PAPER Special Section on Acoustic Diagnosis

A Computer-Aided System for Discrimination of Dilated Cardiomyopathy Using Echocardiographic Images

Du-Yih TSAI†, *Member and* Masaaki TOMITA††, *Nonmember*

SUMMARY In this paper, the discrimination of ultrasonic heart (echocardiographic) images is studied by making use of some texture features, including the angular second moment, contrast, correlation and entropy which are obtained from a gray-level cooccurrence matrix. Features of these types are used as inputs to the input layer of a neural network (NN) to classify two sets of echocardiographic images—normal heart and dilated cardiomyopathy (DCM) (18 and 13 samples, respectively). The performance of the NN classifier is also compared to that of a minimum distance (MD) classifier. Implementation of our algorithm is performed on a PC-486 personal computer. Our results show that the NN produces about 94% (the confidence level setting is 0.9) and the MD produces about 84% correct classification. We notice that the NN correctly classifies all the DCM cases, namely, all the misclassified cases are of false positive. These results indicate that the method of feature-based image analysis using the NN has potential utility for computer-aided diagnosis of the DCM and other heart diseases.

key words: *medical image processing, pattern matching, computer-aided diagnosis, texture features, echocardiographic images*

1. Introduction

The use of 2-dimensional (ultrasonic heart) echocardiographic images has been an important non-invasive means in clinical cardiology [1]. The diagnosis of heart functions using echocardiography [2], [3] is comparably common among various diagnostic methods. However, since the discrimination of normal and abnormal cases largely depends on diagnostician's subjective point of view and his/her experience, the criteria of diagnosis are indeterminate. If quantitative computerized methods which provide a "*second opinion*" for the diagnostician can be developed, then this subjectivity can be reduced and in turn the accuracy in diagnosis is expected to increase.

Investigators have attempted to develop automated means for computer-aided diagnosis in a variety of diseases [4]–[8]. However, little work has dealed with heart diseases in echocardiography. In our previous report [9] we have attempted to develop a computer-aided diagnosis in echocardiography. This method was to discriminate echocardiographic images into normal or abnormal case, namely, dilated cardiomyopathy (DCM) using a feature-based classification approach. The features used in this work included the angular second moment and contrast. We also used a minimum distance (MD) classifier [9], [10] to evaluate the performance of the features. Our preliminary result showed that texture analysis could be a useful tool in diagnosis of heart diseases.

Based on the promising conclusions of our previous investigation, in this study we aim at developing a clinically practical computer-aided system by employing an artificial neural network (NN) for discriminating the DCM. Texture features obtained from the gray-level cooccurrence matrix of an image are used as inputs to the input layer of the NN. The performance of the NN classifier is compared to that of the MD classifier used in our previous study. Furthermore, in order to demonstrate the effectiveness of the use of texture features, the performance of the NN fed with image itself directly to the input layer is also investigated and discussed.

2. Materials and Methods

Figure 1 shows the block diagram illustrating the overall procedure of our approach. Each step is described below.

2.1 Image Data

In this paper, all ultrasonic images were captured from

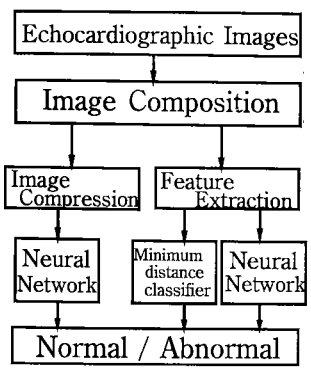


Fig. 1 The block diagram illustrating the overall procedure of our analysis system.

Manuscript received April 28, 1995.
Manuscript revised July 24, 1995.

