Title	Handwritten Chinese character recognition using spatial Gabor filters and self-organizing feature maps
Author(s)	Deng, D; Chan, KP; Yu, YL
Citation	The 1994 International Conference on Image Processing, Austin, TX., 13-16 November 1994. In International Conference on Image Processing Proceedings, 1994, v. 3, p. 940-944
Issued Date	1994
URL	http://hdl.handle.net/10722/45621
Rights	©1994 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

HANDWRITTEN CHINESE CHARACTER RECOGNITION USING SPATIAL GABOR FILTERS AND SELF-ORGANIZING FEATURE MAPS

Da Deng[+], K.P. Chan[+] and Yinglin Yu[*]*

[+]Department of Computer Science, University of Hong Kong
Pokfulam Road, Hong Kong

[*]Institute of Radio-engineering and Automation, South China University of Tech. and Sci.
Guangzhou 510641, PR CHINA

ABSTRACT
So far the bottleneck of Chinese character recognition, especially handwritten character recognition, still lies in the effectiveness of feature-extraction to cater for various distortions and position shifting. In this paper, a novel method is proposed by applying a set of Gabor spatial filters with different directions and spatial frequencies to character images, in an effort to reach the optimum trade-off between feature stability and feature localization. While a classic self-organizing map is used for unsupervised clustering feature codes, a multi-staged LVQ with a fuzzy judgement unit is applied for the final recognition on the feature mapping result.

1. INTRODUCTION
There are still considerable difficulties in handwritten Chinese character recognition[1]. Similar to other pattern recognition problems, the bottleneck to reach a successful system realization, lies in finding a robust and effective scheme of feature extraction and organization, which is able to endure deteriorating factors caused by noise, distortion and position shifting which occur quite often in people's handwritings.

Traditional structural feature extraction methods[1][2], such as local density counting and stroke matching, are rather sensitive to shifting of stroke position and distortions such as crisscross and interruption of character strokes. Sometimes the skeletonization process which is the basis for structural analysis may introduce fatal noises into the result. Global analysis methods, implemented by a number of transforms such as K-L, Walsh etc., on the other hand, turn out more stable features. However, Chinese characters, when compared with digits and English letters, are more complicated patterns with the following special characteristics:
- A large set. There are 3750 characters in the most frequently used Set-1.
- Possessing more local structures corresponding to typical parts and radicals.

Thus both the local structure analysis method and the global transform method have their limitation to achieve an effective representation of handwritten Chinese characters.

In this paper, we introduce a new method of feature extraction for the handwritten Chinese character recognition problem from an early-vision point of view. Instead of doing traditional structural analysis to extract different types of strokes, end-points or cross-points as features, we apply a set of spatial Gabor filters as receptive fields to operate on the handwritten character images of multiple resolution scale. Different responses for different local spatial frequencies and directions are obtained as primary feature codes. We then apply a self-organizing feature map procedure to achieve a clustered topology-preserving map of these primary feature codes. Finally, the recognition code book is obtained through a supervised training LVQ procedure by learning the best mapped IDs of the filtered response of the character image. A fuzzy unit is applied for the final recognition based on the result of multiple codebook matching.

2. GABOR SPATIAL FILTERS

The 2-D Gabor filters ([3][6][7]) are a set of orientation and frequency sensitive band pass filters which have the optimum joint resolution in both the spatial and the spatial-frequency domain. Thus they are suitable for extracting orientation-dependent frequency contents of patterns and have been widely used in problems such as face recognition and texture segmentation.

The spatial Gabor filter kernels are sinusoids modulated by a Gaussian window, taking the form:

$$\psi(f,\theta,x,y) = \exp[if(x\cos\theta + y\sin\theta\,)]$$
$$\cdot \exp[-\frac{x^2 + y^2}{2\sigma^2}] \qquad (1)$$

where x and y are the pixel coordinates in the image, f is the central frequency of the pass band, θ is its spatial orientation, while σ determines the bandwidth of the filter. It is through these parameters that the very shape and spread of the Gabor filters are determined.

In our application the orientation of Gabor filters are sampled upon four values, i.e., $0°$, $45°$, $90°$ and $135°$. This restriction on orientation relieves the computation burden, and is rather reasonable for the case of Chinese character featuring, since most strokes are basically are in these orientations. And for an image of N-pixel width, radial frequencies take the values of $1\sqrt{2}, 2\sqrt{2}, \cdots, (N/4)\sqrt{2}$.

