PERGAMON

Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT

Pattern Recognition 38 (2005) 369–379

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/patcog

Gabor filters-based feature extraction for character recognition

Xuewen Wang, Xiaoqing Ding*, Changsong Liu

State Key Laboratory of Intelligent Technology and Systems, Department of Electronic Engineering, Tsinghua University, Beijing 100084, PR China

Received 17 October 2003; received in revised form 30 August 2004; accepted 30 August 2004

Abstract

A new method using Gabor filters for character recognition in gray-scale images is proposed in this paper. Features are extracted directly from gray-scale character images by Gabor filters which are specially designed from statistical information of character structures. An adaptive sigmoid function is applied to the outputs of Gabor filters to achieve better performance on low-quality images. In order to enhance the discriminability of the extracted features, the positive and the negative real parts of the outputs from the Gabor filters are used separately to construct histogram features. Experiments show us that the proposed method has excellent performance on both low-quality machine-printed character recognition and cursive handwritten character recognition.

Keywords: Character recognition; Gabor filters

1. Introduction

In the field of document image analysis and recognition, researchers have achieved great success in character recognition during the past decades [1]. However, there still exist two challenging problems. The first one is the optical character recognition (OCR) in low-quality images. Some difficulties are from the illumination variance, noise, complex and dirty background. In these cases it is extremely difficult for us to get clean binary character images from the gray-scale ones. This will result in low-recognition accuracy by traditional character recognition methods based on binarized character images. The second one is the cursive handwritten character recognition [2], especially for off-line cursive handwritten Chinese character recognition(OCHCCR) which is considered as one of the most difficult problems in the area of character recognition. The difficulties are from the following four aspects: (1) The Chinese character set is huge. In our daily life, we use about 3755 Chinese characters. (2) The complex structures of Chinese characters. Some Chinese characters have more than 20 strokes. (3) There are a lot of similar characters existing in this huge character set. Some similar characters can only be discriminated by just one stroke. (4) Handwriting individualities, i.e. different people have different handwriting styles. Therefore, we have focused a lot of efforts on OCHCCR in our research work. On the other hand, a lot of existing applications such as wafer OCR, vehicle license plate recognition, bank check image processing, postal address block detection and recognition, camera OCR, and so on, greatly depend on robust character recognition technology with high accuracy. This paper will introduce a new feature extraction method based on Gabor filters for robust character recognition.

As for the problem of OCR in low-quality images, two different approaches have been developed. The first one is

* Corresponding author.
E-mail addresses: wangxw92@tsinghua.org.cn (X. Wang), dingxq@mail.tsinghua.edu.cn (X. Ding), liucs@mail.tsinghua.edu.cn (C. Liu).

0031-3203/$30.00
doi:10.1016/j.patcog.2004.08.004

to binarize the gray-scale image by choosing an appropriate threshold and then do the feature extraction on the binarized image [3–5]. However, the binarization process needs detailed information of image degradation to discriminate the character stroke pixels from the background pixels. For low-quality images, it is almost impossible to get such an accurate degradation model. So the binarization process will inevitably result in information loss and will generate a lot of broken strokes or connected strokes and noise into the binarized image. The other approach, on the other hand, works directly on gray-scale images without binarization [6,7]. Obviously, it avoids the above disadvantages from the binarization step. This approach includes the following categories: (1) topographical feature extraction or edge detection directly from gray-scale images [7]; (2) global feature extraction by discrete cosine transform (DCT) or moment transform [6]; (3) image registration [8]. These methods can obtain some improvements in some cases such as blur character recognition but work poorly in other cases. For example, topographical features have poor performance on images with noise or dirty background. Methods such as DCT, moment transform and image registration are very sensitive to illumination variance and character distortion. The new feature extraction method based on Gabor filters proposed in this paper tries to solve this difficult problem.

As for the second challenging problem, i.e. the off-line handwritten Chinese character recognition, researchers have got great success by applying directional element features (DEF) and modified quadratic discriminant function (MQDF) classifiers [9]. The popular classifier combination [10] and minimum classification error (MCE) training [11] have also been used to solve this problem and have somewhat improved the recognition accuracy. However, newly developed methods such as support vector machines (SVM) [12,13] and tangent distance classifier [14] cannot be easily used for this difficult problem because of the large Chinese character set and complicated structures of Chinese characters. The Gabor filter-based feature extraction method proposed in this paper tries to further improve the recognition performance not only in low-quality images but also for off-line handwritten Chinese characters. It extracts local features directly from the gray-scale images so as to avoid the information loss by the binarization step and to tolerate the stroke variation and character distortion from different handwriting styles. This method applies multi-directional Gabor filters to the extraction of stroke information which builds the basic structures of characters. In this method, the advantages of image analysis based on both spatial domain and frequency domain is well-combined and robust local features are easily extracted which are powerful in discriminating different character strokes from noisy background.

Moreover, an adaptive regulation method is applied to the outputs of Gabor filters so as to get better performance for low-quality images. In our further research, experiments show that the discriminability of histogram features can be greatly improved by separately counting the positive and negative real parts of the Gabor filter outputs. These efforts have achieved significant improvement and are proved to be very effective for both OCR in low-quality images and off-line handwritten Chinese character recognition.

This paper is organized as follows. In Section 2, we will give a brief description of Gabor filters and related research work on Gabor filters. In Section 3, the statistical properties of local structures of Chinese characters will be described, which will be helpful in developing a simple but effective Gabor filter design algorithm. The details of the proposed Gabor filter-based feature extraction method for robust character recognition will be given in Section 4. In Section 4.3, experimental results on both OCR in low-quality images and off-line handwritten character recognition will be given in detail. These results will also demonstrate that the proposed method is powerful and effective in solving the two challenging problems mentioned above.

