Hepatic CT image retrieval based on the combination of Gabor filters and support vector machine

Lijun Jiang (蒋历军), Yongxing Luo (罗永兴), Jun Zhao (赵 俊), and Tiange Zhuang (庄天戈)

Department of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200240

Received October 22, 2007

Content-based image retrieval has been an active area of research for more than ten years. Gabor schemes and support vector machine (SVM) method have been proven effective in image representation and classification. In this paper, we propose a retrieval scheme based on Gabor filters and SVMs for hepatic computed tomography (CT) images query. In our experiments, a batch of hepatic CT images containing several types of CT findings are used for the retrieval test. Precision comparison between our scheme and existing methods is presented.

OCIS codes: 100.2000, 100.2960, 100.6950, 100.7410.
doi: 10.3788/COL20080607.0495.

Content-based image retrieval (CBIR) has been an active research topic for many years. The application of CBIR techniques in medical field has also been proposed for the use in picture archiving and communications system (PACS)[1], case databases[2], and in an even more general sense[3]. In CBIR, texture is the most important information which can be used to characterize an image[4]. Up to now, several methods achieving effective feature extraction have been proposed[5−7]. Among the methods, Gabor filter[7−11] is widely used, which is mainly motivated by two factors[5]: primitives of image representation in vision have a wavelet form similar to Gabor elementary functions (EFs)[12] and Gabor function represents a minimum in terms of the spread of uncertainty in space and spatial frequency[13]. Support vector machine (SVM) method[14−18] is a reliable classification technique, which is based on statistical learning theory and has been introduced for solving pattern recognition problems.

Zhao *et al.*[9] proposed a scheme of hepatic computed tomography (CT) images retrieval using Gabor features. However, the retrieval by this scheme was only based on the distance between the query image and the others, and only part of the Gabor features was used, which induced unsatisfying performance because different classes of images mixed together in feature space. In this paper, we propose a retrieval scheme based on the combination of Gabor features and SVMs for hepatic CT images query. All the features were used in our scheme.

For each hepatic CT image, the region of interest (ROI) is selected manually and the corresponding feature vectors are extracted through the Gabor transform. An image is represented by a 6-feature vector defined by Porat and Zeevi[12]. Then images are classified with the SVM method. Finally, we execute the image retrieval based on the classification. In the experiment, 521 hepatic CT images are used and a CBIR system for hepatic CT images is built based on these images and the method we introduced above. The effectiveness of our method is verified and the comparison between our method and that of Porat and Zeevi is demonstrated.

According to the Gabor approach, an image $\Phi(x, y)$ can be represented as a linear combination of EFs[12]:

$$\Phi(x,y) = \sum_{m_x n_x m_y n_y} a_{m_x n_x m_y n_y} \cdot f_{m_x n_x m_y n_y}(x,y), \quad (1)$$

where $a_{m_x n_x m_y n_y}$ is the coefficient of the order (m_x, n_x, m_y, n_y), representing the relative weight of the respective EF in $\Phi(x, y)$, $f_{m_x n_x m_y n_y}$ is the EF of the order (m_x, n_x, m_y, n_y),

$$f_{m_x n_x m_y n_y}(x,y) = g(x - m_x D_x, y - m_y D_y) \times \exp(i n_x W_x x + i n_y W_y y), \quad (2)$$

where $W_x \cdot D_x \leq 2\pi$, $W_y \cdot D_y \leq 2\pi$ must be satisfied and $g(\cdot,\cdot)$ is a two-dimensional (2D) normalized window function. The function $f_{m_x n_x m_y n_y}(x, y)$ is situated at the point $(x = m_x D_x,\ y = m_y D_y)$ of the Gabor lattice and has a spatial frequency of $(\omega_x = n_x W_x,\ \omega_y = n_y W_y)$. The constants, D_x, D_y and W_x, W_y, are the basic sampling intervals along the spatial and the spatial-frequency axes, respectively. When $g(\cdot,\cdot)$ in Eq. (2) is a Gaussian window function, the Gabor EFs are not orthogonal and thus the coefficients $\{a_{m_x n_x m_y n_y}\}$ are calculated using an auxiliary function $\gamma(\cdot,\cdot)$ which is biorthogonal in a certain sense to the window function $g(\cdot,\cdot)$:

$$a_{m_x n_x m_y n_y} = \iint \phi(x,y) \times \gamma^*(x - m_x D_x, y - m_y D_y) \times \exp(-i n_x W_x x - i n_y W_y y)\mathrm{d}x\mathrm{d}y. \quad (3)$$

It should be noted that when $g(\cdot,\cdot)$ is a square window function, the functions $g(\cdot,\cdot)$ and $\gamma(\cdot,\cdot)$ are identical[19].

