Selecting Optimal Orientations of Gabor Wavelet Filters for Facial Image Analysis *

Tianqi Zhang[1] and Bao-Liang Lu[1,2]

[1]Dept. of Computer Sci. & Eng., Shanghai Jiao Tong University, Shanghai 200240, China
[2]MOE-Microsoft Key Laboratory for Intelligent Computing and Intelligent Systems
Shanghai Jiao Tong University, Shanghai 200240, China
{ztq121, bllu}@sjtu.edu.cn

Abstract. Gabor wavelet-based methods have been widely used to extract representative features for face analysis. However, the existing methods usually suffer from high computational complexity of Gabor wavelet transform (GWT), and the Gabor parameters are fixed to a few conventional values which are assumed to be the best choice. In this paper we show that, for some facial analysis applications, the conventional GWT could be simplified by selecting the most discriminating Gabor orientations. In the selection process, we analyze the histogram of oriented gradient (HOG) of the average face image in a dataset, and eliminate the less significant orientation combinations. Then we traverse the rest combinations and select the best according to classification performance. We find that the selected orientations match the analysis of HOG well, and are therefore consistent with the intrinsic gradient characteristics of human face images. In order to assess the performance of the selected Gabor filters, we apply the proposed method to two tasks: face recognition and gender classification. The experimental results show that our method improves the accuracy of the classifiers and reduces the computation cost.

1 Introduction

Facial image analysis plays an important role in human computer interaction studies because the machine can detect a lot of useful information of their users from the face images, such as identity, age, gender and expression, which makes existing human computer interaction systems more intelligent and easy to use. However, building an automatic system to detect such information is a challenging problem which attracts a great deal of research interests in the computer vision realm. The key to this problem is to extract the most representative features from facial images for classification.

Gabor wavelet transform (GWT) [1] [2] was first introduced as a model of the simple cell in human visual cortex. It derives desirable facial features characterized by spatial frequency, spatial locality, and orientation selectivity to cope with the variations due

* 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).

to illumination and facial expression changes. Gabor-based features have been widely used in many face analysis tasks, such as face recognition [3] [4] and gender classification [5].

However, the huge dimensionality of Gabor feature can be a disaster to statistical learning and pattern classification methods. In order to solve this problem, some subspace methods have been introduced to reduce the dimension of Gabor feature. For face recognition, Liu and Wechsler applied Fisher discrimination model [4], independent component analysis (ICA) [6] and principal component analysis (PCA) [3] with an evenly spaced Gabor filter set of 40 filters (5 scales $\times$ 8 orientations). An alternative way to bypass the dimensionality problem is to use non-statistics based face representation, as Zhang *et al.* did with their local Gabor binary pattern histogram sequence (LGBPHS) method [7] for face recognition. Later Xia *et al.* [8] managed to map the LGBPHS feature into a low-dimensional space with LDA-like method called local Gabor binary mapping pattern (LGBMP), so that classic statistical classifiers like SVM can be used for gender classification.

Although subspace methods significantly reduce the dimension of Gabor features, these Gabor-based methods still suffer from high computational complexity of Gabor wavelet transform and are incompetent for speed-sensitive scenarios. Moreover, the popular Gabor parameters, 5 scales $\times$ 8 orientations, have been assumed to be the best choice in many studies ([3] [4] [6] [7] [8]) without careful discussion and examination on their performance. A few efforts [9] [10] have been made to reduce the number of filters. In [9] the Information Diagram concept was adopted for Gabor parameter selection, and in [10] a minimum set of Gabor filters was arranged in space to cover the frequency plane. However in these two studies, the intrinsic characteristics of face image were not considered.

