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

by Jusak Jusak

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. [6] to perform a time-domain decomposition of an observed signal, $\boldsymbol{x}(\boldsymbol{t})$, into a set of Intrinsic Mode Function (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 $\boldsymbol{x}(\boldsymbol{t})$ is called sifting process. As a result of decomposition process, an observed signal can be expressed as:

$$\boldsymbol{x}(\boldsymbol{t}) = \sum_{k=1}^{K} \boldsymbol{IMF}_k(\boldsymbol{t}) + \boldsymbol{R}_K(\boldsymbol{t}) \qquad (1)$$

where $\boldsymbol{IMF}_k(\boldsymbol{t})$ is the $\boldsymbol{k}$-th IMF, and $\boldsymbol{R}_K(\boldsymbol{t})$ 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 $\boldsymbol{x}[\boldsymbol{n}](\boldsymbol{n} = \mathbf{1}, \mathbf{2}, \cdots, \boldsymbol{N})$ 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, $\boldsymbol{x}[\boldsymbol{n}]$ such that

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

   The $\boldsymbol{w}^i[\boldsymbol{n}](\boldsymbol{i} = \mathbf{1}, \mathbf{2}, \cdots, \boldsymbol{I})$ 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 $\boldsymbol{x}^i[\boldsymbol{n}]$ using the EMD algorithm to yield the $\boldsymbol{IMF}_k^i[\boldsymbol{n}]$, where $\boldsymbol{k} = \mathbf{1}, \mathbf{2}, \cdots, \boldsymbol{K}$ is the $\boldsymbol{k}$-th mode and each residue is obtained as

$$\boldsymbol{R}_k^i[\boldsymbol{n}] = \boldsymbol{R}_{k-1}^i[\boldsymbol{n}] - \boldsymbol{IMF}_k^i[\boldsymbol{n}]. \qquad (3)$$

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

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

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 $\boldsymbol{x}^i[\boldsymbol{n}]$ 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

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

where $\widetilde{\boldsymbol{IMF}}_1[\boldsymbol{n}]$ 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 $\boldsymbol{R}_1[\boldsymbol{n}]$ 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 $\widetilde{\boldsymbol{IMF}}_2[\boldsymbol{n}]$ can be obtained. Then calculate the $\boldsymbol{k}$-th residue for $\boldsymbol{k} = \mathbf{2}, \mathbf{3}, \cdots, \boldsymbol{K}$, decompose the $\boldsymbol{R}_k$ and define the mode $\widetilde{\boldsymbol{IMF}}_{k+1}[\boldsymbol{n}]$. 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 $\boldsymbol{N}_{\text{shift}}$ 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

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

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. $\widetilde{IMF}_1[n]$, $\widetilde{IMF}_2[n]$, $\widetilde{IMF}_3[n]$, 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

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

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

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

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 $\widetilde{IMF}$, 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, $\boldsymbol{p_k \geq 0.2}$ for $\boldsymbol{k = 1, 2, \cdots K_{mur}}$ is applied as a threshold. $\boldsymbol{K_{mur}}$ 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 $\widetilde{IMF}_k[n]$, which has Pearson distance larger than $\boldsymbol{0.2}$ for $\boldsymbol{k = 1, 2, \cdots K_{mur}}$. 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 $\widetilde{IMF}_k[n]$ series that represents extracted murmurs (mostly are those with low subscription index, i.e. $\boldsymbol{k = 1, 2, ..}$) and on the other hand a group of $\widetilde{IMF}_k[n]$ series that constructs fundamental heart sounds signals. Mathematically the two series can be written as follows:

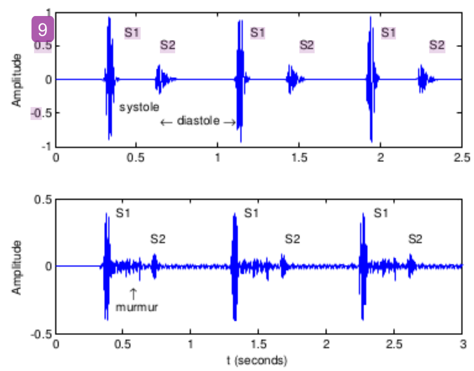

Fig. 1. PCG signal for the normal (above) and abnormal heart sounds with *early systolic murmurs* (bottom).