Let $G(x,y) = \exp(-\frac{x^2 + y^2}{2\sigma^2})$, then we can split Eq.(1) into its real and imaginary parts, i.e.,

$$\psi = \psi_c + j\psi_s\,, \qquad (2)$$

where

$$\psi_c = G(x,y)\cos[2\pi f(x\cos\theta + y\sin\theta)]$$
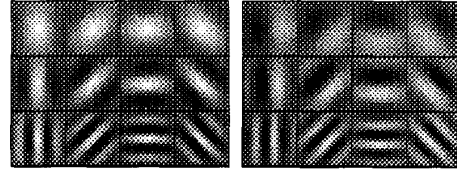$$\psi_s = G(x,y)\sin[2\pi f(x\cos\theta + y\sin\theta)] \qquad (3)$$

Fig. 1. Intensity plot of the Gabor filters with different orientations and radial frequencies. On the left side are the real parts of filter functions, while their imaginary parts are shown on the right.

Some of the Gabor filter patterns are displayed as gray-scaled image shown in Fig. 1. Patterns on the left of the figure are the real parts of the filter kernel, patterns on the right are the imaginary parts. It is easy to see that the real parts of filter functions are even-symmetric, while the imaginary parts are odd-symmetric.

Gabor spatial filters with different central frequencies and orientations are convolved with the image and localized spectrum responses of both amplitude and phase are generated. Usually their amplitude responses are chosen as features since they are more tolerant to distortions in pattern images, as shown in many studies([2], [7]).

3. SOM OF FEATURE CODES

Assume that a set of input vectors x are to be mapped onto a D-dimensional lattice with its reference vectors $\mathbf{w}_q$ indexed by a D-dimensional vector q. The basic self-organizing map (SOM) training algorithm chooses one of the reference vectors which best matches the input vector x, and update the weight vectors with the learning rule:

$$\mathbf{m}_q = \mathbf{m}_q + \alpha h(\mathbf{q},\mathbf{b})(\mathbf{m}_q - \mathbf{x}) \qquad (4)$$

where $\alpha(t)$ is a positive learning rate slowly fading away as training goes on, and $h(\cdot)$ is a Gaussian window of width σ, taking the form:

$$h(q,b) = \exp(\frac{\|\mathbf{q}-\mathbf{b}\|^2}{2\sigma^2}). \qquad (5)$$

941

Simply setting the neighborhood width to zero makes SOM equivalent to the K-means method for data clustering. Comparing with traditional clustering methods, one of the most important features of SOM is its topology preserving ability in mapping data from a high dimensional space to a lower dimensional map lattice. Thus neighboring input vectors are mapped onto neighboring lattice points, which enables a hierarchical processing of the mapping result [5][8].

In our work, however, we take advantage of SOM to achieve a topology-preserving dimension reduction of data features, and then further use the mapping result for a sequence of LVQ procedures, which are a group of supervised training algorithms, to construct a codebook for recognition.

After initialization of a codebook, every time a training vector $\mathbf{x}$ is given, the basic LVQ1 algorithm finds out the nearest codebook vector m_c, i.e.,

$$c = \arg\ \min_i\{\|\mathbf{x} - \mathbf{m}_i\|\}, \qquad (6)$$

and training is performed with the following learning rules:

$$\mathbf{m}_c = \mathbf{m}_c + \alpha(\mathbf{x} - \mathbf{m}_c) \qquad (7)$$

  if $\mathbf{x}$ and $\mathbf{m}_c$ belong to the same class,

$$\mathbf{m}_c = \mathbf{m}_c - \alpha(\mathbf{x} - \mathbf{m}_c) \qquad (8)$$

  if $\mathbf{x}$ and $\mathbf{m}_c$ belong to different classes.

α is the learning rate. For fine-tuning of the code book vectors, more complicated training processes such as LVQ2 and LVQ3 are often carried out in order to enhance the recognition rate.

