# A COMPARATIVE STUDY OF GABOR FEATURE AND GRADIENT FEATURE FOR HANDWRITTEN CHINESE CHARACTER RECOGNITION

**KAI DING, ZHIBIN LIU, LIANWEN JIN, XINGHUA ZHU**

College of Electronic Information, South China University of Technology, Guangzhou Guangdong 510641
E-Mail: lianwen.jin@gmail.com

## Abstract

*Gabor feature and Gradient feature have been proven to be two most efficient features for handwritten character recognition recently. However, few comprehensive comparative researches on the performance of these two methods in large scale handwritten Chinese character recognition (HCCR) were reported in the literature. In this paper, we compare these two methods for large scale HCCR. Some new interesting conclusions were obtained through this study. The result showed that the performance of gradient feature significantly outperforms Gabor feature. Multi-channel Gabor feature can improve the performance of single-channel Gabor feature. We also observed that the recognition accuracy hardly get obvious increasing after the number of directions of Gabor feature achieves 5. That means only 5 directional parameters of Gabor features are necessary to achieve a good enough performance, rather than 8 directions which are widely used in the literature.*

**Keywords:** Gradient feature; Gabor Feature; Feature Extraction; Handwritten Chinese Character Recognition (HCCR)

## 1. Introduction

Handwritten character recognition (HCR) technology is one of the most pivotal parts in Pattern recognition field. Selection of a feature extraction method is probably the single most important factor in achieving high recognition performance. Numerous methods of feature extraction were proposed in reference [1], [2]. Among these features, the directional features, such as directional element features, edge directional features, contour directional features, are widely used and have been previously proved to be very efficient [3,12]. But the best performance of gradient feature [4] outperforms various direction features demonstrated by previous researches [5], [6].

A powerful competitor to the direction feature is Gabor feature, first proposed by Daugman [7] .A Gabor filter is a kind of local narrow band pass filter and selective to both orientation and spatial frequency. It is suitable for extracting the joint information in two-dimensional spatial and frequency domain[8]，and widely applied in the fields such as character recognition, face and texture recognition, etc. Wang et al. [9] compared the Gabor feature with the contour direction (chain code) feature. The results showed that the contour feature performs much inferior than Gabor features for handwritten character recognition.

Gabor feature and gradient feature have some common properties: they are applicable to both binary images and gray-scale images, and are immune to image noise. Liu etc [10] compared these two methods in handwritten digit and printed Japanese character recognition. However, few comprehensive comparative researches have been investigated on the performance of Gabor feature and gradient feature in large scale handwritten Chinese character recognition (HCCR), which is discussed in this paper. Some interesting experimental conclusions are obtained as well.

In reference [10], the results show the performance of Gabor feature and gradient feature for handwritten digit and printed Japanese recognition are similar. But in this paper, we notice that the performance of gradient feature significantly outperforms Gabor feature in large scale handwritten Chinese character recognition. The analysis for this different result is given in section 5.4.

The rest of this paper is organized as follows. Section 2 and Section 3 introduce the methods of the Gabor feature extraction and the gradient feature extraction respectively. In Section 4, elastic meshing methods is presented and all the experiments are described in Section 5. Finally in Section 6 we summarize the paper.

## 2. Gabor feature extraction

### 2.1. Single-channel Gabor feature extraction

The two-dimensional Gabor filter function is presented as follows:

$$G(x,y,\lambda,\delta_x,\delta_y)=\exp\{-\frac{1}{2}\left[\frac{R_1^2}{\delta_x^2}+\frac{R_2^2}{\delta_y^2}\right]\}\bullet\exp\left[i\bullet\frac{2\pi R_1}{\lambda}\right] \quad (1)$$

Where $R_1=x\bullet\cos\varphi_k+y\bullet\sin\varphi_k$ ， $R_2=-x\bullet\sin\varphi_k+y\bullet\cos\varphi_k$ and $\varphi_k=\pi k/D$ $(k=0,1,...D-1)$ is the oscillation orientations, $D$ is the number of directions, $\lambda$ is wave length，and $\delta_x$, $\delta_y$ are the standard deviations (corresponding to scale) of the Gaussian envelope.