†The author is with Gifu National College of Technology, Gifu-ken, 501-04 Japan.
†† The author is with the Second Department of Internal Medicine, Gifu University School of Medicine, Gifu-shi, 500 Japan.

a Toshiba SSH-160A device with a 2.5 MHz (central frequency) transducer (Toshiba Medical Co., Ltd., Tokyo, Japan). The frame rate and scanned mode used were 30 frames/sec and sector phased array, respectively. A logarithmic amplifier for video circuit was also used. A total of 62 samples of echocardiographic images from 31 subjects (2 sample images per subject: an end-systole image and an end-diastole image) was collected at the Hospital of Gifu University School of Medicine. Of the 31 subjects 18 were diagnosed by a highly trained physician as normal hearts and the remaining 13 as abnormal hearts (typical case of DCM). Each image was digitized by an image scanner (GT-600, Epson) at the resolution of 256×256 pixels. Since the original echocardiographic images have 64 gray levels, the scanned images were quantized to the same gray levels.

2.2 Composite Images

In our previous study [9] we noted that the use of composite images could provide higher recognition rate as compared to that of individual images at end diastole and end systole. Therefore, in this study we composed a summation image and a subtraction image for each of the subjects. The summation image $sum(x, y)$ is obtained by adding the gray level of a given pixel of the end-systole image $f(x, y)$ to the gray level of the corresponding pixel of the end-diastole $g(x, y)$, where $0 \leqq x \leqq 255$ and $0 \leqq y \leqq 255$ for a 256×256 image, i.e., $sum(x, y) = f(x, y) + g(x, y)$. Those gray levels greater than 255 are treated as 255. The subtraction image $sub(x, y)$ is obtained by taking the absolute value of the gray-level variation of the end-systole and end-diastole images, i.e., $sub(x, y) = |f(x, y) - g(x, y)|$. By doing so, it is expected to extract more powerful features from the composite images as compared to that from an individual image. Figure 2 shows a typical case of DCM. The images at end-systole and end-diastole states are shown in Figs. 2(a) and (b), and the summation and subtraction images are shown in Figs. 2(c) and (d), respectively.

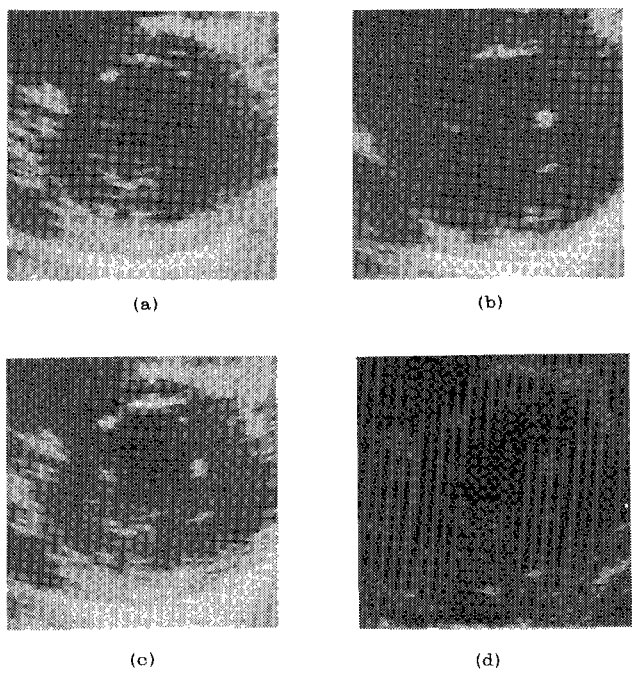


Fig. 2 Images of a typical case of dilated cardiomyopathy. (a) and (b) are images at end-systole and end-diastole, respectively. (c) and (d) are summation image and subtraction image, respectively.

2.3 Cooccurrence Matrix

We generated a gray-level cooccurrence matrix from every summation image and subtraction image. The gray-level cooccurrence matrix is a matrix used to express the correlation of spatial location and gray-level distribution of an image. From it, the local variation of gray levels on an image can be statistically investigated and in turn, enable us to know the manner of change in gray level as a whole. The concept of the cooccurrence matrix is briefly described as follows [9], [11].

Assume that the gray-level values of the two pixels (pair of pixels) at (x, y) and $(x+m, y+n)$ are f_1 and f_2 (see Fig. 3(a)). If the image is quantized by L gray levels, the possible range of the gray levels for each pixels is from 0 to $L-1$, i.e., $0 \leqq f_1 \leqq L-1$; $0 \leqq f_2 \leqq L-1$. In this case, for example, when the values of m and n are constant and when the occurrence frequency of $f_1=1$ and $f_2=2$ is a, then the number a is filled into the corresponding element of the matrix as shown in Fig. 3(b). In this way, a $L \times L$ cooccurrence matrix can be generated.

