ITS
Institut
Teknologi
Sepuluh Nopember

ICTS 2016

PROCEEDINGS OF

2016 International Conference on
Information & Communication
Technology and Systems (ICTS)

October 12th, 2016

Department of Informatics, Faculty of Information Technology
Institut Teknologi Sepuluh Nopember (ITS)

PROCEEDINGS OF

2016 INTERNATIONAL CONFERENCE ON INFORMATION & COMMUNICATION TECHNOLOGY AND SYSTEMS (ICTS)

Surabaya, October 12th, 2016

(ISSN: 2338-185X)
(ISBN: 978-1-5090-1379-1)

Organized by
Department of Informatics
Faculty of Information Technology
Institut Teknologi Sepuluh Nopember (ITS)
Surabaya, Indonesia

2016 INTERNATIONAL CONFERENCE ON INFORMATION & COMMUNICATION TECHNOLOGY AND SYSTEMS (ICTS)

Copyright and Reprint Permission:

Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For reprint or republication permission, email to IEEE Copyrights Manager at pubs-permissions@ieee.org. All rights reserved. Copyright ©2016 by IEEE.

PROCEEDINGS OF

2016 INTERNATIONAL CONFERENCE ON INFORMATION & COMMUNICATION TECHNOLOGY AND SYSTEMS (ICTS)

EXECUTIVE BOARD

Joni Hermana
Rector

Agus Zainal Arifin
Dean of Faculty of Information Technology

Darlis Herumurti
Head of Department of Informatics

Keynote Speakers

Shengdong Zhao (National University of Singapore)
Riyanarto Sarno (Institut Teknologi Sepuluh Nopember)

Scientific Committee

Agus Zainal Arifin
Ahmad Hoirul Basori
Akira Asano
Anders Nilsson Plymoth
Annappa B
Chastine Fatichah
Daniel Petrisor
Darlis Herumurti
Eduardo Pinos
Eraclito de Souza Argolo
Erma Suryani
Faisal Rahutomo
Gulshan Shrivastava
Han-You Jeong
Handayani Tjandrasa
Hazura Haroon
Hengky Susanto
Heru Suhartanto
I Gede Pasek Suta Wijaya
Indah Agustien Siradjuddin
Joko Lianto Buliali
Junfei Qiu
K V Krishna Kishore
Manoochehr Nahvi
Nanik Suciati
Nursabillilah Binti Mohd Ali
Pitoyo Hartono
Raad S. Al-Qassas
Radityo Anggoro
Ravi Subban
Riyanarto Sarno
Rodrigo Campos Bortoletto
Royyana Muslim Ijtihadie
RV Hari Ginardi
Sergio Camorlinga
Shan Sung Liew
Shashikant Shantilal Patil
Siti Rochimah
Supeno Djanali
Suraya Mohammad
Sven Zehl
Taras Maksymyuk
Tohari Ahmad
Tsuyoshi Usagawa
Wenjie Hu

General Co-Chair

Ridho Rahman Hariadi
Rizky Januar Akbar

Organizing Committee

Dwi Sunaryono
Henning Titi Ciptaningtyas
Imam Kuswardayan
Sarwosri

ISSN: 2338-185X
ISBN: 978-1-5090-1379-1

CONTACT ADDRESS
Department of Informatics
Faculty of Information Technology
Institut Teknologi Sepuluh Nopember (ITS) Surabaya
Gedung Teknik Informatika, ITS
Jalan Teknik Kimia, Kampus ITS Sukolilo, Keputih, Surabaya, Indonesia 60111
Tel. (+62)-31-5939214 Fax. (+62)-31-5913804
http://www.if.its.ac.id/

TABLE OF CONTENTS

KN: Keynote Speech

KN.1 Innovation Through Principles
Shengdong Zhao
pp. i

KN.2 Process Model Discovery for Process Aware Information Systems
Riyanarto Sarno
pp. iii-x

T1: Distributed System, Computer Network and Architecture, Network Security, Infrastructure Systems and Services, Ubiquitous System and Infrastructure, Digital Forensic, High Performance Computing, Parallel Programming, Information Security

T1.1 Digital Forensic Analysis of Telegram Messenger on Android Devices
Gandeva Bayu Satrya, Philip T. Daely and Muhammad Arief Nugroho
pp. 1-7

T1.2 Sales Forecasting Using Holt-Winters in Enterprise Resource Planning At Sales and Distribution Module
Vicky Sugiarto, Riyanarto Sarno and Dwi Sunaryono
pp. 8-13

T1.3 Determining Bonus in Enterprise Resource Planning At Human Resource Management Module Using Key Performance Indicator
Dwi Suseno, Riyanarto Sarno and Dwi Sunaryono
pp. 14-19

T1.4 Account Charting and Financial Reporting At Accounting Module on Enterprise Resource Planning Using Tree Traversal Algorithm
Riyanarto Sarno, Ashari Adhitama and Sarwosri
pp. 20-25

T1.5 Dynamics Simulation Model of Demand and Supply Electricity Energy Public Facilities and Social Sector Case Study East Java
Ardhya Putra, Riyanarto Sarno and Erma Suryani
pp. 26-33

T1.6 HTTP Communication Latency Via Cellular Network for Intelligent Transportation System Applications
Michael Ardita, Suwadi Suwadi, Achmad Affandi and E Endroyono
pp. 34-38

T1.7 Clustering of SSH Brute-Force Attack Logs Using k-Clique Percolation
Hudan Studiawan, Baskoro Adi Pratomo and Radityo Anggoro
pp. 39-42

T2: Human-Computer Interaction, Multimedia Application

T2.1 Design and Implementation of Virtual Indonesian Musical Instrument (VIMi) Application Using Leap Motion Controller
Ridho Rahman Hariadi and Imam Kuswardayan
pp. 43-48

T2.2 Augmented Reality Application for Cockroach Phobia Therapy Using Everyday Objects as Marker Substitute
Fiandra Fatharany, Ridho Rahman Hariadi, Darlis Herumurti and Anny Yuniarti
pp. 49-52

T2.3 Implementation of Face Detection and Recognition of Indonesian Language in Communication Between Humans and Robots
Muhtadin and Surya Sumpeno
pp. 53-57

T2.4 Resize My Image: A Mobile App for Interactive Image Resizing Using Multi Operator and Interactive Genetic Algorithm
Anny Yuniarti, Stefanus Anggara and Bilqis Amaliah
pp. 58-62

T3: Pattern Recognition, Intelligent Systems, Knowledge Data Discovery, Bioinformatics/Biomedical Apps, Content-Based Multimedia Retrieval, Remote Sensing, Image Processing, Big Data, Information Retrieval, Robotic, Modelling Simulation, Applied Computing

