Neural Comput & Applic
DOI 10.1007/s00521-011-0647-x

ICONIP2010

Multi-view gender classification using symmetry of facial images

Tian-Xiang Wu · Xiao-Chen Lian · Bao-Liang Lu

Received: 8 April 2011 / Accepted: 7 May 2011

Abstract In this paper, we propose a multi-view gender classification system with a hierarchical framework using facial images as input. The front end of the framework is a classifier, which will properly divides the input images into several groups. To ease the data sparsity problem in the multi-view scenario, facial symmetry is used to reduce the number of views. Moreover, we adopt soft assignment when dividing the input data, which can reduce the errors caused by the boundary effect in hard assignment. Then for each group, we train a gender classifier, called an expert. These experts can be any commonly used classifiers, such as support vector machines or neural networks. In this step, facial components instead of the whole face are used to achieve higher robustness against variations caused by facial alignment, illumination and occlusions. Experimental results demonstrate that our framework significantly improves the performance.

Keywords Hierarchical classifiers · Gender classification · Multi-view facial images · Facial symmetry · Multiple kernel learning

T.-X. Wu · X.-C. Lian · B.-L. Lu
Department of Computer Science and Engineering,
Center for Brain-Like Computing and Machine Intelligence,
Shanghai Jiao Tong University, 800 Dong Chuan Rd.,
Shanghai 200240, China

B.-L. Lu (✉)
MOE-Microsoft Key Laboratory for Intelligent Computing
and Intelligent Systems, Shanghai Jiao Tong University,
800 Dong Chuan Rd., Shanghai 200240, China
e-mail: bllu@sjtu.edu.cn

1 Introduction

Gender classification using facial images is widely used in human-computer interaction, e.g., demographics and visual surveillance. Many approaches [1–8] have been proposed using a pipeline of feature extraction followed by classifiers trained on these features and have achieved great accuracy on various data sets.

However, there are still some problems that need to be solved. One is the multi-view problem, where most of the existing approaches will fail in as they focus on a frontal view only. In this paper, we propose a framework that decomposes the multi-view problem into several single-view subproblems and hence reduces the complexity. Under this framework, all the traditional feature extraction methods and classifiers can be used. The pose is generally decomposed into three rotations: pitch, roll, and yaw (See Fig. 1). To simplify the problem and without loss of generality, we only consider yaw rotation in this paper. 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 is trained to predict which bin the input facial image falls in. We call it the orientation classifier. One problem of the discretization process is the boundary effect. It is unreasonable to simply put an image on a boundary to either side. Therefore, we adopt soft assignment, which is to compute a probability distribution of the estimation of the face orientations. Moreover, we make use of the symmetrical nature of the human face, horizontally flipping the images which face toward the right. By doing this, the number of categories the orientation classifier deals with is reduced to a half, which means the training data for each category is doubled. Then in the second layer, for each bin, we train a classifier which specializes in gender classification of images from

Published online: 18 May 2011

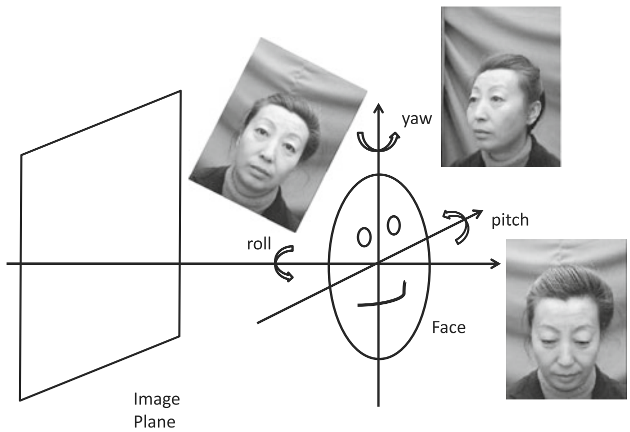

Fig. 1 The pose is decomposed into three rotations: pitch, roll and yaw

that bin. These classifiers are called gender classifiers. For each of them, the existing methods for gender classification can be used. In this paper, multi-resolution local Gabor binary pattern (MLGBP) [9] is used to extract the features followed by support vector machine (SVM) classifiers.

Another problem is face occlusion. Sunglasses, beards, masks, and other noise may make classification difficult. Li et al. [10] have proved that the combination of facial components has robust performance on gender classification. To ease this problem, we use the fusion of facial components instead of using wholistic faces. The combination strategies we evaluated are simple arithmetic operations (sum, maximum, and product), fuzzy integral [11], and the multiple kernel learning (MKL) feature combination method. The former two belong to the classifier combination category and the last one is a feature integration approach.