2. Theory of Gabor filters

Gabor filters have been used extensively in image processing, texture analysis for their excellent properties: optimal joint spatial\spatial-frequency localization and ability to simulate the receptive fields of simple cells in the visual cortex [15–17]. Two-dimensional Gabor filter is a complex sinusoidally modulated Gaussian function with the response in spatial domain (Eq. (1)) and in spatial-frequency domain (Eq. (2)) as follows (Fig. 1):

$$h(x, y; \lambda, \phi, \sigma_x, \sigma_y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\frac{R_1^2}{\sigma_x^2} + \frac{R_2^2}{\sigma_y^2}\right]\right\} \times \exp\left[\mathrm{i} \cdot \frac{2\pi R_1}{\lambda}\right], \tag{1}$$

where

$$R_1 = x \cos\phi + y \sin\phi,$$
$$R_2 = -x \sin\phi + y \cos\phi.$$

$$H(u, v; \lambda, \phi, \sigma_x, \sigma_y) = C \exp\left\{-2\pi^2\left(\sigma_x^2\left(F_1 - \frac{1}{\lambda}\right)^2 + \sigma_y^2 (F_2)^2\right)\right\}, \tag{2}$$

where

$$F_1 = u \cos\phi + v \sin\phi, \quad F_2 = -u \sin\phi + v \cos\phi, \quad C = const.$$

Spatial localization of Gabor filter can be depicted by Δx and Δy which are standard measures of effective spatial widths [15]

$$(\Delta x)^2 = \frac{\int_{-\infty}^{+\infty} h\, h^*(R_1)^2\, \mathrm{d}(R_1)}{\int_{-\infty}^{+\infty} h\, h^*\, \mathrm{d}(R_1)},$$
$$(\Delta y)^2 = \frac{\int_{-\infty}^{+\infty} h\, h^*(R_2)^2\, \mathrm{d}(R_2)}{\int_{-\infty}^{+\infty} h\, h^*\, \mathrm{d}(R_2)} \tag{3}$$

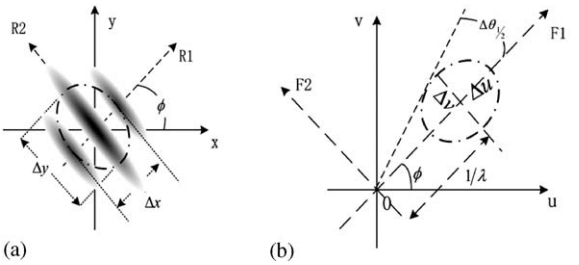

Fig. 1. Top-views of a Gabor filter in spatial domain (a) and spatial-frequency domain (b).

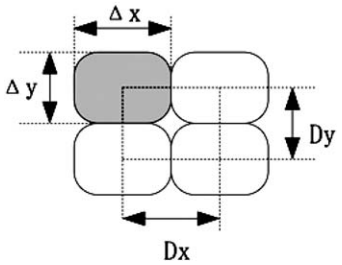

Fig. 2. The relationship of spatial sampling intervals and effective spatial widths of Gabor filters.

then we get

$$\Delta x = \sigma_x/\sqrt{2}, \quad \Delta y = \sigma_y/\sqrt{2}. \tag{4}$$

In Gabor filters, the distances between adjacent filters in an image are called spatial sampling interval Dx and Dy. As shown in Fig. 2, to avoid unexpected image information loss, the relations between effective spatial widths and spatial sampling intervals must follow:

$$Dx \leqslant \Delta x, \quad Dy \leqslant \Delta y. \tag{5}$$

Spatial sampling intervals are important parameters which need careful examination in Gabor filter design. But in previous research work [16], it was ignored and resulted in poor performance for great deal of image information loss.

Spatial-frequency localization of Gabor filter can also be depicted by Δu and Δv which are standard measures of effective bandwidths. Similar to Eq. (3), we get

$$\Delta u = 1/(2\sqrt{2}\pi\sigma_x), \quad \text{and} \quad \Delta v = 1/(2\sqrt{2}\pi\sigma_y). \tag{6}$$

Based on the spatial-frequency bandwidth, we also get another concept called orientation bandwidth [15] (Fig. 1(b))

$$\Delta\theta \approx 2\arcsin((\Delta v/2)/(1/\lambda)) = 2\arcsin(\lambda/(4\sqrt{2}\pi\sigma_y)). \tag{7}$$

In this paper, spatial-frequency localization in two-dimensional space can be expressed in two aspects: line width selectivity and line orientation selectivity, which are illustrated in Fig. 3. Obviously, Gabor filter $h(x, y; \lambda, \phi, \sigma_x, \sigma_y)$ is the most sensitive to lines with the width $\lambda/2$ and the orientation $\phi + \pi/2$.

In the middle of 1990s, some researchers applied Gabor filters to feature extraction for character recognition [16,18]. These methods achieved some progress for noise tolerance. But considering the cost of computation required, the improvement in performance was trivial due to the following reasons: (1) these methods designed Gabor filters according to recognition rates of the system, the vast computation required made it difficult to get adequate parameters of Gabor filters; (2) these methods didn't consider the recognition tasks in low-quality gray-scale images which missed the most advantages of Gabor filters. How to get prospective performance is still a problem needing further investigation.

3. Properties of low-quality character images

Today, so many kinds of low-quality images have to be considered in character recognition tasks, such as natural scene images, vehicle licence plate images, container code images, wafer code images, ID card images, etc. For these images, traditional degraded models for printed document images such as point spread function model [7] and morphology model [19] are not applicable. Because of the complexity of noise distributions and degradation processes, it is extremely difficult to construct a general degradation model here, but we can still summarize some properties from these low-quality images.

