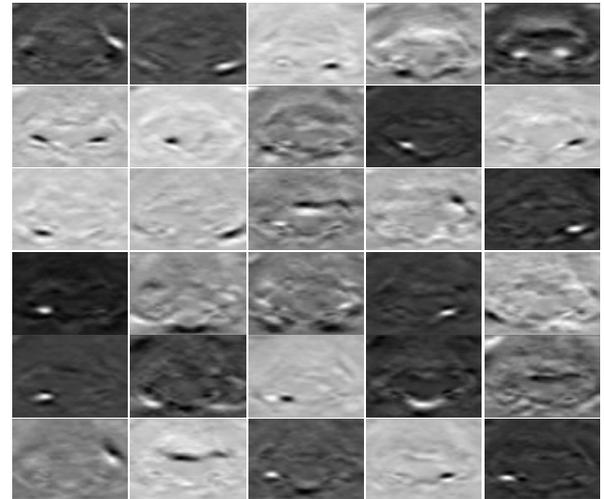


Fig. 3: Nose subspace; the highest 30 ICA basis images

3.2 Detection of facial features

Searching for facial features in the image can be done either directly in the entire image, or rely on the output of a face detector indicating that these features are present in the image. Unfortunately, searching in the whole image is not a suitable method for real time implementations and is more prone to errors. Therefore, a face detector [9] is applied first to locate the face, then searching for regions that contain eye and nose is done in the located face. The located face is normalized to a fixed size of $W_F \times W_F$ pixels. The normalized face is scanned to find the region which might contain eye/nose, the regions data are treated as $w \times h$ dimensional vectors. Each region (vector) is locally preprocessed using norm normalization as in the training step. The subspaces corresponding to the feature point vector in the $w \times h$ dimensional facial feature subspace can be expressed as linear subspaces spanned by multiple ICA vectors. The projection angle θ of an input vector projected onto the ICA subspace represents the extent to which the input vector is analogous to the feature point vector. For verification the value of θ , specifically $\text{Cos}(\theta)$ between

the input vector and each feature point's subspace (ICA basis vectors) is obtained. This angle measure is invariant under linear changes in the contrast of the image and furthermore the cosine similarity measure was previously found to be effective for face processing. The eye or nose region is the input vector that falls into its subspace with the smallest θ or highest similarity. The similarity $S_{\theta_j}(R_j, ICAVectors)$ between the region R_j , $j=1,2, \dots, M$ (number of regions) and ICA basis vectors is calculated using $Cos(\theta)$ of the projection component

$$S_{\theta_j} = Cos^2 \theta_j = \frac{\sum_{i=1}^n \langle V, BaseVector_i \rangle^2}{\|V\|^2} \quad (6)$$

where n is the number of basis vectors that form the feature subspace and $\langle V, BaseVector_i \rangle$ is the inner product between input vector (representing j -th candidate region) and i -th base vector of the ICA subspace. The region R_k with the highest similarity is declared as the target facial feature region.

3.3 Local characteristics facial features

The performance of the method described in section 3.2 can be improved significantly if local characteristics of eye and nose are used besides ICA method. For eye, a variance filter is used. To construct this eye variance filter a set of 30 eye images of left and right eye of same size (i.e. 60 x 30 pixels) as the eye training images is selected randomly of different persons. Each eye image is divided into 3 x 3 non-overlapped subblocks of size 20 x 10 pixels thus, each subblock has its unique features as shown in Fig. (4). The variance of each subblock is calculated, and then the eye variance filter F_e is constructed by calculating average of the variance in the subblock images over all 30 eye images

$$F_e = \frac{1}{N} \sum_{j=1}^N I_{\sigma}^j \quad (7)$$

where I_{σ}^j is the j -th variance filter of the σ subblock eye image and N is the number of used eye images

($N=30$ in this work). To detect the region that most probably contains eye using the proposed eye variance filter, variance vector of each candidate region of the face is calculated in the same manner as in the eye variance filter F_e . Then the correlation between the variance vector of each subblock of the region and the eye variance filter F_e (7) is calculated using this formula

$$R(I_{\sigma}, F_e) = \frac{E[(\xi_{I_{\sigma}} - E(\xi_{I_{\sigma}}))(\xi_{F_e} - E(\xi_{F_e}))]}{\sqrt{D(\xi_{I_{\sigma}}) D(\xi_{F_e})}} \quad (8)$$

where $\xi_{I_{\sigma}}$ and ξ_{F_e} are the concatenated vectors of variance of subblock I_{σ} of the eye candidate region and F_e respectively.