The rest of the paper is organized as follows: Sect. 2 lists some related work on gender classification. In Sect. 3, we describe the main idea in the proposed hierarchical classifiers framework and implementation details. We compare some strategies for the information integration in Sect. 5. Experiment results are presented in Sect. 6. Some conclusions and future work are outlined in Sect. 7.

2 Related work

Gender classification approaches can be divided into two general types. One is based on geometric features. Burton et al. [12] used the distances between points on the face as feature vectors and linear discriminant analysis for gender classification. Brunelli and Poggio [2] obtained 16 kinds of lengths from different organs of the face to train a neural network for classification. Fellous [13] used 24 kinds of lengths for feature vectors.

The other is based on appearance which is a pixel-based approach. Golomb et al. [1] trained a two-layer neural network to perform gender classification on 30×30 pixel images. Tamura et al. [14] showed low-resolution facial images also work well using a neural network. Cottrell and Metcalfe [15] used the data produced by principal component analysis (PCA) as the input of a back propagation neural network and obtained a better result. Moghaddam and Yang [4] demonstrated that support vector machines work much better in gender classification than other classifiers such as classical radial basis function (RBF) networks, ensembles of RBF networks, Fisher linear discriminant, and nearest neighbor. Sun et al. [16] extracted PCA features from faces and selected a subset of features using a genetic algorithm (GA), then used SVMs and linear discriminant analysis (LDA) as classifier. Jain and Huang [17] used independent component analysis (ICA) to extract features from faces, and adopted LDA as classifiers. Lian and Lu [6] introduced local binary pattern (LBP) into the field of gender classification. Zhang et al. [18] used LBP to extract features from Gabor transformed data, a technique that works more like human eyes. Kim et al. [5] applied a Gaussian process method to gender classification. It could automatically determine the hyper-parameters. Wu et al. [19], Sun et al. [20], Baluja and Rowley [7] used the Adaboost framework for gender classification and achieved promising accuracy. Xia et al. [9] adopted a local Gabor binary mapping pattern method to extract face features and used uniform LBP histograms to reduce the feature dimension. Our previous work [8] used a fragment-based hair feature extraction method and a PCA-based facial feature extraction method with a fuzzy integration strategy to combine the outputs of eyes, nose, mouth, and hair classifiers. It greatly improved classification accuracy.

To ease the multi-view problem, Toews and Arbel [21] used relative location information of the organs to infer the most likely positions of faces. Takimoto et al. [22] used local information to facilitate feature extraction around the eyes and mouth, which requires the positions of eyes and mouth to be exactly located in advance. Lian and Lu [6] aligned the facial images based on the position of eyes and applied Local Binary Pattern (LBP) [23] to feature extraction and SVM to gender classification directly. In order to overcome the difficulty introduced by various environments, Ji et al. [24] extracted features both from hair and face and combined them to obtain the classification result. Li et al. [10] compared the performance of different combination strategies, i.e., arithmetic rules and fuzzy integral methods and proposed a method for combining clothes information to improve gender classification accuracy. Ji and Lu introduced a pattern classifier with confidence and applied it to gender classification [25].

3 Hierarchical classifiers

Gender classification algorithms on well-aligned images perform quite well. For multi-view facial images, however, the issue becomes much more complex. The feature space is much larger, and designing an orientation-invariant feature is very difficult. A commonly adopted solution is dividing the feature space into several subspaces according to face orientations, which decomposes the multi-view problem into easier classification tasks on simpler subspaces.

We adopted a two-layered structure. The first layer includes an orientation classifier, which extracts feature vectors from the original image and classifies it into several categories according to its orientation. Then the task of gender classification is passed to the next layer where we make use of experts of gender classification for certain orientations.

As errors in the first layer will propagate to the following steps, the accuracy of the orientation classification is important for the whole problem. One source of error is the boundary effect of hard assignment. For a face whose orientation lies on the boundary of two categories, it is unreasonable to simply put it into one of them. Therefore, we estimate the confidence that one face falls in each orientation category. After obtaining the results from the gender classifiers in the following stage, we compute the weighted-sum using the orientation confidence, which is shown as follows:

$$\mathrm{Prob}(x_i = \mathcal{O}) = \sum_{k=1}^{K} p_i^k \times \mathrm{Prob}(\mathcal{C}_k(x_i) = \mathcal{O}), \tag{1}$$

where $\mathcal{O} \in \{\text{male, female}\}, p_i^k$ is the possibility that the ith face belongs to the kth orientation category and $\mathcal{C}(\cdot)$ is the gender classifier for the kth orientation. In this paper, we use the approach proposed by Huang and Shao [26] for face pose classification.

