Multi-view Gender Classification Using Hierarchical Classifiers Structure

Tian-Xiang Wu[1] and Bao-Liang Lu[1,2,⋆]

[1]Center for Brain-Like Computing and Machine Intelligence
Department of Computer Science and Engineering
[2]MOE-Microsoft Key Laboratory for Intelligent Computing and Intelligent Systems
Shanghai Jiao Tong University
800 Dong Chuan Rd., Shanghai 200240, China
tcwtxster@gmail.com, bllu@sjtu.edu.cn

Abstract. In this paper, we propose a hierarchical classifier structure for gender classification based on facial images by reducing the complexity of the original problem. In the proposed framework, we first train a classifier, which will properly divide the input images into several groups. For each group, we train a gender classifier, which is called expert. These experts can be any commonly used classifiers, such as Support Vector Machine (SVM) and neural network. The symmetrical characteristic of human face is utilized to further reduce the complexity. Moreover, we adopt soft assignment instead of hard one when dividing the input data, which can reduce the error introduced by the division. Experimental results demonstrate that our framework significantly improves the performance.

Keywords: Hierarchical classifiers, Gender classification, Multi-view facial images.

1 Introduction

Gender classification using facial images is widely used in human-computer interaction and the applications depending on it, such as demographics and visual surveillance. Most of the existing approaches do not consider, or design some features which are claimed to be robust to the pose variation of faces. They will fail in practical use facing unconstrained face poses, or say, multi-view faces.

To ease this multi-view problem, Toews and Arbel [1] proposed the idea of relative location information of the organs, which is used to infer the most likely position of the face. The result of gender classification was obtained by combining the results of organs. Takimoto et al. [2] extracted the features around the eyes and mouths which requires the positions of eyes and mouths to be exactly located in advance. In their work, local information is used to facilitate multi-view problems. Lian and Lu [3] aligned the facial images based on the position of eyes, and apply LBP [4] to feature extraction and SVM to gender classification directly.

In this paper, we propose a framework which decomposes the multi-view problem into several single-view subproblems and hence reduces the complexity. Under this

⋆ Corresponding author.

K.W. Wong et al. (Eds.): ICONIP 2010, Part II, LNCS 6444, pp. 625–632, 2010.

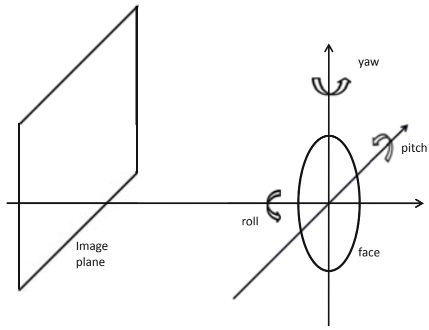

Fig. 1. The pose is decomposed into three rotations: pitch, roll and yaw

framework, any traditional feature extraction methods and classifiers (e.g. SVM, neural network) can be used. The pose is decomposed into three rotations: pitch, roll and yaw (See Fig. 1). To simplify the problem and without lose of generality, we only consider yaw rotation. The extension to the other two is trivial. The framework has two layers. In the first layer, we discretize the continuous angle space into K bins. A classifier whose output is in $\{1, \ldots, K\}$, is trained to predict which bin the input facial image falls in. We call it orientation classifier. One problem of the discretization process is the boundary effect. It is unreasonable to simply put a image on a boundary to either side. Therefore we adopt soft assignment, allowing the partitions overlap on the boundary. Moreover, to reduce the number of categories the orientation classifier should deal with, we make use of the symmetrical characteristic of human face, horizontally flipping the images whose faces toward right. By doing so the accuracy of the orientation classifier is increased. Then in the second layer, for each bin we train a classifier which specializes in gender classification of images from that bin. These classifiers are called gender classifiers.

The rest of the paper is organized as follows: Section 2 describes the main idea in the proposed hierarchical classifiers framework. Section 3 introduces some tricks to improve the accuracy. Section 4 shows the gender classification procedure using our framework. Experiment results are presented in Section 5. Some conclusions and future work are outlined in Section 6.

2 Hierarchical Classifiers

Traditional gender classification algorithms always work well on images with the same pose, since the alignment is easy. For multi-view facial images, the issue becomes much more complex. The feature space is much larger and it is difficulty to design orientation-invariant features. An efficient solution is to divide the feature space into several subspaces according to face orientations, which decomposes the multi-view problem into easier classification tasks on simpler subspaces.

