Study on Microcalcification Detection Using Fisher Discriminant and SVM

Guo Jinghuan[1], Chen Shenglai[2], Ge Ku[3] and Sun Zhaoqian[4]

[1, 3 ,4]*Institute of Information Sciences and Technology, Dalian Maritime University*
[2]*The 28th Research Institute of China Electronics Technology Group Corporation*
[1]*guojh1974@126.com*, [2]*chensl_01@163.com*, [3]*gek88@163.com*,
[4]*sunzhaoqian123@126.com*

Abstract

A hybrid microcalcification detection method based on Fisher discriminant and SVM was presented because signal-to-noise ratio of mammogram image was very low, and microcalcifications were very small and their shape was irregular. Firstly, low frequency information of tissue was removed by wavelet transform in order to reduce the tissue effect to microcalcification segment. Secondly, Fisher discriminant was adopted to find optimum threshold, meanwhile microcalcification was segmented. Lastly, SVM classifier was adopted to recognize true microcalcifications. Experiment results showed that Fisher discriminant could validly segment microcalcifications and the number of false positive targets was less than OSTU's. Detection ratio of our algorithm was about 97%.

Keywords: *Fisher Discriminant, SVM, Microcalcification Detection, Wavelet Transforms*

1. Introduction

The microcalcification was one of the important characterizations in the image of early breast cancer. The microcalcification detection of X-Ray image was very important for breast early prevention, and was one of hotspots in the fields of Image analysis and recognition at present. The microcalcification was very small, irregular, the shape and distribution was different, and the gray of microcalcification image was close to the one of breast tissue image, so it was very hard to detect the microcalcification.

At present, the general method of microcalcification image detection included the statistical methods [1], multi-scale detection method based on wavelet transformation [2], detection method based on morphology [3], and so on. The key technology of the methods was to extract the interesting regions by threshold segmentation method. Because of the lower signal to noise ratio and the very small microcalcification, the classical threshold segmentation method, such as the between class variance method [4], the maximum entropy method and the minimum error method, was not able to effectively segment the microcalcification.

In pattern recognition, the Fisher was an effective criterion to judge the class separation degree [5. 6]. The method could effectively divide the image into the background and the object, and was less affected by the background and the object. A hybrid microcalcification detection method based on Fisher discriminant and SVM was presented. Firstly, the background information was filtered from the X-Ray image by wavelet transform, then to segment the microcalcification object by Fisher discriminant, Lastly, the segmented microcalcification objects were distinguished using SVM classifier [7, 8], and got rid of some false positive target.

2. Microcalcification Segment Based on Fisher Discriminant

2.1. Fisher Discriminant

The Fisher was an effective discriminant to judge the class separation degree. Its aim was to seek for a best discriminating direction. High-dimensional sample was projected to the direction and formed one-dimensional discriminating space. The projection was the best [5] [6]. The discriminated condition ω_1 & ω_2 were represented in Figure 1 in which A belonged to the class of ω_1, B belonged to the class of ω_2, and C and D were the projections of A and B in the direction of Y_1 respectively, E and F respectively were the projections of A and B in the direction of Y_2. From the figure, the projections of ω_1 and ω_2 in the direction of Y_2 was the best. The algorithm for the best discriminating direction was showed as below.

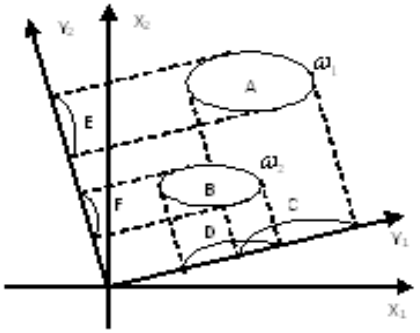


Figure 1. Conditions of Fisher Discriminating

Then, the linear combination of X was expressed as:

$$y = Y^T X \tag{1}$$

When $\|Y\| = 1$, y was the projection of X in the direction of Y.

