World Engineering Congress on Biomedical Engineering and Medical Physics 2006 WC2006, Coex, Seoul, Korea. 27 August – 3 September 2006.

Heart Sound Analysis Using MFCC and Time Frequency Distribution

I. Kamarulafizam, Sh-Hussain. Salleh, J.M. Najeb, A.K. Arif and A. Chowdhury

Centre for Biomedical Engineering (CBE), Faculty of Electrical Engineering, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

Abstract— **This paper presents heart sound analysis method based on Time-Frequency Distribution (TFD) analysis and Mel Frequency Cepstrum Coefficient (MFCC). TFD represents the heart sound in term of time and frequency simultaneously which while the MFCC defines a signal in term of frequency coefficient corresponding to the Mel filter scale. There are 100 normal data and 100 data with disease obtained from the hospital which consists of various kinds of problems including mitral regurgitation and stenosis, tricuspid regurgitation and stenosis, ventricular septal defect and other structural related disease. B-Distribution is chosen from a number of time-frequency analysis methods due its capability to represent the signal in the most efficient way in term of noise and cross term reduction. The advantage of MFCC is that it is good in error reduction and able to produce a robust feature when the signal is affected by noise. SVD/PCA technique is used to extract the important features out of the B-Distribution representation. The coefficient obtained from SVD-PCA and MFCC is later used for classification Artificial Neural Network. The results show that the system is able to produce the accuracy up to 90.0% using the TFD and 80.0% using the MFCC.**

.

Keywords— **Heart Sound, Time Frequency Analysis, Mel-Frequency Cepstrum Coefficient, Singular Value Decomposition, Principle Component Analysis, and Artificial Neural Network.**

I. INTRODUCTION

Auscultation, the act of listening to the sounds of internal organs, is a valuable medical diagnostic tool. Auscultation methods provide the information about a vast variety of internal body sounds originated from the heart, lungs, bowel and vascular disorders. The information acquired by a traditional stethoscope is, however, subjective and qualitative in nature since the differentiation of signals picked up by the sensor during manual interpretation is limited by human perception abilities and varies with personal aptitude and training. This may result in inaccurate or insufficient information due to the inability of the user to discern certain complex, low-level, short duration or rarely encountered abnormal sounds. It is, thus, desirable to enhance the diagnostic ability by processing the auscultation signals electronically and providing a visual display and automatic analysis to the physician for a better comparative study. During the last decade, efforts have been made in this direction and electronic stethoscopes are now available commercially with phono-cardiograph display. The usage of such devices, however, is less common due to the involvement of cumbersome instrumentation and complex/additional skills. Our research at CBE, UTM encompasses a more user friendly and quantitative approach together with automatic signal analysis using intelligent algorithms.

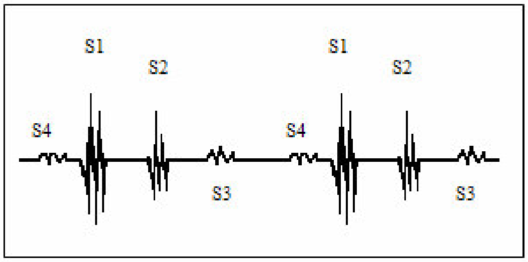

Figure 1: Heart sound.

II. BACKGROUND

A. Time Frequency Analysis

Time Frequency Signal Processing is a method and technique developed for representation, analysis and processing of non-stationary signals, which time and frequency are interrelated. Time representation, *s(t)*, and frequency representation, *S(f)*, are related via the Fourier Transform.

$$s(t)=\int_{-\infty}^{\infty} S(f)e^{j2\pi ft}df \;\; {}_{t}\overset{FT}{\leftrightarrow}{}_{f} \;\; S(f)=\int_{-\infty}^{\infty} S(t)e^{j2\pi ft}.dt \quad (1)$$

Time-frequency representation represent signals in time and frequency which all signals information is accessible and provides a distribution of signal energy versus *t* and *f* simultaneously which called time frequency distribution (TFD). TFD reveals the number of signal components, the time variation of frequency content of the components and order of appearance in time of the different frequencies present.