4 Symmetric trick

Human faces are bilaterally symmetrical. There may be some special characteristics on someone's face which could be the decisive evidence to recognize a certain person. But in the area of gender classification, the special characteristics are not so important. Thus, whether the characteristic is on the left cheek or right would not affect the result of gender classification. In most cases, the face images are exactly bilaterally symmetrical. We can identify a person in an image as a man or a women according to the geometric measurements of the face and organs, the position of the facial components, the style of hair, the texture of the skin, and many other details on the face. A symmetric process will do no harm to the feature extraction of all the above information. In fact, we cannot know if the face in the following image is originally facing left or right (See Fig. 2).

Fig. 2 The female in the image is originally facing right

The bilateral symmetry of the human face is one reason for us to select the yaw rotation from the three possibilities. Because the space of the problem can be cut to a half if we turn all the faces which face right to face left by a horizontal flip. If we just take one of the other two rotations, the flip would not work. The limitation of the facial symmetric trick is that we can only handle the various situations brought by the yaw rotation but not all three possible rotations of the human face. The performance also depends on the accuracy of the classification to tell wether the face is turned toward left or right. Fortunately, this is an easy task (See Sect. 6.4).

Suppose we just consider the yaw rotation of the human face and all the faces are now facing left. The feature space is obviously a half of the original one. As a result, the scales of training data in the expert classifiers of the hierarchical framework increase a lot and the number of the classes into which the first layer is going to classify is also reduced to a half.

5 Combination of facial component classifiers

Even in the same orientation category, the alignment of facial components is still a problem that cannot be ignored when using wholistic features. Thus, Takimoto et al. [22] extracted the features around the eyes and mouth to bypass the complexity of different facial component placements. Motivated by psychological experiments [27] which prove that individual facial components can indicate the gender, we decided to use facial components rather than the whole face. Similar to the bag-of-words framework, we discard the spatial relationship between the components.

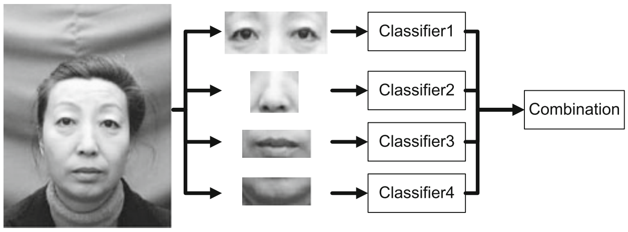

Fig. 3 Gender classification based on facial component combination

We adopt the active shape model (ASM) [28], a statistical model of the shape of a deformable object, to get the locations of eyes, nose, mouth, and chin, and then cut a rectangle of facial components out of the facial image. SVMs with probabilistic outputs [29] are trained based on each facial component in a certain angle category (See Fig. 3).

There are many strategies for combination of classifiers. The strategies are mainly divided into two approaches. One is to use the probabilistic outputs of the classifiers as input to get the result according to some optimized rules. The other receives the feature vectors which are the input of the individual classifiers in the former approach and trains a classifier with all this information. Both of them achieve better performance than the individual classifier does. Here, we mainly evaluate the following three approaches. Simple arithmetic rules and fuzzy integral method are the former while multiple kernel learning presents the later.

5.1 Simple arithmetic combination

The following sum, maximum, and product rules are the simplest computationally efficient rules for combination. ($\mathcal{C}$ indicates the result of certain classifier and $\mathcal{O}$ means the label which is male or female):

$$\mathrm{Prob}_{\mathrm{sum}}(x_i = \mathcal{O}) = \sum_{t=1}^{T} \mathcal{C}_t(x_i; \mathcal{O}) \tag{2}$$

$$\mathrm{Prob}_{\mathrm{max}}(x_i = \mathcal{O}) = \max\{\mathcal{C}_1(x_i; \mathcal{O}), \ldots, \mathcal{C}_T(x_i; \mathcal{O})\} \tag{3}$$

$$\mathrm{Prob}_{\mathrm{prod}}(x_i = \mathcal{O}) = \prod_{t=1}^{T} \mathcal{C}_t(x_i; \mathcal{O}). \tag{4}$$

