Proceedings of the 6th WSEAS International Conference on Applied Informatics and Communications, Elounda, Greece, August 18-20, 2006 (pp106-110)

Medical Image Classification by Supervised Machine Learning

PEI-CHENG CHENG[1,3], BEEN-CHIAN CHIEN[2], and WEI-PANG YANG[4]
[1] Department of Computer Science, National Chiao Tung University,
1001 Ta Hsueh Rd., Hsinchu, Taiwan 30050, R.O.C.
[2]Department of Computer Science and Information Engineering,
National University of Tainan,
33, Sec. 2, Su Line St., Tainan, Taiwan 70005, R.O.C.

[3] Department of Information Management, Ching Yun University
229, Chien-Hsin Rd.,Jung-Li. Taiwan 320. R.O.C.
[4] Department of Information Management, National Dong Hwa University
1, Sec. 2, Da Hsueh Rd., Shou-Feng, Hualien, Taiwan 97401, R.O.C.

Abstract: - In this paper, Support Vector Machine (SVM) was used to learn image feature characteristics for image classification. Several image visual features describe the shape, edge, and texture of image (including histogram, spatial layout, coherence moment and gabor features) have been employed in this paper to categorize the 500 test images into 46 classes.
The result shows that the spatial relationship of pixels is a very important feature in medical image data, because medical image data always have similar anatomic regions (lung, liver, head, and so on).

Key-words: - Medical Image Classification; Support Vector Machine;

1. Introduction

Image retrieval techniques are essentially required due to the enormous amount of digital images produced day by day. Content-based image retrieval using primitive image features is promising in retrieving visually similar images. The QBIC system [2] is one of the well-known content-based systems, and other famous systems are Blob World [3][4], VIPER/GIFT[5], and SIMPLIcity[6]. Various image features and similarity metrics have already been proposed for general images.

Automatic image annotation or image classification is an area of active research in the field of machine learning and pattern recognition. Retrieval systems have traditionally used manual image annotation for indexing and then later retrieving their image collections. However, manual image annotation is an expensive and labor intensive procedure [13]. Here, an automatic approach was proposed to categorize images based on a supervised learning technique. In supervised classification, a collection of training images (labeled images) are given, and the problem is to label a newly encountered, yet unlabeled image. Each instance in the training set contains category or class specific labels and several image feature descriptors in the form of a combined feature vector.

In this paper, the Support Vector Machine (SVM) is used to learn image feature characteristics. Based on the SVM model, several image features that consider, from the viewpoint of human, the invariance in image rotation, shift and illumination invariance are employed in our system. The goal of SVM is to produce a model which will predict target value or category of images with highest probability or confidence in the testing set which are given as input feature vectors to the classification system.

In the experiments, the support vector machine plays as a classifier to categorize the 500 test images into 46 classes. The experiment result shows that proposed image features are useful in the medical image application. The rest of this paper is organized as follows. In Section 2, the employed image features are described. Section 3 illustrates the SVM model that is used to classify the training data. In Section 4, the experiment results are discussed. Finally, Section 5 provides concluding remarks and future directions for medical image retrieval.

2. Image features

In an image retrieval system, image features are extracted from pixels. The extracted features are then used for similarity comparison. For fast response, image features must be concise, and for precision, image features must contain meaningful information to represent the image itself. Image features will directly affect the retrieval result. In this paper we examine several image features to understand which features will have good performance in medical image application.

When designing the image features, to emphasize the contrast of an image and handle images with little illuminative influence, we normalize the value of a pixel before quantization. In [10] we proposed a relative normalization method. First, pixels of image are clustered into four clusters by the K-means clustering method [7]. The four clusters was sort ascendant according to their mean values. The mean of the first cluster was shifted to the value 50 and the fourth cluster was shifted to 200; then, each pixel in a cluster is multiplied by a relative weight for normalization. Let m_{c1} be the mean value of cluster 1 and m_{c4} is the mean value of cluster 4. The normalization formula of pixel p(x,y) is defined in Eq. (1).

$$p(x,y)_{normal} = (p(x,y) - (m_{c1} - 50)) \times \frac{200}{(m_{c4} - m_{c1})} \quad (1)$$

After normalization, we resize each image into 128*128 pixels, and use one level wavelet with Haar wavelet function to generate the low frequency and high frequency sub-images. Process an image using the low pass filter will obtain an image that is more consistent than the original one; on the contrary, processing an image using the high pass filter will obtain an image that has high variation. After wavelet transform, an image get four sub bands of wavelet coefficient that are LL, LH, HL, and HH bands. The high-frequency part highlights the contour of the image. By performing the OR operation for LH, HL, and HH bands, we get the contour of a medical image.