In addition, the relationship between the sample belonging to Class ω_i and the mean value of $X_i = \{x_1, x_2, \cdots\cdots, x_{d_i}\}$ was defined as below

$$m_i = \frac{1}{d} \sum_{X=\omega_i} X \quad , i = 1,2 \tag{2}$$

Similarly, the mean value and the within class variance of y were calculated by the transform (1):

$$u_i = \frac{1}{d} \sum Y^T X = Y^T m_i \quad , i = 1,2 \tag{3}$$

$$\sigma_i^2 = \frac{1}{d} \sum (y - u_i)^2 \, , i = 1,2 \tag{4}$$

After projected, each sample in the one-dimensional space was expected to get spread out more, namely, the more sample spreaded, the larger the mean value of $u_1 - u_2$. And then, it is expected that the samples was concentrated as possible, the less the within class variance, the more the effect. The Fisher discriminant was used to measure the separation, which was expressed as below:

$$J(Y) = \frac{|u_1 - u_2|^2}{\sigma_1^2 + \sigma_2^2} \tag{5}$$

From the expression above, when the variance was the minimum, and the mean value was the maximum, the value of $J(Y)$ reached the maximum. Then, the discriminant direction was the best.

2.2. Microcalcification Segmentation

2.2.1. Low Frequency Information Filtering of Image based on Wavelet Transform: The gray scale of microcalcification was similar to the one of background, and it was hard to segment out the microcalcification by threshold method, but the lower frequency information was filtered from the image by multi-resolution characteristics of wavelet transform to reduce the background influence on the microcalcification.

The wavelet transform was to make clusters wavelet function $\{\psi(x)\}$ approach the signal function $f(x)$, and wavelet function cluster $\{\psi(x)\}$ consists of some basic wavelet functions, which had been operated by image stretching and translation operation in the various wavelet levels. The wavelet transform was defined as below:

$$\begin{cases} \psi_{a,b} = |a|^{\frac{1}{2}} \psi(\dfrac{x-b}{a}) \\ WTf(x) = |a|^{\frac{1}{2}} \displaystyle\int_{-\infty}^{+\infty} f(x)\psi(\frac{x-b}{a})dx \end{cases} \tag{6}$$

In formula 6, a and b were called dilation and translation factors respectively. The resolution characteristics of function $f(x)$ was realized by a and b, According to the different valves of a and b, the time-frequency information were gained to achieve the extraction and to analysis for local information of $f(x)$ in the various wavelet levels.

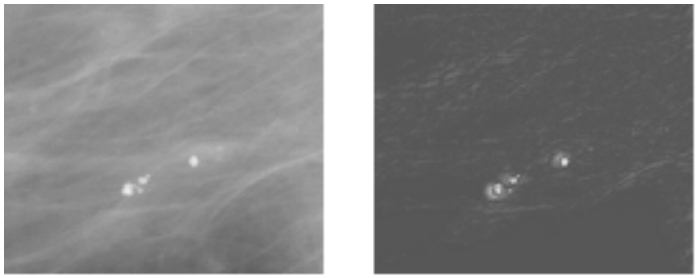
The wavelet decomposition of two-dimensional image function $f(x, y)$ could be expressed as below:

$$\begin{cases} W_\phi(j,m,n) = \dfrac{1}{\sqrt{MN}} \displaystyle\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\phi_{j,m,n}(x,y) \\ W_\varphi^{(i)}(j,m,n) = \dfrac{1}{\sqrt{MN}} \displaystyle\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\varphi_{j,m,n}(x,y); i = \{H,V,D\} \end{cases} \tag{7}$$

Where W_φ represented the approximate coefficients in low pass of image, W^H, W^V and W^D respectively represented high frequency detail coefficients, j represented

resolvable scale, two-dimensional discrete wavelet decomposition was gained with digital filters and down sampling by 2 factor.

The wavelet transform coefficient two-dimensional image, low frequency coefficient, W_ϕ represented the approximate information of the breast tissue image. So, the low-frequency sub-band was set zero to filter the background. The image reconstruction of mammography processed by wavelet low frequency filter was showed in Figure 2. From the Figure 2, most information of breast tissue and blood vessel was filtered. By statistical analysis on a lot of high-frequency image, the high-frequency of microcalcification possessed more gray value compared with the one of the background, which was showed as the lighter point in the figure.



a. Original Image b. Wavelet Reconstruction Image

Figure 2. Image of filtered by wavelet low-frequency

2.2.2. Threshold of Fisher: Because the microcalcification was very small, it was very hard to segment form high frequency image with the general threshold segment. Now Fisher discriminant was introduced from pattern recognition to seek a threshold and segment the microcalcification. The mean value of high-frequency information and the background high-frequency information were supposed to be u_1 and u_2, the variance respectively was σ_1 and σ_2. And the high-frequency was classified to two classes with the threshold value T. The value less than T was the background, and the value greater than or equal to T was the microcalcification. The normalized histograms of high-frequency image was supposed as $h(i)$, where the range of the gradient value $f(x)$ was $0 \sim 255$. The proportion that microcalcification gradient pixel numbers account for the whole image pixels numbers was supposed θ, that is to say, the prior probability for microcalcification and background respectively was θ and $1-\theta$. So the one-dimensional Fisher discriminant function was derived based on the formula 5:

$$J(T)=\frac{|\theta(T)u_1(T)-[1-\theta(T)]u_2(T)|^2}{\theta(T)\sigma_1^2(T)+[1-\theta(T)]\sigma_2^2(T)} \tag{8}$$