5.2 Fuzzy integral

The disadvantage of the above simple arithmetic combination rules is that they assume that all the component classifiers are mutually independent. However, the relationship between classifiers may be a positive or negative correlation. Sugeno [11] introduced the concept of fuzzy integral, which has become increasingly popular for multi-attribute classification. A general fuzzy measure is defined as follows:

Definition 1 A fuzzy measure μ defined on $X = \{x_1, x_2, \ldots, x_n\}$ is a set function μ: $P(X) \rightarrow [0, 1]$ (P(X) indicates the power set of X) satisfying:

(1) $\mu(\varnothing) = 0, \mu(X) = 1,$
(2) $A \subseteq B \Rightarrow \mu(A) \leqslant \mu(B)$

The fuzzy measure we adopt in this paper is the Choquet integral [30].

Definition 2 Let μ be a fuzzy measure on X. The discrete Choquet integral of a function $f : X \rightarrow R^+$ with respect to μ is defined by

$$C_\mu(f(x_1), \ldots, f(x_n)) \triangleq \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))\mu(S_i), \tag{5}$$

where i indicates that the indices have been permuted so that $0 = f(x_0) \leq f(x_1) \leq \cdots \leq f(x_n) \leq 1$ and $S(i) \triangleq \{x_i, \ldots, x_n\}$.

In our previous work [8], we used a fragment-based hair feature extraction method and a PCA-based facial feature extraction method with a fuzzy integration strategy to combine the outputs of eyes, nose, mouth, and hair classifiers. It greatly improved classification accuracy.

5.3 Multiple kernel learning

The two strategies above belong to the classifier combination method, that is, the final decision is made upon the integration of several individual classifiers. Each component classifier is trained to be at the local minimal position but not the minimal position of the composite one. Multiple kernel learning is a representative of feature combination methods. The combination of kernel functions gets better results for some practical requirements such as heterogeneous information and unnormalized data. The composite kernel is defined as follows:

$$K(w) = \sum_{i=1}^{n} w_i K_i, \qquad w_i \geq 0 \tag{6}$$

The composite kernel is a linear combination of the component kernels. This is an easy combination of kernel functions which is used in this article. Some validation data are prepared for training the coefficients of the component kernels. We use the SMO-MKL (sequential minimal optimization-multiple kernel learning) [31] code to train the composite model.

6 Experiment

6.1 Gender classification procedure

Here, we introduce the whole hierarchical classifier-based gender classification system (See Fig. 4).

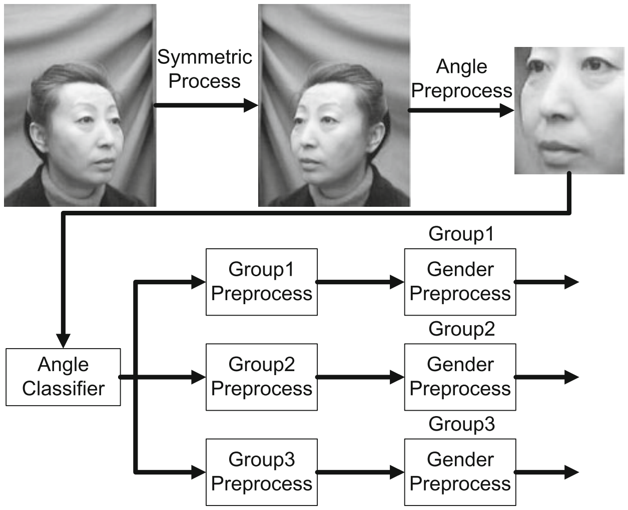

Fig. 4 The gender classification process using hierarchical classifiers structure

Alignment is important for gender classification based on facial images [32]. First, faces are fixed in the center of the result pictures. Facial components are in certain places for feature extraction after alignment. ASM is used to get the locations of eyes and mouth, and then we cut out a rectangle with the facial image.

The bilaterally symmetrical characteristic of the human face is used. The images facing right are turned to face left. Then, the images are classified according to the angle of the faces. 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-cut, using the information on angles which is the label of the category, in order to put most of the face into the picture, and align the organ positions precisely. Some information on the hair is also taken into the images for classification. Feature extraction is done in different categories, and the suitable gender classification is prepared for the facial images.

6.2 Data set

To compare the performance, we select the a gender classification problem based on multi-view facial images in the CAS-PEAL face database [33] (See Table 1). We take all the images labeled “PM” from the “POSE” section. We partition the data to assure that images of a person only appear in either the training or the test set. The total 7273 different-pose facial images are organized into 11 groups. The distributions of training and test data are shown in Table 1.