The B-Distribution is a Quadratic or bilinear Time frequency Distribution algorithm which capable to minimizing cross term [6] and it defined as:

$$B(t,f) = \iint (|\tau| / \cosh^2(u-\tau))^{\beta} z(u+\tau/2) z^*(u-\tau/2) e^{-j2\pi f\tau} du d\tau \quad (2)$$

B. Mel Frequency Cepstrum Coefficient-MFCC

A representation of heart sounds using Mel-Frequency Cepstrum Coefficient (MFCC) would be provided by a set of cepstrum coefficients. These coefficients are the results of a cosine transform of the real logarithm of the short-term energy spectrum expressed on a mel-frequency scale. The MFCC are also an efficient method to extract any kind of features [12]. The number of resulting mel-frequency cepstrum coefficients is practically chosen relatively low, in the order of 12 to 20 coefficients. However, in many cases of MFCC analysis, the 0th coefficient of the 7 MFCC cepstrum is ignored because of its unreliability [13]. In fact, the 0^{th} coefficient can be regarded as a collection of average energies of each frequency bands in the signal that is being analyzed. The energy of heart sound signal is also a very important feature for pattern recognition. Many experiments have shown that the performance can be improved when the energy information is added as another model feature in addition to cepstrums [10]. Mel Frequency Cepstrum Coefficients (MFCC) is also used as a method that analyzes how the Fourier transform extracts frequency components of a signal in the time-domain. In addition, it is a representation defined as the real cepstrum of a windowed short-time signal derived from the Discrete Fourier Transform (DFT) of that signal. The difference from the real cepstrum is that a non-linear frequency, a mel-scale is used. The mapping from linear frequency to Mel frequency is done using an equation as follows

$$\mathrm{Mel}(f) = 2595 \log10 (1 + f/700) \quad (3)$$

C. Singular Value Decomposition-SVD

The singular value decomposition (SVD) has become an important tool in statistical data analysis and signal processing. The existence of SVD was established by the Italian geometer Beltrami in 1873 which was only 20 years after the conception of a matrix as a multiple quantity by Cayley. SVD based technique is introduced to reduce the effect of noise from TFD which deal with singular matrices, or once which are very close to being singular. They are an extension of eigen decomposition to suit non-square matrices. Any matrix may be decomposed into a set of characteristic eigenvector pairs called the component factors, and their associated eigenvalues called the singular value.

In order to extract data dynamically which decompose the X, TF distribution matrix of the power disturbance signals, which $m \times n$ (*time x frequency*) into a set of characteristic. A singular value decomposition of an $m \times n$ matrix X is any factorization of the form:

$$X = U\Sigma V^T \quad (4)$$

where U is an $m \times m$ orthogonal matrix; i.e. has orthonormal columns, V is an $n \times n$ orthogonal matrix and Σ is $m \times n$ an diagonal matrix of singular values with components $\sigma_{ij}=0$ if and $i \neq j$; (for convenience we refer to the *i-th* singular value $\sigma_i=\sigma_{ii}$). Furthermore it can be shown that there exist non-unique matrices U and V such that singular values $\sigma_1 \geq \sigma_2 \geq \ldots\ldots \geq \sigma_n \geq 0$. The columns of the orthogonal matrices U and V are called the left and right singular vectors respectively; an important property of U and V is that they mutually orthogonal [4].

D. Principle Component Analysis-PCA

PCA is generally used when the research purpose is data reduction (to reduce the information in many measured variables into a smaller set of components) and it can minimize the reconstruction error in the sense of least square error then find out the most representative feature. PCA seeks a linear combination of variables such that maximum variance is extracted from the variables. It then removes this variance and seeks a second linear combination which explains the maximum proportion of the remaining variance, and so on. This is called the principle axis method and results in orthogonal (uncorrelated) factors. PCA analyzes total (common and unique) variance. We can see that the SVD is in fact closely related to the PCA. In fact the matrix product $U\Sigma$ is analogous to the matrix Y defined for PCA as:

$$Y = XV = U\Sigma \quad (5)$$

Because both the singular vectors defined for an SVD are square and have orthonormal columns their inverses are given by their transposes. Now the relation in Equation 4 can be expressed $X=U\Sigma V^T$ which is the definition of an SVD. The pairs of eigenvectors are the row in U and the column in V [14].

III. METHODOLOGY

The heart sound is segmented into a small duration of a complete one (1) cycle of heart beat. This is an important procedure which needs to be carefully handled because a poorly handled segmentation process may affect the performance of the feature extractor and the classifier.

In this experiment, the heart sound samples are processed using two (2) methods which are TFD-SVD-PCA and MFCC. In the first procedure, the heart sound sample is transformed using TFD utilizing B-Distribution as a transformation kernel. This will result a time and frequency representation with the high resolution plane (*data length x dft points*). At this stage the features can be observed clearly by looking at the energy at the most dominant frequency with the respect of the time scale. However, selecting the best features is not easy when dealing with a very high dimension of data.

In addition to an appropriate process, it is often essential that dimensionality reduction be performed on the TFD, so that the information is sufficiently compact for presentation to a classifier such as ANN. The main goal of SVD technique is to ensure that as much relevant information as possible is preserved in as few dimensions as possible.