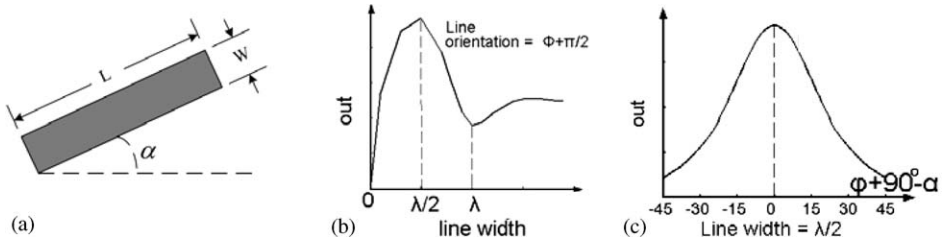

Fig. 3. For a line with length L and width $W(L \gg W)$ as shown in (a), when its width W and orientation α change, the maximum output (only real parts) of Gabor filter $h(x, y; \lambda, \phi, \sigma_x, \sigma_y)$ will change correspondingly as shown in (b) and (c), which suggest the line width selectivity and line orientation selectivity of the Gabor filter.

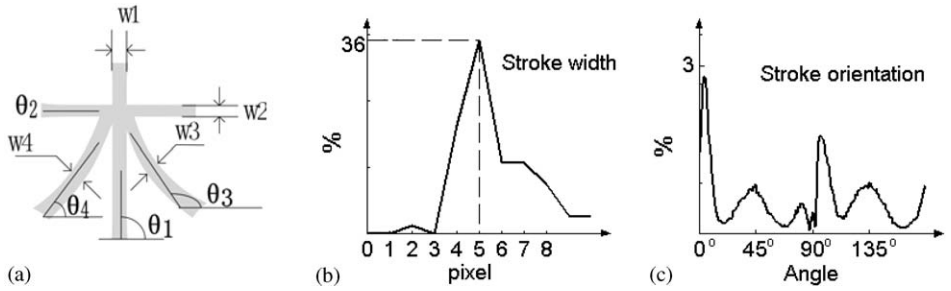

Fig. 4. (a) A demonstration on how to estimate the width and orientation of strokes. Distributions of stroke width (b) and orientation (c) of off-line handwritten Chinese characters after normalized to 64 pixel × 64 pixel.

Generally, low-quality character images are mainly due to the following three reasons:

(1) The degradations in the capture process, such as low resolution, motion blurring, inadequate focus or noise produced in the circuits of devices.
(2) The degradations produced in the printing process, such as the bank checks and ID cards which are printed with complex textures as background.
(3) The environment influence such as non-uniform illumination, shade, or dirt on characters. The images of container code, vehicle licence plate, wafer code and natural scene are examples for this aspect.

To the low-quality images mentioned above, some useful properties can be summarized:

(1) In the interested regions of low-quality character images, character strokes can be discriminated from background by different gray levels or *colors*.
(2) The degradations of low-quality images are mainly distributed in some specific frequency bands. For example, noise, low resolution or blurring degradations mainly disturb high-frequency bands, and non-uniform illuminations primarily disturb low-frequency bands.
(3) Character strokes can be discriminated from the background or noise in the scale space. Generally, the sizes of noise spots are smaller than the line widths of strokes, and the area of image bright or contrast variance because of non-uniform illumination are much bigger than the character strokes.

From our experiences in low-quality character images, the above properties hold true for most of the images. If one character image does not meet any of these requirements, it is a very hard task to recognize such an image even for humans.

For traditional methods for low-quality character image recognition, such as adaptive binarization [5] or DCT feature extraction [6], property (1) and (2) are implied. In this paper, property (3) will be utilized by introducing Gabor filters. To get a further understanding of the role of Gabor filters in our method, some analysis will be done on the local structures of characters.

From the research experiences [9], the styles of strokes, which include stroke positions and stroke orientations, are the intrinsic information to discriminate different characters for English letters, digits, Chinese characters, etc. In low-quality images, the styles of strokes are still stable enough because strokes have a different scale comparing with the degradations, as mentioned in property (3). Even for off-line handwritten characters, after the shape normalization step [9], there still existed astonishing stableness about the styles of strokes. In this paper, we carry out an experiment on Chinese characters for demonstrations.

For Chinese characters images after shape normalizing process such as line density equalization [9], they are all of the given size, for example, 64 pixel × 64 pixel. Then as shown in Fig. 4(a), we can estimate the average width w_i and the orientation θ_i for each stroke. If a stroke is not a straight line, the least square estimation is used to obtain an optimal fitting straight line for this stroke and then the orientation is estimated.

In this experiment, 10 sets of off-line handwritten Chinese character samples (3755 different characters in each set) are used to estimate the distributions of stroke width and orientation. The results are shown in Fig. 4(b) and (c). It is clear that after the characters are normalized to 64 pixel × 64 pixel, the width (W) of most strokes of Chinese characters is near five pixels and the orientations cluster a round four angles $\{0, \pi/4, \pi/2, 3\pi/4\}$.

The above analysis has proven that character images can be depicted perfectly by properties of Gabor filters: spatial\spatial-frequency localization and line width\orientation selectivity. By using a group of Gabor filters, which have a specific scale corresponding with the distribution of stroke width and specific orientations corresponding with the distribution of stroke angles, the intrinsic discriminating information of characters—the styles of strokes—can be robustly extracted from the smeared backgrounds. Then there comes a natural notion that Gabor filters can be designed based on the statistical information of strokes, which will achieve excellent performance without exhausted tests mentioned in above methods. In the next section, we will give a detail introduction for this method.