In this study, we propose a method to reduce the number of Gabor filters while maintaining and even improving classification accuracy. Our work is motivated by the intrinsic gradient distribution of human face images and the redundancy of the Gabor filters when using equally divided orientations. We analyze the histogram of oriented gradient (HOG) of the average face image in a dataset, and eliminate the less significant orientation combinations. Then we traverse and select the rest combinations according to cross validation accuracy. The selection process is quite simple, but also very effective and consistent with face gradient distribution. The performance and generalization of the selected Gabor filters is evaluated in face recognition and gender classification experiments on different datasets. The results show that the selected Gabor filters are optimal and can represent the face images better with less computation cost. We also examined the generalization capability of our optimal Gabor filters by applying the image Euclidean distance (IMED) [11] transform on input images. IMED has been used to gain tolerance to small deformations of face images [12] [13]. The prior selected Gabor filters also show a steady performance after the IMED transform. The details of our selection method is discussed in Sec. 2, and the evaluation methods and results on public datasets are given in Sec. 3.

2 Gabor Filter Selection

The motivation of using Gabor-based feature is mostly biological, since Gabor-like receptive fields have been found in the visual cortex of primates [14] [15] [16]. These neurological studies partially support the evenly spaced Gabor filters that are commonly used in a variety of computer vision tasks. The research in [16] proved that the Gabor wavelet representation is optimal in the sense of minimizing the joint two-dimensional uncertainty in space and frequency. Field also suggested that the evenly tuned orientations and spatial-frequencies of mammalian simple cells are useful to cope with the uncertainty of natural images by preserving a certain level of redundancy and allow information to be transmitted with only a subset of the total number of cells [15].

This kind of redundancy is crucial for mammalian due to the variation in scaling, rotation, shift and illumination of real world objects, but is not necessarily useful for computer vision problems because we usually have aligned and normalized face images as inputs. Therefore most of the rotation uncertainty can be avoided, and the redundancy of Gabor orientations can be reduced. We call it the Gabor filter selection because the reduction is done by selecting the best orientation combination in the parameter space.

2.1 Gabor Wavelet Filters

Based on Gabor wavelet transform, a family of Gabor kernels is defined as[3]:

$$\psi_{\mu,\nu}(z) = \frac{||k_{\mu,\nu}||^2}{\sigma^2} e^{-\frac{||k_{\mu,\nu}||^2 ||z||^2}{2\sigma^2}} \left[e^{(ik_{\mu,\nu}z)} - e^{\left(-\frac{\sigma^2}{2}\right)} \right], \tag{1}$$

where μ and ν define the orientation and scale of the Gabor kernels, $z = (x, y)$, and $k_{\mu,\nu} = k_\nu e^{i\phi_\mu}$ is the wave vector, in which $k_\nu = k_{max}/f^\nu$ and $\phi_\mu = \pi\mu/8$. k_{max} is the maximum frequency and f is the spacing factor between wavelets in the frequency domain. Traditionally, $\nu \in \{0 \dots 4\}$ and $\mu \in \{0, 1, \dots, 7\}$ are used for face images. In the following we will show that evenly divided ϕ_d is redundant and can be improved for better performance.

2.2 Facial Statistics Analysis

The Gabor filters can be considered as orientation and scale tunable edge and line detectors, since they only respond to some specifically oriented textures in some specific scale. In [15], it was suggested that the efficiency of a code method depends on the statistics of the inputs. Therefore by analyzing the intrinsic gradient characteristics of human face images, we can use fewer but specially tuned Gabor filters to acquire more discrimination capability and avoid unnecessary noises.

We borrow the histogram of oriented gradient (HOG) [17] technique to analyze the significant gradient orientations of face images. Given a dataset, we generate its average face A, and then apply the Sobel operators [18] on A to get the horizontal and vertical gradient images G_x and G_y, and the orientation image G_Θ:

$$G_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * A, G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A \tag{2}$$

$$G_\Theta = \arctan(\frac{G_y}{G_x}) \tag{3}$$

The histogram of G_Θ reflects the general orientation distribution of the images in the dataset. To make the Sobel operation region the same size as the Gabor window, the average face A is firstly smoothed by a 3×3 mean filter.