Where,

$$\theta(T)=\sum_0^T h(i)\ ;$$

$$u_1(T)=\frac{\sum_{i=0}^{T} i\cdot h(i)}{\theta(T)}\ ;$$

$$u_2(T) = \frac{\sum_{i=T+1}^{255} i \cdot h(i)}{1-\theta(T)} \ ;$$

$$\sigma_1^2(T) = \sum_{i=0}^{T} [i - u_1(T)]^2 \cdot h(i) \ ;$$

$$\sigma_2^2(T) = \sum_{i=T+1}^{255} [i - u_2(T)]^2 \cdot h(i) \text{。}$$

When T was the best threshold, $J(T)$ reached the maximum value, and the threshold selection criteria was:

$$T^* = Arg \max_{0\langle T \le 255} [J(T)] \qquad (9)$$

3. Microcalcification Detection Based on SVM

The microcalcification of the mammary gland X-Ray image was well positioned by the microcalcification segment based on Fisher discriminant. The detect result showed that there were some false positive object, especially to the image of compact background. So in the paper, SVM was used to reduce the possibility of false positive object.

3.1. Theory of SVM

The basic thought of SVM (Support Vector Machine) was that with nonlinear mapping Φ, the data X of input space X is projected to some a high-dimensional linear space H, and then the optimal linear classification was derived in the mapping characteristic space. The nonlinear transform was achieved by some proper kernel function [9] [10].

Suppose linear separate sample set was (x_i, y_i), where $i = 1, \cdots, n$, ($x \in R^d$), $y \in \{+1, -1\}$, linear discriminant function is $f_{SVM}(\mathbf{x}) = \omega^T \mathbf{\Phi}(\mathbf{x}) + b$, the classifying linear equation is $y(x) = \omega \cdot x + b$, and then the linear function is normalized to make two classes of all samples satisfy the following:

$$yi[(\omega \cdot x_i) + b] - 1 \ge 0; i = 1, \cdots, n \qquad (10)$$

When the classifying interval was equal to $2/\|\omega\|$, and the maximum interval was equivalent to $\|\omega\|$ or the minimum of $\|\omega\|^2$, so the optimal classification was gained when $\|\omega\|$ reached the minimum. The training point at the point H_1 and H_2 was called support vector.

The optimal classification for the formula (10) was changed to dual problem with the optimization method of Lagrange. Namely,

$$\begin{cases} Q(a) = \sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i,j=1}^{n} a_i a_j y_i y_j (x_i \cdot x_j) \\ s.t. \sum_{i=0}^{n} y_i a_i = 0 \quad ai \ge 0, i = 1, \cdots, n \end{cases} \qquad (11)$$

If a^* was the optimum solution,

$$\omega^* = \sum_{i=1}^{n} a_i^* y_i x_i \quad (12)$$

The corresponding optimal classification function was expressed as:

$$f(x) = \text{sgn}\{(\omega^* \cdot x) + b^*\} = \text{sgn}\{\sum_{i=1}^{n} a_i^* y_i (x_i \cdot x) + b^*\} \quad (13)$$

To nonlinear problem, nonlinear transform ϕ could be changed to a linear one in the some high-dimensional space to gain the optimal classification in the transforming space. Namely, an optimal plane was constructed in the linear space by kernel function to complete the linear transform of nonlinear transform. Now then, the object function was changed to:

$$\begin{cases} Q(a) = \sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i,j=1}^{n} a_i a_j y_i y_j K(x_i, x_j) \\ s.t. \sum_{i=0}^{n} y_i a_i = 0 \quad ai \geq 0, i = 1, \cdots, n \\ K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j) \end{cases} \quad (14)$$

The corresponding classifying function was changed to:

$$f(x) = \text{sgn}\{(\omega^* \cdot x) + b^*\} = \text{sgn}\{\sum_{i=1}^{n} a_i^* y_i K(x_i, x) + b^*\} \quad (15)$$

The most commonly kernel functions includes polynomial kernel, Gaussian kernel, Sigmoid kernel. In the practice, the concrete question concrete studied, selected the most proper function and corresponding parameters.