1-4244-1066-5/07/$25.00 ©2007 IEEE

To extract Gabor feature, we assume that $f(x, y)$ is the pixel value at the point $(x, y)$ in two-dimensional character image, whose size is $M\times N$ (in this paper, all the character images are normalized to 64×64). Let the point $(x_m, y_n)$ (obtained by method of elastic meshing described in section 4) denotes the sample point. Then the Gabor feature at the sample point $(x_m, y_n)$ is represented as:

$$\begin{aligned} f_{ga}(x_m, y_n) &= f(x, y) * G(x, y, \lambda, \delta_x, \delta_y) \\ &= \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x, y) G(x - x_m, y - y_n, \lambda, \delta_x, \delta_y) \end{aligned} \tag{2}$$

In practice, its amplitude is used as the feature, that is:

$$F_{ga}(x_m, y_n) = \left| \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x, y) G(x - x_m, y - y_n, \lambda, \delta_x, \delta_y) \right| \tag{3}$$

To obtain the Gabor features of character images, the parameters $\varphi_k$, λ, $\delta_x$, $\delta_y$ must first be determined. It is usually considered $\delta_x = \delta_y = \delta = \pi$ as a constant. The optimal value of $\varphi_k$ (or $D$) and λ will be determined by the following experiments described in Section 5.3.

### 2.2. Multi-channel Gabor feature extraction

As originally proposed by Wang etc [9], as Gabor filter is sensitive to the width of character strokes, and the value of λ reflects the information of stroke width, the parameter λ has a significant impact on the performance of Gabor feature. The Gabor filter proposed in section 2.1 can only be set with a single value of λ, which means this Gabor filter is sensitive only to certain stroke width. Because of the final touch and normalization process, the width of off-line handwritten character's strokes vary greatly（illustrated in figure 5）. Therefore the performance of the single-channel (singe λ value) Gabor filter might not very well for off-line handwritten Chinese character recognition.

To overcome this disadvantage, the multi-channel Gabor feature extraction is proposed: several single-channel Gabor filters with different value of parameter λ are combined into a multi-channel Gabor filter. Therefore multi-channel Gabor filter can be sensitive to several stroke widths and it is expected that its performance would outperforms that of single-channel Gabor feature.

## 3. Gradient feature extraction

To extract Gradient feature, 3×3 Sobel operators (illustrated in Figure 1) are used to obtain the horizontal and vertical gradient at each image pixel respectively.

| | | |
|---|---|---|
| -1 | -2 | -1 |
| 0 | 0 | 0 |
| 1 | 2 | 1 |

| | | |
|---|---|---|
| -1 | 0 | 1 |
| -2 | 0 | 2 |
| -1 | 0 | 1 |

Figure 1 3×3 Sobel operators

Supposed f(x, y) is the grayscale level of point (x, y), the horizontal and vertical grayscale gradient are derived as:

$$\begin{aligned} g_x &= f(x-1, y+1) + 2\bullet f(x, y+1) + f(x+1, y+1) \\ &\quad - f(x-1, y-1) - 2\bullet f(x, y-1) - f(x+1, y-1) \end{aligned} \tag{4}$$

$$\begin{aligned} g_y &= f(x+1, y-1) + 2\bullet f(x+1, y) + f(x+1, y+1) \\ &\quad - f(x-1, y-1) - 2\bullet f(x-1, y) - f(x-1, y+1) \end{aligned} \tag{5}$$

Then we define $L$ directions with an equal interval $2\pi/L$, and decompose the gradient vector $(g_x, g_y)$ into its two nearest directions in a parallelogram manner, as illustrated in Figure 2 (in the figure, $L$ is denoted as 8). This decomposition method is first proposed by Liu etc [11].

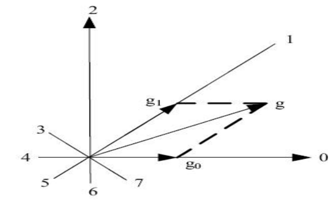


Figure 2 decomposing of the gradient vector

In this way we obtain an $L$-dimensional gradient code at each image pixel. Then we divide the 64×64 normalized image into 8×8 sub-blocks through elastic meshing, the middle of each sub-block is denoted as the sampling point to extract L Gradient features, resulting in a $d$-dimensional Gradient feature vector, denoted as $F_{gr}$, where $d=8\times 8\times L$. Finally a nonlinear transformation Y= $F_{gr}^{\ 0.4}$ is applied on each element of the extracted feature vector to make its distribution more Gaussian-like [9]. In the processes above, the direction value $L$ is the only parameter in the function of Gradient feature. In this paper we set $L$=8.