2.3 Straightforward Filter Selection

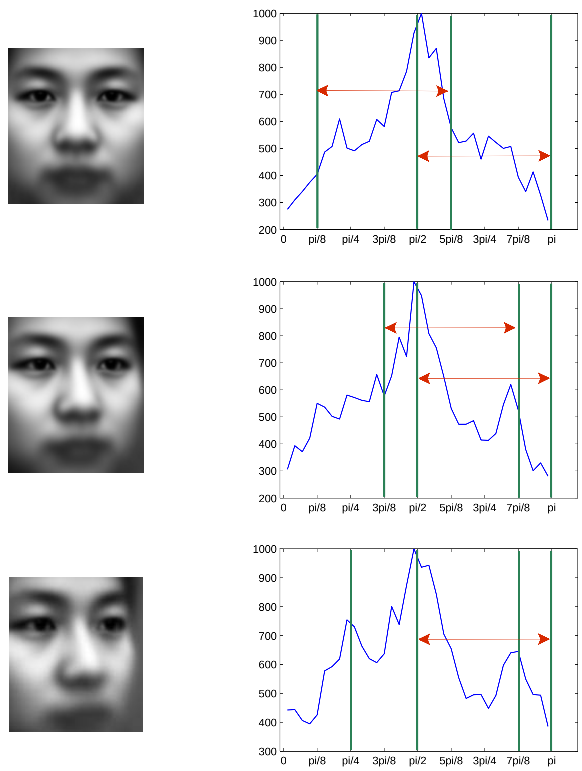

Fig. 1. Three examples of average face images (left) and the corresponding HOGs (right). Datasets used here from top to bottom are C-PM00, C-PM15 and C-PM30, respectively.

The HOG analysis of the average face can reveal the statistics of gradient distribution in a dataset. But HOG analysis is not enough for selecting optimal Gabor filters. On the one hand, the HOG of a dataset does not always have a proper number of obvious peaks. The curve can have various shapes and is hard for us to identify the most significant orientation combination. Thus, the meaning of HOG analysis could only be used as reference. On the other hand, and more importantly, we should also consider the orthogonality of orientations other than the HOG significance. Evidence has been shown that adjacent simple cells demonstrate an orthogonal phase-selective property [19]. And according to [15], as long as the two phase relations are in quadrature (i.e., they differ by 90°), it is not critical what phases are involved.

To take the orthogonality into consideration, we should restrict the selection in an orthogonal orientation space, like the commonly used $[0, \pi/8, 2\pi/8, 3\pi/8, \ldots, 7\pi/8]$. We here use a mixed method to select optimal orientations automatically from this space. Assume the mean intensity of the HOG vector is μ_H. For all k-orientation combinations $\{\delta_k^i | i = 1, 2, \ldots, \binom{8}{k}\}$ in the orientation base, we examine the average HOG intensity of the ith combination, μ_k^i, and narrow our selection base down to set $\Omega = \{\delta_k^i | \mu_k^i \geq \mu_H\}$. In Ω, a traversal is carried out and the combination with the highest classification performance is chosen as the optimal orientations. Ω is usually fairly small and can be traversed in reasonable time. Next we will show that the result of this selection is confirmed by both the shape of facial HOG curve and the classification performance.

Fig. 1 shows the average faces we generate from three different poses (facing front, deflection angle from 0 to 30°) and the HOGs (blue curve) of them. The gradient distribution peaks imply that the most significant gradient changes occur in these orientations. The vertical green lines indicate the best 4 orientations we find from the $\binom{8}{4}$ combination space. Note that lines at '0' are drawn at 'π'. We can see that the green lines are near the peaks of the HOG curve, indicating that Gabor filters specially tuned to these orientations do indeed produce better features for classification. Red double arrows represents the orthogonal relationships between orientations, which coincides with the observations in [19] and [15] about the orthogonal property in adjacent simple cells. One exception in Fig. 1 is pose C-PM30, which sacrifices one pair of orthogonal orientations for better orientation significance. This also proves that our selection method pays overall consideration to both orientation significance and orthogonality.