T3.1 Detecting Source Code Plagiarism on Introductory Programming Course Assignments Using a Bytecode Approach
Oscar Karnalim
pp. 63-68

T3.2 Computer-Aided Screening for Acute Leukemia blood infection using gray-level intensity
Hatungimana Gervais and Darlis Herumurti
pp. 69-74

T3.3 Integration GLCM and Geometric Feature Extraction of Region of Interest for Classifying Tuna
Wanvy Arifha Saputra and Darlis Herumurti
pp. 75-79

T3.4 Face Recognition Based on Extended Symmetric Local Graph Structure
Andhik Yunanto and Darlis Herumurti
pp. 80-84

T3.5 Classification of EEG Signals Using Common Spatial Pattern-Principle Component Analysis and Interval Type-2 Fuzzy Logic System
William Yaputra Budiman, Handayani Tjandrasa and Dini Navastara
pp. 85-89

T3.6 Job-Shop Scheduling Model for Optimization of the Double Track Railway Scheduling (Case Study: Solo-Yogyakarta Railway Network)
Sarngadi Palgunadi, Dian Supraba and Bambang Harjito
pp. 90-95

T3.7 Koi Fish Classification based on HSV Color Space
Dhian Kartika and Darlis Herumurti
pp. 96-100

T3.8 Adaptive Image Compression Using Adaptive Huffman and LZW
Djuned F Djusdek, Hudan Studiawan and Tohari Ahmad
pp. 101-106

T3.9 Indonesia Scholarship Selection Framework Using Fuzzy Inferences System Approach. Case Study: "Bidik Misi" Scholarship Selection
Luther Latumakulita, Fajar Purnama, Tsuyoshi Usagawa, Sary Paturusi and Delta Prima
pp. 107-113

T3.10 Using Google Trend Data in Forecasting Number of Dengue Fever Cases with ARIMAX Method Case Study: Surabaya, Indonesia
Wiwik Anggraeni and Laras Aristiani
pp. 114-118

T3.11 Maximum Entropy Principle for Exudates Segmentation in Retinal Fundus Images
Igi Ardiyanto, Ratna Lestari Budiani Buana and Hanung Adi Nugroho
pp. 119-123

T3.12 Pressure Control of A Wheeled Wall Climbing Robot Using Proportional Controller
Novia Andriani, Yunafi'atul Aniroh, Muhammad Maulida and Miftahul Alfafa
pp. 124-128

T3.13 The Electric Wheelchair Control Using Electromyography Sensor Of Arm Muscle
Rudi Hardiansyah, Yunafi'atul Aniroh, Arista Ainurrohmah and Farida Tyastuti
pp. 129-134

T3.14 Reduce Noise in The Binary Image Using Non Linear Spatial Filtering of Mode
Teady Matius Surya Mulyana
pp. 135-139

T3.15 Heart Murmurs Extraction Using the Complete Ensemble Empirical Mode Decomposition and the Pearson Distance Metric
Jusak Jusak, Ira Puspasari and Pauladie Susanto
pp. 140-145

T3.16 Development of Early Detection of Complication Organ Kidney Disease Caused by Diabetes Mellitus Based on Color Constancy
Agus Prayitno, Adhi Wibawa and Mauridhi Purnomo
pp. 146-149

T3.17 Digital Color Classification for Colorful Cross Stitch Threads Using RGB+Euclidean Distance and LAB+CIE94
David Setiabudi, Sani M. Isa and Bambang Heru Iswanto
pp. 150-156

T3.18 Application Development for Recognizing Type of Infant's Cry Sound
Welly Limantoro, Chastine Fatichah and Umi Laili Yuhana
pp. 157-161

T3.19 Opinion Classification Using Maximum Entropy and K-Means Clustering
Amir Hamzah and Naniek Widyastuti
pp. 162-166

T3.20 Enriching English Into Sundanese and Javanese Translation List Using Pivot Language
Arie Suryani, Isye Arieshanti, Banu Yohanes, Muhammad Subair, Sari Dewi Budiwati and Bagus Rintyarna
pp. 167-171

T3.21 Mixed Vapour Identification Using Partition Column-QCMs and Artificial Neural Network
Eva Agustin, Muhammad Rivai and Achmad Arifin
pp. 172-177

T3.22 BencanaVis Visualization and Clustering of Disaster Readiness Using K Means with R Shiny A Case Study for Disaster, Medical Personnel and Health Facilities Data at Province Level in Indonesia
Renny Kusumawardani, Irmasari Hafidz and Septa Putra
pp. 178-186

T3.23 Dynamics Simulation of Air Passenger Forecasting and Passenger Terminal Capacity Expansion Scenario in Yogyakarta Airport
Bilqis Amaliah, Azizha Zeinita and Erma Suryani
pp. 187-192

T3.24 Feature Extraction Using Statistical Moments of Wavelet Transform for Iris Recognition
Nanik Suciati, Afdhal Anugrah, Chastine Fatichah, Handayani Tjandrasa, Agus Arifin, Diana Purwitasari and Dini Navastara
pp. 193-198

T3.25 Generalized Regression Neural Network for Predicting Traffic Flow
Joko Buliali, Victor Hariadi, Ahmad Saikhu and Saprina Mamase
pp. 199-202

T4: Software Engineering, Formal Methods, E-Learning, Enterprise Information System, Risk Management, Geographic Information System

T4.1 The Development of Method of The Enhancement of Technical Factor (TF) and Environmental Factor (EF) to The Use Case Point (UCP) to Calculate The Estimation of Software's Effort
Sarwosri, Muhammad Jabir Al Haiyan, Aditya Putra Ferza and Mujahid Husein
pp. 203-207

T4.2 Substation Placement Optimization Method Using Delaunay Triangulation Algorithm and Voronoi Diagram In East Java Case Study
Pradipta Ghusti, Riyanarto Sarno and RV Hari Ginardi
pp. 208-213

T4.3 Visual GUI Testing in Continuous Integration Environment
Fachrul Pralienka Bani Muhamad, Riyanarto Sarno, Adhatus Solichah Ahmadiyah and Siti Rochimah
pp. 214-219

T4.4 A Study of Students' Satisfaction Toward Blended Learning Implementation in Higher Education Institution in Indonesia
Sary Paturusi, Arie Lumenta and Tsuyoshi Usagawa
pp. 220-225

T4.5 Sugarcane Variety Identification Using Dynamic Weighted Directed Acyclic Graph Similarity
Adi Heru Utomo, Riyanarto Sarno and RV Hari Ginardi
pp. 226-230