All the regions which have high response to the eye variance filter are selected to generate a small list of possible eye region pairs. If the response of the candidate region to F_e is greater than expected threshold (e.g. =4), this means that the candidate region might contain eye. After that, all the selected regions by applying eye variance filter are then verified using ICA method mentioned in section 3.2 to select only two regions which represent the left and right eye.

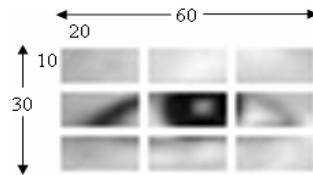


Fig. 4: Variance filter model

On the other hand, for nose three different local characteristics are used as follows:

(i) Similarity of both sides: The left and right sides of nose are similar in a front-view face as shown in Fig. (5a), this property of similarity can be measured using Euclidean distance between both sides.

(ii) Dark-White-Dark (DWD) property: Also, the lower part of nose region is characterized by two dark

nostrils and a light subregion due to the reflection of light on the nose as shown in Fig. (5b). This property can be identified by the average of gray intensity in each subregion, where the average in the two nostrils regions is less than the average of middle lighter subregion containing nose tip.

(iii) The variation in lower/upper parts property:

When the face is rotated some degrees these two properties are despaired and the only clear property is the variation between lower part and upper part as shown in Fig. (5c). This variation can be measured by the variance in each part. Based on this analysis, we search for a certain region among the ten highest regions detected by ICA method which satisfies the properties (i)-(iii). Note that in the case of eye detection ICA method has been applied before local characteristics while in nose detection the opposite has been done.

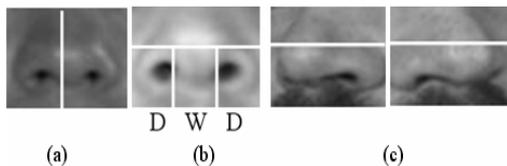


Fig. 5: The local characteristics of nose region

4 Experimental Results

In order to find optimum number of ICA basis vectors for both feature spaces the proposed method was evaluated on XM2VTS database as shown in Fig. (6). The highest successful detection rate of 93.2 % for eye is achieved when the number of ICA components is 20 while for nose 95.8% successful detection rate achieved at 30 vectors, and after that the marginal benefit slows. Also, there is trade-off between running time and ICA dimension. Using fewer basis vectors requires less storage and results in faster matching hence 20 and 30 vectors are sufficient. Therefore, the number of ICA basis vectors is set to the first highest 20 and 30 basis vectors for eye and nose subspaces respectively.

The proposed ICA method is also evaluated on other three databases namely; 1500 images of FERET, BioID, and JAFFE. In this context, the facial features are detected based on the highest response to feature space. The results of this experiment are reported in Table (1).

Because of noise in the images due to lighting condition or facial expression, the detected region is not accurate and thus the whole successful detection rate (i.e. images with correctly detected eye/nose region relative to the whole set of facial database) is not so high, especially on BioID and JAFFE databases. The best results have been obtained on the XM2VTS database because the image quality is very high, the illumination is always uniform, and the image scale is fixed.

The performance of ICA method is improved significantly when the local characteristics of eye and nose are used besides ICA as shown in Table (2). In some databases such as XM2VTS an improvement rate more than 5% has been achieved while in BioID database the improvement is not significant due to the severe conditions in this database.

Compared with other methods [2] achieved detection rate of 91.7% and 85.5% for left and right eye on 180 images without glasses from ORL database, the proposed method achieves an average eye detection rate of 94.3 % on about 4000 images while for nose it can achieve an average detection rate of 96 % outperforming [7] which achieved average detection rate of 78% on 330 images from Pictures of Facial Affect database. Example of successful results is given in Fig. (7).

The main advantages of the proposed method are; the method is very simple. The average execution time of the method on a PC with Core(TM)2 Duo CPU, 2.4 GHz, and 4GB RAM is less than 80 msec. Moreover, detection of nose does not depend on eye detection as in the existing methods. Future research will focus on detecting fiducial points of eye and nose such as eye corners, iris, nostrils or nose tip within the detected region as well as mouth corners.