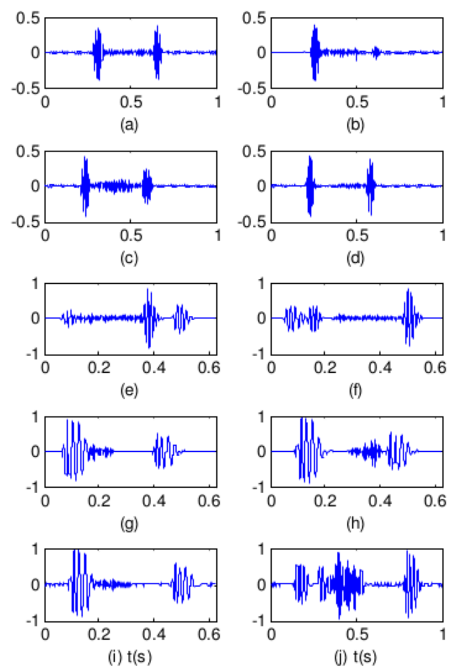

Fig. 2. *Pathological* heart sounds and murmurs. (a)-(f) University of Michigan database, (g)-(j) University of Washington database.

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

and

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

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. $\boldsymbol{p}_{\boldsymbol{k}} \geq \mathbf{0.2}$.

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 $\boldsymbol{\Delta}\mathbf{SNR}$ and is defined as [18]:

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

where $\boldsymbol{\sigma}^2_{e_{\mathbf{before}}}$ is the variance of the heart sound signals before extraction, and $\boldsymbol{\sigma}^2_{e_{\mathbf{after}}}$ 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 $\boldsymbol{\Delta}\mathbf{SNR}$s as a function of $\boldsymbol{\beta}$ for each of heart sound signals are shown in Table 1. It is apparent that low $\boldsymbol{\Delta}\mathbf{SNR}$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 $\boldsymbol{\Delta}\mathbf{SNR}$s can be attained by setting $\boldsymbol{\beta} = \mathbf{0.1}$.

Let us now examine the usefulness of the $\boldsymbol{\Delta}\mathbf{SNR}$ parameter in Eq. (10) for measuring performance of murmurs extraction method in this paper. Small $\boldsymbol{\Delta}\mathbf{SNR}$ means close similarity between original signal and the extracted signal. For example, an extreme condition is $\boldsymbol{\Delta}\mathbf{SNR} = \mathbf{0}$, 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 $\boldsymbol{\Delta}\mathbf{SNR} \gg \mathbf{0}$, it signifies huge disparity between the $\boldsymbol{\sigma}^2_{e_{\mathbf{before}}}$ and the $\boldsymbol{\sigma}^2_{e_{\mathbf{after}}}$ implying the decomposition algorithms cause deformation of the original signal.

TABLE 1. MURMURS EXTRACTION FOR VARIOUS HEART SOUND SIGNALS IN TERMS OF $\boldsymbol{\Delta}\mathbf{SNR}$ 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 $\boldsymbol{\Delta}\mathbf{SNR}$ in the range of $\mathbf{0.04dB} < \boldsymbol{\Delta}\mathbf{SNR} < \mathbf{0.4dB}$ using both decomposition algorithms. It should be noted that in the designated range, the bigger value of $\boldsymbol{\Delta}\mathbf{SNR}$ 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 $\boldsymbol{\Delta}\mathbf{SNR}$. 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, $\boldsymbol{K_{\mathrm{mur}} = 3}$. According to Eq. (8), the outcome of summing up the first three levels of modes is the extracted murmurs, $\boldsymbol{Y_{\mathrm{mur}}[n]}$, 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, $\boldsymbol{Y_{\mathrm{FS}}[n]}$, 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 $\boldsymbol{\Delta\mathrm{SNR} = 0.145}$dB, lower than that of the $\boldsymbol{\Delta\mathrm{SNR}}$ provided by the CEEMD algorithm. In this case, lower $\boldsymbol{\Delta\mathrm{SNR}}$ 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 $\boldsymbol{\Delta\mathrm{SNR} = 0.272}$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, $\boldsymbol{K_{\mathrm{mur}} = 5}$. 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 $\boldsymbol{Y_{\mathrm{mur}}[n]}$ as it was defined in Eq. (8), meanwhile, summing up together the sixth until the fourteenth mode resulting in fundamental heart sound signal, $\boldsymbol{Y_{\mathrm{FS}}[n]}$, 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 $\boldsymbol{\Delta\mathrm{SNR} = 0.204}$.

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.

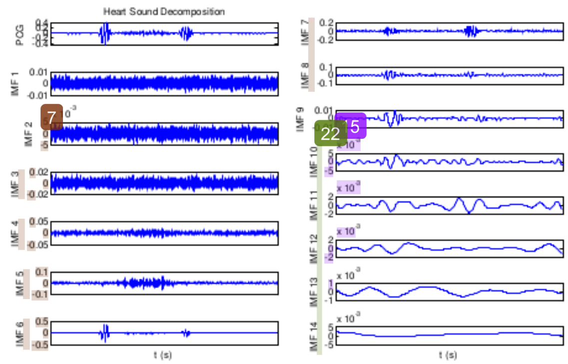

Fig. 3. Decomposition of the *Mid-systolic murmurs* signal using the CEEMD algorithm for $\boldsymbol{\beta = 0.5}$.