Porat and Zeevi defined six localized features to analyze the texture of an image. The six features are dominant localized frequency (denoted by F), variance of the dominant localized frequency (VF), dominant orientation (T), variance of the dominant orientation (VT), mean of the localized intensity level (L), and variance of

the localized intensity level (VL)[19]:

$$F_{m_x m_y} = \frac{\sum_{n_x=1}^{N-1}\sum_{n_y=1}^{N-1}\left|a_{m_x n_x m_y n_y}\right|\sqrt{n_x^2+n_y^2}}{\sum_{n_x=1}^{N-1}\sum_{n_y=1}^{N-1}\left|a_{m_x n_x m_y n_y}\right|}, \quad (4)$$

$$VF_{m_x m_y} = \frac{\sum_{n_x=1}^{N-1}\sum_{n_y=1}^{N-1}\left|\sqrt{n_x^2+n_y^2}-F_{m_x m_y}\right|}{N^2}, \quad (5)$$

where the number of spectral components (along the frequency coordinates), N, is determined by the sampling rate of the digitized image.

$$T_{m_x m_y} = \frac{\sum_{n_x=1}^{N-1}\sum_{n_y=1}^{N-1}\left|a_{m_x n_x m_y n_y}\right|\theta(n_x,n_y)}{\sum_{n_x=1}^{N-1}\sum_{n_y=1}^{N-1}\left|a_{m_x n_x m_y n_y}\right|}, \quad (6)$$

$$VT_{m_x m_y} = \frac{\sum_{n_x=1}^{N-1}\sum_{n_y=1}^{N-1}\left|\theta(n_x,n_y)-T_{m_x m_y}\right|}{N^2}, \quad (7)$$

where $\theta(n_x, n_y)$ is defined as $\tan[\theta(n_x, n_y)] = (n_y/n_x)$ for $n_x \neq 0$ and $\tan[\theta(n_x, n_y)] = \pi/2$ for $n_x = 0$.

$$L_{m_x m_y} = \frac{1}{K}\sum_{x,y\in A(m_x,m_y)} I(x,y), \quad (8)$$

$$VL_{m_x m_y} = \frac{1}{K}\sum_{x,y\in A(m_x,m_y)}\left|I(x,y)-L_{m_x m_y}\right|, \quad (9)$$

where $A(m_x, m_y)$ is the set of K pixels belonging to the area defined by the window function centered according to m_x, m_y, and $I(x, y)$ is the intensity function.

In SVM method, the data are mapped into a higher dimensional input space, and an optimal separating hyperplane is constructed in this space. This basically involves solving a quadratic programming problem. Kernel functions and parameters are such chosen that a bound on the Vapnik-Chervonenkis (VC) dimension is minimized. Besides the linear case, SVMs based on polynomials, splines, radial basis function networks, and multilayer perceptrons have been successfully applied. Being based on the structural risk minimization principle and capacity concept with pure combinatorial definitions, the quality and complexity of the SVM solution do not depend directly on the dimensionality of the input space[20].

Given a training set of instance-label pairs (x_1, y_1), $(x_2, y_2), \cdots, (x_n, y_n)$, where $x_i \in R^n$, $y \in \{-1, 1\}^l$, the SVMs require the solution of the following optimization problem:

$$\min_{w,b,\xi}\frac{1}{2}w^{\mathrm{T}}w + C\sum_{i=1}^{l}\xi_i, \quad (10)$$

subject to

$$y_i\left[w^{\mathrm{T}}\phi(x_i)+b\right] \geq 1-\xi_i, \quad (11)$$

$$\xi_i \geq 0. \quad (12)$$

Here training vectors x_i are mapped into a higher (maybe infinite) dimensional space by the function Φ. Then SVM finds a linear separating hyperplane with the maximal margin in this higher dimensional space. $C > 0$ is the penalty parameter of the error term. Furthermore, $K(x_i, x_j) = \phi(x_i)^{\mathrm{T}}\phi(x_j)$ is called the kernel function. Though new kernels are being proposed by researchers, the following kernels are most commonly used:

linear:

$$K(x_i, x_j) = x_i^{\mathrm{T}}x_j, \quad (13)$$

polynomial:

$$K(x_i, x_j) = \left(\gamma x_i^{\mathrm{T}}x_j + r\right)^d, \quad \gamma > 0, \quad (14)$$

radial basis function (RBF):

$$K(x_i, x_j) = \exp\left(-\gamma\left\|x_i - x_j\right\|^2\right), \quad \gamma > 0. \quad (15)$$

In our experiment, we took use of the code from LIBSVM[21]. We first extracted the Gabor features from the hepatic CT images used in our experiments. Then the images were classified using the SVMs. Finally, we proceeded to execute the image retrieval based on the classification.