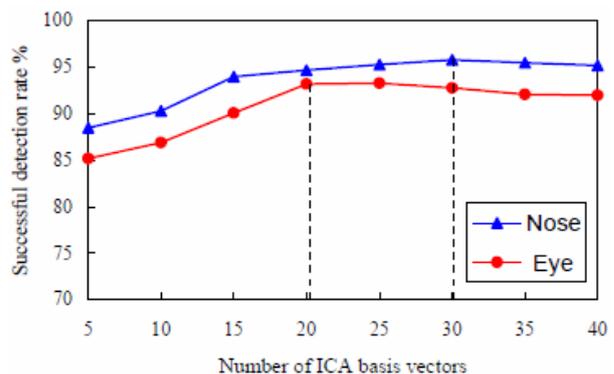


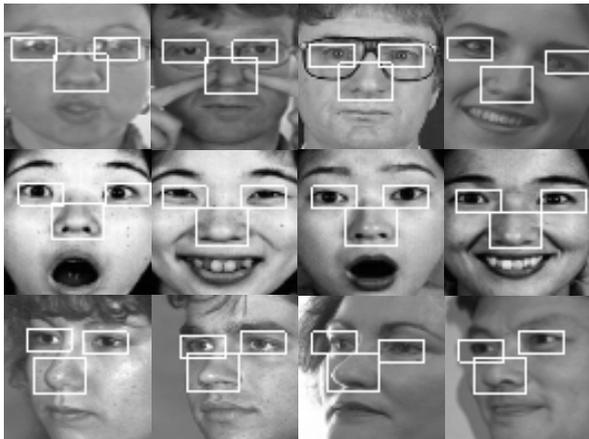
Fig. 6: Selection the optimum number of basis vectors

Table 1: Performance of ICA only on different databases

Feature	XM2VTS	FEERT	BioID	JAFFE
Eye	93.2%	91.3%	87%	85.3%
Nose	95.8%	94.3%	91.8%	91.5%

Table 2: Performance of ICA besides local characteristics

Feature	XM2VTS	FEERT	BioID	JAFFE
Eye	98.4%	97.1%	91.6%	90.2%
Nose	99%	96%	92.3%	96.2%

**Fig. 7:** Example of successful results

Conclusions

In this paper, an automatic method to detect eye and nose location from facial images based on the response of face regions to feature subspaces is presented. Eye and nose spaces are estimated using independent components analysis basis vectors. In order to further improve the performance of the method, we proposed a subregion-based framework that depends on local characteristics of facial features regions. The efficiency of the method is evaluated in different databases stressing variety of imaging conditions. It has been shown by experimental results that the proposed method can accurately detect the eye/nose with high detection rate comparable with existing methods.

Acknowledgments

The authors would like to acknowledge the financial support of the Egyptian government for this work under Grant No.1/13/101-2006-2007. The authors would also like to thank the owners of used databases XM2VTS, FERET, BioID, and JAFFE for experiments of this work.

Reference

- [1] Asteriadis S., Nikolaidis N., Pitas, I., Facial feature detection using distance vector fields, *Pattern Recognition*, 42, 1388-1398, 2009.
- [2] Ryu Y.-S., Oh S.-Y., Automatic extraction of eye and mouth fields from a face image using eigenfeatures and ensemble networks, *Applied Intelligence*, 17, pp. 171-185, 2002.
- [3] Kroon B., Maas S., Boughorbel S., Hanjalic A., Eye localization in low and standard definition content with application to face matching, *Computer Vision and Image Understanding*, 113 (8), 921-933, 2009.
- [4] Dan W. H., Qiang J., In the eye of the beholder: A survey of models for eyes and gaze, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 32(3), pp. 478-500, 2010.
- [5] Sankaran, P., Gundimada, S., Tompkins, R.C., Asari, V.K., Pose angle determination by faces, eyes and nose localization, *IEEE CVPR'05*, 161-167, 2005.
- [6] Campadelli, P., Lanzarotti, R., Fiducial point localization in color images of face foregrounds, *Image and Vision Computing*, 22, 863-872, 2004.
- [7] Gizatdinova Y., Surakka V., Feature-based detection of facial landmarks from neutral and expressive facial images, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 28(1), 135-139, 2006.
- [8] Hyvarinen A., Oja E., Independent component analysis: algorithms and applications, *Neural Networks*, 13, pp. 411-430, 2000.
- [9] Viola P., Jones J. M., Robust real-time face detection, *International Journal of Computer Vision*, 57(2), pp. 137-154, 2004.