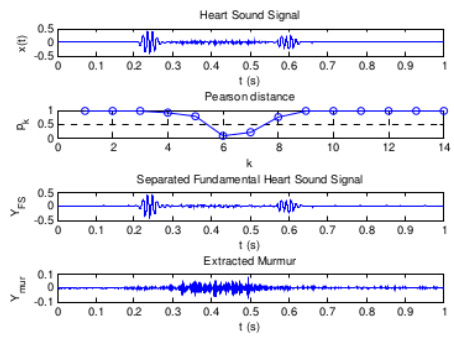

Fig. 4. *Mid-systolic murmurs* signal, Pearson distance, separated fundamental heart sound signal and extracted murmurs from the CEEMD for $\boldsymbol{\beta = 0.5}$, $\boldsymbol{\Delta\mathrm{SNR} = 0.204}$dB.

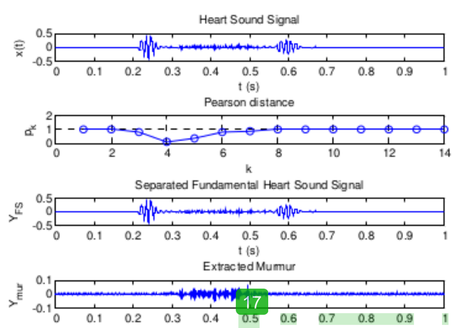

Fig. 5. *Mid-systolic murmurs* signal, Pearson distance, separated fundamental heart sound signal and *extracted* murmur from the EEMD for $\boldsymbol{\beta = 0.5}$, $\boldsymbol{\Delta\mathrm{SNR} = 0.145}$dB.

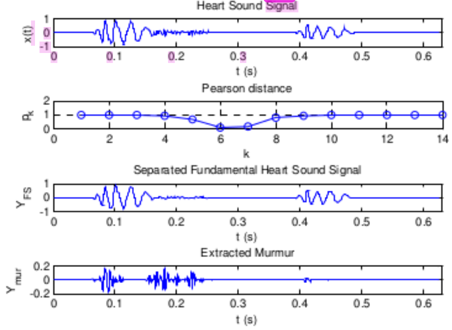

Fig. 6. *Early Aortic Stenosis murmurs* signal, Pearson distance, separated fundamental heart *sound* signal and extracted murmurs from the CEEMD for $\boldsymbol{\beta} = \mathbf{0.1}$, ΔSNR $= \mathbf{0.271}$dB.

Besides the Δ**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 in 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 Δ**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.

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

ORIGINALITY REPORT

17%	10%	15%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Jusak Jusak, Heri Pratikno, Vergie Hadiana Putra. "Internet of Medical Things for cardiac monitoring: Paving the way to 5G mobile networks", 2016 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT), 2016 Publication	4%
2	www.semanticscholar.org Internet Source	1%
3	strathprints.strath.ac.uk Internet Source	1%
4	Ira Puspasari, Weny Indah Kusumawati, Eka Sari Oktarina, Jusak Jusak. "A New Heart Sound Signal Identification Approach Suitable for Smart Healthcare Systems", 2019 2nd International Conference on Applied Engineering (ICAE), 2019 Publication	1%
5	link.springer.com	

Internet Source

1%

6 Kedadouche, M., M. Thomas, and A. Tahan. "A comparative study between Empirical Wavelet Transforms and Empirical Mode Decomposition Methods: Application to bearing defect diagnosis", Mechanical Systems and Signal Processing, 2016.
Publication

1%

7 Sankaran, Adarsh. "Unveiling the multiscale teleconnection between Pacific Decadal Oscillation and global surface temperature using time-dependent intrinsic correlation analysis : MULTISCALE TELECONNECTION BETWEEN PDO AND GLOBAL TEMPERATURE", International Journal of Climatology, 2016.
Publication

1%

8 MARCELO A. COLOMINAS, GASTÓN SCHLOTTHAUER, MARÍA E. TORRES, PATRICK FLANDRIN. "NOISE-ASSISTED EMD METHODS IN ACTION", Advances in Adaptive Data Analysis, 2013
Publication