## 4. Elastic meshing

Meshing method is the region partition for character images with imaginary grids. If the character images are equally partitioned by linear grids, the meshing is denoted as linear meshing (illustrated in Figure 3(a)). However, as the styles of characters with different writer are varied greatly, the performance of linear meshing method is not very well. To overcome this disadvantage, the global elastic meshing [12] and local meshing (illustrated in Figure 3(b) and (c), which equally global or local distribute the black pixels in each row and column respectively, are constructed.

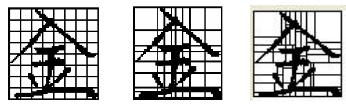
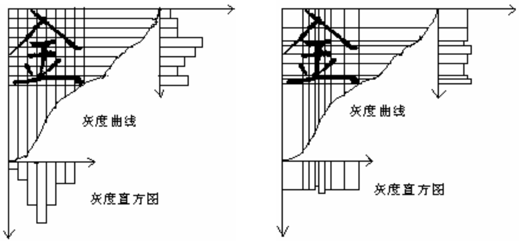

(a) (b) (c)

Figure 3. Linear meshing and global elastic meshing.
(a) 8×8 linear meshing
(b). 8×8 overall elastic meshing
(c) 8×8 local elastic meshing

The histograms of the image processed by linear meshing method and elastic meshing method are illustrated in figure 4. As it's shown in the figure, elastic meshing method equally compartmentalize the distribution of histogram, therefore, it can shape and standardize the images of the handwritten character and intolerant the character's local deformation and stretching, which are caused by different writing styles.




4-a：histogram of linear meshing    4-b：histogram of global elastic meshing

Figure 4 histograms of linear meshing and elastic meshing

## 5. Experiments

### 5.1 Experimental data

The experimental data is HCL2000 database, which is collected by Beijing University of Posts and Telecommunications for China 863 project. It contains 3,755 frequently used simplified Chinese characters, and all the character images are normalized to 64×64 pixels. A part of experimental samples is illustrated in figure 5.

Figure 5 a part of samples in HCL2000

To reduce the running time and improve efficiency, the numbers of samples in different experiment are distinct. For example, we use 100 sets for training and 30 sets for testing in the experiments for optimal parameters determination for Gabor feature. To obtain more reliable experimental results in the pivotal experiments, such as performance comparison between Gabor feature and Gradient feature, we use 300 sets for training and 100 sets for testing respectively. It is worth noting that, because the training sets and testing sets in different experiments are distinct, the accuracies of recognition, obtained by the same method, are also alien in different experiments. However, it will not affect the comparative study of the experiments and the conclusions of this paper.

### 5.2 Performance comparison against different meshing methods

In this section, we will compare the performance of these three methods described in section 4 under Gabor feature extraction and Gradient feature extraction. We use 100 sets for training and 30 sets for testing. For Gradient features, the direction number we used is 8 and the original dimension of the feature vector is 512, and we use LDA (Linear Discriminat Analysis) to reduce the dimension to 256, and the classifier is a minimum distance classifier. The result is illustrated in table 1. For Gabor feature, we don't use LDA to reduce dimension. Alternatively, we extract two kinds of Gabor features with the direction number of 8 and 4, the original dimensions are 512 and 256, respectively. The result is given in table 2.

Table 1 performance comparison against different meshing methods under Gradient features

| Meshing | Original feature | LDA feature |
|---|---|---|
| linear meshing | 76.3106% | 80.4954% |
| global elastic meshing | **90.8078%** | 93.4735% |
| local elastic meshing | 90.3980% | **94.083%** |

Table2 performance comparison against different meshing methods under Gabor features

| Meshing | 256-dimensional | 512-dimensional |
|---|---|---|
| linear meshing | 73.6618% | 75.6201% |
| overall elastic meshing | 85.7221% | 87.5667% |
| local elastic meshing | **85.9771%** | **88.0639%** |

From table 1, we can clearly see that the performance of elastic meshing significantly outperforms linear meshing and local elastic meshing is a litter better than overall elastic meshing under the Gradient feature extraction with LDA transformation. From table 2, we can also draw the same conclusion that local elastic meshing is the best in these three methods, no matter the dimension is large or small.

The reason is that the linear meshing can not regulate the deformation and displacement of strokes. Due to the good performance of local elastic, it is applied in all of the following experiments.