In the present study, we used the following conditions to generate the gray-level cooccurrence matrix.

(a) Number of gray levels: A cooccurrence matrix of 64×64 size can be obtained from a 64 gray-level image. In order to save computation time, we decompress the gray levels to 16 in our study. Experimental results showed that the matrix size of 16×16 was adequate.

(b) Direction (θ): Generally, it is better to generate

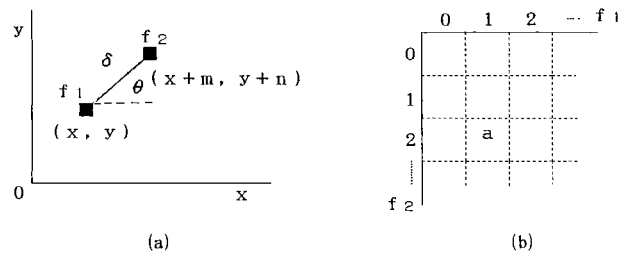


Fig. 3 (a) Explanation of how to construct a gray-level co-occurrence matrix; (b) Layout of a gray-level cooccurrence matrix.

gray-level cooccurrence matrices from all directions. However, it is considerably time consuming. Therefore the directions need to be confined. In our previous work [9] we chose the direction of 90° with respect to the horizontal direction for computation. In the present study we expand the directions to 0°, 45°, 90° and 135° to obtain more features from the cooccurrence matrix.

(c) Distance (δ): In this study $\delta=5$ pixels was used, because we experimentally found that this value was optimal.

2. 4 Texture Features Used

From the generated gray-level cooccurrence matrix as described in Sect. 2.3, 14 features can be calculated for each image, namely, angular second moment, contrast, correlation, variance, inverse different moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy, information measure of correlation A, information measure of correlation B, and maximal correlation coefficient. All of these features were experimentally evaluated for their ability to discriminate between normal and abnormal cases, and the four ($Q1$–$Q4$) with the greatest discriminatory power were selected. They are as follows.

(1) Angular Second Moment (ASM):

$$Q_1=ASM=\sum_{f_1}\sum_{f_2}P(f_1, f_2)^2, \tag{1}$$

where, $P(f_1, f_2)$ is the probability obtained by dividing the number of (f_1, f_2) element in the matrix by the total number of occurrence in the matrix.

(2) Contrast (CON):

$$Q_2=CON=\sum_{f_1}\sum_{f_2}(f_1-f_2)^2P(f_1, f_2). \tag{2}$$

(3) Correlation (COR):

$$Q_3=COR=\frac{\sum_{f_1}\sum_{f_2}f_1f_2P(f_1, f_2)-\mu_1\mu_2}{\sigma_1\sigma_2}, \tag{3}$$

where,

$$\mu_1=\sum_{f_1}f_1\sum_{f_2}P(f_1, f_2)$$

$$\mu_2=\sum_{f_2}f_2\sum_{f_1}P(f_1, f_2)$$

$$\sigma_1=\sum_{f_1}(f_1-\mu_2)^2\sum_{f_1}P(f_1, f_2)$$

$$\sigma_2=\sum_{f_2}(f_2-\mu_2)^2\sum_{f_1}\sum P(f_1, f_2)$$

(4) Entropy (ENT):

$$Q_4=ENT=-\sum_{f_1}\sum_{f_2}P(f_1, f_2)\ln\{P(f_1, f_2)\}. \tag{4}$$

It is still not very clear that these features concretely describe what kinds of physical properties of the tissue, respectively. However, in the sense that different tissue or different quality of image provides different feature values, these statistic values can be used to represent texture features of echocardiographic images.

2. 5 Artificial Neural Networks

In this paper we employed a 3-layer (one hidden layer) NN [12], [13] for classification of normal and abnormal hearts. A back-propagation algorithm was used for learning (or training). Learning occurs by adjusting the values of weighting coefficients to minimize the mean squared error between the desired and actual values of the output units. The desired value is the output component of the input/output pairs in the training sets.

