ELectronics EXpress

IEICE Electronics Express, Vol.6, No.10, 587–593

Fusion of multiple facial regions for expression-invariant gender classification

Li Lu[a)] and Pengfei Shi

Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, Shanghai, China, 200240

a) *lulihappy@sjtu.edu.cn*

Abstract: A novel gender classification method is presented which fuses information acquired from multiple facial regions for improving overall performance. It is able to compensate for facial expression even when training samples contain only neutral expression. We perform experimental investigation to evaluate the significance of different facial regions in the task of gender classification. Three most significant regions are used in our fusion-based method. The classification is performed by using support vector machines based on the features extracted using two-dimension principal component analysis. Experiments show that our fusion-based method is able to compensate for facial expressions and obtained the highest correct classification rate of 95.33%.

Keywords: gender classification, two-dimension principal component, support vector machines

Classification: Science and engineering for electronics

References

[1] J. Yang, D. Zhang, A. F. Frangi, and J. Y. Yang, "Two dimensional PCA: a new approach to appearance-based face representation and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.* vol. 24, no. 1, pp. 131–137, 2004.

[2] Moghaddam B. and Yang M. H., "Learning Gender with Support Faces," *IEEE Trans. Pattern Analysis and Machine Intelligence* vol. 24, no. 5, pp. 707–711, 2002.

[3] Gao, W., Cao, B., Shan, and S. G., et al., "The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations," *technical report of JDL*, (2004), [Online] http://www.jdl.ac.cn/peal/files/IEEE_SMC_A_gao_CAS-PEAL.pdf.

[4] J. Yang, D. Zhang, Y. Xu, J. Y. Yang, "Two-dimensional discriminant transform for face recognition," *Pattern Recognition.* vol. 38, no. 7, pp. 1125–1129, 2005.

[5] FERET Database, [Online] http://www.itl.nist.gov/iad/humanid/feret/, 2004.

DOI: 10.1587/elex.6.587
Received February 22, 2009
Accepted April 17, 2009
Published May 25, 2009

IEICE Electronics Express, Vol.6, No.10, 587–593

[6] Chang, C. C., Lin, C. J. "LIBSVM: a Library for Support Vector Machines," [Online] http://www.csie.ntu.edu.tw/-cjlin/papers/libsvm.ps.gz.

1 Introduction

As an interesting yet challenging problem, recognizing the gender of face images is a rather new research topic in computer vision. Automatic gender classification could be of important value in human-computer interaction, such as personal interaction. However, the expression variations is an inevitable issue in the development of a practical gender classification system.

Most computational models of gender classification use global information (the whole face image) giving equal weight to all areas of the face, irrespective of the importance of internal features. Intuitively, we argue that smaller facial regions, if judiciously selected, would be less sensitive to expression variations and may lead to better overall performance. Our work on gender classification is one of the first attempts to report a detailed evaluation of the significance of different facial regions for gender perception. Considering the significance of facial regions, we propose a fusion-based method, combining the classification results of three facial regions, for improving the robustness to facial expressions. The two-dimension principal component analysis (2DPCA) [1] is applied to the facial regions and the corresponding principal images were used as the feature vectors. The classification is performed by using support vector machines (SVMs), which had been shown to be superior to traditional pattern classifiers in gender classification problem [2]. The experimental result shows that this method performs better than the method based on the whole face image.

The rest of this paper is organized as follows: In Section 2, we evaluate the significance of different facial regions. The feature extraction and classification are described in Section 3 and 4. The experimental results are presented in Section 5, and the conclusion is drawn in Section 6.

2 Significance of different facial regions

Our investigation involved the assessment of seven different facial regions for gender classification. These facial regions are the whole face (including hairline), the internal face, the upper region of face, the lower region of face, the left eye, the nose, and the mouth, as illustrated in Figure 1 (a). Given a face image, the images of the seven facial regions are obtained. A 480×360 gray scale image is converted to a normalized whole face image and a normalized internal face image. The geometric normalization gives the same pixel distance between eye locations to all faces. The histogram equalization is used to reduce the effect of illumination variations. Five smaller facial regions, the upper region of face, the lower region of face, the left eye, the nose, and the mouth, are obtained from the whole face image. We apply

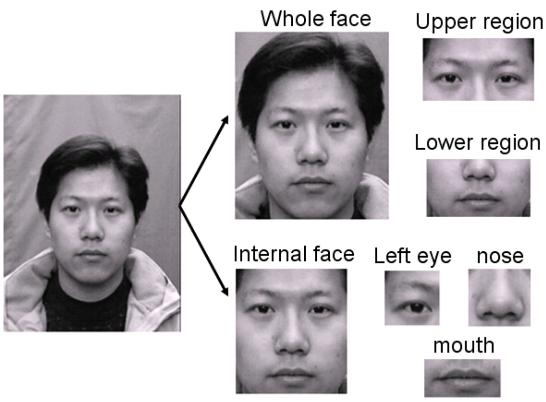
DOI: 10.1587/elex.6.587
Received February 22, 2009
Accepted April 17, 2009
Published May 25, 2009

IEICE Electronics Express, Vol.6, No.10, 587–593

SVMs [2] on facial region images to evaluate the significance of different facial regions.

