AUTOMATED PATTERN CLASSIFICATION FOR PCG SIGNAL BASED ON ADAPTIVE SPECTRAL K-MEANS CLUSTERING ALGORITHM

[1]Abbas K. Abbas , [2]Rasha Bassam

[1]Biomedical engineering Dept. RWTH-Aachen University, [2]Aachen university of Applied Sciences

INTRODUCTION

PCG Pattern classification, also known as auscultation pattern recognition, was one of the efficient computer-based methods applied to medical decision making. PCG Pattern recognition generally is interpreted in two ways. The most general definition includes recognition of patterns in any type of PCG dataset and is called uniform PCG pattern classification this discriminate peaks of heart sounds as excitation source for circulation hemodynamic, and other is called adaptive pattern clustering which magnify and observe the spectral characteristics associated with PCG waveform turbulences and differentiate them as clinical diagnostic indices. Fig.1 shows how the four heart sounds are correlated to the electrical and mechanical events of the cardiac cycle.

PCG CLASSIFICATION TECHNIQUE

This work reports robust results with phonocardiogram PCG-signal pattern classification. Linear prediction analysis and basic agglomerative clustering techniques were applied to extract the spectral pattern from phonocardiogram signals, a relatively new technique. In this examination, 35 PCG samples are classified correctly, except for seven samples; and 24 PCG samples correctly, except for three samples. The characteristics for each class are well extracted and the results of spectral classifications are obviously robust. The efficiency of PCG spectral features classification has been confirmed experimentally to be integrated in automated auscultation computer aided diagnosis (AuCAD) systems. Discrimination of abnormal S_1, S_2 and S_3 peaks was succeed with BAC-algorithm and k-mean based data clustering technique. The more specific interpretation algorithms are limited to finding patterns in PCG signals or other related biosignal activities. This work covers the new techniques applied in basis of pattern classification for mitral regurgitation PCG signals to investigate different hemodynamic turbulences and stochastic blood flow patterns associated with cardiac circulation

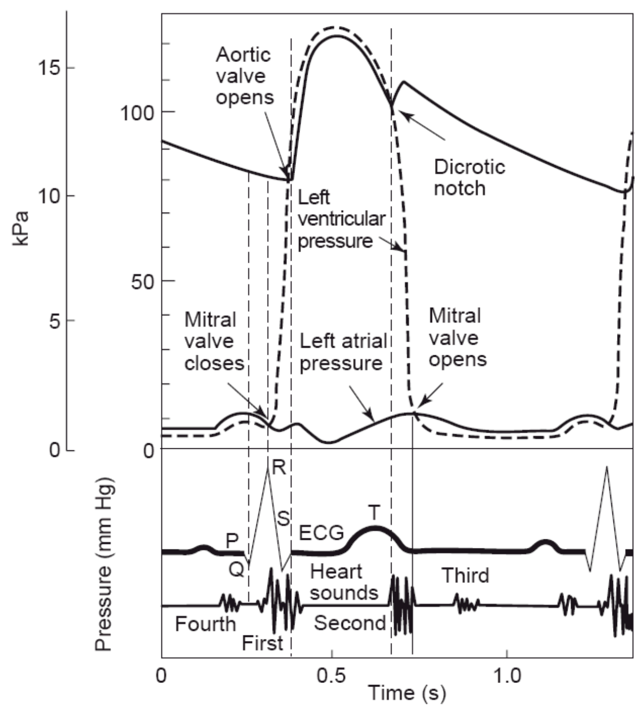

Fig.1 Correlation of the four heart sounds with the electrical and the mechanical events of the cardiac cycle [1].