T4.6 Business Process Model Similarity Analysis Using Hybrid PLSA and WDAG Methods
Indra Gita Anugrah and Riyanarto Sarno
pp. 231-236

T4.7 Optimization Solar Farm Site Selection Using Multi-Criteria Decision Making Fuzzy AHP and PROMETHEE: Case Study in Bali
Kadek Aldrin Wiguna, Riyanarto Sarno and Nurul Fajrin Ariyani
pp. 237-243

T4.8 Student Perceptions of Virtual Programming Lab on E-Learning Class at University of Sam Ratulangi
Sumenge Tangkawarouw Godion Kaunang, Tsuyoshi Usagawa, Sary Paturusi, Alwin Sambul, Glanny Mangindaan and Brave Sugiarso
pp. 244-248

T4.9 Fraud Detection on Event Logs of Goods and Services Procurement Business Process Using Heuristics Miner Algorithm
Dewi Rahmawati, Muhammad Ainul Yaqin and Riyanarto Sarno
pp. 249-254

T4.10 The Development of Quality Gates Instrument for e-Learning Implementation
Feby Artwodini Muqtadiroh, Hanim Maria Astuti and Rian Triadi
pp. 255-261

T4.11 Designing a Gamification for Monitoring Surabaya City Development
Nur Rakhmawati and Bagus Fibrianto
pp. 262-265

T4.12 Poverty Classification Using Analytic Hierarchy Process and K-Means Clustering
Sarwosri, Dwi Sunaryono, Rizky Januar Akbar and Risky Setiyawan
pp. 266-269

SPONSORED BY :

9781509013791

Heart Murmurs Extraction Using the Complete Ensemble Empirical Mode Decomposition and the Pearson Distance Metric

Jusak Jusak, Ira Puspasari, Pauladie Susanto
Dept. of Computer Engineering
Institut Bisnis dan Informatika Stikom Surabaya
Surabaya, Indonesia
jusak@stikom.edu, ira@stikom.edu, pauladie@stikom.edu

***Abstract*— Signal processing for pathological heart sound signals can be considered as a fundamental part of the whole process in tele-auscultation systems. In this paper, we employ the CEEMD and the EEMD algorithm to decompose various pathological heart sound signals in the form of phonocardiograph (PCG) signals. Following the decomposition process, we subsequently extract murmurs from the targeted heart sound signals using our proposed technique that based on the Pearson distance metric. Performance analysis of the decomposition algorithms as well as the extraction method is evaluated in terms of delta SNR that signifies variance comparison of targeted signal before and after murmurs extraction. It can be concluded that in general pathological heart sound signals that have been decomposed by the CEEMD algorithm followed by the Pearson distance metric for murmurs extraction, provide the finest murmurs extraction than the EEMD. Additionally, the EEMD algorithm involves smaller number of modes to form the extracted murmurs signal as compared to the CEEMD algorithm. However, employing the CEEMD algorithm produces higher number of shifting procedures causing higher computational complexity than the EEMD algorithm.**

Keywords—phonocardiograph, heart sound, heart murmurs, empirical mode decomposition.

I. INTRODUCTION

Based on the latest data released by the World Health Organization (WHO) in 2014, deaths caused by cardiovascular disease in 2012 has reached 17.5 million deaths, or 46% of the total number of non-communicable diseases deaths in the world [1]. In addition to that, in another WHO report states that in 2020, it is estimated that the coronary heart disease will be the major killer diseases in countries throughout Asia-Pacific [2]. Based on those facts, there can be seen urgent need for assisted technologies that will be able to counteract or at least to do early detection for the diseases. To anticipate terminal ill (that mostly leads to deaths) initiated by the cardiovascular diseases, some works have been proposed to utilize the so-called online observation and detection models. These are commonly termed as tele-auscultation systems [3-4]. The tele-auscultation systems proved to be useful specifically in the remote areas where the presence of cardiovascular experts are void.

Recent studies show some promising results in the heart sounds signal processing by means of decomposing signal into a set of intrinsic mode functions (IMFs) using the Empirical Mode Decomposition (EMD) method proposed by Huang et al. in 1998 [5]. The advantages of analyzing heart sounds signals by IMFs have been shown, for example, in [6] to study signal recording that are often contaminated with spike noise produced by measurement instruments and in [7] where EMD was used to improve the spectrum estimates of heart rate variability. Furthermore, authors in [8] suggested the use of EMD for separating heart sounds signals and murmurs by extracting a multi-component signal into a set of mono-component signals, called the IMFs and then selecting the most appropriate IMFs to represent the undistorted heart sound signals. To overcome the scale separation problem in the EMD, therefore, a noise-assisted signal was introduced in [9]. The new method is called Ensemble Empirical Mode Decomposition (EEMD), which describes new IMF components as the mean of an ensemble of trials of signals, each consisting of the signal with addition of white noise series [10-11]. Advances in the study of EEMD development lead to a Complete Ensemble Empirical Mode Decomposition (CEEMD) that was argued to have significant improvement in terms of algorithm complexity reduction [12,13].

In this paper, we will employ the CEEMD and the EEMD algorithms to decompose various pathological heart sound signals taken from University of Michigan and University of Washington databases in the form of phonocardiograph (PCG) signals. These PCG signals provide information about cardiac valve function producing murmurs signals. Following the decomposition process, we subsequently extract murmurs from the targeted heart sound signals using our proposed technique that based on the Pearson distance metric. Performance analysis of the decomposition algorithms as well as the extraction method is evaluated in terms of **ΔSNR** (delta SNR) that signifies variance comparison of measured signal before and after murmurs extraction.

The work is supported by Indonesian Minister of Research and Higher Education through the competitive grant 2014-2015 number 058/SP2H/P/K7/KM/2015.

978-1-5090-1381-4/16/$31.00 ©2016 IEEE

The remaining of the paper is organized as follows: Section 2 elaborates algorithms for signal decomposition that will be utilized in this paper. Sections 3 explains our proposed murmurs extraction scheme. Performance evaluation of murmurs extraction scheme for various pathological heart sound signals is presented in Section 4. Finally, conclusion will be drawn in the last section of this paper.

II. Empirical Mode Decomposition and Its Development

The EMD is a method proposed by Huang et. al. [11] to perform a time-domain decomposition of an observed signal, $\boldsymbol{x(t)}$, into a set of Intrinsic Mode Functions (IMFs). According to the authors, a signal to be considered as an IMF must satisfy two basic conditions: (i) the whole length of an IMF should retain the number of extrema and number of zero-crossings either equal or differ at most by one, (ii) the mean value of the upper envelope defined by the local maxima and the lower envelope defined by the local minima is zero at any location. The algorithm for producing the $\boldsymbol{I}$-empirical modes from $\boldsymbol{x(t)}$ is called sifting process. As a result of decomposition process, an observed signal can be expressed as:

$$\boldsymbol{x(t) = \sum_{k=1}^{K} IMF_k(t) + R_K(t)} \qquad (1)$$

where $\boldsymbol{IMF_k(t)}$ is the $\boldsymbol{k}$-th IMF, and $\boldsymbol{R_K(t)}$ is the trend-like final residue.