6.3 Implementation

The images in the data set are at many different angles. We want to use this prior knowledge to separate them into categories. However, as we have pointed out, too many categories will introduce complexity into the division problem. Therefore, some of the images at different angles must be put into the same category. We use the symmetric property of the human face to reduce the total number of angles from 11 to 6. For example, PM + 30 and PM − 30 combine to be PM ± 30, and PM + 45 and PM − 45 become PM ± 45.

The division illustrated in Fig. 5(a) looks quite natural. Given such a division, neither of the two corresponding classifiers would be able to solve the angle between the two regions. As a result, a classification error in the first layer probably leads to erroneous answers in the second layer, especially near the division of the two neighboring regions.

Table 1 Description of training and test data selected from the CAS-PEAL face database

Data set	Description	Total	Male	Female	Training	Test male	Test female	Test total
CAS-PEAL	PM − 67	101	79	22	11*2	67	11	79
	PM − 45	1,039	595	444	306*2	289	138	427
	PM − 30	938	516	422	295*2	221	127	348
	PM − 22	101	79	22	11*2	67	11	79
	PM − 15	938	516	422	295*2	221	127	348
	PM + 00	1,039	595	444	306*2	289	138	427
	PM + 15	938	516	422	295*2	221	127	348
	PM + 22	101	79	22	11*2	67	11	79
	PM + 30	938	516	422	295*2	221	127	348
	PM + 45	1,039	595	444	306*2	289	138	427
	PM + 67	101	79	22	11*2	67	11	79
	Total	7,273	4,165	3,108	4,284	2,023	996	2,989

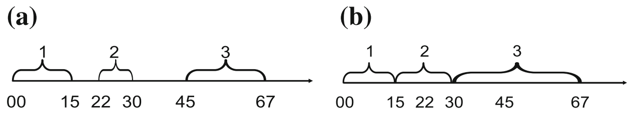

Fig. 5 Angle region division for gender classification: **a** naive way, 6 angles are divided equally into three groups; **b** the way to get all angles in to improve accuracy

In this paper, we use the division in Fig. 5(b). Images of PM00 and PM $\pm$ 15 are used for training the first classifier. Images of PM $\pm$ 15, PM $\pm$ 22, and PM $\pm$ 30 are for the second classifier. And PM $\pm$ 30, PM $\pm$ 45, and PM $\pm$ 67 are for the third. This will also help to reduce the risk of the accumulative error introduced by the classification error in the first layer. Since we want to compare the performance with and without symmetric processing, the second and third angle categories are each broken into two parts which are positive ones and negative ones when symmetric process is not performed. At that time, we have five angle categories. For example, the second angle category is broken like this: PM + 15, PM + 22, and PM + 30 for one part, PM − 15, PM − 22, and PM − 30 for another.

Facial components are cut up according to the position labeled by ASM. Each facial component is trained for an individual model which is prepared for the combination strategies. SVMs with probabilistic outputs [29] are used here. Simple combination strategies (sum, maximum, and product rule), fuzzy integral, and multiple kernel learning are used to combine the facial component information into the result of gender classification. As a baseline, for the gender classifier in each angle category, we also use the whole faces.

For feature extraction, we use multi-resolution local Gabor binary pattern (MLGBP) to extract the features of each facial image. The MLGBP features, which are the input of the SVM classifiers, are derived by combining multi-resolution analysis, Gabor characteristic, and uniform LBP histograms [9, 34].

The experiments also consider the performance of the symmetric process in tough conditions. We check the following three different cases(See Fig. 6): (a) Gaussian noise: add Gaussian white noise (with zero mean and 0.01 variance) to sample images. (b) occlude facial regions randomly: four subregions of thes face are randomly selected to be occluded with black rectangles with random sizes (ranging from 25 × 25 pixels to 35 × 35 pixels). (c) Occlude facial components randomly: one of eye, nose, mouth, and chin component regions is randomly selected to be occluded with a rectangle generated in the same.

The classifiers in this paper are SVMs with RBF kernels. All experiments were performed on a Pentium quad-core CPU (2.83 GHz) PC with 8 GB RAM.