PCG CLASSIFICATION TECHNIQUE

PCG spectra contain vast number of definite harmonics categories that would be useful to be identified as clustering scheme in any data classification algorithms, Majority of these spectra although it belongs to specific valvular pathologies have a distinct energy (intensity) level in STFT plot, this would be attractive point to consider such variation as clustering point, this considerably will oriented the classifier algorithm to stable entropy value. The spectral information characteristics of PCG lie within a frequency band of (54-520 Hz) and this band depend on digital stethoscopic interface and resolution of data converters in instrumentation platform. This criteria is the basis for our pattern classification technique in which the dependency on frequency (spectral) characteristics. Block diagram for overall spectral classification system is demonstrated in Fig.2 [1]

Several patterns can be derived from the vector input PCG signal, in which they processed with a specific FIR-filter length. The most recognized patterns in PCG are systolic and diastolic and presystolic and post-diastolic peaks of sound (S1, S2, S3, and S4) which is shown in Fig.1, most of cardiologist prefer the base-diagnosis on 2 categories of PCG, S1 and S2, so that they can discriminate the hemodynamics turbulences in appropriate method, the spectra stamp can be oriented in 3 schema (supraspectra, infraspecta and mid-spectra) in which are represent the intensity harmonics for PCG waveform, correlation between two intensity peaks of PCG gives a defined index for clustering profile $M_{j\text{-}PCG}$ of PCG signal which in turn apply a segmental cluster for input vector[2].

Systolic and diastolic murmur frequencies are classified according to the frequency band containing the largest power value in the tenth (s) of the systole/diastole corresponding to the found maximum values of SI/DI. If the largest power value is found in one of the two lowest frequency bands (containing frequencies below 125 Hz), the murmur is classified as a low-frequency murmur. If the largest power value is found in one of the eight highest frequency bands (containing frequencies above 250 Hz), the murmur is classified as a high-frequency murmur. If the none of the above is the case, the murmur is classified as a medium-frequency murmur [5].

1st –step: obtain PCG spectral information.

This result obtained by using Db-wavelets decomposition techniques for a set of PCG signal as below.

$$y[n] = (x_{PCG} * g)[n] = \sum_{k=-\infty}^{\infty} x_{PCG}[k] g[n-k] \text{..... (1)}$$

Extracting the PCG diastolic low frequency components

$$y_{low-PCG(diastolic)}[n] = \sum_{k=-\infty}^{\infty} x_{PCG}[k] g[2n-k] \text{.....(2)}$$

And for PCG systolic high frequency components

$$y_{high-PCG(systolic)}[n] = \sum_{k=-\infty}^{\infty} x_{PCG}[k] h[2n-k] \text{.... (3)}$$

Based on the characteristic features extracted from the heart sound signal, the nature of the heart sound can be identified using pattern recognition techniques. A number of pattern recognition and classification schemes have been implemented for the analysis of heart sounds. Classical pattern recognition techniques include the Gaussian– Bayes classifier and the K-nearest neighbor classifier (k-mean clustering). The Gaussian–Bayes classifier is the most popular parametric technique of supervised pattern recognition. It is considered optimal when the probability density functions (p.d.f) of the patterns in the feature space are known (a pattern is defined as an N-dimensional vector composed of N features) [3]. The K-nearest neighbor classifier is a nonparametric approach, which is useful when the probability density functions are difficult to estimate or cannot be estimated [4]. The nearest neighbor method is an intuitive approach based on distance measurements, motivated by the fact that patterns belonging to the same class should be close to each other in the feature space. Joo et al. demonstrated the diagnostic potential of a Gaussian–Bayes classifier for detecting degenerated bioprostheses implanted in the aortic position [4].

2nd Step: apply k-mean clustering for derived spectra.

Taking the momentum equation for input vector $x_{PCG\text{-}signal}$ to separate the spectral pattern derived from Db-wavelets decomposition algorithm.

$$M_{j\,\text{PCG-signal}} = \left(x_{PCG} - \mu_j\right)^T \sum_j^{-1} \left(x_{PCG} - \mu_j\right) \text{..... (4)}$$

$$S_{W\,PCG-signal} = \sum_{i=1}^{C} \sum_{x \in X_i} \left(x_{PCG} - m_i\right) \cdot \left(x_{PCG} - m_i\right)^T \text{..(5)}$$

Where $M_{j\ PCG}$ is the class momentum in k-space and it is constructed first to set the separation line for each class, and $S_{W\ PCG\text{-}signal}$ is the segmental pattern derived after integrating the momentum equation (4).

