CONTENT-BASED RETRIEVAL OF CALCIFICATION LESIONS IN MAMMOGRAPHY

Song Li-xin, Chang Rui-feng,Wang Qian

Harbin University of Science and Technology, Harbin 150040, China
T: 86-451-86396216; F: 86-451-86396201; E:lixins99@yahoo.com.cn

Abstract

In order to assist doctors to detect micro- calcifications，about the similar lesions retrieval problem of mammographic micro-calcification cluster, we develop a new algorithm with multi-feature fusion and relevance feedback based on the study of single feature and feature fusion using single distance measure image retrieving techniques, this method adopts multi-distance measure to calculate the similarity directing at different features. Experiment is based on mammography image database which contains 250 mammography images and each image contains calcification cluster, we verified the retrieval performance by the precision - recall ratio (PVR) of single feature, feature fusion and relevance feedback. Experimental results show that the method has a better retrieval result than these methods which based single feature and feature fusion which using single distance measurement.

Key words: *mammography image, calcification lesions, Content-based Image Retrieval, feature fusion, relevance feedback, multi-distance measure*

1. INTRODUCTION

With the number of medical images increasing，the capability to retrieval relevant images from large database is becoming more and more important. Although the progress made in general area of image retrieval in recent years, its success in biomedicine thus far has been quit limited [1]. In addition, mammography images have difference from general images, they have some characters, such as high gray-scale resolution, more visually similar. Therefore, it has more difficult to retrieve similar images in medical sense from database, it need to understand sufficiently the knowledge of diagnostic imaging, pathology and physiology [2]. It is an in-depth study problem that how to retrieve calcification lesions according to the characteristics of the mammography images, so that the retrieval result can help medical diagnosis.

At present, there are some works have explored the use of CBIR in mammographic calcification lesions retrieval. For example, El-Naqa et al. [3] proposed an approach to the retrieval of digital mammograms using micro-calcification clusters. They explored the use of neural networks and support vector machines, in a two-stage hierarchical learning network to predict perceptual similarity from similarity scores collected in human-observer studies. Chia-Hung Wei et al. [4] proposed the method that using six relevance feedback algorithms, which fall in the category of query point movement, for improving system performance. Although CBIR has been many applications proposed for several applications abroad, one encounters a “semantic gap” between the quantitative features used to represent the images and the interpretation of the images by users who are experts in the domain of application [5]. This leads to the need to guide the retrieval algorithm by incorporating the user’s judgment of similarity and the relevance of each retrieved [6, 7].

Because simple distance measure based on a single method of similarity measure calcification of breast lesions can not meet the similarity of retrieval results in the medical sense [3], in this paper, we propose a new similarity measurement that uses multi-distance, which was determined by he optimal query rules, it solved the defect of single distance existing, and combined with the user's relevance feedback to adjust the characteristics component weight dynamically, thus the similarity of medical sense was enhanced.

2. IMAGE RETRIEVAL ALGORITHM

2.1. Feature Extraction

Conventional content-based retrieval system aimed at improving the visual similarity between retrieval images and the query images, but the mammography images are visually similar, so the process of feature selection should not simply base on visual sense similarity. Therefore, doctors will be more inclined to see the same types of images as similar images, this is the medical sense similarity.

In this paper, feature selection bases on the following principles: if a feature was effective in classification, it was also effective in the retrieval [8]. At the same time, taking into account the performance and characteristics of calcifications, and in accordance with the theory of computer image recognition and machine vision, we use the following search features:

(1) Gray-Scale Features: We extracted the mean, variance, kurtosis, skewness, entropy, and energy characteristics as the gray features.
(2) Texture Features: We combine the Tamura [9] and Gabor [10] features to compose the texture feature.
(3) Shape Feature: The shape feature is composed of the seven invariant moments [11] and five other features [3]: Cross sectional area, Compactness, Eccentricity, Density, Solidity.

Due to the ranges of feature components between each feature are different, we should normalize these feature components using Gauss normalization to make them have the same range [-1, 1].

The specific method is as following:

$$f_{i,j} = (f_{i,j} - \mu_i) / 3\sigma_i \quad (1)$$

Where $f_{i,j}$ denotes the feature component of feature f_i, μ_i and σ_i denote the mean and the standard deviation of f_i.