The matching of the selected orientations to the HOG peaks indicates that this kind of statistical characteristics can be useful for classification. In fact, according to our gender classification experiment, the two best orientations already contain as much gender information as all 8 orientations do.

3 Experiments

3.1 Set-up

In this section we assess the performance of the selected Gabor filters by both face recognition and gender classification experiments. The framework of the feature extraction method used for comparison is described in this section. For face recognition, we follow the Gabor-based PCA method used in [3], and for gender classification we adapt the more recent framework using LGBP features [5,8] that originated with the work of [7].

In the pre-processing step, with eye positions located, a face image is firstly cropped and normalized. And the result is then fed to the selected Gabor filters to produce Gabor magnitude pictures (GMPs). For face recognition, the GMPs are down-sampled by a factor ρ and normalized according to [3] and finally fed to PCA to generate lower-dimensional features. While for gender classification, GMPs is converted to LGBP images with $LBP_{8,1}^{u2}$ operator, and then each is divided into 10×10 non-overlapping rectangular regions, so that they can be mapped to LGBMP features, as proposed by Xia *et al.*[8].

The optimal orientations used in Sec. 3 are selected prior to the experiment by the unsupervised process mentioned in Sec. 2 according to gender classification cross validation accuracy on CAS-PEAL[20] dataset. Later the same orientation set was used for face recognition experiment on FERET[21] dataset in Sec. 3.3 and on the IMED[11] transformed images in Sec. 3.4 to show the generalization of the optimal orientations selected by our method.

3.2 Dataset

For frontal face recognition, we conduct our experiment on FERET[21] database with 600 frontal face images belonging to 200 subjects. In every three images that correspond to the same subject, one is taken under different illumination and another one in different facial expression. We randomly chose two images of one subject for training and leave the rest one image for testing. This results in a training set of 400 images and a test set of 200 images.

For pose-angled gender classification, we use both FERET frontal and CAS-PEAL[20] databases to compose multiple datasets of different poses. CAS-PEAL is a large-scale face database that currently contains 30,864 facial images with 9 different poses (vertically facing up, middle and down, each coupled with a horizontal deflection angle of 0°, 15°and 30°), and some sample images are given in Fig. 2. For each pose and also for frontal images in FERET we randomly chose about 600 training images and leave about 400 images for testing. The detailed numbers are shown in Table. 1.

	Trainging	Test
Dataset	Male/Female	Male/Female
C-PD00	311/311	284/134
C-PD15	296/296	220/127
C-PD30	295/295	220/127
C-PM00	282/282	285/134
C-PM15	310/310	221/127
C-PM30	295/295	221/127
C-PU00	311/311	284/134
C-PU15	296/296	220/127
C-PU30	296/296	220/127
FERET	282/282	307/121

Table 1. Training and test data for gender classification, and the male/female composition of each set is given. 'C-PD00' means CAS-PEAL subset, facing down with a deflection angle of 0°.

Fig. 2. Normalized sample images from CAS-PEAL dataset in different poses. From top to bottom: C-PU, C-PM and C-PD; from left to right: angle 0°, 15°and 30°.

3.3 Face Recognition

For comparison purpose, we implemented the Gabor-based PCA method from [3] using 5 scales and 8 orientations (5×8) of Gabor filters, and then replaced them with the independently selected filter set (5×4). Fig. 3 shows the performance of these two different sets of Gabor filters. We can see that the recognition rate of selected Gabor filters is slightly better than that of the full 5×8 set. In particular, the selected Gabor filters achieves 100% correct face recognition accuracy using 365 features, while the full 5×8 set achieves 98.5%. This comparison demonstrates both the effectiveness and generalization performance of selected Gabor filters.

Another advantage of selecting Gabor filters is that it can reduce the dimension and computing time of raw Gabor features, which is very helpful for speed-sensitive applications. Subspace methods like PCA can benefit from lower-dimensional features as well, since solving the eigenvalue problem of a huge matrix is extremely time-consuming. For the $5 \times 8 = 40$ filters, the dimension of Gabor features could be as much as 10,000 even after down-sampling, which will result in a PCA coefficient matrix of 10000×10000. But with selected Gabor filters, we can reduce the size to about 5000×5000. The summarized comparison is in Table 2, and the numbers may be approximative.

