Using Two-Dimensional Gabor Filters for Handwritten Digit Recognition

Tien C. Bau
Department of Electrical Engineering and Computer Science
University of California, Irvine, CA 92697
tbau@uci.edu

Abstract

A feature extraction method for handwritten digit using Gabor filterbank is proposed. I designed a Gabor filterbank which uniformly covers the frequency domain with each filter captures the energy of a localized frequency. The obtained Gabor feature vectors are subjected to the task of classification and feature selection. A comparison against pixel features is presented. Mixing pixel and Gabor features for feature selection and classification is also evaluated.

1 Introduction

Handwritten digits recognition has been very successful in recent years. Applications such as automatic form processing, zip-code recognition have been widely adopted. Before any handwritten digits can be recognized as a meaningful digits, they have to be processed in many steps: scanning to grayscale image, convert to binary image, feature extraction, classification then post decision making. For a handwritten recognition system to be successful, the selection of a feature extraction method is critical in getting a high recognition rate. Several methods have been proposed[1][2], they are: direct matching, zoning, geometric moment invariants, zernike moments, Fourier descriptors, … etc. Each of these methods may have their advantages in one or more applications. Gabor filters have received considerable attention in image processing. The Gabor functions achieve optimal joint localization in the original and transform domains[5]. The Gabor functions are also closely related to the human visual system and texture interpretation[5][6]. Gabor filters can be used to extract components corresponding to different scales and orientations from images. 2D Gabor filters have been used for texture segmentation[3][7], texture analysis[8] and color texture recognition [4].

2 2D Gabor filters

Gabor filters have been used in many image/texture recognition/detection researches [4]. A 2D spatial Gabor filter [3] is defined in the radiance domain by

$$g(x, y) = a(x, y)c(x, y) \quad (1)$$

where

$$a(x, y) = \frac{1}{(2\pi)\ \sigma_x \sigma_y} e^{-0.5\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)} \quad (2)$$

is the Gaussian component and

$$c(x, y) = \cos\left(2\pi\left(F_x x + F_y y\right)\right) \quad (3)$$

is the sinusoidal component. The variables, x and y are the spatial variables. The standard deviations (σ_x,σ_y) in (2) describe the size of the Gaussian envelope and define the scale of the filter along the spatial and spectral axes. (F_x,F_y) represents the frequency of the sinusoidal component and thus the center frequency of the filter in the 2D frequency domain. The orientation of the filter is defined as the unit vector from the origin to the center frequency (F_x,F_y) of the filter. Figure 1 shows Gabor filterbank of 3 scales and 9 orientations (total 27 filters) used in this paper. Note that the light and dark gray shadow mark positive and negative values of the filter respectively. Besides from the aforementioned 27 filters, one more filter is added for orientation (0,0) and scale (σ_x,σ_y)=(∞,∞) truncated at sample's size (i.e. DC) to make the final filterbank of 28 filters.

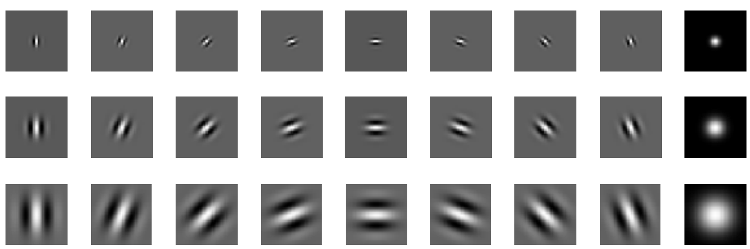

Figure 1. 2D Gabor filterbank of 3 scales and 9 orientations.

3 Experiments

3.1 Dataset

I use 8-bit grayscale images of "0" through "9" (size 28x28 each) obtained from

http://www.cs.toronto.edu/~roweis/data/mnist_all.mat

There are about 6K training examples of each class; 1K test examples. However, due to computational limitation, only 200 samples in each training class (total 2000), and 100 samples in each test class (total 1000) are used in this paper.