This allows almost all the characteristics value falling into the range of [-1, 1], the values outside the range are set to -1 and 1, so we assure all the values are in [-1, 1].

2.2. Image Retrieval based on feature fusion and multi-distance similarity measure

We integrate the gray-scale feature, shape feature and texture feature of the mammography images which contain micro-calcification cluster to retrieve the calcification lesions. At the same time, we adopt different distance measurement to compute similarity directing at lack of single distance measure.

As the important degree of every feature is different in retrieval, therefore, we should adjust the weight of these features before computing similarity. The specific method is as follows:

(1) all the weight $W = [W_i, W_{ij}]$ will be initialized $W0$ according to the number of the extracted features and feature components they include, so that all the features and characteristics of components have the same weight: $W_i = W0_i = 1/L$, $W_{ij} = W0_{ij} = 1/J_i$, where L denotes the number of image features, J_i denotes a characteristic feature of the number of components.

(2) Calculating the similarity between the query image and database image:

$$D(f_i) = \sum_j W_{ij} D_i(r_{ij}) \text{, } D = \sum_i W_i D_i(f_i) \quad (2)$$

Where D_i denotes the similarity among features, D denotes the total similarity.

2.2.1. Distance measure methods

At present, the common methods of similarity measure are as follows:

(1) Minkowski distance:

$$D(q,p) = (\sum_m (q_m - p_m)^L)^{1/L} \quad (3)$$

it is the Euclidean distance when $L = 2$.

(2) The histogram intersection distance:

$$D_{hi}(q,p) = 1 - \sum_{m=0}^{M-1} (h_q[m], h_p[m]) \Big/ |h_q|$$
$$|h| = \sum_{m=0}^{M-1} h[m] \quad (4)$$

(3) Quadratcic distance:

$$D(p,q) = \sqrt{(p_m - q_m)^T A (p_m - q_m)} \quad (5)$$

where $A = [a_{ij}]$ denotes the similarity matrix among features.

(4) Canberra Distance

$$D(q,p) = \sum_{m=0}^{M-1} |q_m - p_m| / (|q_m| + |p_m|) \quad (6)$$

2.2.2. The determination of the optimal similarity measure

We definite the optimal method of similarity measure in accordance with the optimal query rule to meet the similarity of mammographic calcification lesions retrieval in medical sense. The basic idea of the optimal query rule is:

(1) Every feature all has some measure methods, these combinations compose a collection, in the collection, every element expresses a query rule.
(2) For every query rule, compute the similarity distance between the query image and all the images in the database. Sort the distance in ascending order and construct a length- $2N$ rank list (N is the number which the user specifies how many retrieval images to be returned). Every element in the list is the image id, the image is similar to the query image
(3) Set the list that obtains in some query rule to reference list, the first N images in the reference are the retrieval results. At the same time, we obtain the corresponding list in accordance with the other query rules, define a rank function, it expresses rank number of the image in the list, if the images in the retrieval results is in the list, the function value is the rank number of the image in the list, otherwise, assign $2N+1$ to the function value.
(4) For the each image in the reference list, compute the overall rank numbers, these rank numbers are obtained by every query rule, then establish a length- N combined rank list, which contains the overall most similar N images, then return them to the user.
(5) The ranks for the retrieved images might not be the same as the user's perception, the user sends back a modified feedback rank list.
(6) Compute the rank difference in every list (the rank sum of the absolute difference of the first N images before and after the reorder by the user), the smaller the difference the better the query rule.

By the optimal query rule which is introduced above, we get the optimal similarity measure: gray feature adopts quadratcic distance, texture feature adopts Canberra distance, shape feature adopts the Euclidean distance.

2.2.3. The Normalization Among Features

The range of the similarity distance which obtained by different similarity measurement is not different, it causes the assignment imbalance of the weights, so we should normalize the distance. The method is as follows:

$$D_{i,j} = [1+(\frac{D_{i,j}-\mu_{iD}}{3\sigma_{iD}})] \Big/ 2$$

(7)

$D_{i,j}$ is the similarity distance based on the feature part $f_{i,j}$ between the query image and all the images in the database, μ_{iD} and σ_{iD} denote the mean and the standard deviation respectively of the similarity distance vector.