The EEMD was proposed in [15] to overcome the mode mixing problems occurring in the EMD method. In the EEMD, each of the $\boldsymbol{k}$-th IMF component is considered as a mean of $\boldsymbol{IMF_k}$ generated over an ensemble of trials of $\boldsymbol{I}$ numbers of the original signal plus white Gaussian noise series with finite variance. For convenience, here we define the original signal as $\boldsymbol{N}$-sample heart sounds signal and denote as $\boldsymbol{x[n](n = 1, 2, \cdots, N)}$ obtained through process of recording an acoustic wave from a stethoscope. The signal is commonly called as phonocardiography (PCG) signal. Step-by-step of the EEMD algorithm can be described as follows:

1. We add a white Gaussian noise series to heart sounds signal, $\boldsymbol{x[n]}$ such that

$$\boldsymbol{x^i[n] = x[n] + \beta w^i[n].} \qquad (2)$$

   The $\boldsymbol{w^i[n](i = 1, 2, \cdots, I)}$ indicates $\boldsymbol{I}$ realizations of the zero mean unit variance white Gaussian noise series, while $\boldsymbol{\beta}$ is the controlled noise amplitude (noise standard deviation)[15].
2. Decompose the $\boldsymbol{x^i[n]}$ using the EMD algorithm to yield the $\boldsymbol{IMF_k^i[n]}$, where $\boldsymbol{k = 1, 2, \cdots, K}$ is the $\boldsymbol{k}$-th mode and each residue is obtained as

$$\boldsymbol{R_k^i[n] = R_{k-1}^i[n] - IMF_k^i[n].} \qquad (3)$$

3. Find the average of the $\boldsymbol{IMF_k^i[n]}$ to get the IMF component, $\boldsymbol{\overline{IMF}_k[n]}$, according to

$$\boldsymbol{\overline{IMF}_k[n] = \frac{1}{I}\sum_{i=1}^{I} IMF_k^i[n].} \qquad (4)$$

The key feature of the CEEMD that makes it more sophisticated than the EEMD is in terms of the residue generation. See Eq. (3). As can be seen in the above algorithm, each $\boldsymbol{x^i[n]}$ in the EEMD is decomposed independently and hence the residue is obtained accordingly. In contrast to that, the CEEMD algorithm calculates a unique first residue as

$$\boldsymbol{R_1[n] = x[n] - \widehat{IMF}_1[n],} \qquad (5)$$

where $\boldsymbol{\widehat{IMF}_1[n]}$ denotes the first decomposition mode of the CEEMD, it is derived the same way as in the EEMD. Following the generation of the first residue, a new ensemble of $\boldsymbol{R_1[n]}$ with addition of white Gaussian noise series for $\boldsymbol{I}$ realizations is obtained. It is then continued with computation the first EMD mode of each element. By taking average of these first EMD modes, the $\boldsymbol{\widehat{IMF}_2[n]}$ can be obtained. Then calculate the $\boldsymbol{k}$-th residue for $\boldsymbol{k = 2, 3, \cdots, K}$, decompose the $\boldsymbol{R_k[n]}$ and define the mode $\boldsymbol{\widehat{IMF}_{k+1}[n]}$. This procedure is repeated until the obtained residue can no longer be decomposed. The repetition procedure is commonly known as shifting process, which is identified by $\boldsymbol{N_{shift}}$ signifying total number of shifting process of $\boldsymbol{K}$ modes.

Having all of the modes generated by the CEEMD in hand, the observed signal can be expressed in terms of modes and residue as

$$\boldsymbol{x[n] = \sum_{k=1}^{K} \widehat{IMF}_k[n] + R[n],} \qquad (6)$$

where the equation in itself features complete and exact reconstruction of the observed signal.

III. Murmurs Extraction Scheme

Fig. 1 depicts PCG signals for normal and abnormal heart sounds signals. The most fundamental heart sounds are the first and the second (S1 and S2, respectively) sounds shown in the top figure. The first heart sound, S1, associated with cardiac vibrations produced by the closure of mitral and tricuspid valves. On the other hand, the second heart sound, S2, related to cardiac vibrations produced by the closure of the aortic and pulmonic valves. The S1 and S2 components of normal heart sound signal are clearly seen in the first row plot of Fig. 1. It approximately spans for 0.8 seconds period for 1 cycle normal heart sounds.

The two fundamental components of the heart sound signals that are perturbed by systolic murmurs are depicted on the bottom part of Fig. 1. These murmurs are produced by long vibrations that occur during systole period. These vibrations result from turbulent blood flow through a partially obstructed opening mitral or tricuspid valves that are found between the ventricular and the aortic chambers [14]. The presence of murmurs in the series of heart sounds can be an indication of abnormalities resided in the heart.

Advanced studies around murmurs characteristics show that murmurs are nonstationary in nature [15]. Murmurs' pitch, or frequency varies from low-pitched to high-pitched. They also exhibit sudden frequency changes. Hence, murmurs extraction requires advanced methods and algorithms in order to produce a meaningful set of separated signals as a tool for cardiologist in interpreting the cardiac dysfunction. In this study, murmurs extraction will be devised into two main procedures, i.e., signal decomposition and murmurs extraction.

PCG signals decomposition will be accomplished utilizing the most advanced CEEMD and the EEMD algorithms where original signals are decomposed into a set of IMFs. In most practical applications, the CEEMD generates noise-like series such as murmurs at lower subscription index of the IMFs, e.g. $\widetilde{\boldsymbol{IMF}}_1[\boldsymbol{n}]$, $\widetilde{\boldsymbol{IMF}}_2[\boldsymbol{n}]$, $\widetilde{\boldsymbol{IMF}}_3[\boldsymbol{n}]$, while the rest of the IMFs (at higher subscription index) contain lower frequency of the original signal. In this way, the CEEMD algorithm distributes the frequency heart sound components across different IMFs.