In this paper we propose a two-layered classifier (See Fig. 2). The first layer includes a classifier, which extracts the feature vectors from the original image and classifies it into several categories according to its orientation. Then the task of gender classification

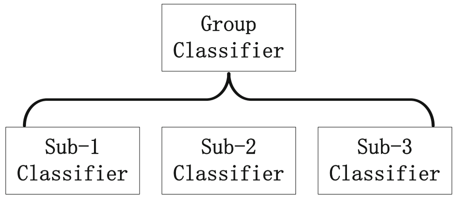

Fig. 2. Hierarchical classifiers structure

is passed to the next layer where we make use of experts of gender classification for certain orientations. It is obviously that the accuracy of the classifier in the first layer is important for the whole problem. Classification error in the first layer leads to error answers in the second layer. So this method is suitable when the initial classification problem has a high precision.

3 Angle Categories Selection

In this section, we show some technologies that can be used in angle category selection to improve the accuracy of gender classification.

Since the faces of human beings are bilaterally symmetric, the images in which the persons face right turn to be the ones facing left after a horizontal flip. If we get the information of the face direction in the images, the original space of input images should be reduced by a half. An easy classifier is trained for the direction classification in this paper to reduce the complexity.

The hierarchical classifiers structure is suitable for the problem whose initial classification problem has a high precision. Classification error of the initial separation leads to error answers in the individual classification. We found in the experiment that the images whose angles are near the dividing line of the two angle categories are easily to be misclassified. The error answer caused by the misclassification is due to the lack of information in the certain individual classifier. We get another trick which is to add the samples whose angles are near the dividing line into the training data of the two neighboring categories. The detail is refer to Section 5.2.

By using the combination of results generated in the gender classifiers, the risk in the first classification layer is apportioned.

$$result_i = \sum_{k=1}^{K} p_i^k \times result_i^k \tag{1}$$

$result_i$ is the possibility of the ith sample to be a male, and $result_i^k$ is the result from the sub classifiers. p_i^k means the possibility for the ith sample to belong to the category. Moreover, Weighted Relevance Aggregation Operator(WRAO) [5] helps hierarchical model with hierarchical fuzzy signatures to work better.

In our problem, the classification is the first layer is not so hard. Huang and Shao [6] use SVM to achieve perfect performance on face pose classification problem on the standard FERET data base. With tolerance on the dividing region of the neighboring

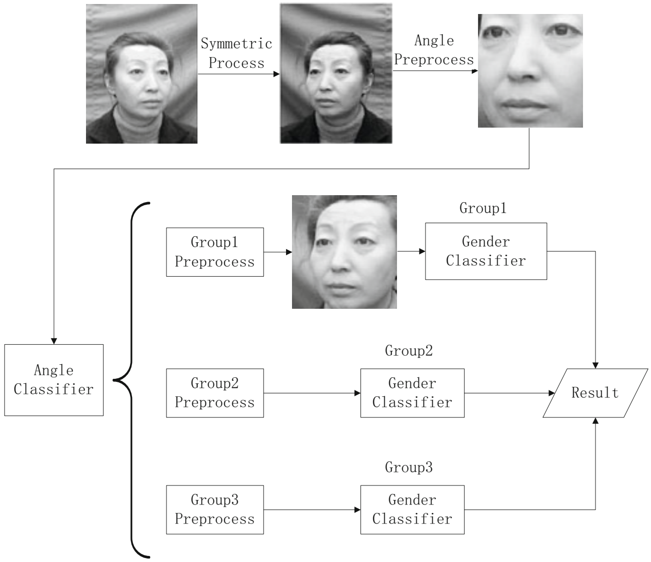

Fig. 3. The gender classification process using hierarchical classifiers structure

categories, the uncertainty in the first layer won't make trouble for the gender classification.

4 Gender Classification Procedure

In this section, we introduce the whole hierarchical classifier based gender classification system(See Fig. 3).

Alignment is important for gender classification based on facial images [7]. First, faces are fixed in the center of the result pictures. Facial components are in the certain places for feature extraction after alignment. To obtain the facial components, we adopt Active Shape Model (ASM) [8], a statistical model of the shape of the deformable object, to get the locations of eyes and mouth, and then cut the rectangle out of the facial image.

The bilaterally symmetrical characteristic of human face is make use of. The images facing right are turned left. Then, the images will be classified into some classes according to the angle of the face. Images in different angle classes are taken to their own gender classifiers. Now we have converted the original problem to gender classifications based on facial images of fixed angle, a well-studied problem with many good approaches.

Gender classification processes in the different categories are similar. The facial images are re-cutting, using the information of angles which is the label of the category, in order to put most of the human face into the picture, and align the organ positions precisely. Some information of hair is also taken into images for classification. Feature extraction is done in different categories and the suitable gender classification is prepared for the facial images.