In the experiment, five types of hepatic CT manifestations were included: the multi-focal nodule (MFN), the uniformly low attenuation (ULA), the low attenuation with infiltration (LAI), the lipiodol retention (LR), and images of normal people (NRM). There were totally 521 CT images employed in the experiment, including 108 MFN, 101 ULA, 102 LAI, 109 NRM, and 101 LR, which were collected from the Eastern Hepatobiliary Surgery Hospital of the Second Military Medical University. The size of the images was 512×512 pixels and the gray level was stored at 12 bits per pixel. The ROI was designated manually and the size of ROI was 64×64 pixels, as shown in Fig. 1.

For each hepatic CT image, the ROI was selected manually and the corresponding feature vectors were extracted through the Gabor transform. An image was represented by a 6-feature vector $[F, VF, T, VT, L, VL]$, as shown by Eqs. (4)—(9). In experiment, the Porat and Zeevi's features scheme was used for the description of the liver CT images. The parameters involved in the computation of Gabor filters were: $W_x = W_y = 2\pi/64$, $D_x = D_y = 64$, $n_x, n_y = 1, \cdots, 6$, $m_x = m_y = 1$.

The Gabor features of the CT images using the scheme described by Porat and Zeevi is demonstrated in Fig. 2. It should be noted that only a small part of images, which were randomly selected from the images involved in the experiment, are depicted in Fig. 2 for clarity.

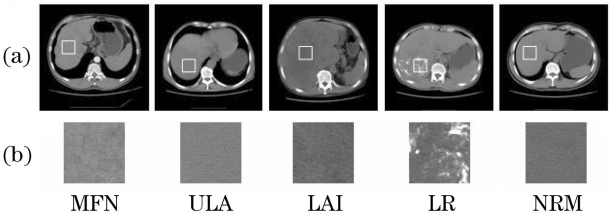

Fig. 1. (a) Hepatic CT images and (b) corresponding ROIs.

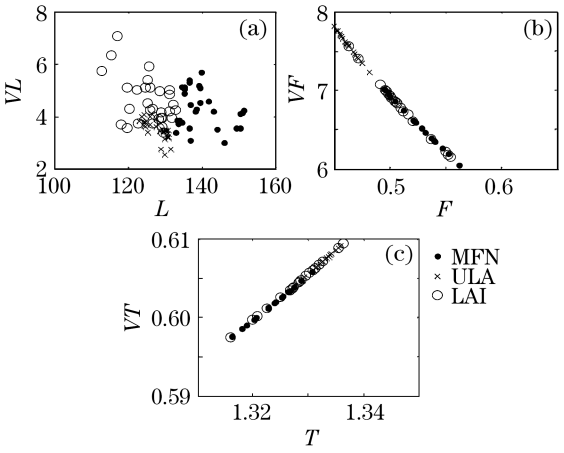

Fig. 2. Gabor features of the hepatic CT images.

Table 1. Sensitivity and Specificity of SVM Classifier Based on Gabor Features

	MFN	ULA	LAI	LR	NRM
Sensitivity	0.9259	0.9505	0.7745	1	0.9633
Specificity	0.9758	0.9738	0.9761	1	0.9782

After the CT images were classified based on the Gabor features with the SVM, the images were divided into four parts. We randomly selected one part as testing data and the other three parts as training data. The sensitivity and specificity of image classification are listed in Table 1. Sensitivity and specificity are defined as follows:

$$\text{sensitivity} = \frac{\text{positive items classified as positive}}{\text{all positive items}} = \frac{TP}{TP+FN}, \quad (16)$$

$$\text{specificity} = \frac{\text{negative items classified as negative}}{\text{all negative items}} = \frac{TN}{FP+TN}, \quad (17)$$

where TP, TN, FP and FN denote true positive, true negative, false positive, and false negative, respectively.

After classifying images with the support vector machines method, we executed the image retrieval based on the classification. It should be noted that the Euclidean distance was used for similarity measure and all features were normalized before the distance was calculated.

The effectiveness of the scheme is described by the average retrieval precision, which is the mean of the precision of all the query images within a certain type of texture. The precision used here is defined as

$$\text{precision} = \frac{\text{number of relevant images retrieved}}{\text{number of images retrieved}}. \quad (18)$$

The performance of our scheme is demonstrated in Figs. 3 and 4, where the retrieval performance of Porat and Zeevi's (without the classification by SVM) is also given for comparison.

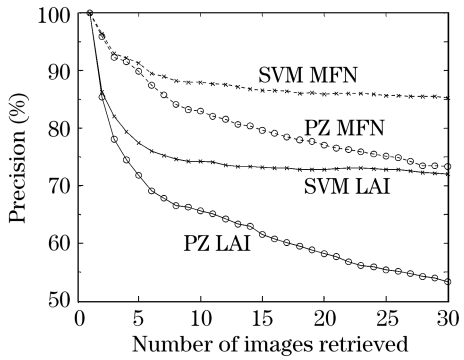

Fig. 3. Relationship between precision and number of images retrieved for the classes of MFN and LAI. PZ means the results of Porat and Zeevi's scheme; SVM means the results of SVM scheme.