Figure 4 illustrates the schematic diagram of the classification scheme. The NN consisted of 16 input units. Input data, which were represented by 16 extracted features, namely, 4 statistical measures (Q_1–Q_4) $\times$4 directions ($\theta=0°$, 45°, 90° and 135°) =16 features, were supplied to the input units. The NN was trained to output 1 for a normal case and 0 for an abnormal case. After the NN was trained, its classification of a given input was deemed to be "normal" if the NN's output was >0.9 (sometimes we called it confidence level=0.9) and "abnormal" if the output was <0.9.

In this study we randomly selected 9 normal and 6 abnormal echocardiograms from the 31 samples and

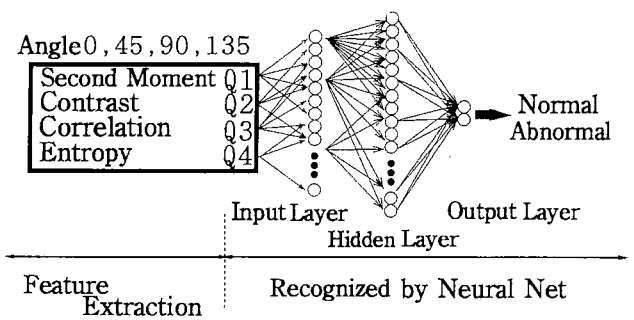


Fig. 4 Schematic diagram of the feature-based classification method using a neural network.

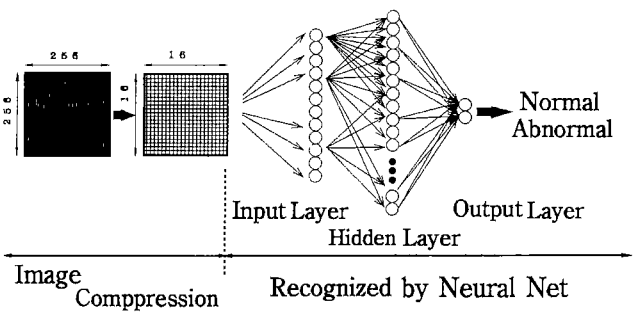


Fig. 5 Schematic diagram for image-based classification method using a neural network. The gray level of each pixel on the compressed image corresponds to an input unit.

called them "*set A*," and called the remaining (9 normal and 7 abnormal) hearts "*set B*." We used set A as learning data for training and used set B as test data for classification, and vice versa. The NN was trained with 10000 learning iterations. Each learning iteration corresponded to the entry of a complete set of training data.

In order to verify the superiority of the use of texture features, we also investigate the performance of the NN fed with image's gray values directly to the input layer. Figure 5 illustrates the procedure of this approach. Due to the memory capacity limitation of our NN, we compressed the image size to 16×16 from the original image of 256×256. In this case, the gray level of each pixel on the compressed image corresponds to an input unit. Therefore the number of input units is $16\times16=256$. The training process and the parameters used in the NN were the same as that employed in the above described feature-used method.

2.6 Pattern Classification Technique

To compare the classification performance of the NN and pattern-matching technique for discriminating the DCM, the Euclidean distance which is basically used to find the decision boundary between two pattern classes, was employed in the MD classifier. We expanded this 2-dimensional pattern vectors to 4-dimensional pattern vectors corresponding to the above-mentioned four texture features.

3. Results and Discussion

We first examined the effect of the number of hidden units on NN performance. Tables 1 and 2 summarize the results for the summation images and the subtraction images, respectively. In the case of summation images, a 80.3% of average recognition rate can be obtained when the hidden layer having 50 units were used, while 77.1% of recognition rate can be obtained for the remaining cases as shown in Table 1. As for the case of hidden units greater than 50, no increase in recognition rate was found. We noticed from the results that the recognition rates for set A and set B were significantly different and favorable to "set B." The reason is considered due to that set B consists of greater number of *difficult* cases as compared to set A, although the two sets were randomly grouped. In the case of subtraction images, a maximum of 93.6% average recognition rate can be obtained when the number of hidden units is greater than 40 as shown in Table 2. The table also indicates that the recognition rates for the two sets are the same. This implies that the texture features extracted from the subtraction images are powerful and effective for discriminating normal and abnormal cases even they are difficult ones. Thus higher recognition rate can be obtained when the subtraction images are used. If we ease the confidence level to 0.8 from 0.9, then the average recognition rates for the summation image and subtraction image are 93.4% and 96.7%, respectively. We noticed that the NN correctly classified all the DCM cases, namely, all the misclassified cases are of false positive (the normal case is classified as abnormal case).