We normalize all the similarity distance which obtained by different similarity measurement, it makes sure that all the distance have the same importance.

2.3. Relevance Feedback Algorithm

If we only use the bottom features of the image to retrieve the calcification lesion, it can not always meet the use's "semantic", we introduce relevance feedback algorithm in order to reduce the semantic gap between user and the retrieval result. The basic idea is: return the first N retrieved images to the user, the user estimates the relevant degree between each returned image and the query image according to the requirements and the subjective views. We adjust dynamically the feature weight by the feedback information, so as to improve the medical similarity.

The Specific method is as follows:

$RT = [RT_1, RT_2, \cdots, RT_N]$ is the first retrieval result based on all features. The collection: $RT^{ij} = [RT^{ij}{}_1, RT^{ij}{}_2, \cdots, RT^{ij}{}_N]$ is composed of the first N (in the experience, N=32) images which are retrieved according to a certain feature component r_{ij}, the user judge the relevance of each image that is in the RT^{ij}, we let Score denotes the collection of feedback scores which given by user.

$$score_l \begin{cases} =1, \text{ relevant} \\ =0, \text{ not sure} \\ =-1, \text{ not relevant} \end{cases}$$

Next, we calculate the weights. First of all, we initialize the weights: $W_{ij} = 0$, then calculate the weights as following:

$$W_{ij} = \begin{cases} W_{ij} + score_l & RT_l^{ij} \in RT \\ W_{ij} + 0 & RT_l^{ij} \notin RT \end{cases}$$

If $W_{ij} < 0$, W_{ij} will be set 0. Finally, we normalize the weights: $W_{ij} = W_{ij} / \sum W_{ij}$.

We can carry out a new round retrieval by adjusting the weights as the above method introduced, it will be repeated until the user is pleased with the retrieval result.

3. EXPERIMENT AND RESULT ANALYSIS

The images used in the experiment are from the mammographic image database of the University of South Florida. The image contents were approved by clinical diagnosis and pathology. According to the doctor marked to the calcification cluster region, we extract 250 regions of interest (ROIs) which all contain calcification cluster. There are 145 ROIs which are malignant calcifications cluster, 105 ROIs are benign, the size of each ROI is 256 × 256 pixels.

We return the first five images to user as the retrieval result in the experiment. Because doctor are more inclined to consider the images which have the same kind lesions as the similar images, so we evaluate the retrieval performance according to this character.

We use query method of QBE (Query By Example), and select any image from the database as the query image to verify the retrieval performance by four methods. The experimental result shows the method that we proposed is superior to the traditional method which uses single-distance to measure similarity. Figure 1 is the result based on texture feature, Figure 2 is the result based on shape feature, Figure 3 is the result based on feature fusion using single distance, Figure 4 is the result based on feature fusion using multi-distance, Figure 5 is the result based on feature fusion using multi-distance and relevance feedback.

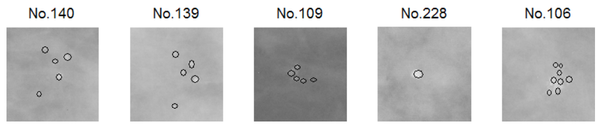

Fig. 1. The retrieval result based on texture feature

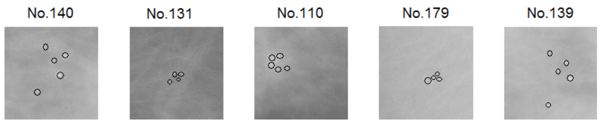

Fig. 2. The retrieval result based on shape feature

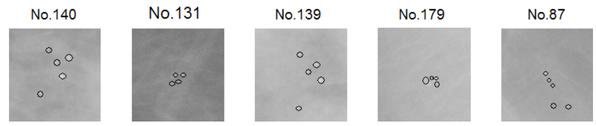

Fig. 3. The retrieval result based on single distance measure feature fusion

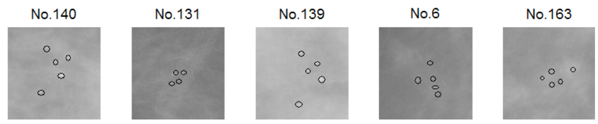

Fig. 4. The retrieval result based on multi-distance measure feature fusion

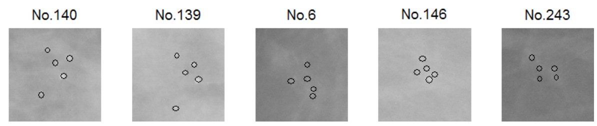

Fig. 5. The retrieval result based on feature fusion using multi-distance and relevance feedback

We choose 3 images from the database as the sample images to evaluate the precision-recall ratio of different methods quantificationally. See Table1.