5 Experiment

5.1 Data Set

To compare the performance, we select the gender classification problem based on multi-view facial images in the CAS-PEAL face database [9](See Table 1). We take all the images labeled ”PM” from the ”POSE” section. We put different people into the training and test table, i.e., we partition the set according to people instead of single images so that two different images of a person cannot stay in both training or test set. This would help avoid over-fitting, or the similarity of the two images will take unnecessary information to the test set. The total 7273 different-pose facial images are so organized into 11 groups, such that within each of them, the number of the training samples is 70% the number of female facial images, which are fewer than male images. We let the percentage be 50% if there are not many in that group.

Table 1. Description of training and test data based on facial images

Data Set	Description	Total	Male	Female	Training	Test
	PM-67	101	79	22	11*2	79
	PM-45	1039	595	444	306*2	427
	PM-30	938	516	422	295*2	348
	PM-22	101	79	22	11*2	79
	PM-15	938	516	422	295*2	348
CAS-PEAL	PM+00	1039	595	444	306*2	427
	PM+15	938	516	422	295*2	348
	PM+22	101	79	22	11*2	79
	PM+30	938	516	422	295*2	348
	PM+45	1039	595	444	306*2	427
	PM+67	101	79	22	11*2	79
	TOTAL	7273	4165	3108	4284	2989

5.2 Implementation

The images in the data set are of many different angles. We want to use this prior knowledge to separate them into some categories. However, as we have pointed out, too many categories will introduce complexity to the division problem. So some of the images in different angles must be put into one category. We use the symmetric property of human face to reduce the total amount of angles from 11 to 6. In this case, 3 categories are enough.

The division illustrated in Fig. 4(a) looks quite natural. Given such a division, none of the two corresponding classifiers would be able to solve the angle between the two regions. For we use the structure of hierarchical classifiers, classification error in the first layer probably leads to error answers in the second layer, especially near the division of the two neighboring regions. In this paper, we use the division in Fig. 4(b). Images of PM00 and PM15 are used for training the first classifier. Images of PM15, PM22 and PM30 are for the second classifier. And PM30, PM45 and PM67 are for the

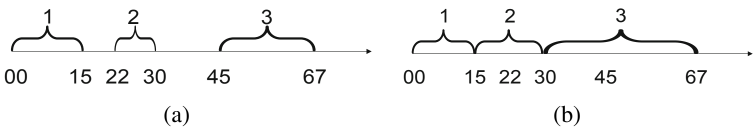

Fig. 4. Angle region division about gender classification: (a)naive way, 6 angles are divided equally into three groups; (b)the way to get all angles in and improve accuracy

Table 2. Gender classification with or without symmetric process

Method	Accuracy (RBF kernel)	Accuracy (linear kernel)
Symmetric accuracy	99.63%	99.60%
With symmetric process	93.34%	92.64%
Without symmetric process	92.31%	91.50%

third. The training data of the division of the two neighboring regions are trained in both corresponding classifiers. The images whose angles are near the division can be classified correctly in both corresponding classifiers. This will also help to reduce the risk of the accumulative error introduced by the classification error in the first layer.

For feature extraction, we use multi-resolution local Gabor binary pattern (MLGBP) to extract the features of each facial image. The MLGBP [10] feature, which is the input of the SVM classifiers, is derived by combining multi-resolution analysis, Gabor characteristic and uniform LBP histograms [11]. All experiments were performed on a Pentium fourfold CPU (2.83GHz) PC with 8GB RAM.

All the classifiers in this paper are Support Vector Machines, lib-svm v.2.86 in detail. RBF kernel and linear kernel are used for comparison.

5.3 Result

The effect of symmetry is shown in Tab. 2. We make use of the symmetry of human face to reduce the originally 11 angles to 6. The accuracy of the whole process including symmetric transformation and gender classification is better than that without the process. It means the symmetric transform provides the classification with less complexity and more accuracy but less harm.