Experiments are carried out using 800 frontal face (400 females and 400 males) gray scale images with neutral expression. The images are from the CAS-PEAL [3] database. Figure 1 (b) shows the average classification rate percentage of the facial regions.

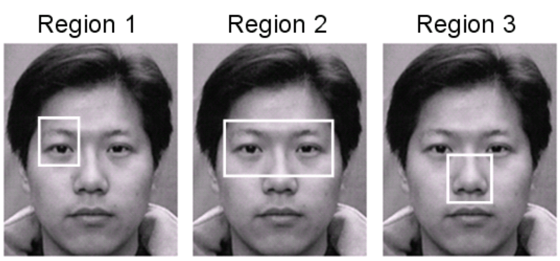
Obviously, a relatively small facial region is more robust to expression variations than a whole face. Hence, one can improve the gender classification performance by using local facial regions. Instead of using the entire normalized texture image as most gender classification methods do, we ex-



(a)

Facial Region	Classification Rate
Whole Face	92.50
Internal Face	91.67
Upper Region	92.90
Lower Region	80.50
Left Eye	91.17
Nose	88.00
Mouth	82.00

(b)



(c)

Fig. 1. Evaluate the significance of different facial regions and determine three most significant regions.

DOI: 10.1587/elex.6.587
Received February 22, 2009
Accepted April 17, 2009
Published May 25, 2009

IEICE Electronics Express, Vol.6, No.10, 587–593

tract feature vectors from three local face regions, namely left eye, upper region of face, and nose (see Figure 1 (c)). These areas tend to change less for expression variations.

3 Feature extraction with 2DPCA

We employ 2DPCA on the facial regions, and the resultant eigen-images are used as the feature vectors. 2DPCA is a novel transform technique derived from the principal component analysis (PCA) technique, directly extracts features from image matrices [1, 4]. 2DPCA is much more efficient than PCA, requiring less memory and having a lower computational cost and has obtained promising experimental results in the areas of feature extraction and dimension reduction.

Given a training set $\{X_1, X_2, \cdots, X_N\}$, 2DPCA first uses all training samples to construct the total sample covariance matrix G_t

$$G_t = \frac{1}{N}\sum_{i=1}^{N}\left(X_i - \bar{X}\right)^T \left(X_i - \bar{X}\right) \tag{1}$$

where X_i is the i th training sample, $\bar{X}$ is the mean sample matrix of all training sample matrix, and N is the number of training samples. The projection axes of 2DPCA, $w_1, \cdots, w_d$, can be obtained by solving the algebraic eigenvalue problem $G_i w_i = \lambda_i w_i$, where w_i is the eigenvector corresponding to the i th largest eigenvalue of G_t. Finally we obtain the 2DPCA projector W_{opt}

$$W_{opt} = [w_1, \cdots, w_d] \tag{2}$$

and use

$$Y = XW_{opt} \tag{3}$$

to extract the low-dimensional feature matrix of an image matrix X.

4 Classification with SVMs

SVMs is a supervised learning technique from the field of machine learning, it is applicable to both classification and regression. Consider the training data set $D = \{(x_i, y_i)\}_{i=1}^{l}$ of labeled training samples, $x_i \in R^d$, where d denotes the dimensionality of the training samples, and $y_i \in \{-1, +1\}$ is the associated label. The training vectors x_i are mapped into a high dimensional feature space by the function ϕ. The SVM finds an optimal linear separating hyperplane that separates all the projected training samples $\phi(x)$ with the maximal margin in this high dimensional feature space. The hyperplane is defined by:

$$f(x) = \sum_{i=1}^{l}\lambda_i y_i k(x, x_i) + b^* \tag{4}$$

where $k(x, x_i) = \phi(x_i)^T \phi(x_j)$ is called the kernel function, b^* is a bias. To construct an optimal hyperplane is equivalent to finding all the nonzero λ_i and is formulated as a quadratic programming problem with constrains. In order to avoid the curse of dimensionality problem, the inner product in

DOI: 10.1587/elex.6.587
Received February 22, 2009
Accepted April 17, 2009
Published May 25, 2009

IEICE Electronics Express, Vol.6, No.10, 587–593

the high dimensional feature space is replaced by a simpler kernel function according to Mercer's theorem, i.e., $\phi(x_i)^T \phi(x_j) = k(x, x_i)$.

We first employs 2DPCA on three facial regions, left eye region, upper region of face, and nose region, to obtain the feature vectors, then apply SVMs on the regions respectively, and finally obtain the classification result using a consensus decision.

5 Experimental results

We compared the fusion-based gender classification method with the single region based method on FERET face database [5], and compared three different approaches for gender classification on CAS-PEAL database.

5.1 Experiment on the FERET

We use 800 adult frontal face grey scale images (400 males and 400 females). The variations of those 800 images are mainly in pose, facial expression and lighting condition. All the face images are warped to the same scale, orientation and position. Histogram equalization was applied to the extracted face images to normalize for different lighting conditions.