<1%

9 www.cinc.org
Internet Source

<1%

10 worldwidescience.org
Internet Source

<1%

11	Ozgoren, M.. "Flow structure in the downstream of square and circular cylinders", Flow Measurement and Instrumentation, 200608 Publication	<1%
12	ZHAOHUA WU, NORDEN E. HUANG. "ENSEMBLE EMPIRICAL MODE DECOMPOSITION: A NOISE-ASSISTED DATA ANALYSIS METHOD", Advances in Adaptive Data Analysis, 2011 Publication	<1%
13	journals.plos.org Internet Source	<1%
14	Karin Schiecke, Matthias Wacker, Franz Benninger, Martha Feucht, Lutz Leistritz, Herbert Witte. "Advantages of signal-adaptive approaches for the nonlinear, time-variant analysis of heart rate variability of children with temporal lobe epilepsy", 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2014 Publication	<1%
15	F. Sabbaghian-Bidgoli, J. Poshtan. "Fault Detection of Broken Rotor Bar Using an Improved form of Hilbert–Huang Transform", Fluctuation and Noise Letters, 2018 Publication	<1%

16	www.academypublisher.com Internet Source	<1%
17	Submitted to National University of Singapore Student Paper	<1%
18	M. Sheraz Ahmad, Junaid Mir, Muhammad Obaid Ullah, Muhammad Laiq Ur Rahman Shahid, Muhammad Adnan Syed. "An efficient heart murmur recognition and cardiovascular disorders classification system", Australasian Physical & Engineering Sciences in Medicine, 2019 Publication	<1%
19	P.F. Hsieh. "Empirical mode decomposition for dimensionality reduction of hyperspectral data", Proceedings 2005 IEEE International Geoscience and Remote Sensing Symposium 2005 IGARSS 05, 2005 Publication	<1%
20	Submitted to University of Auckland Student Paper	<1%
21	www.etis.ee Internet Source	<1%
22	Xuan Song, Shuxiang Guo, Baofeng Gao, Zhenyu Wang. "Motion recognition of the bilateral upper-limb rehabilitation using sEMG based on ensemble EMD", 2014 IEEE	<1%

International Conference on Mechatronics and Automation, 2014
Publication

---

23 "Web, Artificial Intelligence and Network Applications", Springer Science and Business Media LLC, 2020 <1%
Publication

---

24 scholars.cityu.edu.hk <1%
Internet Source

---

25 repository.maranatha.edu <1%
Internet Source

---

26 Bhadhan Roy Joy, Amy Amara, Arie Nakhmani. "Transform with no Parameters Based on Extrema Points for Non-stationary Signal Analysis", Circuits, Systems, and Signal Processing, 2017 <1%
Publication

---

27 "Machine Learning, Image Processing, Network Security and Data Sciences", Springer Science and Business Media LLC, 2020 <1%
Publication

---

28 "Intelligent Data Engineering and Automated Learning – IDEAL 2018", Springer Science and Business Media LLC, 2018 <1%
Publication

---

29 repository.dinamika.ac.id
Internet Source

<1%

30 www.mdpi.com
Internet Source
<1%

31 www.ripublication.com
Internet Source
<1%

32 Daoud Boutana, Messaoud Benidir, Braham Barkat. "On the selection of Intrinsic Mode Function in EMD method: Application on heart sound signal", 2010 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2010), 2010
Publication
<1%

33 cwww.intechopen.com
Internet Source
<1%

34 "Optical and Wireless Convergence for 5G Networks", Wiley, 2019
Publication
<1%

35 Zhiqiang Liao, Haihong Tang, Shuuji Miyazaki, Daisuke Matsuo, Yuji Yonekura, Ho Jinyama. "An automatic bearing fault signal filtering and diagnosis method based on empirical mode decomposition and C0 complexity", The Proceedings of the Symposium on Evaluation and Diagnosis, 2018
<1%

Publication

36	L. N. Sharma. "Teager-Kaiser energy operator in Hilbert space for heart sound segmentation", 2016 International Conference on Biomedical Engineering (BME-HUST), 2016 Publication	<1%
37	open.library.ubc.ca Internet Source	<1%
38	teses.usp.br Internet Source	<1%
39	Tary, Jean Baptiste, Roberto Henry Herrera, Jiajun Han, and Mirko van der Baan. "Spectral estimation-What is new? What is next?", Reviews of Geophysics, 2014. Publication	<1%
40	Myung K. Park. "Physical Examination", Elsevier BV, 2008 Publication	<1%
41	Gabriel Rilling. "", IEEE Transactions on Signal Processing, 1/2008 Publication	<1%
42	"Proceedings of Second International Conference on Electrical Systems, Technology and Information 2015 (ICESTI 2015)", Springer Science and Business Media LLC, 2016 Publication	<1%

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		