Table 1 Recognition rates, when the summation images are used for feature extraction.

No. of hidden units	Recognition rate (%)		
	Set A	Set B	Average
5	66.7	87.5	77.1
10	66.7	87.5	77.1
20	66.7	87.5	77.1
30	66.7	87.5	77.1
40	66.7	87.5	77.1
50	66.7	93.8	80.3

Table 2 Recognition rates, when the subtraction images are used for feature extraction.

No. of hidden units	Recognition rate (%)		
	Set A	Set B	Average
5	86.7	68.8	77.8
10	86.7	81.3	84.0
20	86.7	93.8	90.3
30	86.7	87.5	87.1
40	93.3	93.8	93.6
50	93.3	93.8	93.6

Table 3 summarizes the performance of the NN fed with compressed image itself to the input layer. It is apparent from the table that the recognition rate of the direct input method is poorer than that of the features-used method. This demonstrates the superiority of the use of texture features.

The recognition rates for the MD classifier were 83.9% and 80.6% for the summation and subtraction images, respectively. As described earlier that we used 4-dimensional pattern vectors (Q_1-Q_4 at $\theta=90°$) for classification. It is interesting to note that the MD classifier gave better performance than the NN classifier fed with compressed image itself to the input

Table 3 Recognition rates, when the compressed, summation or subtraction image itself is used as input to the input layer of the NN.

No. of hidden units	Recognition rate (%): summation image		
	Set A	Set B	Average
40	60.0	68.8	64.4
50	60.0	68.8	64.4

No. of hidden units	Recognition rate (%): subtraction image		
	Set A	Set B	Average
40	66.7	75.0	70.9
50	66.7	75.0	70.9

layer, although the NN with texture features is superior to it. We also experimentally found that the recognition rate decreased, when 2-dimensional pattern vectors (2 features are selected from the 4 features) were used for classification.

The recognition rate of 93.6% determined by our purposed NN classifier using texture features is comparable to that by several-year experienced physicians. Therefore, we consider that the recognition rate obtained using our method is satisfactory, although further improvement is required. The computation time for generating the 16 features is approximately 5 minutes by our PC-486 personal computer. However, we anticipate that the computation time can be much reduced if a workstation is used. As for the NN-based classification, a near real-time implementation can be achieved after training process.

In general, good features normally lead to a good performance for the NN. The results of our previous study [9] support that the use of composite images could provide higher discrimination rate as compared to that of individual images. This implies that the four features obtained from the composite images carry more useful information than that obtained from the individual ones. However, this argument still remains room for further investigation.

It is recognized that the summation and subtraction images may provide not only the texture features indicating the nature of the tissue but also the information of shape variation of the heart. Therefore the effect of this shape variation on the texture can not be ignored, although we have not taken it into account. This point of issue will be addressed in future work.

As described earlier that the aim of this study is to develop a clinically practical computer-aided system. However, it is also recognized that the discrimination rate is very much influenced by the quality of images. It is expected that the discrimination rate may increase if image quality is enhanced using a further improved ultrasonic equipment.

4. Conclusions

In this paper we have proposed a NN method using texture features to classify echocardiographic images. This method enables the classification to achieve a 93.6% average recognition rate. This result suggests that the method of feature-based image analysis using the NN has the potential to become clinically useful for computer-aided diagnosis of the DCM and other heart diseases. Future work increasing sample sets for further feasibility test on the proposed method is needed.

Acknowledgement

The authors would like to thank Mr. D. Fukuoka, a student of Gifu National College of Technology (presently, Gifu University), for his help with technical assistance.