AUTOMATED PCG DATA ACQUISTION

Interpretation of PCG spectra pattern was done through an EDA-Exploratory Data Analysis platform-MATLAB® with spectral generating by Autosignal® ;where the vectors of PCG data fed into k-mean-agglomerative cluster kernel and extracted the main classes for S1 and S2 waveform. The computation algorithm present a stable operation through classification of 15 PCG (10 male and 5 female) samples width S.D. =0.023 and 37-yr old average.

Results obtained from classification methods are shown in fig.3 where the chaotic patterns have been improved in identification clinical value for auscultation diagnostics, robust and stable clustering methods have been approved by k-mean technique. The clustering with k-mean method seems to be more robust in response to PCG patterns table.1 shows the performance analysis result; fig.4 shows the difference between 7 cluster graphs before applying *k*-clustering and after. Pattern orientation has been improved with this method. Definition of each spectra cluster within agglomerative and *k*-mean aspect in illustrated well, and it constitutes a good basis for effective pattern identification. An automated k-mean classifier system developed based on multiple recording stethoscope system acquired PCG from traditional site of auscultation from the body chest and back [3].

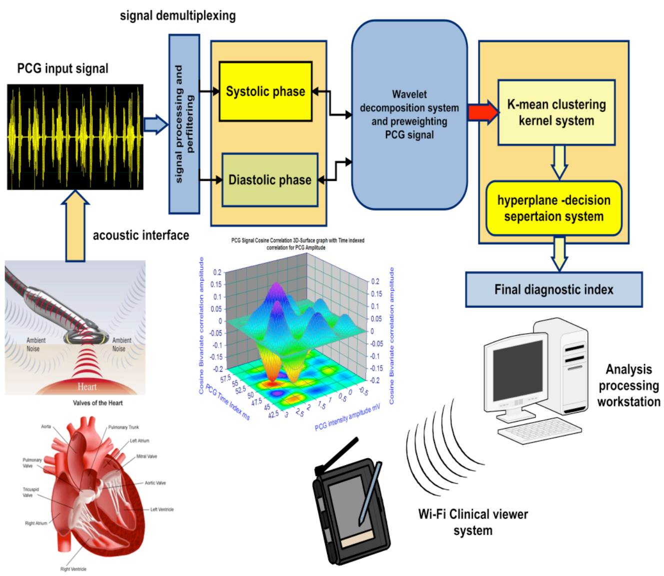

Fig.2 Block diagram of Automated PCG classifier system with k-mean algorithm and wireless connectivity to clinical mobile workstation.

PCG spectral estimation is demonstrated in Fig.3. Where the 3 dominant cased for systolic and diastolic murmurs are observed.

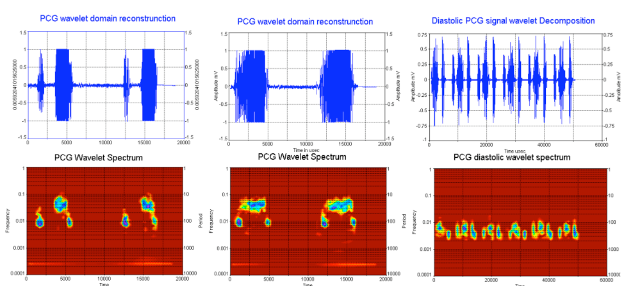

Fig.3 PCG pattern spectra derived from wavelet decomposition and to be entered to k-mean classification algorithm.