Table 1 the precision and recall based various feature and similarity measurement

feature	Image A		Image B		Image C	
	Precision	Recall	Precision	Recall	Precision	Recall
texture	0.77	0.34	0.75	0.46	0.69	0.62
shape	0.79	0.42	0.77	0.52	0.72	0.65
Mixed feature using single distance	0.81	0.58	0.79	0.65	0.76	0.72
Mixed feature using multi-distance	0.84	0.64	0.82	0.74	0.79	0.78
RF after 5 times	0.89	0.81	0.86	0.83	0.83	0.85

We can see from Table 1, the method based on multi-distance similarity measure has a higher precision-recall ratio, the effect is superior clearly to the methods based on single feature and feature fusion using single distance to measure the similarity. After five times feedback, the recall and precision is improved markedly. But take into account the retrieval time, the feedback times should not too much. The retrieval result is closer to the user request by the feature fusion and relevance feedback.

4. CONCLUSION

We can see that the retrieval performance of the calcification lesion is related to the extracted features and the method that measures the similarity by the experiment. The method based on feature fusion and relevance feedback that we have adopted enhanced the validity of the retrieval technology of the mammographic calcification lesion, and we adopted using multi-distance to measure the similarity which determined by the optimal query rule, the precision ratio is improved. Due to features are extracted are more, the computing speed is slower, we should consider the characteristics optimization problems, at the same time, in order to reduce the function of user demand and the gap between systems, we should combine the low-level features and semantic features, thereby, it provide better assisted diagnosis of mammographic calcification lesions to doctor.

Reference

[1] S. Wong. CBIR in medicine: still a long way to go. IEEE Workshop on content-Based Access of Image and Video Libraries, Santa Barbara, CA, June, 1998.

[2] Shao Hong，Cui Wen cheng，Zhao Hong. Research of content- based image retrieval technology [J]. Mini-Micro Systems, 2003, 24(10):1845-1848.

[3] El- Naqa I, Yang Y, Nishikawa RM et al. A similarity learning approach to content-based image retrieval: application to digital mammography [J]. IEEE Trans Med Imaging, 2004, 23: 1233–1244.

[4] Chia-Hung Wei, Chang-T, Sun Li. Calcification Descriptor and Relevance Feedback Learning Algorithms for Content-Based Mammogram Retrieval [C]. Springer-Verlag Berlin Heidelberg, Susan M. Astley et al. (Eds.): IWDM 2006, LNCS 4046, 2006: 307-308.

[5] Müller H, Michoux N, Bandon D. A review of content-based image retrieval systems in medical applications clinical benefits and future directions [J]. Int J Med Inform 2004,73:1–23.

[6] Baeza-Yates RA, BA. Modern information retrieval. Addison- Wesley, Wokingham.1999.

[7] Ortega- Binderberger M, Mehrotra S Relevance feedback in multimedia databases[C]. In: Furht B, Marques O(eds) Handbook of video databases—design and applications. CRCPress, Boca Raton, 2003: 511–536.

[8] Ren-chao Jin, Bo Meng, En-ming Song. Content-based Image Retrieval for Computer-aided Detection of Mammographic Masses [J]. Mini- Micro Systems,2008, 29(2), 326- 327.

[9] Quan-ming Zhou，Guo-hua Geng，Na Wei. Content-Based Image Retrieval Technology [M]. 2007, 5: 100- 101.

[10] Zhi-jie Qiao, Jia-fu. Jiang Texture Image Retrieval Based on Gabor Wavelet Transform [D]. 2006: 22- 23.

[11] Fan Ya-chun, Geng Guo-hua, Zhou Quan-ming. Shape-Based Retrieval Using Invariant Moments and Edge Direction. Mini- Micro Systems, 2004, 25(4):659-660.