References

[1] C.Y. Han, K.N. Lin, W.G. Wee, R.M. Mintz, and D.T. Porembka, "Knowledge-based image analysis for automated boundary extraction of transesophageal echocardiographic left-ventricular images," IEEE Trans. Med. Imaging, vol. 10, pp. 602-610, 1991.

[2] J. Koyama, S.S. Beppu, N. Tanaka, S. Kawano, T. Shimizu, T. Shishido, H. Matsuda, and K.S. Miyatake, "Usefulness and limitation of digital echocardiography in detecting wall motion abnormality during stress echocardiography," Jpn. J. Med. Ultrasonics, vol. 21, pp. 695-701, 1994.

[3] H. Murayama, F. Kimura, Y. Miyake, S. Tsuruoka, M. Motoyasu, K. Sekioka, and T. Nakano, "Evaluation of phasic local wall motion by serial frame subtraction," Jpn. J. Med. Ultrasonics, vol. 22, pp. 395-398, 1995.

[4] S. Katsuragawa, K. Doi, and H. MacMahon, "Image feature analysis and computer-aided diagnosis in digital radiography: Detection and characterization of interstitial disease in digital chest radiographs," Med. Phys., vol. 15, pp. 311-319, 1988.

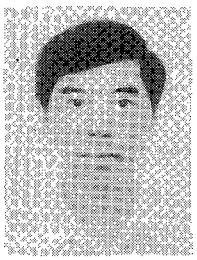
[5] F.-F. Yin, M. L. Giger, K. Doi, C.E. Metz, C. J. Vyborny, and R.A. Schmidt, "Computerized detection of masses in digital mammograms: Analysis of bilateral subtraction images," Med. Phys., vol. 18, pp. 955-963, 1991.

[6] S. Sanada, K. Doi, and H. MacMahon, "Image feature analysis and computer-aided diagnosis in digital radiography: Automated delineation of posterior ribs in chest images," Med. Phys., vol. 18, pp. 964-971, 1991.

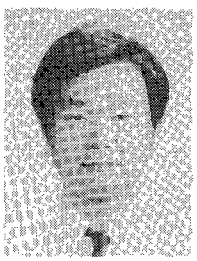
[7] P. Caligiuri, M.L. Giger, M.J. Favus, H. Jia, K. Doi, and L.B. Dixon, "Computerized radiographic analysis of osteoporosis: Preliminary evaluation," Radiology, vol. 186, pp. 471-474, 1993.

[8] C.-M. Wu, Y.-C. Chen, and K.-S. Hsieh, "Texture features for classification of ultrasonic liver images," IEEE Trans.

Med. Imaging, vol. 11, pp. 141–152, 1992.
[9] D.-Y. Tsai and M. Tomita, "Feature-based image analysis for classification of echocardiographic images," IEICE Trans. Fundamentals, vol.E78-A, pp. 589–593, 1995.
[10] R.C. Gonzalez and R.E. Woods, "Digital Image Processing," Addison-Wesley, New York, 1992.
[11] M. Takagi and H. Shimoda, "Handbook of Image Analysis," Tokyo University Publication, Tokyo, 1991.
[12] D.E. Rumelhart, G.E. Hinton, and R.S. Williams, "Learning representations by back-propagation errors," Nature, vol. 323, pp. 533–536, 1986
[13] R.P. Lippmann, "An introduction to computing with neural nets," IEEE ASSP Magazine, pp. 4–22, April, 1987.



Du-Yih Tsai was born in Taichung, Taiwan in 1948. He received the B.S. degree in Electrical Engineering from National Cheng-Kung University, Taiwan in 1970, M.E. degree in Electrical Engineering from Gifu University, Japan in 1979, and Ph.D. degree from Nagoya University, Japan in 1985. He is currently an Associate Professor in the Department of Electrical Engineering, Gifu National College of Technology, Japan. His research interests include imaging processing and pattern recognition.



Masaaki Tomita received the M.D. degree in Gifu University School of Medicine, Japan in 1980. He was a Research Associate at Medical University of South Carolina, USA from 1988–1990. His research interests include the assessment of cardiac function using echocardiography. He is a member of the Japanese Circulation Society.