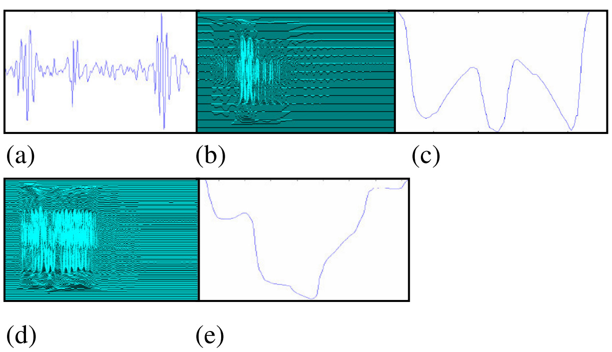

Figure 2 (a) Time domain heart sound signal, (b) Time-frequency representation of a normal heart sound, (c) SVD-PCA representation of a normal heart sound signal. (d) Time-frequency representation of an abnormal heart sound. (e) SVD-PCA representation of an abnormal heart sound signal

As for the second method, the segmented heart sound sample is transformed using MFCC. This experiment is carried out based on the initial parameter of 12 coefficients with 240 samples for each frame and with 80 samples overlap between each frame. MFCC has a property where the transformation scale is linear for the frequency below 1 kHz and expand logarithmically above 1 kHz. This property works well for speech processing as it suit the human auditory spectral band. However for the heart sound analysis, the higher filter bank scale will not affect much as most of the heart sound component exist below 1 kHz [15]. This experiment is carried out to verify the property of MFCC in terms of filter bank analysis and transformation,

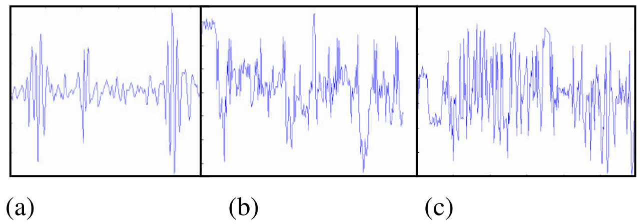

Figure 3 (a) Time domain heart sound signal, (b) MFCC coefficient of a normal heart sound. (c) MFCC coefficient of an abnormal heart sound

IV. RESULTS

The output features of both methods are fed to Multi Layer Back Propagation Artificial Neural Network for performance comparison.

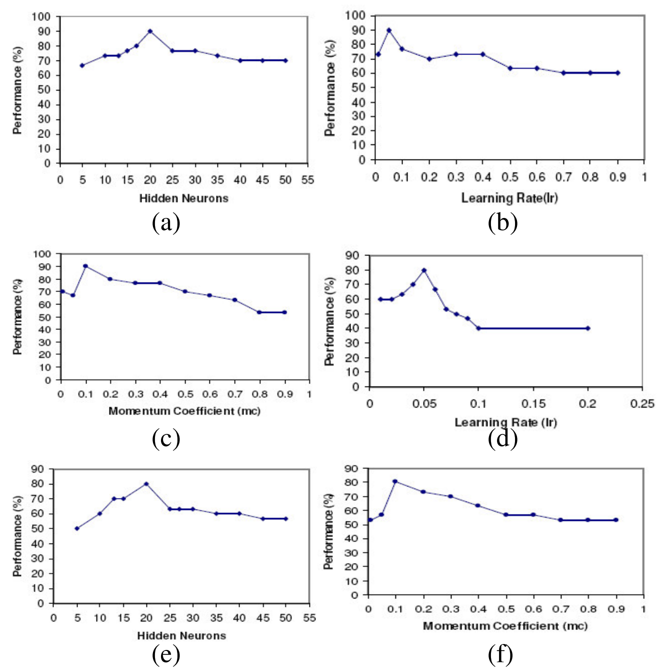

Figure 4 (a) Verification on hidden node for TFD, (b) Verification on learning rate for TFD. (c) Verification on momentum term for TFD. (d) Verification on hidden node for MFCC. (e) Verification on learning rate for MFCC. (f) Verification on momentum term for MFCC.

Figure 4(a) – 4(c) shows the results of verification technique using TFD. They showed that the highest performance they can achieved is up to 90% at hidden nodes = 20, learning rate = 0.05 and momentum coefficient = 0.1. Figure 4(d) – 4(f) shows the results of verification technique using MFCC. They showed that the highest performance they can achieved is only up to 80% at hidden nodes = 20, learning rate = 0.05 and momentum coefficient = 0.1.