After the generation of IMFs using the CEEMD algorithm, an exact criterion is certainly needed to distinguish IMFs that contain murmurs from those that contain fundamental signals. For simplicity of the overall scheme, we proposed to employ Pearson's correlation criteria to differentiate IMFs and to select them automatically for revealing either murmurs or fundamental signals content out of the original signals. The correlation criteria is defined as

$$\boldsymbol{d_k} = \frac{\mathbf{cov}(\boldsymbol{x}[\boldsymbol{n}],\widetilde{\boldsymbol{IMF}}_{\boldsymbol{k}}[\boldsymbol{n}])}{\sqrt{\mathbf{cov}(\boldsymbol{x}[\boldsymbol{n}]).\mathbf{cov}(\widetilde{\boldsymbol{IMF}}_{\boldsymbol{k}}[\boldsymbol{n}])}} \qquad (7)$$

where $\boldsymbol{x}[\boldsymbol{n}]$ is the original heart sound signal and $\boldsymbol{d_k}$ is the correlation coefficient that is associated with the $\widetilde{\boldsymbol{IMF}}_{\boldsymbol{k}}[\boldsymbol{n}]$. It is clear from Eq. (6) that the correlation coefficient is retrieved from cross-correlation between original signals and the $\widetilde{\boldsymbol{IMF}}_{\boldsymbol{k}}[\boldsymbol{n}]$. Hence, the correlation coefficient, $\boldsymbol{d_k}$, represents degree of similarity between original signals and each of the $\boldsymbol{k}$-th $\widetilde{\boldsymbol{IMF}}$. We further describe this degree of similarity in terms of Pearson distance that is written as

$$\boldsymbol{p_k} = \mathbf{1} - |\boldsymbol{d_k}| \qquad (8)$$

The Pearson distance can be interpreted as follows: smaller value of Pearson distance signifies a close distance (close similarity) between the originals signals and the $\boldsymbol{k}$-th $\widetilde{\boldsymbol{IMF}}$, in contrast, the highest value of Pearson distance explains that the two series are totally different.

Besides that, the Pearson distance, $\boldsymbol{p_k}$, in Eq. (7), is also expected to serve as a proper threshold for IMF modes separation. In this work, $\boldsymbol{p_k} \geq \mathbf{0.2}$ for $\boldsymbol{k} = \mathbf{1}, \mathbf{2}, \cdots \boldsymbol{K}_{\mathbf{mur}}$ is applied as a threshold. $\boldsymbol{K}_{\mathbf{mur}}$ is the number of modes that is involved in the construction of the extracted murmurs. The main task of the algorithm is to select series of $\widetilde{\boldsymbol{IMF}}_{\boldsymbol{k}}[\boldsymbol{n}]$, which has Pearson distance larger than $\mathbf{0.2}$ for $\boldsymbol{k} = \mathbf{1}, \mathbf{2}, \cdots \boldsymbol{K}_{\mathbf{mur}}$. Those modes denotes the murmurs signal while the rest of the modes are the fundamental heart sound signal. Therefore, final results are a group of $\widetilde{\boldsymbol{IMF}}_{\boldsymbol{k}}[\boldsymbol{n}]$ series that represents extracted murmurs (mostly are those with low subscription index, i.e. $\boldsymbol{k} = \mathbf{1}, \mathbf{2}, ..$) and on the other hand a group of $\widetilde{\boldsymbol{IMF}}_{\boldsymbol{k}}[\boldsymbol{n}]$ series that constructs fundamental heart sounds signals. Mathematically the two series can be written as follows:

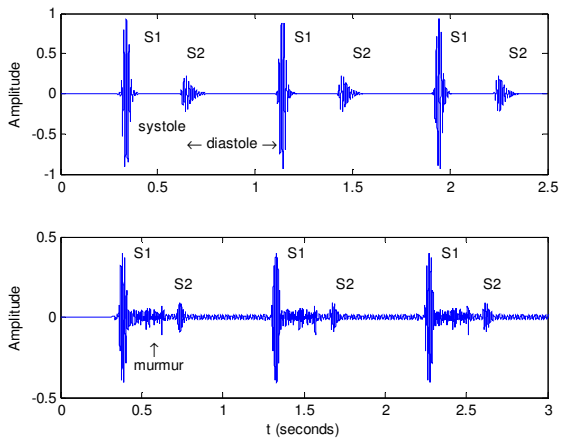

Fig. 1. PCG signal for the normal (above) and abnormal heart sounds with *early systolic murmurs* (bottom).

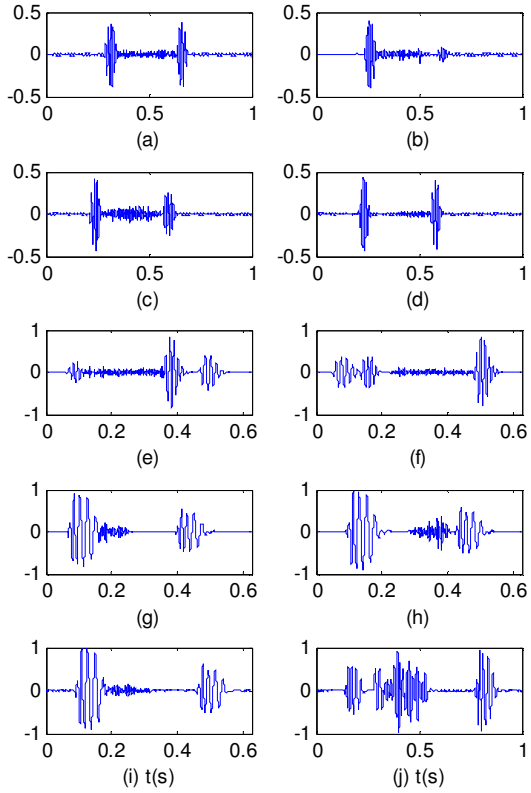

Fig. 2. *Pathological* heart sounds and murmurs. (a)-(f) University of Michigan database, (g)-(j) University of Washington database.

$$\boldsymbol{Y}_{\mathbf{mur}}[\boldsymbol{n}] = \sum_{k=1}^{K_{\mathbf{mur}}} \widetilde{\boldsymbol{IMF}}_k[\boldsymbol{n}] \qquad (8)$$

and

$$\boldsymbol{Y}_{\mathbf{FS}}[\boldsymbol{n}] = \sum_{k=K_{\mathbf{mur}}+1}^{K} \widetilde{\boldsymbol{IMF}}_k[\boldsymbol{n}] \qquad (9)$$

where $\boldsymbol{Y}_{\mathbf{mur}}$ signifies the extracted murmurs and $\boldsymbol{Y}_{\mathbf{FS}}$ denotes the separated fundamental heart sound signal. The variable $\boldsymbol{K}_{\mathbf{mur}}$ is selected in such a way that threshold criterion described above is fulfilled, i.e. $\boldsymbol{p}_k \geq \mathbf{0.2}$.