3.4 Gender Classification

We employ the LGBMP method [8] as the baseline of our experiments, and the parameters are set to 5 scales and 8 orientations. Then we select the best $1, 2, \ldots, 7$ orientations from the full 8 family to show the advantage of selected filters. Fig. 4 shows that the best accuracy comes with 4 or 5 orientations, which corresponds with the number of

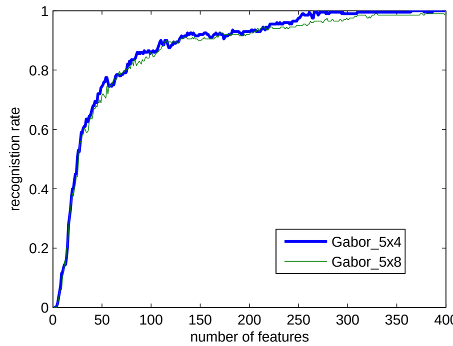

Fig. 3. Face recognition performance of the PCA method with two different sets of Gabor filters: Gabor_5×4 and Gabor_5×8.

significant peaks in the HOG of the corresponding dataset. In fact, the 4 or 5 orientations are also a trade-off between keeping information and avoiding noises. Too few filters can not give enough information for classification, while too many filters may bring too much noises. As mentioned before, Fig. 4 indicates that we can use only 25% of the original features while maintaining a competitive performance. That is from a dimension of 4000 to only 1000, for our $5 \times 8 \times 10$ baseline. And the computing time can also be reduced greatly to nearly real-time. The detailed comparison is shown in Table. 2. Fig. 4 also demonstrates that the effectiveness of our filter selection method is not affected by IMED transform.

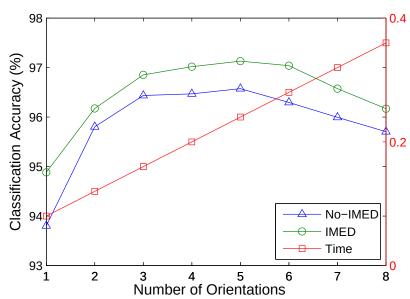

Fig. 4. The average accuracy over all 10 datasets using different orientation numbers and with/without IMED transform. The approximate computing time is also given.

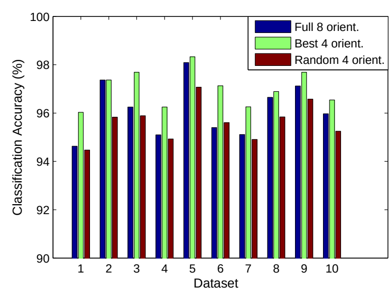

Fig. 5. The accuracy with different filter sets on each dataset. 1:FERET; 2:C-PD00; 3:C-PD15; 4:C-PD30; 5:C-PM00; 6:C-PM15; 7:C-PM30; 8:C-PU00; 9:C-PU15 and 10:C-PU30.

Fig. 5 shows the effect of orientation selection on every dataset we used. The average accuracy of randomly chosen orientations is also given for comparison. An interesting fact is that orientation selection brings less improvement on C-PD00, C-PM00 and C-PU00, which are also the datasets with less significant HOG peaks. Actually when the face deflection angle is 0°, the image is horizontally symmetrical and so is the HOG curve, therefor the characteristic orientations may not be as obvious as the ones of other datasets.

Experiment	No. of Filters	Feature Dim.	Time	Recog.Rate/Accuracy
Face Recognition	$5 \times 8 = 40$	10000	0.36s	98.5%
	$5 \times 4 = 20$	5000	0.18s	100%
Gender Classification	$5 \times 8 = 40$	4000	0.4s	96.17%
	$5 \times 4 = 20$	2000	0.2s	97.02%
	$5 \times 2 = 10$	1000	0.1s	96.17%

Table 2. Performance comparison of optimal Gabor orientations and full Gabor family. Time is measured on a Intel Pentium 4 2.8GHz PC.