3.2. Feature Extraction

The microcalcification in the mammary gland image was usually manifested as some small outliers, grey level of which was slightly larger than surrounding tissues, and had larger high-frequency information. The mircocacification grey value was different from the background. So the microcalcification characteristic was selected from the feature such as the grey, the geometry, the texture and high-frequency information. It was showed as below:

Table 1. Feature of Microcalcification

Classification	Feature
Grey	mean, variance, skewness and kurtosis
Geometry	area, compactness
Texture	Entropy, energy
High-frequency	mean and variance of wavelet high-frequency coefficient

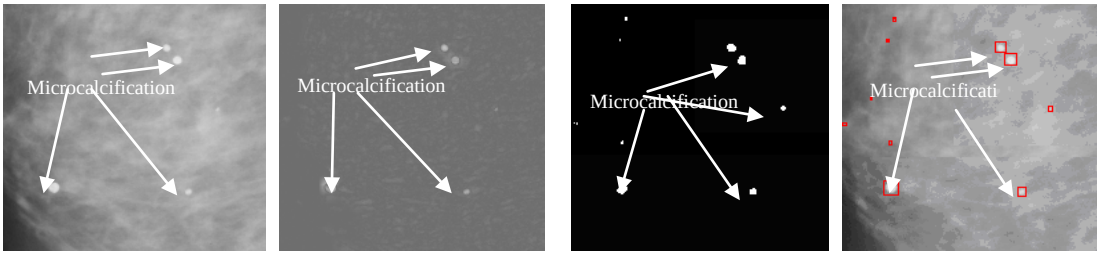
4. Experimental Result

The method presented in the paper was verified experimentally using MIAS image database. The database included 322 pieces of clinical trials X-Ray image, where 207 pieces were normal and 115 pieces were abnormal. The normal image did not contain diseases images, and abnormal images contained one or more diseases images. All images sizes of image database were 1024×1024. 116 microcalcification points in 15 pieces of breast images from MIAS were selected as the subjects in the experiments.

Lab 1 was to verify the ability of Fisher discriminant to segment microcalcification. During the experiment, db4 was selected as bi-orthogonal compact support wavelets, and wavelet transform series was 3 because db4 wavelet was high correlation to the tiny structures microcalcification, and was easier to detect microcalcification. The result from microcalcification segmentation based on Fisher discriminant was showed in the Figure 3. From Figure 3, it was showed that the microcalcification was characterized by the brighter point in the high-frequency image, and then microcalcification was segmented from the binary image by Fisher discriminant threshold segment, but there were a few false positive objects existing. In the experiments, 347 objects were segmented by Fisher discriminant, where there were 231 false positive objects.

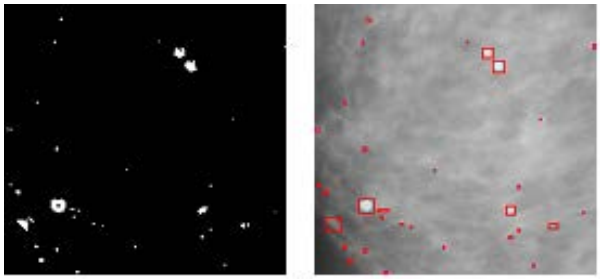
In the Lab 1, Fisher discriminant was compared with the OSTU, and the experiment result of the OSTU was showed in Figure 4. The OSTU also could segment microcalcificaiton from the binary image segmented by the OSTU segment, but the number of the binary image by the OSTU was a lot more than the one by Fisher discriminant. That could bring great difficulties to further discrimination microcalcification. So the Fisher discriminant was a better method to segment microcalcification. The gained threshold was more precise, and the Fisher discriminant had the less false positive objects.