Three facial regions, the left eye, the upper region of face and the nose, are extracted and normalized for each of the training images. For each facial region we obtain 800 training cases (400 females and 400 males). We extract feature vectors from region images using 2DPCA. The average classification rate was estimated with five-fold cross validation (CV) — i.e., a five-way data set split, with 4/5th used for training and 1/5th used for testing, with four subsequent nonoverlapping rotations. For each facial regions, the average size of the training set was 640 (320 males and 320 females) and the average size of the test set was 160 (80 males and 80 females). We use a SVM classifier with RBF kernel. The SVMs come from LIBSVM [6] and all parameters are selected using cross validation via parallel grid research. Figure 2 illustrated

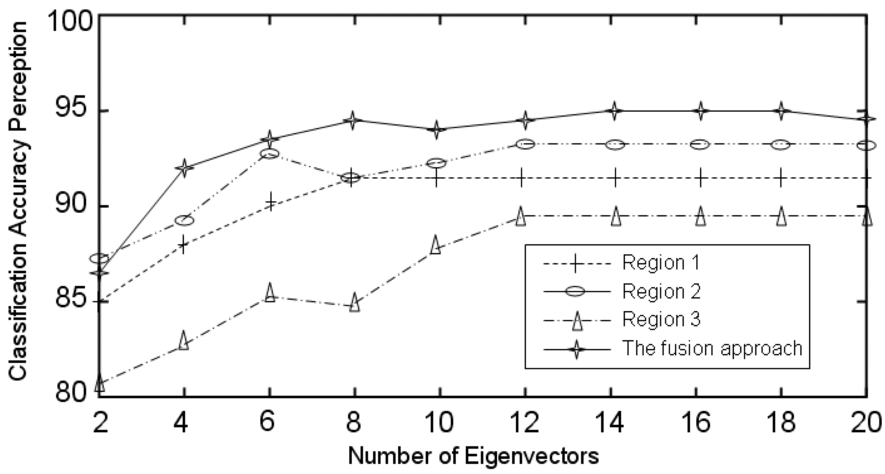


Fig. 2. The gender classification accuracy perception of Region 1 to 3 (the left eye, the upper face region, and the nose) using 2DPCA and a SVM classifier.

DOI: 10.1587/elex.6.587
Received February 22, 2009
Accepted April 17, 2009
Published May 25, 2009

the gender classification accuracy perception of Region 1 to 3 (the left eye, the upper face region, and the nose) and the proposed fusion-based method.

As shown in Figure 2, all three regions produced high classification rates, indicating each of them contains a high amount of gender information. The upper region of face yields the highest classification rate followed the left eye. The fusion-based approach, i.e. fusion of the classification results of three regions, obtains the highest correct classification rate of 94.83%.

5.2 Experiment on the CAS-PEAL

The CAS-PEAL contains face images with a variety of expressions such as neutral, smile, surprised, frown, close eyes, and open mouth. We use 600 front gray scale face images with neutral expression from 600 subjects (300 males and 300 females) as training samples. The whole face region and three smaller facial regions are extracted and normalized for each of the training images. For each kind of facial region we obtain 600 training cases (300 females and 300 males). We use 1800 front gray scale face images with various expressions, neutral, smile, surprised, frown, close eyes, and open mouth, from 360 subjects as test samples.

In the gender classification process, three facial regions, the left eye, the upper region and the nose, are extracted and normalized from the test face image. We compare three approaches for gender classification. All the approaches use a SVM classifier with RBF kernel. In the first approach, the feature vectors are extracted using 2DPCA. In the second approach, PCA is used to extract feature vectors. In the third approach, the feature vectors are constructed from the gray scale images directly.

Table I shows the highest classification accuracies of all the approaches on the CAS-PEAL database. As shown in Table I, for each facial region, the 2DPCA plus SVM approach is more efficient than the other two approaches. Based on the same approach, the fusion-based method obtains the best classification rate. Overall, the fusion-based method using 2DPCA plus SVM obtained the highest correct classification rate of 95.33%.

Table I. The highest gender classification accuracies (%) of all the approaches on the CAS-PEAL.

	Region 1	Region 2	Region 3	Whole face	Fusion of Regions
2DPCA+SVM	90.72	91.56	89.00	90.44	95.33
PCA+SVM	89.22	91.28	88.39	88.50	92.78
SVM	89.50	90.11	88.17	87.55	92.50

6 Conclusion

We have developed a novel gender classification approach which integrates multiple regions of a facial image. The proposed approach overcomes the limitations of the whole face algorithm. The performance improvement is

DOI: 10.1587/elex.6.587
Received February 22, 2009
Accepted April 17, 2009
Published May 25, 2009

IEICE Electronics Express, Vol.6, No.10, 587–593

due to the combination schemes which generate the final classification result with a higher accuracy rate than those based on a single region or a whole face region. Experimental results also indicate that the feature extraction method using 2DPCA is more efficient than PCA or feature vector constructed from the gray scale image directly.

Acknowledgments

This work is supported by National Natural Science Foundation of China under grant No.60775009.

DOI: 10.1587/elex.6.587
Received February 22, 2009
Accepted April 17, 2009
Published May 25, 2009