6.4 Result

The effect of symmetry is shown in Table 2. The bold value in the table shows the best performance in each small data set. The later tables are displayed in the same style. We make use of the symmetry of the human face to reduce the originally 11 angles to 6 and 5 angle categories to 3 as mentioned in Sect. 6.3. The accuracy of classifying the facial images as facing left or right is 99.63%. Here, we compare the overall gender classification accuracies using the hierarchical classifiers framework. From angle classification to the final gender classification, the overall performance of classifiers which use facial components such as eye, nose, mouth, and chin or the whole face for feature extraction and combination strategies such as max, sum, product, and fuzzy integral are compared. If an incorrect classification result is received in the first layer of the

Table 2 Overall performance of gender classification with or without symmetric process

Method	With symmetric process (%)	Without symmetric process (%)
Eye	**90.63**	90.40
Nose	81.60	**82.03**
Mouth	**82.80**	82.23
Chin	**81.57**	79.76
Max	**93.71**	92.64
Sum	**93.88**	93.84
Product	**94.41**	94.01
Fuzzy	**93.47**	93.44
MKL	**95.68**	93.98
Face	**97.89**	95.65
Directly	**93.34**	92.31

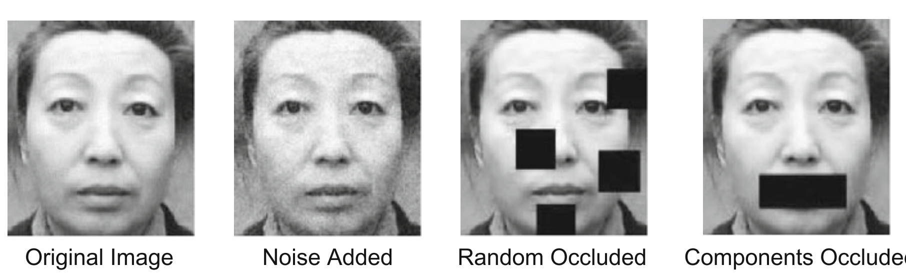

Fig. 6 Images with Gaussian white noise and artificial occlusions

hierarchical classifiers framework, which means the facial image is classified into the wrong angle category, we still treat the image as if it is really in that category and still get the gender classification result. "Directly" means we train the model using all the facial images without the hierarchical classifiers framework. The accuracy of the whole process including the symmetric transformation and gender classification is better than that without the process. This means the symmetric transformation provides the classification with less complexity and more accuracy and does no harm. In other methods, a symmetric process also helps in improving performance more or less. MKL performs the best in the overall accuracy comparison and the product rule ranks second. In tough conditions, the symmetric process also works in most cases which is shown in Table 3.

For the hierarchical classifiers framework, the results show that the additional classification step does not won't harm the accuracy but increases it. In the experiment, we find most of the testing data that are misclassified in the angle classification are facial images lying near the dividing line and being classified into the neighboring category. The selected classifiers are still suitable for these images and prepare better classification. Therefore, the accuracy of the gender classification with symmetric process and angle classification is close to the performance of the expert classifiers in their fields (See Table 4). In the original problem, we put all the images to the gender classifier regardless of the various angles, which is mentioned as "all angles in one model." It is a really difficult problem, revealed by the large amount of total support vectors in the model and the low accuracy. While the expert classifiers using angle information for the whole face get much better results. Also, the individual models can be trained in parallel to save much time.

Table 4 Gender model training using the entire facial images in different angle categories

Angle category	Accuracy (%)	Total SVs
PM00, PM ± 15	97.51	416
PM ± 15, PM ± 22, PM ± 30	98.26	423
PM ± 30, PM ± 45, PM ± 67	97.78	586
PM00, PM ± 15	96.35	361
PM + 15, PM + 22, PM + 30	97.55	255
PM − 15, PM − 22, PM − 30	97.68	244
PM + 30, PM + 45, PM + 67	98.71	249
PM − 30, PM − 45, PM − 67	97.42	250
All angles in one model	92.31	1,480

Expert classifiers which extract features only from facial components know less information about gender and they have worse performance than the ones knowing the whole face(See Table 5). All the results are obtained in the corresponding categories in the test data set. Eyes are the facial component that give the best performance. This indicates that eye regions provide the most discriminative information for gender, while others are comparatively less useful. The experiments show the fact that facial components actually carry the information of gender.

Then different combination strategies such as simple arithmetic (max, sum, and product rules), fuzzy integral, and multiple kernel learning are compared. Multiple kernel learning, Fuzzy integral, and product rules are better in experiments (See Table 6). Simple arithmetic strategies use the least time to combine the results from facial component classifiers. Multiple kernel learning has a large parameter space to search and needs more time. Although combinations of classifiers may lose some performance in classification, they are able to handle tough environments including illumination problem, occlusions, and some noise.