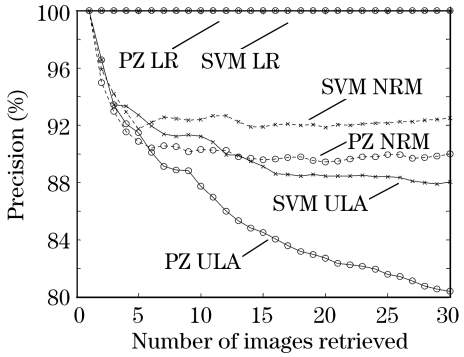

Fig. 4. Relationship between precision and number of images retrieved for the classes of LR, ULA, and NRM.

From the results, we find that the classification sensitivity and specificity of our scheme are satisfying except the class of LAI. The effectiveness of the image retrieval based on Gabor features with the SVM classifier (our scheme) is much better than the one without the SVM classifier. However, the performance of the LAI class is not as well as that of the other classes.

In this paper, we propose a retrieval scheme based on Gabor features and SVM method for hepatic CT images query. Experimental results show that the effectiveness of the image retrieval based on the Gabor features with the SVM classifier (our scheme) is much better than the one without the SVM classifier. However, the performance of the LAI image class is not satisfying, which is mainly because of the LAI's widely scattering and mixing with other types of images in feature space. So other features representing the image texture is needed to be explored.

The authors thank Dr. C. Zhao for offering the hepatic CT images. This work was supported by the Joint National Natural Science Foundation of China under Grant No. 30770589. L. Jiang's e-mail address is jianglijun@sjtu.org.

References

1. Y. Tan, J. Zhang, Y. Hua, G. Zhang, and H. K. Huang, Proc. SPIE **6145,** 614515 (2006).
2. A. Rosset, O. Ratib, A. Geissbuhler, and J.-P. Vallée, Radiographics **22,** 1567 (2002).

3. H. Müller, A. Rosset, J.-P. Vallée, and A. Geissbuhler, Proc. SPIE **5371,** 99 (2004).
4. C. G. Zhao, H. Y. Cheng, Y. L. Huo, and T. G. Zhuang, in *Proceedings of 26th Annual International Conference of the IEEE EMBS* 1491 (2004).
5. B. S. Manjunath and W. Y. Ma, IEEE Trans. Pattern Anal. Machine Intell. **18,** 837 (1996).
6. R. Manthalkar, P. K. Biswas, and B. N. Chatterji, Pattern Recogn. Lett. **24,** 2455 (2003).
7. S. Li and J. Shawe-Taylor, Pattern Recogn. Lett. **26,** 633 (2005).
8. C. S. Sastry, M. Ravindranath, A. K. Pujari, and B. L. Deekshatulu, Pattern Recogn. Lett. **28,** 293 (2007).
9. C. Zhao, H. Cheng, and T. Zhuang, Chin. Opt. Lett. **2,** 383 (2004).
10. Y. Chang, F. Peng, G. Zhu, Y. Liu, and Y. Hu, Chinese J. Lasers (in Chinese) **30,** 282 (2003).
11. X. Huang, Z. Yang, and W. Liu, Information and Electronic Engineering (in Chinese) **4,** 22 (2006).
12. M. Porat and Y. Y. Zeevi, IEEE Trans. Pattern Anal. Machine Intell. **10,** 452 (1988).
13. J. G. Daugman, J. Opt. Soc Am. A **,** 1160 (1985).
14. B. E. Boser, I. M. Guyon, and V. N. Vapnik, in *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory* 144 (1992).
15. C. Cortes and V. Vapnik, Machine Learning **20,** 273 (1995).
16. S. Li, Y. Zhang, L. Dong, S. Chang, and J. Shen, Acta Photon. Sin. (in Chinese) **35,** 304 (2006).
17. S. Li, Y. Han, Y. Zhang, S. Chang, and J. Shen, Acta Opt. Sin. (in Chinese) **26,** 147 (2006).
18. X. Zhang, Acta Automatica Sinica (in Chinese) **26,** 32 (2000).
19. M. Porat and Y. Y. Zeevi, IEEE Trans. Biomed. Eng. **36,** 115 (1989).
20. J. A. K. Suykens and J. Vandewalle, Neural Processing Lett. **9,** 293 (1999).
21. C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines" (2001). http://www.csie.ntu.edu.tw/∼cjlin/libsvm/index.html (Apr. 10, 2007).