Table 1: Result analysis using TFD and MFCC

Item	TFD,SVD/PCA	MFCC
Feature Dimension	256	170
Training Data	50	50
Testing Data	50	50
Hidden Layer	20	20
Learning Rate	0.05	0.05
Momentum Term	0.1	0.1
Accuracy	90%	80%

Table 1 shows the result obtained by the analysis using TFD and MFCC. It is obvious that analysis using TFD gives a better performance based on ANN. This is due to the ability of TFD to represent the heart sound in various frequencies as a non-stationary signal. Transformation using MFCC maps the time domain signal into frequency domain according to the Mel filter scale. Although MFCC offers a detail analysis of a short duration of the heart sound with the overlapping property, it is still unable to represent the signal as good as the time frequency.

V. Conclusions

This paper presents the most widely used signal analysis methods which are time frequency analysis and Mel-Frequency Cepstrum Coefficient with the application to heart sound analysis. The proposed method is able to discriminate the normal heart sound from the abnormal heart sound. The result shows that the time frequency analysis performs better compared to the MFCC in term of providing the best resolution in frequency domain representation. The analysis using TFD preserved the uniqueness of the input signal in term of time and frequency. It enhanced the signal information while suppressing the unnecessary noise and interference.

References

1. J.Y.Lee,T.J Won, Jeong And S.W.Nam, "Classification Of Power Disturbances Using Feature Extraction In Time Frequency Plane", 2002.
2. Z. Chen, Senior Member, Ieee, And P Urwin, "Power Quality Detection & Classification Using Digital Filter," Ieee, 2001.
3. Y.J Shin, Antony C.Parsons, Edward J.Powers, W.M. Gardy, "Time-Frequency Analysis of Power System Disturbance Signals for Power Quality", pp 402-407,1999
4. Michael E.Wall, Andreas Rechtsheiner, Luis M. Rocha, "Singular Value Decomposition and Principal Component Analysis", 2003
5. Boashash, B. (1992). Time Frequency Signal Analysis: Methods andApplications. Melbourne, Australia: Longman Cheshire.
6. Boashash, B. and Sucic, V. (2000). A Resolution Performance Measure for Quadratic Time-Frequency Distributions. Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing. August 14 -16. 584 -588.
7. Barkat B. and Boashash B. (2001). A High-Resolution Quadratic Time–Frequency Distribution for Multicomponent Signals Analysis. IEEE Transactions on Signal Processing. 49(10). October 2001. 2232-2239
8. Cohen L. (1989). Time-Frequency Distributions: A Review. Proceedings of the IEEE. 77(7). July 1989. 941-981.
9. Daliman, S. and Sha'ameri, A. Z. (2003). Time-Frequency Analysis of Heart Sounds using Windowed and Smooth Windowed Wigner-Ville Distribution. Proceedings Seventh International Symposium on Signal Processing and its Application. 2. July 1-4. 625-626.
10. Molau, S., Pitz, M., Schluter, R. and Ney, H. (2001). Computing Mel- Frequency Cepstral Coefficients on the Power Spectrum. Proceedings ICASSP '01) 2001 IEEE International Conference on Acoustics, Speech and Signal Processing. May 7-11. 73-76.
11. Sucic, V., Barkat B. and Boashash, B. (1999). Performance Evaluation of the B-Distribution. Fifth International Symposium on Signal Processing and its Applications. August 22-25. Brisbane, Australia: IEEE, 267-270.
12. Davis, S. B. and Mermelstein, P. (1980) Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences. IEEE Transactions on Acoustics, Speech and Signal Processing. 28(4). August 1980. 357-366.
13. Picone, J. W. (1993). Signal Modeling Techniques in Speech Recognition. Proceedings of the IEEE. 81(9). September 1993. 1215-1247.
14. Wall, M. E., Rechtsteiner, A. and Rocha, L. M. (2003). Singular Value Decomposition and Principal Component Analysis. in D. P. Berrar, W. Dubitzky, and M. Granzow, A Practical Approach to Microarray DataAnalysis, Kluwer: Norwell. 91-109.
15. Jozef Wartak, M.D. (1972). Phonocardiology: Integrated Study of Heart Sound and Murmurs, Harper and Row Publisher.
16. Hotelling, H. (1933), "Analysis of a Complex of Statistical Variables into Principal Components," Journal of Educational Psychology, 24, 417-441, 498-520.

Address of the corresponding author:

Author: I.Kamarulafizam
Institute: Centre for Biomedical Engineering, Universiti Teknologi Malaysia
Street: -
City: Johor Bahru
Country: Malaysia
Email: mr_fizam@yahoo.com