### 5.3 Optimal parameters determination for Gabor feature

As illustrated in section 2, there are two parameters for Gabor feature, say, $\varphi_k$ (or $D$) and $\lambda$. In this section we will determine the optimal parameters from the experiments.

As proposed in reference [9], the optimal value of wave length is twice the width of the strokes. But as showed in figure 5, the widths of strokes vary greatly. So we have to choose the optimal value through the experiment.

We set $D$=8, which means $\varphi_k=\pi k/D$ $(k=0,1,\cdots 7)$.，and consider $\lambda$ as a variable, and choose the optimal value by using 70 sets for training and 30 sets for testing. The direction number is 8 and the original dimension is 512. We also used LDA to reduce the dimension to 256. The classifier is a minimum distance classifier. The result is illustrated in Table 3.

As is showed in table 3 the optimal value of $\lambda$ is 9 and the width of character strokes in the database is about 4. So this result is in accord with the conclusion in reference [9].

In another experiment, we set $\lambda=9$ and then use 300 sets for training and 100 sets for testing. The original dimension is $64\times D$, and we also use LDA to reduce the dimension to half of the original dimension. The result is illustrated in table 4.

Table 3 comparison against different $\lambda$ under Gabor feature

| $\lambda$ | Original feature | LDA feature |
|---|---|---|
| 4 | 57.156% | 76.064% |
| 5 | 61.193% | 78.189% |
| 6 | 73.751% | 83.270% |
| 7 | 82.057% | 86.765% |
| 8 | 85.268% | 88.248% |
| 9 | **85.923%** | **88.493%** |
| 10 | 85.608% | 88.478% |
| 11 | 84.522% | 88.091% |
| 12 | 82.477% | 87.302% |

Table 4 comparison against different $D$ under Gabor feature

| $D$ | Original feature | LDA feature |
|---|---|---|
| 1 | 50.486% | 54.249% |
| 2 | 72.538% | 77.533% |
| 3 | 76.261% | 81.497% |
| 4 | 85.272% | 88.483% |
| 5 | **88.401%** | **90.651%** |
| 6 | 88.628% | 91..070% |
| 7 | 88.909% | 91.559% |
| 8 | 88.835% | 91.598% |
| 9 | 88.859% | 91.685% |
| 10 | 88.860% | 91.689% |

As shown in table 4, the accuracy of recognition is increased with the increase of $D$, no matter LDA is used or not. But the increase rate is diverse with different $D$. When $D \geqslant 5$, the increase rate is very small, especially for original feature. So considering the feature dimension, computation and accuracy of recognition, we confirm that only 5 directional parameters of Gabor features are necessary to achieve a good enough performance, rather than 8 directions which are widely used in the literature.

### 5.4 Performance comparison of Gradient feature against single-channel and multi-channel Gabor feature

We use 300 sets for training and 100 sets for testing. The direction number is 8 and the original dimension is 512, and we also used LDA to reduce the dimension to 256. The classifier is a minimum distance classifier. We extract one-channel Gabor feature, three-channel Gabor feature (the wave lengths of these three filters are 9, 6 and 12, and the directional parameter of $D$ is set to 5) and Gradient feature respectively, and then compare their performance with each other. The results are illustrated in table 5.

Table 5 Comparison of Gabor feature extraction with Gradient feature

| Feature | Original feature | LDA feature |
|---|---|---|
| One-channel Gabor | 88.835% | 91.598% |
| Three-channel Gabor | 90.7290% | 93.6720% |
| Gradient | **92.729%** | **95.652%** |

As shown in table 5, the recognition accuracy of Gradient feature is about 4% higher than one-channel Gabor feature and about 2% higher than 3-channel Gabor feature, no matter LDA is used or not. Compared with single-channel Gabor feature, the accuracy of recognition is about 2% higher for multi-channel Gabor feature. However, the computation will increase rapidly in the case of multi-channel Gabor feature.