IV. Performance Evaluation

In order to examine performance of our proposed scheme in extracting murmurs throughout the heart sound signals, we will apply the algorithms to 10 various heart sound signals that are perturbed by murmurs. The signals were taken from University of Michigan Department of Medicine and University of Washington Department of Medicine databases. Both are freely available on the Internet. The pathological heart sounds and murmurs are depicted in Fig. 2. Fig. 2 shows various pathological heart sound signals with fundamental components S1 and S2 as well as murmurs signal.

In this study, performance of the algorithms are evaluated in terms of difference in variance of the targeted signal before and after murmurs extraction. This performance metric is denoted by $\boldsymbol{\Delta SNR}$ and is defined as [18]:

$$\boldsymbol{\Delta SNR} = \mathbf{10}\log\left(\frac{\sigma^2_{\boldsymbol{e}_{\mathbf{before}}}}{\sigma^2_{\boldsymbol{e}_{\mathbf{after}}}}\right). \qquad (10)$$

where $\sigma^2_{\boldsymbol{e}_{\mathbf{before}}}$ is the variance of the heart sound signals before extraction, and $\sigma^2_{\boldsymbol{e}_{\mathbf{after}}}$ is the variance of the heart sound signals after removal of the murmurs.

For both decomposition algorithms, the CEEMD and the EEMD, we employ an ensemble size of $\boldsymbol{I} = \mathbf{100}$ realizations of zero mean unit variance white Gaussian noise series with increment noise amplitude from $\boldsymbol{\beta} = \mathbf{0.05}$ to $\mathbf{0.8}$ with a $\mathbf{0.05}$ step.

The $\boldsymbol{\Delta SNR}$s as a function of $\boldsymbol{\beta}$ for each of heart sound signals are shown in Table 1. It is apparent that low $\boldsymbol{\Delta SNR}$s can be achieved by setting the noise amplitude, $\boldsymbol{\beta} = \mathbf{0.5}$ for signal decomposition using the CEEMD algorithm. Similarly, for signal decomposition using the EEMD algorithm, small $\boldsymbol{\Delta SNR}$s can be attained by setting $\boldsymbol{\beta} = \mathbf{0.1}$.

Let us now examine the usefulness of the $\boldsymbol{\Delta SNR}$ parameter in Eq. (10) for measuring performance of murmurs extraction method in this paper. Small $\boldsymbol{\Delta SNR}$ means close similarity between original signal and the extracted signal. For example, an extreme condition is $\boldsymbol{\Delta SNR} = \mathbf{0}$, in this case it can be interpreted that the targeted heart sound signal is exactly the same as the extracted signal, hence in this case it is said that the algorithm fails to extract murmurs. Comparably, for $\boldsymbol{\Delta SNR} \gg \mathbf{0}$, it signifies huge disparity between the $\sigma^2_{\boldsymbol{e}_{\mathbf{before}}}$ and the $\sigma^2_{\boldsymbol{e}_{\mathbf{after}}}$ implying the decomposition algorithms cause deformation of the original signal.

TABLE 1. MURMURS EXTRACTION FOR VARIOUS HEART SOUND SIGNALS IN TERMS OF $\boldsymbol{\Delta SNR}$ AND NOISE AMPLITUDES, $\boldsymbol{\beta}$.

Heart sound signal	ΔSNR (dB)	
	CEEMD	EEMD
Holo-systolic murmurs, $\boldsymbol{\beta} = \mathbf{0.5}$	0.0876	0.0455
Early-systolic murmurs, $\boldsymbol{\beta} = \mathbf{0.5}$	0.0725	0.0091
Mid-systolic murmurs, $\boldsymbol{\beta} = \mathbf{0.5}$	0.2043	0.1446
Late-systolic murmurs, $\boldsymbol{\beta} = \mathbf{0.5}$	0.0497	0.0245
S3 and Holo-systolic murmurs, $\boldsymbol{\beta} = \mathbf{0.5}$	0.274	0.1227
S4 and Mid-systolic murmurs, $\boldsymbol{\beta} = \mathbf{0.5}$	0.088	0.0506
Early aortic stenosis murmurs, $\boldsymbol{\beta} = \mathbf{0.1}$	0.2718	3.4387
Late aortic stenosis murmurs, $\boldsymbol{\beta} = \mathbf{0.1}$	0.2521	3.1271
Benign murmurs, $\boldsymbol{\beta} = \mathbf{0.1}$	0.2691	2.5332
Mitral Stenosis, $\boldsymbol{\beta} = \mathbf{0.1}$	3.9827	2.7106

Therefore, both of those extreme cases are something that we have to avoid in the heart sound signal analysis. Our study showed that perfect murmurs extraction can be achieved by keeping $\boldsymbol{\Delta SNR}$ in the range of $\mathbf{0.04dB} < \boldsymbol{\Delta SNR} < \mathbf{0.4dB}$ using both decomposition algorithms. It should be noted that in the designated range, the bigger value of $\boldsymbol{\Delta SNR}$ is considered the better. It indicates that variance of the separated fundamental heart sound signal after murmurs extraction is smaller than that of signal with small $\boldsymbol{\Delta SNR}$. In other words, the signal contains less murmurs than the other.

As a first example, we will decompose the *Mid-systolic murmurs* signal from University of Michigan database using the CEEMD for $\boldsymbol{\beta} = \mathbf{0.5}$. It can be appreciated from Fig. 3 that decomposition using the CEEMD produces 14 levels of IMFs in which the first until the fifth modes represent higher frequency of the original signal. It clearly seen that lower frequency of the signal is enclosed in the sixth to the fourteenth modes. Concerning murmurs' characteristics described above, we can interpret intuitively that the murmurs might be resided from the first to the fifth modes. This conclusion is confirmed by measuring the Pearson distance defined as in Eq. (7) and is plotted in the second row plot of Fig. 4.

The separated fundamental heart sound signal as well as the extracted murmurs can be seen clearly in the third and the fourth row plot of Fig. 4. Based on the figure, it is clearly understood that the CEEMD algorithm combined with the Pearson distance method completely remove murmurs from the heart sounds signal.

In the same way, a *Mid-systolic murmurs* signal decomposition and extraction was carried out using the EEMD algorithm. The noise amplitude, $\boldsymbol{\beta}$, was set to 0.5 the same value that we used for the CEEMD algorithm. Decomposition of the *Mid-systolic murmurs* heart sound signal results in 14 levels of modes (the picture does not include in this paper). However, we

can see from the second row plot of Fig. 5 that the Pearson distance decreases to a value that is smaller than $\boldsymbol{p_k = 0.2}$ at the fourth mode ($\boldsymbol{p_4 = 0.081}$), implying that murmurs might be conceived by the first, the second and the third modes. Hence, $\boldsymbol{K_{mur} = 3}$. According to Eq. (8), the outcome of summing up the first three levels of modes is the extracted murmurs, $\boldsymbol{Y_{mur}[n]}$, showed in the last row plot of Fig. 5. On the other hand, aggregating the fourth until the fourteenth row plot of the figure gives us the separated fundamental heart sound signal, $\boldsymbol{Y_{FS}[n]}$, as seen in the third row plot of Fig. 5.