2.1 Histogram layout

Histogram [9] is a prime image feature for image retrieval. Histogram is invariant in image rotation. It is easy to implement and has good results in color image indexing. Because a radiotherapy medical image only consists of gray-level pixels, spatial relationship becomes very important. Medical images always contain particular anatomic regions (lung, liver, head, and so on); therefore, similar images have similar spatial structures. For spatial relationship, an image was divided into nine sections and the histogram of each section was calculated respectively.

The LL band was used for histogram layout and coherence analysis. To get the gray-spatial histogram, we divide the LL band image into nine areas. The gray values are quantized into 16 levels for computational efficiency. In the histogram layout, a gray value may be quantized into several bins to improve the similarity between adjacent bins. We set an interval range δ to extend the similarity of each gray value. The histogram layout estimates the probability of each gray level that appears in a particular area. The probability equation is defined in Eq. (2), where δ is set to 10, p_j is a pixel of a given image, and m is the total number of pixels in this image. In our implementation, a total of 144 bins was used for the histogram layout.

$$h_{c_i}(I) = \frac{\sum_{j=1}^{m} \frac{[p_j - \frac{\delta}{2}, p_j + \frac{\delta}{2}] \cap c_i}{\delta}}{m} \quad (2)$$

2.2 Coherence Moment

One of the problems to devise an image representation is the semantic gap. The state-of-the-art technology still cannot reliably identify objects. The coherence moment feature attempts to describe the features from the human's viewpoint in order to reduce the semantic gap.

We cluster an image into four classes by the K-means algorithm. After clustering an image into four classes, The number of pixels (COH_κ), mean value of the gray values (COH_μ) and standard variance of the gray values (COH_ρ) in each class was calculated. For each class, connected pixels are grouped in the eight directions as an object. If an object is bigger than 5% of the whole image, it was viewed as a big object; otherwise it is a small object. Big objects (COH_o) and small objects (COH_v) are caculated in each class, and use COH_o and COH_v as parts of image features.

Since we intend to know the reciprocal effects among pixels, we use the smoothing method to reflect the influence of neighbor pixels of the image. If the spatial distribution of the pixels in two images is similar, they will also be similar after smoothing. If their spatial distributions are quite different, they may have a different result after smoothing. After smoothing, we cluster an image into four classes and calculate the number of big objects (COH_τ) and small objects (COH_ω). Each pixel will be influenced by its neighboring pixels. Two close objects of the same class may be merged into one object. Then, we can analyze the variation between the two images before

and after smoothing. The coherence moment of each class forms a seven-feature vector, (COH_κ, COH_μ, COH_ρ, COH_o, COH_v, COH_τ, COH_ω). The coherence moment of an image is a 56-feature vector that combines the coherence moments of the four classes.

2.3 Gray Correlogram.

The contour of a medical image contains rich information. A broken bone in the contour may be different from a healthy one. Thus we proposed a representation that can estimate the partial similarity of two images and can be used to calculate their global similarity.

We analyze the image pixels by modified correlogram algorithm. The definition of the correlogram is in Eq. (3). Let D denote a set of fixed distances $\{d_1, d_2, d_3,\ldots, d_n\}$. The correlogram of an image I is defined as the probability of a color pair (c_i, c_j) at a distance d.

$$\gamma^d_{c_i,c_j}(I) = \Pr_{p_1\in c_i, p_2\in I}\{p_2 \in c_j \| p_1 - p_2 \| = d\} \quad (3)$$

For computational efficiency, the autocorrelogram is defined in Eq. (4)

$$\lambda^d_{c_i}(I) = \Pr_{p_1\in c_i, p_2\in I}\{p_2 \in c_i \| p_1 - p_2 \| = d\} \quad (4)$$

The contrast of a gray image dominates human perception. If two images have different gray levels they still may be visually similar. Thus the correlogram method cannot be used directly.

Our modified correlogram algorithm works as follows. First, we sort the pixels in descending order. Then, we order the results of the preceding sorting by ascendant distances of pixels to the center of the image. The distance of a pixel to the image center is measured by the L2 distance. After sorting by gray value and distance to the image center, we select the top 20 percent of pixels and the gray values higher than a threshold to estimate the autocorrelogram histogram. The threshold was set zero in this work. Any two pixels have a distance, and we estimate the probability that the distance falls within an interval. The distance intervals are {(0,2), (2,4), (4,6), (6,8), (8,12), (12,16), (16,26), (26,36), (36,46), (46,56), (56,76), (76,100)}.We calculate the probability of each interval to form the correlogram vector.