3.2 Features

All sample images are converted into three different vectors for three cases study, they are pixel, Gabor and pixel+Gabor vectors. For pixel vector, I treat each pixel as one feature. A pixel vector therefore contains 28x28=784 features. Gabor feature is obtained from summing up the energy of one Gabor filter convoluting with one sample. A 28-feature Gabor vector can be produced by applying a bank of 28 filters to a sample. In this paper, Gabor+Pixel vector is also evaluated, which is the augmentation of these two types of vectors into one 812-feature vector

3.3 Model training and feature selection

For model training and testing, I will assume that the feature vectors for the samples in a digit class c are normally distributed with mean vector μ_c and covariance matrix Σ_c. To determine how far a digit sample s is away from digit class c, the Mahalanobis distance is used

$$M(s,c) = \sqrt{(x_s - \mu_c)^T \Sigma_c^{-1} (x_s - \mu_c)} \tag{4}$$

where:

Σ : the covariance matrix.

$x_s = (x_s(1), x_s(2), \ldots, x_s(D))$, the features vector for sample s. D is the total number of features in a features subset,

$\overrightarrow{\mu_c} = (\mu_c(1), \mu_c(2), \ldots, \mu_c(D))$, the class mean vector for class c

To determine which features are most effective, two methods will be used to identify them. They are: Fisher LDA, Greedy search (forward selection). Fisher LDA considers maximizing the following objective function

$$J(v_k) = \frac{v_d^T \Sigma_B v_d}{v_d^T \Sigma_W v_d} \tag{5}$$

where

Σ : the "between classes scatter matrix" and

Σ_W : the "within classes scatter matrix".

v_d : the feature under evaluation. d=1~D, the total number of features.

The fact that when evaluating one feature at a time, the implicit assumption is that feature is independent by itself. In this case the Σ and Σ_W can be simplified to their diagonal as

$$S_W = \sum_c \sum_{s \in c} (x_s - \mu_c)(x_s - \mu_c)^T \tag{6}$$

$$S_B = \sum_c n_c (x_s - \mu)(x_s - \mu)^T \tag{6}$$

where n_c is the number of samples in class c, and μ is the overall samples' mean. Features with maximum $J(v_d)$ value will be chosen first. The first 28 features are evaluated in this paper. Forward selection algorithm uses stepwise optimal algorithm for an approximately optimal features subset. This algorithm builds features subset by adding new features one at a time until total features reaches a predetermined number. It first finds out the single feature which ,when used alone, will give the best classification rate. Then we add a new feature from the remaining features that leads to a maximum increase in the classification rate. The addition of new features is repeated until the total numbers of features reaches 28.

3.4 Classification method

The obtained model with parameter (μ ,σ) is applied to the test dataset (1000 samples, 100 samples per class). For an unknown sample, the Mahalanobis distance (MD) between the sample's vector and the mean vector for each one of the ten known classes is calculated using (4).

3.5 Classification results

Figure 2 and Figure 3 plot the classification results for model training and testing on their corresponding training and testing dataset. Key numbers are summarized in Table 1. Results from experiment with all features are listed as comparison. With classification result for Gabor+Pixel features subset reaches 90%, it clearly shows that when combine Gabor features with pixel features, the discerning capability is far better than any one of them used alone.

Table 1. Classification result for training and test data

Experiment	(μ_c,Σ_c) All Features			(μ_c,Σ_c), Fisher-LDA			(μ_c,Σ_c), Greedy		
	Train	Test	D'	Train	Test	D'	Train	Test	D'
Pixel (784 pixels)	10%	10%	784	70%	57%	23	95%	81%	28
Gabor (28 features)	87%	74%	28	92%	83%	27	98%	90%	27
Pixel+Gabor	10%	10%	812	93%	84%	28	98%	90%	28