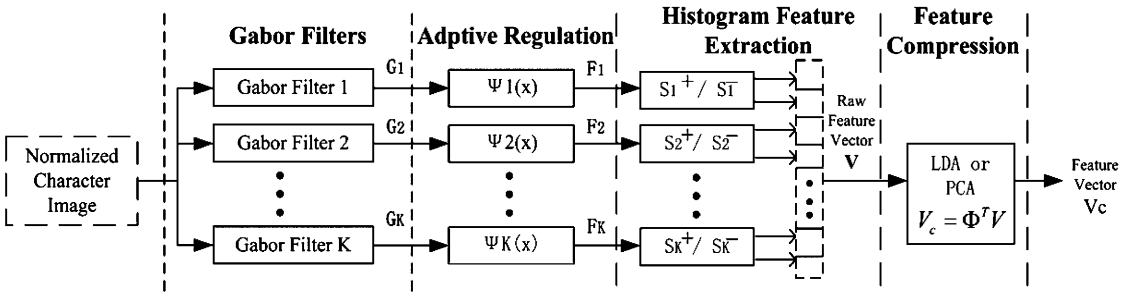

Fig. 5. The flowchart of Gabor filters-based feature extraction method.

4. Gabor filters-based feature extraction method

As shown in Fig. 5, Gabor filters-based feature extraction method can be divided into nearly four sequential parts: (1) design and apply multi-directional Gabor filters to extract the stroke information from the feeding shape normalized character image; (2) adaptively regulate the outputs of Gabor filters to achieve better performance against non-uniform illumination and background noise; (3) extract histogram features by counting the positive and negative real parts of the outputs from the Gabor filters separately, from which the discriminability of histogram features can be greatly improved; (4) using LDA or PCA algorithm to compress the histogram features to get more powerful recognition features. After obtaining the recognition features, a classifier will be incorporated to get the final results.

4.1. Gabor filters design

As shown in Fig. 4, because the strokes of characters have a specific scale and orientation distributions, Gabor filters can be designed based on the statistical information of strokes without the tedious blind try as suggested in previous research work [16,18]. In our method, Gabor filter design process can be divided into two steps: firstly, the statistical information of strokes will be used to determine the scale and the orientations of Gabor filters (corresponding to parameter λ and ϕ), at the same time the value ranges of parameter σ_x and σ_y are estimated; secondly, the concept of entropy correlation coefficient [8] is applied to get the optimal values of σ_x and σ_y, then the maximum bound of spatial sampling interval Dx and Dy are derived following Eq. (5), thus we get the desired Gabor filters $h_m(x, y; \lambda_m, \phi_m, \sigma_{xm}, \sigma_{ym}, D_{xm}, D_{ym})$ $(1 \leqslant m \leqslant K)$. Here we will describe the design process by taking an example of Chinese characters.

After normalizing to 64 pixel × 64 pixel, the width (W) of most strokes of Chinese characters is near 5 pixels and the orientations cluster round four angles. Moreover, most of the intervals of strokes (S) exceed five pixels. As discussed in the section above, an array of Gabor filters $\{h_k\}_{k=1,2\ldots K}$ must be placed in every spatial sampling point of images; each h_k corresponds to one specific stroke orientation. Therefore, some parameters of Gabor filters can be determined directly:

$$\lambda = 2W = 10, \quad K = 4,$$
$$\{\phi_k\}_{k=1\ldots4} = \left\{-\frac{\pi}{2}, -\frac{\pi}{4}, 0, \frac{\pi}{4}\right\}, \tag{8}$$

and for convenience, we can set

$$\Delta x = \Delta y \leqslant S, \quad D = D_x = D_y \leqslant \Delta x = \Delta y, \quad \Delta\theta \leqslant \pi/4. \tag{9}$$

Using Eq. (4) and Eq. (7), we obtain

$$\sigma = \sigma_x = \sigma_y \in [1.47, 7.07]. \tag{10}$$

Then we apply the concept of entropy correlation coefficient (ECC) [8] to acquire the value of σ. ECC is a normalized measure of mutual information, which measures the degree of matching between images. To two images A and B with entropies $H(A)$, $H(B)$ and joint entropy as $H(A, B)$, mutual information as $I(A, B) = H(A) + H(B) - H(A, B)$, ECC is defined as follows:

$$ECC(A, B) = 2 \cdot I(A, B)/(H(A) + H(B)). \tag{11}$$

In this paper, the purpose of applying Gabor filters is to discriminate strokes of different orientations, which suggests the outputs $\{G_k\}_{k=1\ldots4}$ of Gabor filters should be decorrelated. Following this idea we can define *Average ECC* (AECC) on character images

$$AECC = \frac{1}{6}\sum_i \sum_{j>i} ECC(G_i, G_j), \quad i, j = 1 \ldots 4. \tag{12}$$

Optimal σ should minimize *AECC* or make it below a given threshold. As shown in Fig. 6, *AECC* achieves the minimum when $\sigma = 5.6$ and changes extremely little afterwards, which makes it clear that $\sigma = 5.6$ is a proper selection for our application. Simultaneously, we get $D \leqslant 4$ by using Eq. (5) and Eq. (9) then the design of Gabor filters is finished.

Several experiments on handwritten Chinese character recognition have been carried out to demonstrate the performance of the selected parameters. Fig. 7(a) shows that the

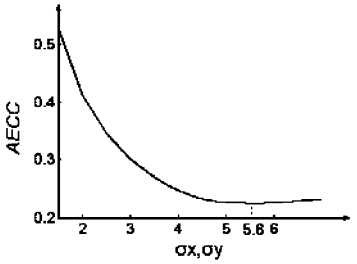

Fig. 6. The curve of AECC of Chinese characters corresponding with parameter σ.

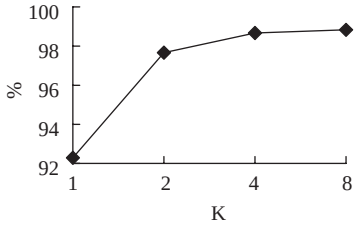

Fig. 9. Handwritten Chinese character recognition rates correspond with the number of orientations used in Gabor filters.