The classifications in small fields show great advantage over the one in the large complex space(See Tab. 3). The result shows the additional classification step won't

Table 3. Gender classification in different angle categories

Angle Category	Accuracy (RBF)	Total SV	Accuracy (linear)	Total SV
PM00,PM15	97.51% (1095/1123)	416	97.51% (1095/1123)	346
PM15,PM22,PM30	98.26% (1523/1550)	423	97.94% (1518/1550)	352
PM30,PM45,PM67	97.78% (1670/1708)	586	97.48% (1665/1708)	348
Classify directly	92.31% (2759/2989)	1480	91.50% (2735/2989)	840

Table 4. Gender classification using hierarchical classifiers structure

Method	Accuracy (RBF)	Total SV	Accuracy (linear)	Total SV
Angle classify	98.43% (2942/2989)	1161	98.43% (2942/2989)	965
Classify directly	92.31% (2759/2989)	1480	91.50% (2735/2989)	840
Whole System	97.89% (2926/2989)	–	97.59% (2917/2989)	–

harm the accuracy but increase it(See Tab. 4). In the experiment, we find most of the testing data which are misclassified in the angle classification are the facial images laying near the dividing line and being classified into the neighboring category. The selected classifiers are still suitable for these images and prepare better classification. The angle categories selection helps to solve the main problem in hierarchical classifiers framework. So the accuracy of the gender classification with symmetric process and angle classification is close to the performance of the expert classifiers in their fields.

5.4 Complexity Analysis

As we know, the time complexity of a standard SVM QP solver is $O(M^3)$, where M denotes the number of training samples. In our hierarchical classifiers framework, we cut the training samples into K groups, where K is the number of classifiers in the second layer of the structure. In each group, the corresponding classifier only needs to deal with its own training samples, so they can be trained in parallel, meaning that the running time could be improved to $O((\frac{M}{K})^p)$. Even if we run the training in serial, it will only take $O(K(\frac{M}{K})^p)$. Thus time complexity is reduced in both situations.

During recognition, the time eater is to calculate the kernel of test and support vectors especially in high dimension space. So we suppose the time complexity of SVM is $O(v)$, where v is the number of support vectors. The statistic shows that the sub-problems need much less time than the original problem.

6 Conclusions and Future Work

We have proposed a novel framework for gender classification based on multi-view facial images, i.e., hierarchical classifiers. The most important advantage of our framework over traditional SVM is that prior knowledge is used to get the input image to the expert of that field who can be easily get trained and give an answer. Experimental results show the effectiveness of our framework. A future extension of our work is to use the combination of some classifiers of organs or other parts to make our classifier more robust and improve accuracy.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (Grant No. 60773090 and Grant No. 90820018), the National Basic Research Program of China (Grant No. 2009CB320901), the National High-Tech Research Program of China (Grant No. 2008AA02Z315), and the Science and Technology Commission of Shanghai Municipality (Grant No. 09511502400).

References

1. Toews, M., Arbel, T.: Detection, localization, and sex classification of faces from arbitrary viewpoints and under occlusion. IEEE Transactions on Pattern Analysis and Machine Intelligence 31, 1567–1581 (2009)
2. Takimoto, H., Mitsukura, Y., Fukumi, M.: Robust Gender and Age Estimation under Varying Facial Pose. IEE J. Trans. 127(7), 1022–1029 (2007)
3. Lian, H., Lu, B.: Multi-view gender classification using local binary patterns and support vector machines. In: Proceedings of the Third International Symposium on Neural Networks, pp. 202–209 (2006)
4. Ahonen, T., Hadid, A., Pietikainen, M.: Face recognition with local binary patterns. In: Pajdla, T., Matas, J(G.) (eds.) ECCV 2004. LNCS, vol. 3021, pp. 469–481. Springer, Heidelberg (2004)
5. Mendis, B., Gedeon, T., et al.: Generalised weighted relevance aggregation operators for hierarchical fuzzy signatures. In: International Conference on Computational Intelligence for Modelling Control and Automation (2006)
6. Huang, J., Shao, X., Wechsler, H.: Face pose discrimination using support vector machines (SVM). In: Proc. of 14th International Conference on Pattern Recognition (ICPR 1998) (1998)
7. Makinen, E., Raisamo, R.: Evaluation of Gender Classification Methods with Automatically Detected and Aligned Faces. IEEE Trans. Pattern Analysis and Machine Intelligence 30(3), 541–547 (2008)
8. Cootes, T., Taylor, C., Cooper, D., Graham, J., et al.: Active shape models-their training and application. Computer vision and image understanding 61(1), 38–59 (1995)
9. Gao, W., Cao, B., Shan, S., et al.: The cas-peal large-scale chinese face database and baseline evaluations. Technical report of JDL, `http://www.jdl.ac.cn/~peal/peal_tr.pdf`
10. Lian, H., Lu, B.: Multi-view gender classification using multi-resolution local binary patterns and support vector machines. International Journal of Neural Systems 17(6), 479–487 (2007)
11. Xia, B., Sun, H., Lu, B.: Multi-view gender classification based on local Gabor binary mapping pattern and support vector machines. In: IEEE International Joint Conference on Neural Networks, pp. 3388–3395 (2008)