Table 3 Overall performance of gender classification with noise or occlusion

Noise	Method	With symmetric process (%)	Without symmetric process (%)
Noise	Max	**84.21**	83.34
	Sum	**81.43**	78.96
	Product	**83.87**	82.10
	Fuzzy	**81.93**	81.70
	MKL	**91.10**	87.52
Occlusion	Max	**87.86**	86.82
	Sum	**87.79**	87.19
	Product	**89.33**	88.16
	Fuzzy	**83.97**	83.71
	MKL	**89.70**	89.26
Organ occlusion	Max	**87.89**	83.84
	Sum	**88.32**	86.32
	Product	**88.99**	86.02
	Fuzzy	82.90	**83.81**
	MKL	**89.03**	88.93

6.5 Complexity analysis

As is well-known, 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 $O\left(\left(\frac{M}{K}\right)^p\right)$. Even if we run the training sequentially, it will only take $O\left(K\left(\frac{M}{K}\right)^p\right)$. Thus, time complexity is reduced in both situations.

Table 5 Gender model training with facial components in different angle categories

Angle Category	Eye (%)	Nose (%)	Mouth (%)	Chin (%)
PM00, PM ± 15	**92.88**	81.30	82.19	84.24
PM ± 15, PM ± 22, PM ± 30	**91.74**	84.13	83.03	84.32
PM ± 30, PM ± 45, PM ± 67	**90.75**	82.03	83.31	79.80
PM00, PM ± 15	**92.34**	82.19	82.37	83.79
PM + 15, PM + 22, PM + 30	**92.00**	82.19	84.00	82.84
PM − 15, PM − 22, PM − 30	**89.81**	83.74	83.48	81.42
PM + 30, PM + 45, PM + 67	**90.98**	82.08	81.03	78.22
PM − 30, PM − 45, PM − 67	**91.33**	80.21	80.56	75.88

Table 6 Performance comparison of combination strategies in test data set

Angle category	Max (%)	Sum (%)	Product (%)	Fuzzy (%)	MKL (%)
PM00, PM ± 15	96.17	96.05	**96.75**	94.77	95.74
PM ± 15, PM ± 22, PM ± 30	95.64	95.14	**95.98**	**95.98**	95.42
PM ± 30, PM ± 45, PM ± 67	92.72	93.39	**93.77**	91.74	92.97
PM00, PM ± 15	95.70	95.59	**96.28**	94.66	94.30
PM + 15, PM + 22, PM + 30	95.14	95.31	**95.98**	**95.98**	95.61
PM − 15, PM − 22, PM − 30	93.47	96.82	**95.98**	95.48	95.87
PM + 30, PM + 45, PM + 67	92.64	93.69	94.29	92.34	**95.31**
PM − 30, PM − 45, PM − 67	90.54	91.14	91.44	91.59	**92.15**

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. Feature extraction for facial components needs much less time than the whole face because of the small size of the images. Models can be trained fog each component in parallel. The combination of the classifiers really costs just a little. So the combination of facial components also saves time for gender classification.

7 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. Combinations of facial components are used for feature extraction instead of the whole face in order to get robust gender classifiers. We use the bilateral symmetry of the human face to simplify gender classification. Experimental results show the effectiveness of our framework and the facial symmetry trick. Complex backgrounds will adversely affect our feature extraction method. The future work is to handle this difficulty.

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

References

1. Golomb B, Lawrence D, Sejnowski T (1990) Sexnet: a neural network identifies sex from human faces. Proc Adv Neural Inf Process Syst 3:572–579
2. Brunelli R, Poggio T (1992) Hyperbf networks for gender classification. Proceedings of the DARPA Image Understanding Workshop, pp 311–314
3. Gutta S, Wechsler H (July 1999) Gender and ethnic classification of human faces using hybrid classifiers. In: Neural Networks, 1999. IJCNN '99. International Joint Conference on, vol 6, pp 4084–4089
4. Moghaddam B, Yang M (2002) Learning gender with support faces. IEEE Trans Pattern Anal Mach Intell 24(5):707–711
5. Kim H, Kim D, Ghahramani Z, Bang S (2006) Appearance-based gender classification with gaussian processes. Pattern Recogn Lett 27(6):618–626
6. Lian H, Lu B (2006) Multi-view gender classification using local binary patterns and support vector machines. Proceedings of the third international symposium on neural networks, pp 202–209
7. Baluja S, Rowley H (2007) Boosting sex identification performance. Int J Comput Vis 71(1):111–119