Table 1
Handwritten Chinese character recognition rates correspond with the number of scales used in Gabor filters

Number of Scales	1	2
Number of Gabor filters	4	8
Recognition rates	98.77	98.88

recognition rate drops dramatically when the spatial sampling interval D exceeds four, but when D is no greater than four, it remains on the top level; and at the same time, small changes of wavelength λ (Fig. 7(b)) and variance σ (Fig. 7(c)) around the selected values affect the rates very little. Example shown in Fig. 8 also proves that Gabor filters of such parameters can work effectively as discussed above.

Two more experiments are carried out to investigate the performance when changing the number of scale or orientation used in Gabor filters, the results are shown in Fig. 9 and Table 1. In multi-scale experiment (Table 1), another four Gabor filters in a much small scale ($\lambda = 2, \sigma = 1$) are used to simulate the traditional DEF features. From the results, it is clear that, when more Gabor filters beyond the above mentioned are used, the performance improves little at the cost of dramatically increasing computations.

4.2. Adaptive regulation of outputs of Gabor filters

In different images or different parts of an image, the outputs of Gabor filters may be affected tremendously by bright or contrast variance. Then additional efforts must be put to regulate the outputs to further reduce the outputs from noise or background. In traditional methods, complex image enhancement or filtering algorithms were used to accomplish this task. For methods based on Gabor filters, simple but effective solutions have been developed. In previous research, Jain and Farrokhnia [17] regulated the outputs non-linearly via sigmoid functions $S(t)$, which saturated the larger outputs of Gabor filters, and made Gabor filters as blob detectors. Unfortunately, $S(t)$ will amplify the small outputs from noise or background simultaneously, then damage the line width selectivity. Therefore, we modified $S(t)$ to $S'(t)$ to avoid this problem (Fig. 10(a))

$$S(t) = \tanh(\alpha t) = (e^{2\alpha t} - 1)/(e^{2\alpha t} + 1), \tag{13}$$

$$S'(t) = \tanh(\alpha(t - \chi)) + \beta, \tag{14}$$

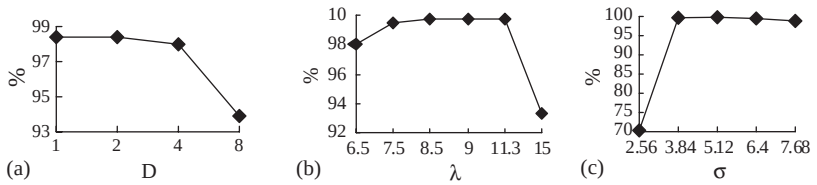

Fig. 7. Handwritten Chinese character recognition rates correspond with spatial sampling interval D (a), wavelength λ (b) and variance σ (c). In this experiment, the testing sets in (a) are a little different from (b) and (c).

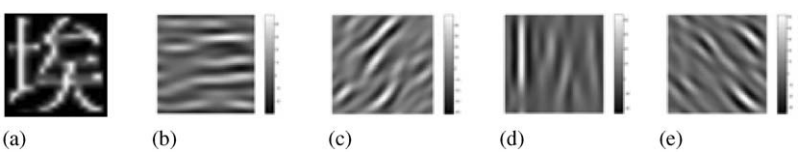

Fig. 8. Real parts of outputs from Gabor filters. (a) A low-resolution image of 17 pixel × 17 pixel is normalized to 64 pixel × 64 pixel. (b)–(e) are real parts of outputs from Gabor filters which extract different orientation strokes from (a).

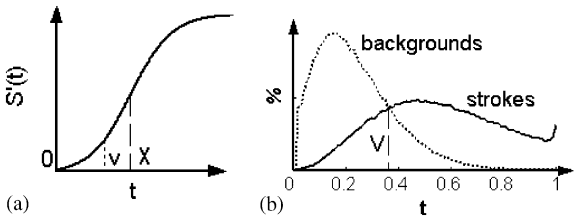

Fig. 10. (a) Curve of $S'(t)$. (b) Distributions of outputs (positive real parts) of Gabor filters from stroke pixels and background pixels.

Table 2
Recognition results of Chinese ID cards

		ID Cards	Chinese characters	Digital characters
Total number		41	104	633
Error number	Simple regulation	6	2	4
	Adaptive regulation	3	3	0

where $t \in [0, 1]$ is the real part of the output after the simple regulation which scale the outputs into the given dynamic range. Only positive real parts are discussed here; a similar method can be apply to negative real parts.

Given the distributions of the outputs of the Gabor filters from stroke pixels and background pixels as shown in Fig. 10 (b), $S'(t)$ can be constructed following such constrains:

$$S'(0) = 0, \quad \left.\frac{\mathrm{d}S'}{\mathrm{d}t}\right|_{t=v} = 1, (v < \chi),$$
$$\alpha^* = \arg\min_{\alpha} \sum_{w=0}^{\lambda} [L(w) - L_\alpha(W)]^2, \quad \alpha \geqslant 1, \tag{15}$$

where $L(w)$ is the original line width selectivity function (Fig. 3(b)) and $L_\alpha(w)$ is the function after adaptive regulation.

$S'(t)$ achieves illumination invariance, noise and background suppression, and maintenance of the line width selectivity, which has been validated by a preliminary experiment on Chinese ID card recognition task.

Because of the degradation from texture, low-quality printing and daily wear, Chinese ID card recognition is a mission impossible for traditional methods. A recognition system based on the proposed method has been developed and validated by small sets of testing samples. Results shown in Table 2 suggest the advantage of the proposed adaptive regulation algorithm which pulls the read rate up to 90%.