4. EXPERIMENT IMPLEMENTATION

Handwritten Chinese characters for training and recognition in our experiment are from a database we collected locally. There are 40 samples for each of the 180 character categories, and each sample is of 64x64 pixel in size. Some of them are shown in Fig. 2. Preprocessing procedures, such as normalization according to the mass centre of the image, and fuzzing which grayscales the binary image, are applied. These preprocessing procedures are performed to reduce the variance of font shape and position shift of local parts of character patterns. We use 20 sample images of each Chinese handwritten character as training
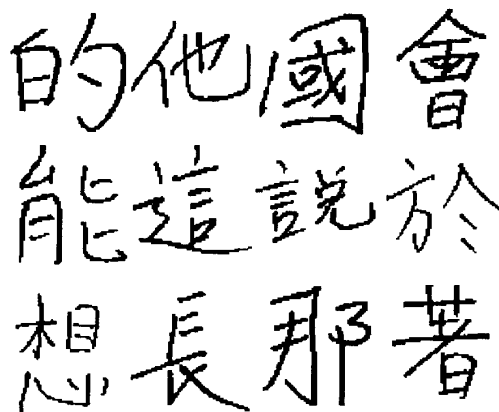

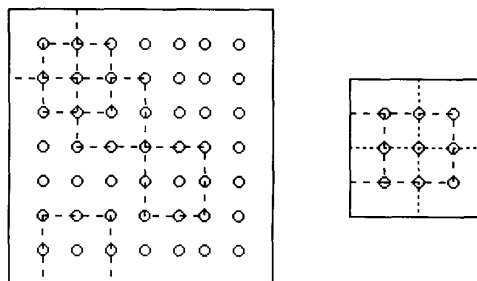Fig. 2. Samples of handwritten Chinese characters. All are of 64x64 size.

Fig. 3. Overlapped lattices which the Gabor filters operate on. They are 16x16 squares shown with the dashed lines. Numbers of lattice centre for resolution 1 & 2 are 49 and 9, respectively.

patterns and the other 20 for testing. Some of the character samples used in the experiment are shown in Fig. 2.
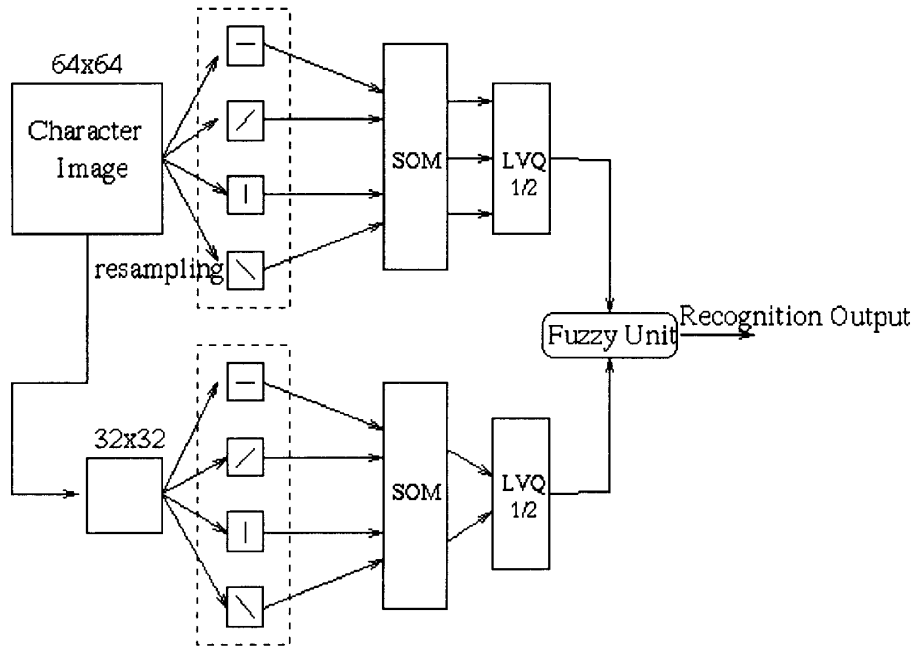
Fig. 4. Diagram of the whole system.

4.1 Leaning Phase 1: Self-organization of feature codes

Feature codes are generated by a set of 16x16 Gabor filters with 3 frequencies on 4 orientations. The filters are applied to half-overlapping 16x16 lattices of the handwritten character image on two resolution stages, i.e. 64x64 and 32x32, as shown in Fig. 3, and the magnitude of the convolution with every set of the Gabor filters gives a group of 12-dimensional feature vectors respectively.