Lab 2 was to verify the ability of SVM classifier to segment microcalcification. During the experiment, the features of 347 objects were extracted to form 347 x 10 dimension characteristic vectors, and selected 247 objects as the training samples and 100 objects as the testing samples. Gauss radial primary kernel function $K(x, x_i) = \exp\{-g \mid x - x_i \mid^2\}$ was selected as the SVM kernel function and cross-validation method was used to train the samples. In final, the optimum parameters of SVM classifier were gained, $c = 25$, $g = 0.5$. To verify the advantage of the SVM classifier, the classifier was compared with the ANN in the microcalcification recognition method. The results of the two methods were showed in the Table 2 as below.



a. Original mage b. Wavelet high-Frequency Image c. Binary image segment by Fisher d. Segment result

Figure 3. Segment Result of Fisher



a. Binary image by OSTU b. Segment result

Figure 4. Segment Result of OSTU

Table 2. Compared Result between SVM and ANN

	SVM	ANN
Detection Rate	96.97%	93.94%
False Positive Rate	1.49%	2.99%

5. Concluding Remarks

Because signal-to-noise ratio of mammogram image was very low, microcalcifications were very small, their shape was irregular and the image gray of microcalcification was similar to the gray of breast tissue image, a hybrid microcalcification detection method based on Fisher discriminant and SVM was presented. Firstly, high frequency information of tissue was gained by wavelet transform, meanwhile Fisher discriminant was adopted to find optimum threshold and microcalcification was segmented to extract some characteristics in the gray, geometry and texture and so on. Then SVM classifier was adopted to recognize true microcalcifications and further remove false positive objects. The experiment results showed that Detection ratio of our algorithm by the method was about 97%, and the method could be applied to clinical diagnosis.

Acknowledgements

Study on CT Image Automatic Segment based on Machine Learning, Young Key Teachers Foundation Projects of Dalian Maritime University (2012QN031)

Study on network technology of intelligent navigation, Fundamental Research Funds for the Central University (3132013308)

References

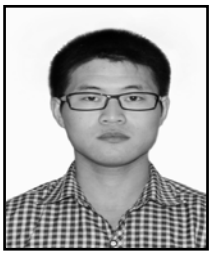
[1] D. David, “Automation computer detection of clustered microcalcifications in digital mammograms”, Physics in medicine and Biology, vol. 35, no. 8, **(1990)**, pp. 1111-1118.
[2] H. Moradmand, S. Setayeshi and H. Khazaei Targhi, “Comparing Methods for segmentation of Microcalcification Clusters in Digitized Mammograms”, International Journal of Computer Science Issues, vol. 8, **(2011)**, pp. 1694-0814.
[3] J. Yang and D. S. Park, “Detecting Region-of-Interest(ROI) in Digital mammogram by using Morphological Band pass Filter”, IEEE international conference on multimedia and expo, Taipei, Taiwan, vol. 6, **(2004)** June.
[4] N. Otsu, “A Threshold Selection Method from gray level histogram”, IEEE Trans. System, Cybern, SMC, vol. 8, **(1979)**, pp. 62-66.
[5] Z. Xuegong, “Pattern recognition (Third Edition)”, Beijing, Tsinghua University Press, **(2010)**, pp. 88-93.
[6] W. Shuhuan and T. Yinggan, “Multi-Threshold Image Segment Based on Fisher”, Laser & Infrared, vol. 38, no. 7, **(2008)**, pp. 742-744.

[7] L. Ting, W. Xianbin and Q. Jinjuan, "Multiscale SAR image segmentation using support vector machines", Proceedings of the 2008 Congress on Image and Signal Processing, USA:IEEE, **(2008)**, pp. 706-709.
[8] W. Honglei, O. Zongying and Z. Jianxin, "Fingerprint image segmentation based on Support vector machines", Journal of system simulation, vol. 19, no. 5, **(2007)**, pp. 2362-2365.
[9] V. Vapink, "The Nature of Statistical Llearning Theory", New York, Springer Verlag, **(2000)**.
[10] O. Chapelle, P. Haffner and V. N. Vapnik, "Support vector machines for histogram-based image classification", IEEE Transactions on Neural Networks, vol. 10, no. 5, **(1999)**, pp. 1055-1064.

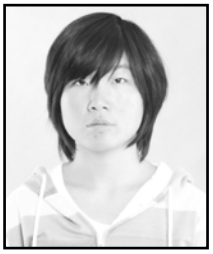
Authors

Guo Jinghuan, (1974), female, Liaoning Changtu Country, Doctor, Associate Professor. Research on Image Processing and Pattern Recognition

Chen Shenglai, (1978), male, Zhejiang Haining Country, Doctor, Senior Engineer. Research on Image Processing and Intelligence Fusion



Ge Kun, (1989), male, Shandong Heze, Master in reading, Research on Web Service Somposition and Development of Agent Aystem based on JADE



Sun Zhaoqian, (1990-), female, Shandong Jining, Master in reading, Research on Software Test and Software Fault location