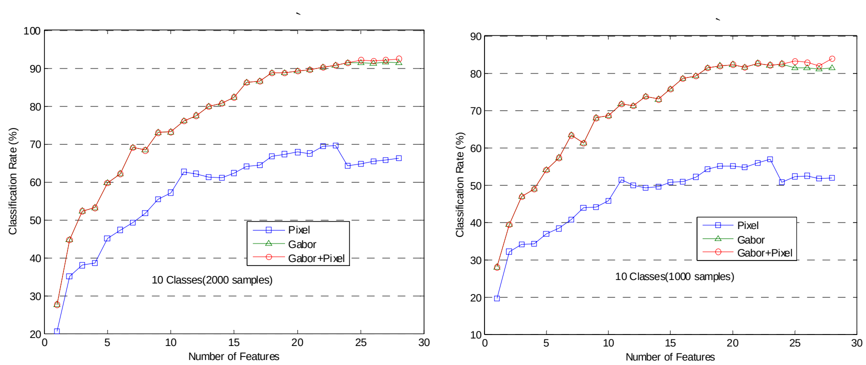

Figure 2. Classification result with model (μ_c,Σ_c), feature selection by Fisher LDA.

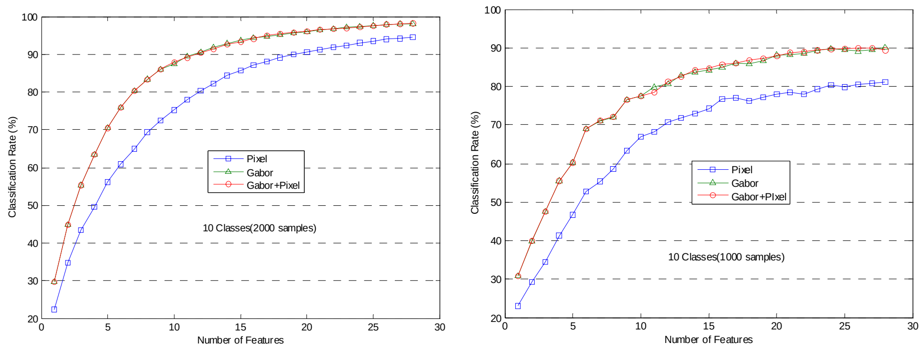

Figure 3. Classification result with model (μ_c, Σ_c), feature selection by Greedy seach.

4 Summary and conclusion

I have presented an algorithm for handwritten digits recognition by using spaitial Gabor features. I designed a Gabor filterbank capable of capturing the energy of spatial variations at different orientations and scales. Three different combinations of feature vectors (Pixel, Gabor, and Gabor+Pixel) were tested in classification experiments. The results show that the Gabor features successfully separate classes with less features used than those used in Pixel features. I have also demonstrated features selection using greedy search algorithm gives a satisfactory result for the dataset used in this paper. The combined Pixel and Gabor features are shown to be effective for handwritten digits recognition.

References

[1] Trier, O. D., Jain, A. K. and Taxt, T. "Feature extraction methods for character recognition: A survey", Pattern Recognition, Vol. 29, No. 4, pp. 641-662, (1996).

[2] Wierer, J.; Boston, N.," A Handwritten Digit Recognition Algorithm using Two-Dimensional Hidden Markov Models for Feature Extraction," Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on, Volume 2, 15-20 April 2007 Page(s):II-505 - II-508.

[3] Weldon, T. and Higgins, W. E., "Designing multiple gabor filters for multitexture image segmentation," Optical Engineering 38(9), 1478-1489 (1999).

[4] Jain, A. and Healey, G., "A multiscale representation including opponent color features for texture recognition," 7(1), 124-128 (1998).

[5] Daugman, J. G. "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," J. Opt. Soc. Am. A 2(7), 1160-1169, 1985.

[6] Clausi, D. A. and Jernigan, M. E. "Designing Gabor filters for optimal texture separability," Pattern Recognition 33(11), 1835-1849, 2000.

[7] Jain, A. K. and Farrokhnia, F. "Unsupervised texture segmentation using Gabor filters," Pattern Recognition 24(12), 1167-1186, 1991.

[8] Bovic, A. C., Clark, M. and Geisler, W. S. "Multichannel texture analysis using localized spatial filters," IEEE Trans. Pattern Anal. Mach. Intell. 12(1), 55-73, 1990.