In our simulation, combination of the EEMD algorithm for signal decomposition and Pearson distance method for murmurs extraction gives $\boldsymbol{\Delta SNR = 0.145}$dB, lower than that of the $\boldsymbol{\Delta SNR}$ provided by the CEEMD algorithm. In this case, lower $\boldsymbol{\Delta SNR}$ implies some murmurs cannot be removed perfectly by the EEMD algorithm. See the third row of Fig. 6 and compare with the third row of Fig. 5. It is obvious that the separated fundamental heart sound signal in Fig. 5 is smeared by the murmurs. Using the two Fig.s, we can observe visually that the CEEMD algorithm outperforms the EEMD algorithm.

As a second example, we present murmurs extraction of the *Early Aortic Stenosis* signal from the University of Washington database. Signal decomposition is done by employing both the CEEMD and the EEMD algorithms with noise amplitude, $\boldsymbol{\beta = 0.1}$. After finalizing iterations, the CEEMD decomposes the targeted signal into 14 levels of modes, while the EEMD decomposes it in 13 levels of modes. As a result of reconstruction process utilizing the Pearson distance criteria, signal that was decomposed by CEEMD algorithm produces $\boldsymbol{\Delta SNR = 0.272}$dB.

In Fig. 6, evolution of the Pearson distance against the $\boldsymbol{k}$-th $\boldsymbol{\widehat{IMF}}$ shows the threshold of the Pearson distance that we used i.e., $\boldsymbol{p_k \geq 0.2}$ is surpassed for the first time at the sixth mode, $\boldsymbol{p_6 = 0.065}$. Hence, $\boldsymbol{K_{mur} = 5}$. This means to say that the sixth mode is not the murmurs signal. Therefore, we can observe that the murmurs can be reproduced by adding the first until the fifth modes of the decomposed signals resulting in $\boldsymbol{Y_{mur}[n]}$ as it was defined in Eq. (8), meanwhile, summing up together the sixth until the fourteenth mode resulting in fundamental heart sound signal, $\boldsymbol{Y_{FS}[n]}$, as it was defined in Eq. (9). Utilization of the CEEMD algorithm for signal decomposition and the Pearson distance method for murmurs extraction can achieve $\boldsymbol{\Delta SNR = 0.204}$.

As a summary, pathological heart sound signals that have been decomposed by the CEEMD algorithm provide the finest murmurs extraction than the EEMD. The only peculiarity is occurred at extraction of the *Mitral Stenosis* heart sound signal where for both decomposition schemes cannot produce proper murmurs removal. This is mainly because amplitude of the murmurs signal is as high as or even higher than the S1 and S2 components of fundamental heart sound signal. See Fig. 2(j). In this case, the algorithms cannot differentiate between the murmurs and the S1 and S2 components properly.

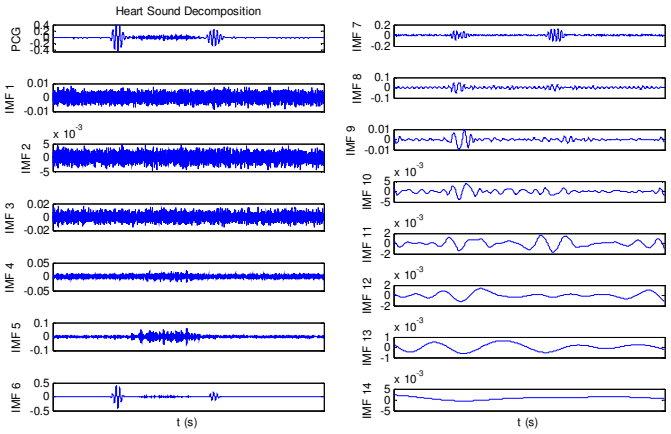

Fig. 3. Decomposition of the *Mid-systolic murmurs* signal using the CEEMD algorithm for $\boldsymbol{\beta = 0.5}$.

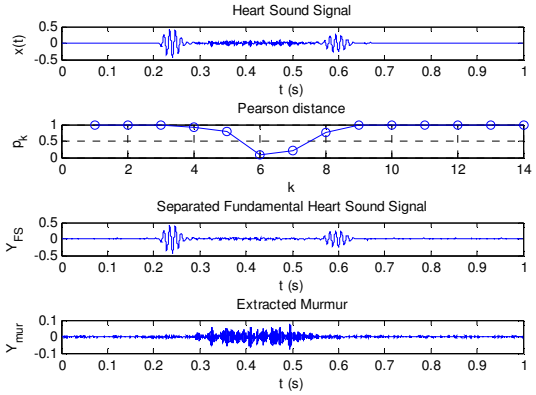

Fig. 4. *Mid-systolic murmurs* signal, Pearson distance, separated fundamental heart sound signal and extracted murmurs from the CEEMD for $\boldsymbol{\beta = 0.5}$, $\boldsymbol{\Delta SNR = 0.204}$dB.

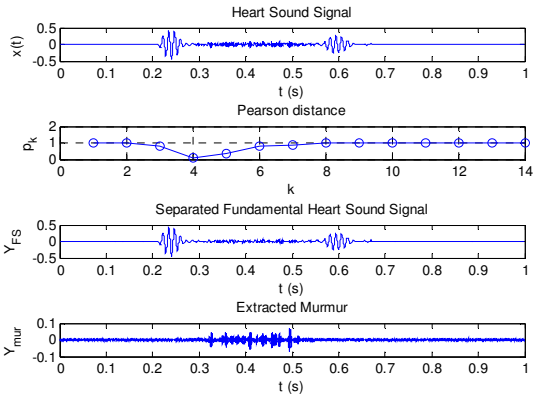

Fig. 5. *Mid-systolic murmurs* signal, Pearson distance, separated fundamental heart sound signal and *extracted* murmur from the EEMD for $\boldsymbol{\beta = 0.5}$, $\boldsymbol{\Delta SNR = 0.145}$dB.