4.3. Histogram features extraction

After the regulation, histogram features can be extracted from the outputs of Gabor filters. The two-dimensional output planes are divided into $N \times N$ overlapped blocks as shown in Fig. 11, each block $r(x, y)$ is of size $M \times M$. In each block, histogram features are calculated separately by positive and negative real parts of the outputs weighted by a Gaussian function as shown

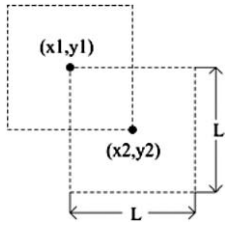

Fig. 11. Extraction of histogram features.

$$F^+_{x,y} = \sum_{(m,n)\in r(x,y)} G(m-x, n-y) \times \max(0, F_K(m,n)), \tag{16}$$

$$F^-_{x,y} = \sum_{(m,n)\in r(x,y)} G(m-x, n-y) \times \min(0, F_K(m,n)), \tag{17}$$

where $G(x, y) = \exp\{-(x^2 + y^2)/(2\tau^2)\}/(2\pi)$, F_k are real parts of the outputs from the Gabor filter with orientation ϕ_k. N, M and τ can be set to 8, 8 and 6, respectively, when Chinese characters are normalized to 64 pixel × 64 pixel.

When extracting histogram features, only the real parts of the outputs from Gabor filters are utilized because making use of the imaginary parts of the outputs introduce no more advantage in recognition rates as verified in Section 5.3.

It differs from previous work [16–18] which make use of energy magnitudes or absolute value of real parts. As shown in Fig. 8, in the output planes, the details of strokes are characterized well by one large positive lobe and two small negative lobes on either side. A more theoretic explanation for such a treatment is that it makes use of more phase information from Gabor filters than using just energy magnitudes or absolute value of real parts, so our method improves the discriminability of histogram features. Furthermore, Box–Cox transforms [20] can be applied to make the distributions of histogram feature vectors more Gaussian-like.

After the extraction of histogram features following the above steps, we can apply LDA or PCA feature compression algorithms [21] to reduce the dimension of feature vectors and, at the same time, to improve the generalization ability of classifiers and reduce the computational requirement for recognitions. In this paper, the features extracted by the proposed method are called as Gabor features.

5. Experiments and results

In this section, several experiments including seven kinds of features are carried out to verify our proposed method.

Table 3
Description of features involved in following experiments

Feature	Description
GAB	Gabor features proposed in this paper
GABA	Extract features from absolute values of real parts of outputs from Gabor filters
GABM	Extract features from energy magnitudes of outputs from Gabor filters
GABN	Extract features from both real parts and imagine parts of outputs from Gabor filters
DCT	DCT transform-based method proposed in [6]
HGAB	Method proposed by Hamamoto [16]
DEF	Directional histogram features [9,21]

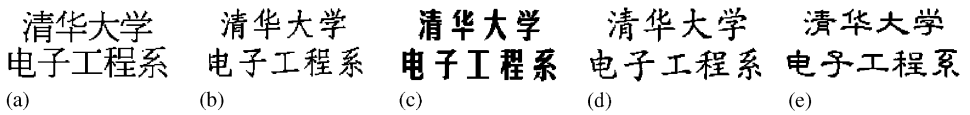

Fig. 12. Chinese characters in different fonts: (a) Songti, (b) Fangsong, (c) Heiti, (d) Kaiti, (e) Lishu.

In Table 3, HGAB and DEF feature are extracted from binary images; GAB, GABM, GABA, GABN and DCT features are all extracted from gray-scale images. Image enhancement must be applied for DCT feature but not for GAB, GABM, GABA and GABN features. The MQDF classifier proposed in [9,21] is incorporated in these experiments.

5.1. Experiments on low-resolution character images

An experiment is carried out on printed Chinese character images, 3755 character classes in commonly used fonts such as Songti, Fangsong, Heiti, Kaiti, Lishu, etc, (see Fig. 12) are included in this experiment. 200 sets of printed samples (3755 characters in each set; in this paper, every sample set of Chinese includes 3755 characters with different class labels) are scanned with 300 dpi to make the training samples, together with the artificial samples which are down-sampled from the real 200 training sets; the testing samples include another 582 sets of printed samples which are scanned with 75 to 300 dpi. For character images scanned with 75 dpi, the typical character size is 17 pixel $\times$ 17 pixel which is difficult for traditional methods. In this experiment, we apply GAB, DCT and DEF features to construct the recognition systems and the results are given in Table 4. In further experiments, the performances of GAB and DEF features in different character sizes are compared and the results are plotted in Fig. 13. It reveals that the Gabor features perform nearly the same as DEF or DCT in high-resolution cases but gets surprising superiority in low-resolution cases.

5.2. An experiment on degraded document images

In this experiment, document images are scanned with 300 dpi, which contain 11,688 Chinese characters, then artificial Gaussian noises $N(0, \sigma^2)$ are added as displayed in Fig. 14. The results are shown in Table 5 which suggests that Gabor features also have the best performance for images degraded by noise.

Table 4
Recognition rates of printed Chinese characters in different resolutions

	Resolution (Dpi)	GAB	DCT	DEF
Songti	75–100	99.28	98.96	97.71
Fangsong	150–200	99.87	99.87	99.59
Heiti	250–300	99.82	99.91	99.82
Kaiti	75–100	96.70	86.61	89.47
	250–300	99.40	98.31	99.75
Lishu	75–100	74.41	55.87	47.45
	150–200	95.60	85.37	86.55
	250–300	98.18	93.74	95.74

5.3. Experiments on handwritten characters

Here we carry out two experiments on off-line handwritten character recognition for Chinese and digits (Fig. 15). In the experiment on Chinese characters, 500 sets of Chinese handwritten samples are used, 300 sets are selected randomly for training and others for testing. Recognition rates of seven kinds of features (see Table 3) are shown in Table 6. The performance of GAB on handwritten characters is close to DEF and superior to other kinds of features.