STATISTICAL PERFORMANCE ANALYSIS

Performance analysis of the k-mean clustering showing a stable and robust result in comparison to supervised classification method such neural network and higher order statistics HOS , as shown in Fig.5 where the SIR value of clustered signal approach to 1.68 as compared to value in ANN-based classifier or HOS method , following table 1.1 shows this comparison as well , consistent clustering in k-mean can be observed as decreasing the mean p-value of significance probability and increase no of cluster in the class plane domain , this will forced the kernel to stabilize the clustering process ,and adapt itself to a convergence point regarding a multiple input vector of PCG signals. Ability to isolate different cluster from numerous number of received PCG spectra [7].

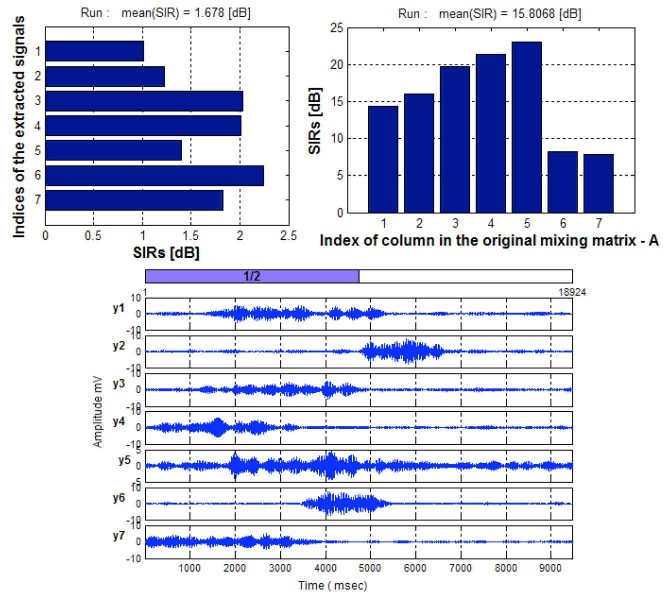

Fig.4 performance analysis of extracted PCG signal and SIR value for overall classification matrix

Performance index for k-mean clustering has been noted as a higher index for PCG data clustering. Viability to robust clustering and a well-defined pattern adding superiority to k-mean clustering techniques, by which cardiologist have to be considered in artificial intelligence-based PCG diagnostic modes and intensity-oriented classifiers system.

Basically, k-mean clustering computation time is comparably longer that with agglomerative based clustering, this is due to the high categorization line with k-mean and insufficient spectral correlation between intensity peaks of murmurs in PCG. Optimization for each method needs intensive work on biomedical spectral analysis. Redundancies pattern associated with this technique raised from incomplete training sets or random noise spectra associated with acquired PCG signal, dynamic identification of diastolic and systolic murmurs can be corrected from associated noise by using FIR filtering implemented in the automated classification path to remove any parasitic acoustic noise which may occurred during data acquisition procedure[8].

Table 1: Performance index for PCG clustering methods

PCG clustering index				
Clustering Method	***p-value***	***SIR index***	***Mean of PCG signal***	***No .of cluster identified***
K-Mean	0.0123	1.682	0.542	11
ANN-RBN	0.0167	1.732	0.732	8
HOS	0.0189	1.788	0.931	9
Basic Aggl.	0.0154	1.892	0.963	7
Model based	0.1923	2.013	1.038	7

Table.2 illustrates the classification performance test of k-mean clustering method. Different physiological parameters of the multiple cycles PCG recorded for 12 second PCG trace have been classified. In this table the most robust classified pattern is the diastolic and abnormal blood flow turbulence which are associated with acute and chronic coronaries arteries disease and the relative decrease in performance for systolic pattern referred to the coherent noise parasitized in PCG recording and this may eliminate the classifier form being recognized the mean value of PCG amplitude. Finally the further averaging and filtering using a composed wavelets decomposition and PCA separation for optimized the extracted coefficients for PCG signals, clinical validation should take around as minimum 130 cases for approved techniques as automated auscultation diagnostics module in heart sound pathology [8].