The two sets of feature vectors are clustered with SOM onto two 8x8 2-D lattices. Vectors on these 2 lattices form two codebooks as the feature mapping basis for the next phase.

4.2 Learning Phase 2: Supervised learning of mapping ID codes

Training samples are reused to generate local feature vectors. Each feature code is then mapped onto the lattices and the ID index of the best-matching lattice node is obtained. Then mapping IDs of the same character, coming out of two resolution stages, form two input vectors, and go through two LVQ1 training procedures to construct codebooks for the category the training character sample belongs to. Codebooks are further fine-tuned with LVQ2 algorithm to enhance their accuracy.

4.3 Final recognition with a fuzzy judgement unit

After training, the two stages working with different spatial resolution generate two different codebooks of different code vectors. A fuzzy judgement unit is applied to pick out the character category candidates which best match these code vectors.

Assume that for a testing character image, the feature mapping codes are c_1 and c_2. And for category i, the vectors in codebooks with minimum distances to the feature mapping codes,

are w_1 (from codebook 1) and w_2(from codebook 2). So the likelihood that the testing sample belongs to category i, i.e. μ_1 and μ_2, can be calculated as

$$\mu_1 = \exp(\frac{\|\mathbf{w}_1 - \mathbf{c}_1\|^2}{\sigma_1^2}) \qquad (9)$$

$$\mu_2 = \exp(\frac{\|\mathbf{w}_2 - \mathbf{c}_2\|^2}{\sigma_2^2}) \qquad (10)$$

and the final score of category i is

$$s_i = \frac{\alpha\,\mu_1 + \beta\,\mu_2}{\alpha + \beta} \qquad (11)$$

Then the final fuzzy judgement unit picks out the category label with the maximum matching score as recognition result. In our experiment, we take $\sigma_1 = 10, \sigma_2 = 3, \alpha = 3, \beta = 1$. The block diagram of the whole system is shown in Fig. 4 .

After training with samples out of the 180 categories, an average recognition rate of 94.2% is gained for testing samples.

5. CONCLUSION

In this paper, we propose an alternative way to tackle the feature extraction problem in handwritten Chinese character recognition from an early-vision point of view, in which a set of Gabor filters are applied as receptive fields for feature extraction. This leads to the effect of reaching more stability in local feature detection to cope with distortion and variance of human handwritings. After self-organization maps of the feature codes, the recognition result is obtained through a fuzzy judgement unit, which combines the judgements from feature mappings on different resolution stages. Although presently our experiment are still limited on a relatively small scale with a set of only 180 characters, the performance of the system is quite promising for further research.

REFERENCES

[1] S.Mori, et. al., Historical Review of OCR Research and Development, Proc. of IEEE, vol. 80, No. 7, pp. 1029-2058, 1992.

[2] F.Cheng, et. al., Research on Chinese OCR in Taiwan, Int. Jour. of Pattern Recognition and Artificial Intelligence, vol. 5, No. 1&2, 1991.

[3] J.G. Daugman, "Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression", IEEE Trans. on ASSP, vol. 36, No. 7, pp. 1169-1179, 1988.

[4] B. MacLennan, Gabor Representation of Spatiotemporal Visual Images, Tech. Report CS-91-144, Univ. of Tennessee, 1991.

[5] J. Lampinen, E. Oja, "Clustering Properties of Hierarchical Self-Organizing Maps", Jour. of Math. Imaging and Vision, vol. 2, pp. 261-272, 1992.

[6] J. Buhman, J. Lange, C. von der Malsburg, "Distortion Invariant Object Recognition by Matching by Hierarchical Graphs", Proc. IEEE IJCNN-89, vol. 1, pp155-159, Washington DC, 1989.

[7] A.K. Jain, F. Farrokhnia, "Unsupervised Texture Segmentation using Gabor Filters", Pattern Recognition, Vol. 24, No. 12, pp. 1167-1186, 1991.

[8] H. Ritter, T. Martinetz, K. Schulten, Neural Computation and Self-organizing Maps, Addison-Wesley Publishing Company, 1991.

[9] P. Tavan, H. Grubmueller, H. Kuehnel, "Self-organization of associative memory and pattern classification: recurrent signal processing on topological feature maps", Biol. Cybern. 64, pp.95-105, 1990.