In the experiment on digits, MNIST [22] is used as the database in which 60,000 characters are treated as training samples and 10,000 characters as testing samples. In Table 7 the results of Gabor features and DEF features are given, the best result on MNIST reported publicly is also given

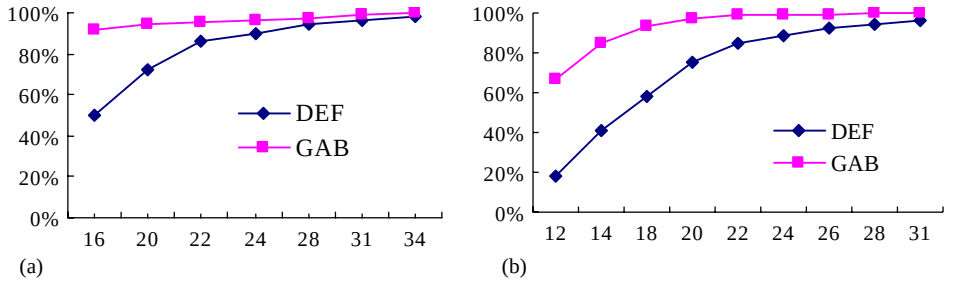

Fig. 13. For GAB and DEF features, the recognition rates of different sizes of printed Chinese characters in different fonts such as: (a) Songti; (b) Kaiti.

Fig. 14. Demonstrations of character images degraded by Gaussian noises: (a) the clean image; the images degraded by Gaussian noise as (b) $N(0, 2.5)$, (c) $N(0, 12.5)$, (d) $N(0, 25)$, (e) $N(0, 38.3)$.

Table 5
Recognition rates of corrupted images (σ, %)

	Gaussian noise (mean = 0)				
	0	2.5	12.5	25	38.3
GAB	99.44	98.92	98.38	96.56	92.53
DCT	99.33	98.00	95.54	86.11	71.65
DH	99.55	92.05	72.90	30.85	8.28

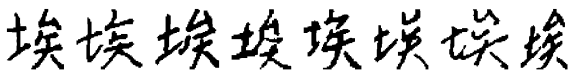

Fig. 15. Some samples of a handwritten Chinese character.

Table 6
Recognition rates of handwritten Chinese characters (%)

	First candidate	Top 10 candidates
DEF	98.90	99.95
GAB	98.77	99.90
GABN	98.50	99.90
GABA	97.63	99.58
GABM	96.53	99.51
HGAB	95.27	99.43
DCT	94.89	98.02

which applies the invariant SVM algorithm [13], our proposed method achieve satisfactory result with a much simpler classifier—MQDF classifier [9].

Table 7
Testing error rates on MNIST (%)

Method	Error rate
Invariant SVM	0.55
DEF+MQDF	0.76
GAB+MQDF	0.87

MNIST can be got from URL: http://yann.lecun.com/exdb/mnist/.

5.4. An experiment on character images degraded by rotation

Rotation is a kind of degradation which will affect the performance of character recognition systems in practice. In this experiment, we will evaluate the robustness of Gabor features in rotated handwritten Chinese character images. Ten sets of samples are randomly selected from the handwritten Chinese testing sets of Section 5.3, then all the character images are rotated by an angle θ, the rotation directions (clockwise or anti-clockwise) are randomly decided. The classifiers derived in Section 5.3 is used to get the recognition results which are shown in Fig. 16. It reveals that when the rotation angles exceed 10°, the performance of Gabor features is a little worse than DEF features.

The slight inferiority of Gabor features in rotated character image recognition can be explained well by the line orientation selectivity as suggested in Section 2. From the parameters discovered in Section 4.1 and with Eq. (7), we can easily estimate the orientation bandwidth of the Gabor filters is $2\Delta\theta_{1/2} \approx 11°$; But theoretically, the orientation bandwidth of DEF features is 45°[9]. So when the line

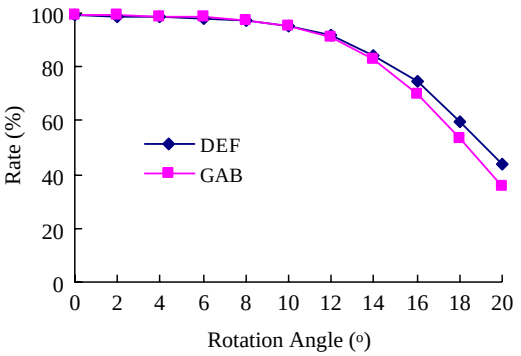

Fig. 16. Recognition rates for GAB and DEF features in different rotation angles.

selectivity of Gabor filters builds the fundamental robustness of Gabor features to resist noise or other kinds of degradations, it also unfortunately brings a little shortcoming into the recognition systems as proven by this experiment.

Another problem is that Gabor features need more computations than DEF features because of great amounts of convolutions involved in it. Fortunately, the speed of Gabor features extraction can be accelerated by processing Gabor filters in frequency domain, the DFT\IDFT algorithm can also be further accelerated by code optimizations. For off-line handwritten Chinese character recognition, after the accelerations, the typical speed of the system using Gabor features is 118 characters/s in a computer with a 1.7 GHz PIV CPU and 512 MB memory, the speed of the system using DEF features is 134 characters/s with the same classifier and computer setup.

6. Conclusions

In this paper we propose a Gabor filter-based feature extraction method for character recognition. The task of Gabor filter design is accomplished by a simple but effective method. Special efforts have been put to enhance the discriminability of histogram features and improve the performance to low-quality character images. Experiments have shown that our method performs excellently for images with noise, complex background or stroke distortions, it can be applied to printed or handwritten character recognition tasks in low-quality gray-scale or binary images. In our work in several challenging applications such as Wafer OCR, vehicle licence plate recognition, container code recognition, etc, Gabor features have got inspiring successes.