Liu etc [10] compared these two methods in handwritten digit and printed Japanese character recognition. They used two handwritten digit databases and one printed Japanese database, the result is that one-channel Gabor feature performs a littler inferior (the difference of the accuracy is less than 0.5 percent) to Gradient feature on two handwritten digit databases, but better on the printed Japanese database; Two-channel Gabor feature improves the performance of the one-channel and performs better than Gradient feature on one of the handwritten digit database but inferior on the other one. These results are different from our experiment results. The reason is manifold. Firstly, the number of characters and similar characters of handwritten Chinese character are much larger than that of handwritten Japanese characters and printed digits. Secondly, there are much more writing styles in Chinese characters, especially in offline handwritten characters. And Gabor feature is sensitive to the width of the strokes, so the performance will be declined when the width of strokes vary greatly in samples. Though multi-channel Gabor feature overcome this disadvantage in a certain extent, but it is still not good enough as Gradient feature.

## 6. Conclusions

Gabor feature and Gradient feature have been proven to be two most efficient features for handwritten character recognition. In this paper, a comparative study of Gabor feature against Gradient feature for handwritten Chinese character recognition was conducted. We observed that: (1). Gabor feature performs much inferior to Gradient feature for handwritten Chinese character recognition, because it is sensitive to the stroke width. Though the multi-channel Gabor feature can overcome this disadvantage and performs better than single-channel Gabor feature, the increase of the feature dimension, computation and time cost is larger. (2) For the recognition accuracy hardly increases much when the value of the direction parameter in Gabor feature reaches 5, we confirm that only 5 orientation parameters are necessary to achieve good enough performance, which is different from the value in most of other papers. (3). Elastic meshing methods can regulate the deformation and displacement of strokes and obtain much better performance than linear meshing methods. From this comparative study, we suggest that Gradient feature is a good choice for large-scale handwritten Chinese character recognition.

Although some interesting conclusions were drawn in this paper, it is just an experiential study on the performance comparison between Gradient feature and Gabor feature. More theoretical investigation and systematic analysis need to be carried out in the future research.

**Acknowledgements**

This paper is supported by the Fund of New Century Excellent Talents of MOE under Grant (NO. NCET-05-0736), the University Foundation of Microsoft Research Asia (NO. OPP-2006-Researcb-10)

**References**

[1] Qivind Due Trier, et al. Feature Extraction Methods for Character Recognition: A Survey [J]. Patter Recognition, 1996, 29(4): 641-662

[2] R. Plamondon, S. N. Srihari Online and Offline Handwriting Recognition: A Comprehensive Survey [J]. IEEE Trans. on PAMI, 2000, 22 (1): 63-81.

[3] H. Fujisawa, C.-L. Liu, Directional pattern matching for character recognition, *Proc. 7th ICDAR*, Edinburgh, Scotland, 2003, pp.794-798.

[4] Hailong Liu, Xiaoqing Ding, Handwritten Character Recognition Using Gradient Feature and Quadratic Classifier with Multiple Discrimination Schemes, In Proc. 8th ICDAR, 2005, vol.1: 19-23

[5] C.-L. Liu, K. Nakashima, H. Sako, H. Fujisawa, Handwritten digit recognition: benchmarking of state-of the-art techniques, *Pattern Recognition*, 36(10): 2271-2285, 2003.

[6] C.-L. Liu, K. Nakashima, H. Sako, H. Fujisawa, Handwritten digit recognition: investigation of normalization and feature extraction techniques, *Pattern Recognition*, 37(2): 265-279, 2004.

[7] Daugman. J. G. Two Dimensional Spectral Analysis of Cortical Receptive Field Profiles [J]. Vision Research, 1980, 20:847-856

[8] Lee, T. S, Image Representation Using 2D Gabor Wavelets. IEEE Trans. Pattern Analysis and Machine Intelligence Vol. 18.(1996) 959-971

[9] Xuewen Wang, Xiaoqing Ding, Changsong Liu, Optimized Gabor Filter Based Feature Extraction For Character Recognition, ICPR, 2002, 223-226。

[10] Cheng Lin Liu, Masashi Koga, Hiromichi Fujisawa, Gabor Feature Extraction for Character Recognition: Comparison with Gradient Feature, In Proc. 8th ICDAR,2005, vol.1: 121-125

[11] Liu Cheng-Lin, Nakashima, Kazuki, H, Sako, and H.Fujisawa, "Handwritten digit recognition: investigation of normalization and feature extraction techniques", *Pattern Recognition,* Vol. 37, No. 2, pp. 265-279, 2004.

[12] Jin Lianwen, Wei Gang, Handwritten Chinese Character Recognition with Directional Decomposition Cellular Features, Journal of Circuit, System and Computer, 1998, 8(4): 517-52