2.4 Gabor texture features

Gabor filter is widely adopted to extract texture features from images for image retrieval [1], and has been shown to be very efficient. Gabor filters are a group of wavelets, with each wavelet capturing energy at a specific frequency and a specific direction. Expanding a signal using this basis provides a localized frequency description, therefore capturing local features/energy of the signal. Texture features can then be extracted from this group of energy distributions. The scale (frequency) and orientation tunable property of Gabor filter makes it especially useful for texture analysis.

The Gabor wavelet transformation W_{mn} of Image I(x,y) derived from Gabor filters according to [1] is defined in Eq. (5)

$$W_{mn}(x,y) = \int I(x,y) g^*_{mn}(x-x_1, y-y_1)dx_1dy_1. \quad (5)$$

The mean μ_{mn} and standard deviation σ_{mn} of the magnitude $|W_{mn}|$ are used for the feature vector, as shown in Eq. (6).

$$\mu_{mn}(x,y) = \iint |W_{mn}(x,y)|\, dxdy, \text{ and } \delta_{mn} = \sqrt{\iint (|W_{mn}(x,y)| - u_{mn})^2\, dxdy} \quad (6)$$

.

This image feature is constructed by μ_{mn} and σ_{mn} of different scales and orientations. Our experiment uses four (S=4) as the scale and six (K=6) as the orientation to construct a 48 feature vectors $\bar{f}$, as shown in Eq. (7).

$$\bar{f} = [u_{00}, \delta_{00}, u_{01}, \delta_{01}, \ldots\ldots, u_{35}, \delta_{35}]. \quad (7)$$

2.5 Relative vector

Relative vector is proposed for illumination invariant represent the content of medical image. Any two neighbor pixel p_i and p_j have a relationship; the relation R is defined as fallows:

R(p_i, p_j)={ 0 if gray_value(p_i)=gray_value(p_j)
1 if gray_value(p_i)>gray_value(p_j)
2 if gray_value(p_i)<gray_value(p_j)
}

As pixel(x,y) of image, the relative vector(RV) is described as following.

Each pixel references four directions (up, down, right, left) to generate a relative vector RV.

RV(pixel(x,y))=
<R(pixel(x,y),pixel(x-1,y)),R(pixel(x,y),pixel(x+1,y)),
R(pixel(x,y),pixel(x,y+1)),R(pixel(x,y), pixel(x,y-1))>

All pixel of image will be translated into a relative vector, RV$\in$ {[0-2],[0-2],[0-2],[0-2]}. The relative vector total have 81 classes, we can translate the relative vector RV(pixel(x,y)) into an integer for fast computational. We hash the relative vector into one of 81 classes by Eq. (8).

Let RV(pixel(x,y))=<h, I, j, k>,
Class(RV(pixel(x,y)))=h*3^3+i*3^2+j*3^1+k (8)

The representation based on the relativity

relation between pixels to denote an image will have illumination invariant character. Second, as above definition each pixel references four neighbors (up, down, right, left) to generate the RV. The relative vector we proposed will also contains local spatial information.

Up to now, an Image will be translated into an illumination invariant representation. Each pixel will fall in a class. We calculate the occurrence frequency of each class as the glob feature of an image. As an image, the final signature representation totally has 81 vectors.

3. Classification method

In this work, Support Vector Machine (SVM) [11] was used to learn image feature characteristics. SVM is an effective classification method. Its basic idea is to map data into a high dimensional space and find a separating hyper plane with the maximal margin. Given a training set of instance-label pairs (x_i, y_i), i=1,…,k where $x_i \in R^n$ and $y \in \{1,-1\}^k$, the support vector machine optimizes the solution of the following problem:

$$\underset{w,b,\phi}{Min}\left(\frac{1}{2}w^T w + C\sum_{i=1}^{k}\phi_i\right) \text{ and } y_i(w^T\psi(x_i)+b) \geq 1-\phi_i, \quad (9)$$

$$\phi_i \geq 0$$

Training vectors x_i are mapped into a higher dimensional space by the function ψ. Then SVM finds a linear separating hyper plane with the maximal margin in this higher dimensional space. C>0 is the penalty parameter of the error term. Furthermore, $K(x_i,x_j) \equiv \psi(x_i)^T\psi(x_j)$ is called the kernel function. In this paper we use LIBSVM [12] to classify the training data with a radial basis function or a polynomial function as the kernel function. The radial basis function (RBF) and the polynomial function used is defined in Eq. (10) and Eq. (11), respectively, where γ , r, and d are kernel parameters.

$$K(x_i,x_j) = \exp(-\gamma \| x_i - x_j \|^2), \gamma > 0. \quad (10)$$

$$K(x_i,x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0. \quad (11)$$

4. Experiment and Result

In this experiment, there are 2500 images in total, where 2000 of them are used as the training dataset and the rest 500 images are used for testing. We use Error rates to evaluate the efficient of proposed method. The error rate is noted that each 0.2% corresponds to 1 misclassification, because the test images has total of 500 images to be classified.