8. Lian X, Lu B (2009) Gender classification by combining facial and hair information. Proceedings of advances in neuro-information processing, pp 647–654
9. Xia B, Sun H, Lu B (2008) 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
10. Li B, Lian X, Lu B (2011) Gender classification by combining hair, clothing and face organ classifiers. to appear in Neuralcomputing
11. Sugeno M (1974) Theory of fuzzy integrals and its applications
12. Burton A, Bruce V, Dench N (1993) What's the difference between men and women? Evidence from facial measurement. PERCEPTION-LONDON-22:153–153
13. Fellous J (1997) Gender discrimination and prediction on the basis of facial metric information. Vis Res 37(14):1961–1973
14. Tamura S, Kawai H, Mitsumoto H (1996) Male/female identification from 8 Ũ 6 very low resolution face images by neural network. Pattern Recogn Lett 29(2):331–335
15. Cottrell G, Metcalfe J (1990) EMPATH: Face, emotion, and gender recognition using holons. In: Proceedings of the 1990 conference on advances in neural information processing systems 3, Morgan Kaufmann Publishers Inc. 571
16. Sun Z, Bebis G, Yuan X, Louis S (2003) Genetic feature subset selection for gender classification: a comparison study. In: Applications of computer vision, 2002. (WACV 2002). Proceedings. Sixth IEEE Workshop on, IEEE, pp 165–170
17. Jain A, Huang J (2004) Integrating independent components and linear discriminant analysis for gender classification. In: Automatic face and gesture recognition, 2004. Proceedings. Sixth IEEE international conference on, IEEE, pp 159–163
18. Zhang W, Shan S, Gao W, Chen X, Zhang H (2005) Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition. Tenth IEEE international conference on computer cision 1
19. Wu B, Ai H, Huang C (2003) Lut-based adaboost for gender classification. Proceedings of international conference on audio- and video-based biometric person authentication, pp 104–110
20. Sun N, Zheng W, Sun C, Zou C, Zhao L (2006) Gender classification based on boosting local binary pattern. Proceedings of 3rd International Symposium on Neural Networks, vol 2, pp 194–201
21. Toews M, Arbel T (2009) Detection, localization, and sex classification of faces from arbitrary viewpoints and under occlusion. IEEE Trans Pattern Anal Mach Intell 31:1567–1581
22. Takimoto H, Mitsukura Y, Fukumi M (2007) Robust gender and age estimation under varying facial pose. IEEJ Trans 127(7): 1022–1029
23. Ahonen T, Hadid A, Pietikainen M (2004) Face recognition with local binary patterns. Comput Vis Proc 3021:469–481
24. Ji Z, Lian X, Lu B (2009) Gender classification by fusion of face and hair feature. State of the Art in Face Recognition 215–230
25. Ji Z, Lu B (2011) A support vector machine classifier with automatic confidence and its application to gender classification. To appear in Neuralcomputing
26. Huang J, Shao X, Wechsler H (1998) Face pose discrimination using support vector machines (svm). In: Proceedings of 14th international conference on pattern recognition (ICPR98)
27. Brown UE, Perrett D (1993) What gives a face its gender. Perception 22:829–840
28. Cootes T, Taylor C, Cooper D, Graham J et al (1995) Active shape models-their training and application. Comput Vis Image Underst 61(1):38–59
29. Wu T, Lin C, Weng R (2004) Probability estimates for multiclass classification by pairwise coupling. J Mach Learn Res 5:975–1005
30. Murofushi T, Sugeno M (1989) An interpretation of fuzzy measures and the Choquet integral as an integral with respect to a fuzzy measure. Fuzzy Sets Syst 29(2):201–227
31. Vishwanathan SVN, Sun Z, Theera-Ampornpunt N, Varma M (December 2010) Multiple kernel learning and the SMO algorithm. In: Advances in Neural Information Processing Systems
32. Makinen E, Raisamo R (2008) Evaluation of gender classification methods with automatically detected and aligned faces. In: IEEE Trans. Pattern Analysis and Machine Intelligence, vol 30(3), pp 541–547
33. Gao W, Cao B, Shan S et al (2008) The cas-peal large-scale chinese face database and baseline evaluations. IEEE T Syst Man Cy A 38(1):149–161
34. Lian H, Lu B (2007) Multi-view gender classification using multi-resolution local binary patterns and support vector machines. Int J Neural Syst 17(6):479–487