13-1 MVA2011 IAPR Conference on Machine Vision Applications, June 13-15, 2011, Nara, JAPAN

Eye and Nose Fields Detection From Gray Scale Facial Images

M. Hassaballah[1,2], Kenji Murakami[1], and Shun Ido[1]
[1]Department of Computer Science, Ehime University, 790-8577, Japan
[2]Department of Mathematics, Faculty of Science, South Valley University, Qena, Egypt
Email: m.hassaballah@svu.edu.eg, murakami@cs.ehime-u.ac.jp, ido@cs.ehime-u.ac.jp

Abstract

Detection of facial features such as eye, nose, and mouth is an important step for many subsequent facial image analysis tasks such as face recognition. In this paper, we introduce a method to detect eye and nose fields from gray scale facial images. The Independent Components Analysis (ICA) is utilized to learn the appearance and shape of the facial feature. The regions with high response to the ICA basis vectors are chosen as the target facial features. For improving the performance further, the local characteristics of eye and nose are adapted besides ICA. Experiments on different databases show the promising results of the proposed method.

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 [2]. Ryu and Oh [3] 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. [4] 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). Besides, some approaches consider eye detection as a two-class or multi-class pattern classification problem [5]. The detailed survey [6] of the most recent eye detection methods 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 [7]. In these methods, any error in eye localization will affect on detection of nose, moreover, eye detection as a first step is not a trivial task. Other approaches detect nose depending on the reflection of light on the nose tip or using projection methods [8]. However, all existing projection methods do not consider complex conditions such as illumination and pose variation, thus they will fail under these imaging conditions. Using different approach, Gizatdinova and Surakka [9] introduced a feature based method in which the information on local oriented edges was utilized to compose edge maps of the image at several levels of resolution. The method 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.

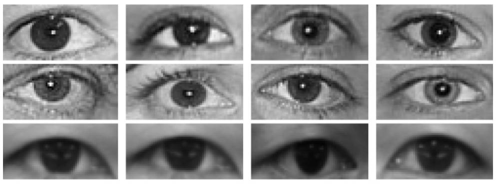
In spite of considerable amount of previous work on the subject, detection of facial features will remains a challenging problem because the shape and texture of facial features vary widely under changing expression, head pose and illumination. Therefore, new robust methods are required. In this paper, 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. 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 rest of the paper is organized as follows. The detailed description of the proposed facial features detection method is presented in Section 2. In Section 3, experimental results of the proposed method are discussed. Finally, the conclusion is given in Section 4.

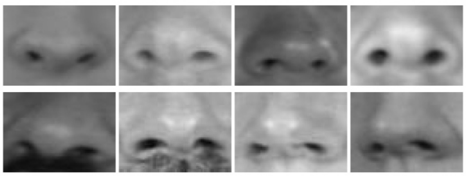
2 Details of the Proposed Method

2.1 Computation of ICA features spaces

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 [10] is implemented to estimate the ICA basis vectors which will form the feature space. The number of basis vectors for each space is optimized experimentally. Hence, 20 vectors are found sufficient for eye space, while 30 vectors are sufficient for nose space. The eye and nose spaces 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

Figure 1. Facial features training data samples.

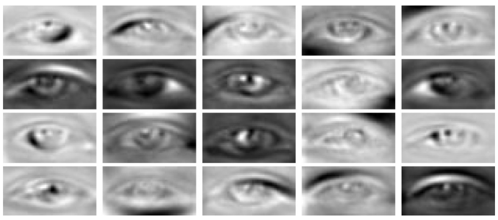


Figure 2. Eye space; the highest 20 ICA basis vectors.

2.2 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 [11] 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 x W_F pixels. The normalized face is scanned to find the region which might contain eye/nose, the regions data are treated as w x 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 x h dimensional facial feature space 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 Cos(θ) between the input vector and each feature point's space (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 space with the smallest θ or highest similarity. The similarity $S_{\theta_j}(R_j, ICAvectors)$ between the region R_j, $j = 1, 2, ...M$ (number of regions) and ICA basis vectors is calculated using Cos(θ) of the projection component

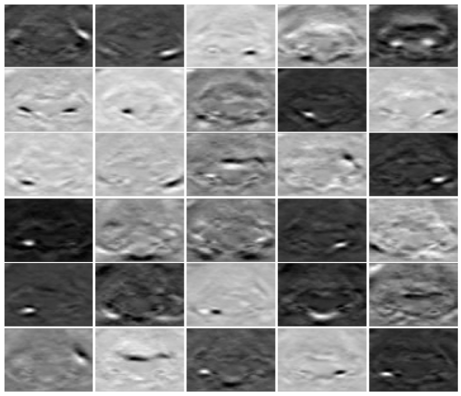


Figure 3. Nose space; the highest 30 ICA basis vectors.

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

where n is the number of basis vectors that form the feature space 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 S_{θ_k} is declared as the target facial feature region.

2.3 Local characteristics of facial features

The performance of the method described in section 2.2 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

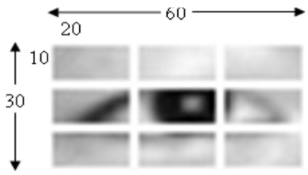


Figure 4. Variance filter model.

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

$$F_e = \frac{1}{N} \sum_{j=1}^{N} V_\sigma^j \tag{2}$$

where V_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. Then the correlation between the variance vector of each subblock of the region and the eye variance filter (2) is calculated using this formula

$$R(I_\sigma, F_e) = \frac{E[(\xi_{I_{\sigma i}} - E(\xi_{I_{\sigma i}}))\,(\xi_{F_e} - E(\xi_{F_e})]}{\sqrt{D(\xi_{I_{\sigma i}})\ D(\xi_{F_e})}} \tag{3}$$

where $\xi_{I_{\sigma i}}$ and ξ_{F_e} are the concatenated vectors of variance of subblock $I_{\sigma i}$ 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 2.2 to select only two regions which represent the left and right eye.