In the image features used in our experiment. The histogram layout feature divides an image into nine sections contain spatial information. The coherence moment considers the image rotation and shift, but cannot carry much spatial information. Gabor method describes the texture feature and relative vector have illumination invariant characters.

In the experiments, all features are considered has the best result, with error rate 16.3%. The Histogram layout has error rate 18.3%. Relative vectors has error rate 19.2%. Combined all features will improve the result of experiments. In medical image data, the spatial distribution of pixels is very significant. The coherence moment contains the least spatial information, thus it has the worst result.

One experiment for a nearest neighbor classifier that scales down images to 64*64 pixels and use the Euclidean distance for comparison has error rate 32.8%, which means that 164 images are misclassified. This experiment shows that the SVM method has better performance than the Euclidean distance metric.

Image features	SVM kernel function	error rate %
Histogram layout: 144 vectors	radial basis function	18.3
All features: 341 vectors	radial basis function	16.3
All features: 341 vectors	polynomial	21.4
relative vectors: 81	radial basis function	19.2
Coherence Moment: 56 vectors	radial basis function	22.5

Table 1: The experiment results of different features.

5. Conclusion

In this paper, several image features are examined for medical image data. The medical image application is unlike general-propose images. In general purpose images, the representation always consider the invariance in image rotation, zooming and shift. Medical images have more stable camera settings than general purpose images; therefore, the spatial information becomes very important in medical images, and we must improve the representation regarding spatial relation in this kind of images.

The support vector machine as a classifier is very efficient, but it seems that the SVM lacks the ability of feature selection. In the future, we plan to develop the feature selection technology for the SVM to

improve the performance.

References:

[1] B. S. Manjunath and W. Y. Ma. "Texture features for browsing and retrieval of large image data" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (Special Issue on Digital Libraries), Vol. 18(8), August 1996, pp. 837-842.

[2] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, P. Yanker, Query by Image and Video Content: The QBIC system, *IEEE Computer* 28 (9), 1995, pp. 23-32.

[3] C. Carson, M. Thomas, S. Belongie, J. M. Hellerstein, and J. Malik, Blobworld: A system for region-based image indexing and retrieval, in: D. P. Huijsmans, A. W. M. Smeulders (Eds.), *Third International Conference On Visual Information Systems* (VISUAL' 99), no. 1614 in Lecture Notes in Computer Science, Springer{Verlag, Amsterdam, The Netherlands, 1999, pp. 509-516.

[4] S. Belongie, C. Carson, H. Greenspan, and J. Malik, Color and texture based image segmentation using EM and its application to content-based image retrieval, in: *Proceedings of the International Conference on Computer Vision (ICCV'98)*, Bombay, India, 1998, pp. 675-682.

[5] D. M. Squire, W. Muller, H. Muller, and J. Raki. Content-Based Query of Image Databases, Inspirations from Text Retrieval: Inverted Files, Frequency-Based Weights and Relevance Feedback. *In Scandinavian Conference on Image Analysis*, Kangerlussuaq, Greenland, June 1999, pp.143-149.

[6] J. Z. Wang, J. Li, and G. Wiederhold. SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 23(9), September 2001, pp. 947-963.

[7] J., Han, M., Kamber,: Data Mining: Concepts and Techniques. *Academic Press*, San Diego, CA, USA 2001.

[8] D. Keysers, W. Macherey, H. Ney, and J. Dahmen. Adaptation in Statistical Pattern Recognition using Tangent Vectors. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 26(2), February 2004, pp. 269-274.

[9] Swain M.J. and D. H. Ballard, "Color Indexing", *International Journal of Computer Vision,* Vol. 7, 1991, pp. 11-32.

[10] Pei-Cheng Cheng, Been-Chian Chien, Hao-Ren Ke, and Wei-Pang Yang, *"KIDS's evaluation in medical image retrieval task at ImageCLEF 2004",* Working Notes for the CLEF 2004 Workshop September, Bath,UK 2004 , pp. 585-593.

[11] Boser, B., I. Guyon, and V. Vapnik, "A training algorithm for optimal margin classifiers," *In Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, 1999.

[12] Chang, C.-C. and C.-J. Lin, LIBSVM: a library for support vector machines. http://www.csie.ntu.edu.tw/~cjlin/libsvm, 2001.

[13]Chang, E. Kingshy, G. Sychay, G. Gang, CBSA: Content-based Soft Annotation forMultimodal Image Retrieval Using Bayes Point Machines., *IEEE Transactions on Circuitsand Systems for Video Technology*, 13, 2003, pp26–38.