Table 2: Performance index for PCG k-mean clustering

Classification test performance for K-mean clustering				
	Corrected PCG pattern	***Uncorrected PCG pattern***	***Total No. of PCG pattern***	***% performance of identified PCG***
Systolic pattern	81	6	87	93
Diastolic pattern	64	3	67	95.5
Murmurs	78	5	83	91.7
Abnormal blood flow pattern.	46	3	49	93.9

CONCLUSION

Application of k-mean and agglomerative clustering showing high stability to different PCG signal categories and this indeed play a vital role in quantifying and identifying different hemodynamic cases through pattern classification technique. Well define separation line has been identified and robust S1 and S2-spectra clustering was achieved and variable pathological cases were tested, including different mitral valve regurgitation and systolic hemodynamic abnormalities , stochastic blood flow patterns. Improvement of the clustering system makes through successive repeatable iteration for each pattern on *k*-mean and agglomerative clustering technique by minimizing the distance function of clustered pattern. Clinical information about the health of heart valves is contained in a single cycle of PCG signal. Hence it is very important to identify single cycle for analysis of defects. Current state of art techniques use a reference signal like ECG or Carotid blood pressure pulse, to obtain a single cycle of the PCG signal. We have proposed a novel method to segment and classified PCG signal into single cycle using wavelet-debauches decomposition and k-means clustering, which works well when the heart rate is uniform for the entire sequence of PCG signal recording. The segmentation algorithm has shown 91.22% of success. The 29 wavelet features of each segmented cycle were reduced to 11 using basic agglomerative clustering. These reduced feature sets were classified by the k-mean clustering into 5 classes. Classification performance of 95.32% was obtained. It is concluded that classification of segmented PCG signals obtained without using a reference signal can be achieved through an automated k-mean clustering of PCG spectra in single or multiple heart cycle.

ACKNOWLEDGEMENTS

We acknowledge biomedical engineering department at RWTH-Aachen and Aachen University of applied sciences and German Academic Exchange Service DAAD, for their support and assistance for this project to get on track of success.

REFERENCE S

[1] Criley, J. M., D. Criley, and C. Zalace. "The physiological origins of heart sounds and murmurs: the unique interactive guide to cardiac diagnosis". Boston: Blaufuss Medical Multimedia, 1995.

[2] Debbal, S. M. and F. Bereksi-Reguig."Analysis of the second heart sound using continuous wavelet transforms". J. Med. Eng. Technol. 28:151–156, 2004.

[3] El-Segaier, M., O. Lilja, S. Lukkarinen, L. Sornmo, R. Sepponen, and E. Pesonen. "Computer-based detection and analysis of heart sound and murmur". Ann. Biomed. Eng. 33:937–942, 2005.

[4] Guo, Z., L. G. Durand, H. C. Lee, L. Allard, M. C. Grenier, and P. D. Stein." Artificial neural networks in computer-assisted classification of heart sound in patients with porcine bioprosthetic valves". Med. Biol. Eng. Comput. 32:311–316, 1994.

[5] Leatham A. Auscultation of the Heart and Phonocardiography, 2nd Edition. London: Churchill Livingstone, 1998. 4th edition.

[6] Li, X., M. Parizeau, and R. Plamondon. "Training hidden Markov models with multiple observations-a combinatorial method". IEEE Trans. Pattern Anal. Mach. Intel. 22:371–377, 2000.

[7] Liang, H., S. Lukkarinen, and I. Hartimo. Heart sound segmentation algorithm based on heart sound envelogram. Comput. Cardiol. 24:105–108, 1997.

[8] Brock A, Pateros M, and Clifford P, "Characterization of PCG sound as index for hemodynamic disturbances: review study", IFBME 2000, Springer.

Author:
Abbas Kader Abbas M.Sc. Medical Engineering (2004)
Biomedical Engineering Dept. RWTH Aachen University
E-mail: abbas.kader@rwth-aachen.de
abbas.khudair@gmail.com
Rasha Bassam M.Sc. Medical Engineering (2007)
Aachen university of Applied Sciences
Biomedical Engineering Dept.
E-mail: rashabassamss84@gmail.com