Further research should be focused in two topics: (1) for specific tasks, construct the degradation models to artificially generate degraded samples to alleviate the shortage of training samples which is a painful problem in most practical applications; (2) algorithms for Gabor filters design should be developed via statistical learning methods. For a better understanding of Gabor filters in other recognition fields, such as face recognition [23], iris recognition [24], etc, more attention should be paid on the latter topic.

Acknowledgements

This work was supported by 863 Hi-tech Plan (Project 2001AA114081) and National Natural Science Foundation of China (Project 60241005). We would like to thank Dr. Youbin Chen and Dr. Jing Zheng for their contribution to this paper.

References

[1] G. Nagy, Twenty Years of Document Image Analysis in PAMI, IEEE. Trans. PAMI 22 (1) (2000) 38–62.
[2] R. Plamondon, S.N. Srihari, On-line and off-line handwriting recognition: a comprehensive survey, IEEE. Trans. PAMI 22 (1) (2000) 63–84.
[3] O.D. Trier, A.K. Jain, Goal-directed evaluation of binarization methods, IEEE Trans. PAMI 17 (12) (1995) 1191–1201.
[4] W. Niblack, An Introduction to Digital Image Processing, Prentice-Hall, Englewood Cliffs, NJ, 1986, pp. 115–116.
[5] H. Kamada, K. Fujimoto, High speed, high-accuracy binarization method for recognizing text in images of low spatial resolutions, Proceedings of the ICDAR'1999, Bangalore, India, 1999.
[6] X. Wang, X. Ding, et al., A gray-scale image based character recognition algorithm to low quality and low resolution Images, Document Recognition and Retrieval VIII, Electronic Imaging 2001, San Jose, CA, USA.
[7] L. Wang, T. Pavlidis, Direct gray-scale extraction of features for character recognition, IEEE Trans. PAMI 15 (10) (1993) 1053–1066.
[8] J.P.W. Pluim, J.B.A. Maintz, et al., Mutual information matching in multi-resolution contexts, Image Vision Comput. 19 (2001) 45–52.
[9] F. Kimura, K. Takashina, et al., Modified quadratic discriminant functions and the application to Chinese character recognition, IEEE Trans. PAMI 9 (1) (1987) 149–153.
[10] R.P.W. Duin, The combining classifier: to train or not to train, Proceedings of the ICPR'2002, Quebec City, Canada, 2002.
[11] B.H. Juang, W. Chou, C.H. Lee, Minimum classification error rate methods for speech recognition, IEEE. Trans. Speech Audio Process. 5 (3) (1997) 257–265.
[12] V.N. Vapnik, Statistical Learning Theory, Wiley, New York, 1998.
[13] D. DeCoste, B. Schoelkopf, Training invariant support vector machines, Mach. Learn. J. 46 (1–3) (2002) 161–190.
[14] P.Y. Simard, et al., Transformation invariance in pattern recognition—tangent distance and tangent propagation, Int. J. Imaging Syst. Technol. 11 (3) (2001) 181–194.
[15] J.G. Daugman, Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters, J. Opt. Soc. Am. A 2 (7) (1985) 1.
[16] Y. Hamamoto, S. Uchimura, et al., A Gabor filter-based method for recognizing handwritten numbers, Pattern Recog. 31 (4) (1998) 395–400.

[17] A.K. Jain, F. Farrokhnia, Unsupervised texture segmentation using Gabor filters, Pattern Recog. 24 (12) (1991) 1167–1186.
[18] Y. Ge, Q. Huo, Offine recognition for handwritten Chinese characters using Gabor features, CDHMM modeling and MCE training, Proceedings of the ICASSP'2002, Orlando, USA, 2002.
[19] T. Kanungo, R.M. Haralick, H.S. Baird, et al., Statistical, nonparametric methodology for document degradation model validation, IEEE. Trans. PAMI 22 (11) (2000) 1209–1223.
[20] D.C. Hoaglin, F. Mosteller, et al., Understanding Robust and Exploratory Data Analysis, Wiley, New York, 1983.
[21] J. Zhang, X. Ding, Multi-Scale feature extraction and nested-subset classifier design for high accuracy handwritten character recognition, Proceedings of the ICPR'2000, 2000.
[22] Y. Lecun, et al., Gradient-based learing applied to document recognition, Proc. IEEE 86 (11) (1998) 2325–2344.
[23] J. Zhang, Y. Yan, M. Lades, Face recognition: eigenface, elastic matching, and neural nets, Proc. IEEE 85 (9) (1997) 1423–1435.
[24] Richard P. Wildes, Iris recognition: an emerging biometric technology, Proc. IEEE. 85 (9) (1997) 1348–1363.

About the Author—XUEWEN WANG was born in Hubei, China, in 1973. He received his B.S. degree and Ph.D. degree from the Department of Electronic Engineering of Tsinghua University, Beijing, China, respectively in 1997 and 2003. Now he works for Mitek Systems, Inc. (San Diego, CA, USA) as a staff scientist. His current research interests include character recognition, biometrics, object detection and recognition.

About the Author—PROF. XIAOQING DING graduated from Dept. of Electronic Engineering, Tsinghua University, in 1962. She is a Principal Professor and Ph.D. supervisor in Dept. of Electronic Engineering, Tsinghua University. She has published more than 250 papers and is the co-author of four books. Her research interests include Image Processing, Pattern Recognition, Character Recognition, Document Analysis and Recognition, Computer Vision, etc. As a principal investigator, Professor Ding has researched and developed a series of excellent Chinese character recognition systems. She has owned more Honors and Awards of *China National Scientific and Technical Progress Award*s.

About the Author—CHANGSONG LIU is an associate professor in the Department of Electronics engineering at Tsinghua University. He received the BS degree in both mechanics engineering and electronics engineering in 1992, Master degree in electronics engineering in 1995 from Tsinghua university, China. His fields of interests includes document image processing, image compression, pattern recognition, nature language processing.