4 Conclusion

This paper presents a novel criterion to select the most discriminating Gabor filters for facial image analysis. We first exclude the less significant orientation combinations according to HOG analysis and then traverse the rest to choose the best combinations. The result of this method has been shown to match the analysis of image HOG well, and is therefore consistent with the intrinsic gradient characteristics of human face images. We evaluated our selected Gabor filters with face recognition and gender classification tasks, and the experimental results show that the selected Gabor filters are optimal and can represent the face images better using less computational time.

References

1. Daugman, J.: Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. Journal of the Optical Society of America A **2** (1985) 1160–1169
2. Marcelja, S.: Mathematical description of the responses of simple cortical cells. Journal of the Optical Society of America **70** (1980) 1297–1300
3. Liu, C.: Gabor-based kernel PCA with fractional power polynomial models for face recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence **26** (2004) 572–581
4. Liu, C., Wechsler, H.: A Gabor feature classifier for face recognition. In: Proceedings of IEEE International Conference on Computer Vision. Volume 2. (2001)
5. Lian, H., Lu, B.: Multi-view gender classification using local binary patterns and support vector machines. Lecture Notes in Computer Science **3972** (2006) 202
6. Liu, C., Wechsler, H.: Independent component analysis of Gabor features for face recognition. IEEE Transactions on Neural Networks **14** (2003) 919–928
7. Zhang, W., Shan, S., Gao, W., Chen, X., Zhang, H.: Local Gabor binary pattern histogram sequence (LGBPHS): A novel non-statistical model for face representation and recognition. In: Proceedings of ICCV. (2005) 786–791

8. Xia, B., Sun, H., Lu, B.: Multi-view gender classification based on local Gabor binary mapping pattern and support vector machines. In: Proceedings of IEEE International Joint Conference on Neural Networks. (2008) 3388–3395
9. Moreno, P., Bernardino, A., Santos-Victor, J.: Gabor parameter selection for local feature detection. In: Proceedings of IbPRIA, Springer (2005)
10. Wu, H., Yoshida, Y., Shioyama, T.: Optimal Gabor filters for high speed face identification. vectors **3** (2002) 5
11. Wang, L., Zhang, Y., Feng, J.: On the euclidean distance of images. IEEE transactions on pattern analysis and machine intelligence **27** (2005) 1334–1339
12. Li, J., Lu, B.: An adaptive image Euclidean distance. Pattern Recognition **42** (2009) 349–357
13. Zhu, S., Song, Z., Feng, J.: Face recognition using local binary patterns with image Euclidean distance. In: Proceedings of SPIE. Volume 6790. (2007) 67904Z
14. Burr, D., Morrone, M., Spinelli, D.: Evidence for edge and bar detectors in human vision. Vision Research **29** (1989) 419–431
15. Field, D.: Relations between the statistics of natural images and the response properties of cortical cells. Journal of the Optical Society of America A **4** (1987) 2379–2394
16. Jones, J., Palmer, L.: An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. Journal of Neurophysiology **58** (1987) 1233
17. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Volume 1., Citeseer (2005) 886
18. Sobel, I., Feldman, G.: A 3x3 isotropic gradient operator for image processing. Presentation for Stanford Artificial Project (1968)
19. Pollen, D.A., Ronner, S.F.: Phase relationships between adjacent simple cells in the cat. Science **212** (1981) 1409–1411
20. Gao, W., Cao, B., Shan, S., Chen, X., Zhou, D., Zhang, X., Zhao, D.: The CAS-PEAL large-scale Chinese face database and baseline evaluations. IEEE Transactions on Systems Man and Cybernetics Part A Systems and Humans **38** (2008) 149
21. Phillips, P., Wechsler, H., Huang, J., Rauss, P.: The FERET database and evaluation procedure for face-recognition algorithms. Image and Vision Computing **16** (1998) 295–306