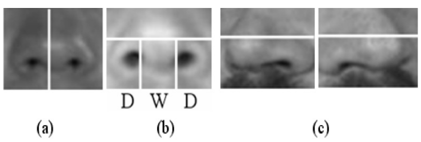


Figure 5. The local characteristics of nose region.

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.

3 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 feature spaces 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 successful detection rate (i.e. images with correctly detected eye/nose region relative to the whole set of facial database) is not so high, but it is reasonable. The lowest detection rate is on BioID and JAFEE databases because the variation on lighting conditions and facial expressions in these two databases is relatively higher than that in the other databases. The best performance is obtained on the XM2VTS database because the image quality is very high and the illumination is always uniform.

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

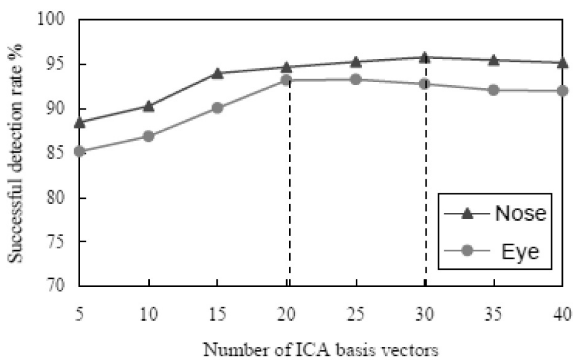


Figure 6. Selection the optimum number of basis vectors.

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 on different databases.

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

some databases such as XM2VTS an improvement rate more than 5% has been achieved, while in the 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. Also, compared with method [8] achieved detection rate of 95.7% on 750 images only of XM2VTS database, the proposed method achieves 98.4% on the whole database (Table (2)). While, for nose it can achieve an average detection rate of 96% outperforming [9] 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.

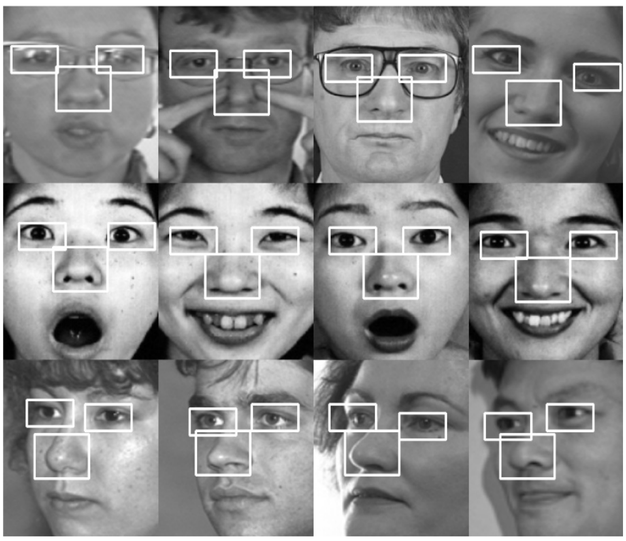


Figure 7. Examples of successful results.

4 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 spaces is presented. Eye and nose feature 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 and nose fields with high detection rate comparable with existing methods.

Acknowledgment

The authors would like to thank Egyptian government for supporting this work under Grant No.1/13/101-2006-2007. The owners of all databases that are used in the experiments in this research namely; FERET, BioID, XM2VTS, and JAFFE are gratefully acknowledged.

References

[1] S. Asteriadis, N. Nikolaidis, I. Pitas: Facial feature detection using distance vector fields, *Pattern Recognition*, vol. 42, no. 7, pp. 1388-1398, 2009.

[2] M. Hassaballah, T. Kanazawa, S. Ido, S. Ido: Efficient eye detection method based on grey intensity variance and independent components analysis, *IET Computer Vision*, vol. 4, no. 4, pp. 261-271, 2010.

[3] Y.-S. Ryu, S.-Y. Oh: Automatic extraction of eye and mouth fields from a face image using eigenfeatures and ensemble networks, *Applied Intelligence*, vol. 17, no. 12, pp. 171-185, 2002.

[4] B. Kroon, S. Maas, S. Boughorbel, A. Hanjalic: Eye localization in low and standard definition content with application to face matching, *Computer Vision and Image Understanding*, vol. 113, no. 8, pp. 921-933, 2009.

[5] P. Wang, Q. Ji: Multi-view face and eye detection using discriminant features, *Computer Vision and Image Understanding*, vol. 105, no. 2, pp. 99-111, 2007.

[6] W. H. Dan, J. Qiang: In the eye of the beholder: A survey of models for eyes and gaze, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 32, no. 3, pp. 478-500, 2010.

[7] P. Sankaran, S. Gundimada, R.C. Tompkins, V.K. Asari: Pose angle determination by faces, eyes and nose localization, *IEEE CVPR'05*, pp. 161-167, 2005.

[8] P. Campadelli, R. Lanzarotti: Fiducial point localization in color images of face foregrounds, *Image and Vision Computing*, vol. 22, no. 11, pp. 863-872, 2004.

[9] Y. Gizatdinova, V. Surakka: Feature-based detection of facial landmarks from neutral and expressive facial images, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 135-139, 2006.

[10] A. Hyvarinen, E. Oja: Independent component analysis: algorithms and applications, *Neural Networks*, vol. 13, pp. 411-430, 2000.

[11] P. Viola, J. M. Jones: Robust real-time face detection, *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137-154, 2004.