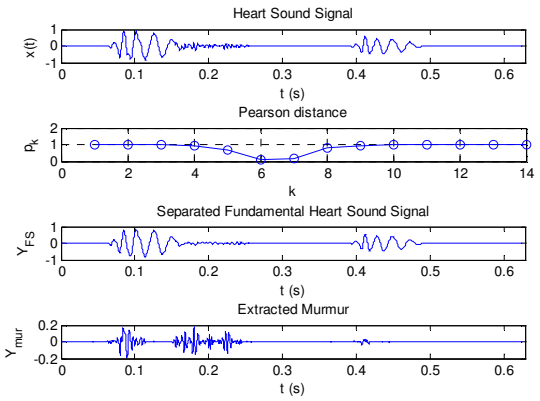

Fig. 6. *Early Aortic Stenosis murmurs* signal, Pearson distance, separated fundamental heart *sound* signal and extracted murmurs from the CEEMD for $\boldsymbol{\beta = 0.1}$, $\boldsymbol{\Delta SNR = 0.271}$dB.

Besides the $\boldsymbol{\Delta SNR}$, particular attention should also be put on the number of modes involved in the construction of the extracted murmurs. Our studies show that for the same value of $\boldsymbol{\beta}$, the decomposition using the EEMD algorithm involves smaller number of modes to form the extracted murmurs signal as compared to the CEEMD algorithm. In other words, $\boldsymbol{K_{mur}}$ of the EEMD decomposition is smaller than the $\boldsymbol{K_{mur}}$ that is generated from the CEEMD decomposition. This conclusion is applied to all targeted heart sound signals in our experiment. This is an indication that the EEMD can only decompose high frequency part of the heart sound signals into smaller number of modes, while the CEEMD efficiently decompose the high frequency component into higher number of modes resulting in more detail and precise construction of the murmurs signals.

The last observation considers the number of total shifting procedures generated by the two algorithms in the study. It can be shown that for various heart sound signals and noise amplitudes, the CEEMD algorithm produces higher number of shifting procedures than the EEMD algorithm. Higher number of shifting procedures directly implies the number of computation needed by the algorithm. This can be considered as the only drawback in employing the CEEMD algorithm for the purpose of murmurs extraction. However, in the case where quality of murmurs extraction becomes a priority and computation complexity can be sacrificed (considering the advance of computation speed today), the CEEMD algorithm will be a reasonable choice for signal decomposition and murmurs extraction from the pathological heart sound signals.

V. CONCLUSIONS

In this study, murmurs extraction from various heart sound signals has been carried out. The murmurs extraction procedure can be partitioned into two main processes, i.e., signal decomposition and murmurs extraction. It can be concluded that in general pathological heart sound signals that have been decomposed by the CEEMD algorithm followed by the Pearson distance metric for murmurs extraction, provide the finest murmurs extraction than the EEMD. Performance evaluation was done by employing the $\boldsymbol{\Delta SNR}$ metric. Additionally, the EEMD algorithm involves smaller number of modes to form the extracted murmurs signal as compared to the CEEMD algorithm. As a result, the CEEMD algorithm was able to decompose the high frequency component into higher number of modes resulting in more detail and precise construction of the murmurs signals. However, employing the CEEMD algorithm produces higher number of shifting procedures causing higher computational complexity than the EEMD algorithm. Nevertheless, considering the advance of computation speed today, the CEEMD algorithm will be a reasonable choice for signal decomposition and murmurs extraction.

REFERENCES

[1] WHO, *Global Status Report on Noncommunicable diseases (NCD) 2014*, Switzerland, 2014.

[2] WHO, *Health in Asia and the Pacific, World Health Organization-Western Pacific & South-East Asia*, 2008.

[3] F. Hu, S. Lakdawala, Q. Hao and M. Qiu, "Low-Power, Intelligent Sensor Hardware Interface for Medical Data Pre-Processing," *IEEE Transaction on Information Technology in Biomedicine*, Vol. 13, No. 4, pp. 656-663, May 2009.

[4] J. Jusak and I. Puspasari, "Wireless Tele-Auscultation for Phonocardiograph Signal Recording Through the Zigbee Networks," *Proceeding of IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, Bandung, Indonesia, August 27-29, 2015.

[5] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. Ch. Yen, C. C. Tung and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," in *Proc. of the Royal Society of London,* vol. 454, pp. 903–995, 1998.

[6] T.H. Hung, C.C. Chou, W.C. Fang, A.H.T Li, Y.C. Chang, B.K Hwang and Y.W. Shau, "Time Frequency Analysis of Heart Sound Signals Based on Hilbert Huang Transformation," *Proceeding of IEEE International Symposium on Consumer Electronics (ISCE) 2012,* Harrisburg, PA, 4-6 June 2012.

[7] L. Helong, S. Kwong, L. Yang, D. Huang and D. Xiao, "Hilbert-Huang Transform for Analysis of Heart Rate Variability in Cardiac Health," IEEE/ACM Transaction on Computational Biology and Bioinformatics, Vol. 8, No. 6, 2011, pp. 1557-1567.

[8] D. Boutana, M. Benidir and B. Barkat, "Segmentation and Time-Frequency Analysis of Phatological Heart Sound Signal using the EMD Method," *Proceeding of 22nd European Signal Processing Conference (EUSIPCO)*, Lisbon, 1-5 Sept 2014.

[9] Z. Wu and N.E. Huang, "Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method," *Advances in Adaptive Data Analysis, Vol. 1, No. 1*, January 2009, pp. 1-41.

[10] Flandrin, P., G. Rilling and P. Gonçalvès, Empirical mode decomposition as a filter bank. *IEEE Signal Processing Lett.*, Vol. 11, 2004, pp. 112-114.

[11] R.K. Niazy, C.F. Beckmann, J.M. Brady and S.M. Smith, "Performance Evaluation of the Ensemble Empirical Mode Decomposition," *Advances in Adaptive Data Analysis, Vol. 1, No.2*, April 2009, pp. 231-242.

[12] M.E. Torres, M.A. Colominas, G. Schlotthauer and P. Flandrin, "A Complete Ensemble Empirical Mode Decomposition with Adaptive Noise," *IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP) 2011*, Prague, 22-27 May 2011.

[13] M.A. Colominas, G. Schlotthauer, M.E. Torres and P. Flandrin, "Noise-Assisted EMD Methods in Action," *Advances in Adaptive Data Analysis*, Vol. 4, No. 4, 2012.

[14] J. S. Coviello, *Auscultation Skills: Breath & Heart Sounds 5th Ed.*, Philadelphia: Lippincott Williams & Wilkins, 2014.

[15] J.A. Jimenez, M.A. Becerra and E.D. Trejos, "Heart Murmur Detection using Ensemble Empirical Mode Decomposition and Derivations of the Mel-Frequency Cepstral Coefficients on 4-Area Phonocardiographic Signals," *Computing in Cardiology Conference (CiCC) 2014*, Cambridge, Sept. 2014, pp. 493-496.