

A Robust Algorithm for Segmentation of Heart Signal Sound using Tunable Quality Wavelet Transform

As observed in the previous chapter, the proposed motion noise cancellation method and DWT based denoising algorithm suppresses the external noises efficiently and also murmurs in most of the cases. However, the need of improvisation was observed in the DWT based method for few pathological cases where the murmur overlaps the FHS significantly. This shortcoming of the method was due to the limitation of wavelet, in which the Q-factor (ratio of centre frequency to bandwidth) cannot be tuned according to the oscillatory behaviour of the FHS. To address the issue, in this chapter, we present a robust algorithm for the segmentation of PCG signal using Tunable Quality Wavelet Transform (TQWT), which provides the ability to tune the Q-factor and, hence, adequate separation of FHS and murmur can be achieved.

5.1 INTRODUCTION

As discussed in the previous Chapter, the primary task in the heart sound signal analysis is its segmentation [Moukadem et al., 2013]. However, segmentation is a challenging task in view of the following issues [Kao and Wei, 2011; Patidar and Pachori, 2013].

- Variable length of the cardiac cycle, induced due to heart rate variability and due to arrhythmia [Oliveira et al., 2014].
- Variation in the amplitude of the components from beat-to-beat [Jiang and Choi, 2006].
- Contamination of noise which may occur due to motion or speaking of the subject and environmental noises. These noise sources exhibit a very broad range of spectral characteristics, duration and loudness.
- In the case of abnormality, the presence of murmur sounds and extra peaks such as S3 and S4 may overlap the FHS in time and frequency, both [Yuenyong et al., 2011]. Moreover, the frequency range of the FHS also varies from one pathological case to another.

As a consequence of the above-mentioned challenges, various segmentation algorithms for the heart sound signal have been proposed in the literature. Autocorrelation based methods have been used to predict the length of the cardiac cycle and then the signal is segmented [Kao and Wei, 2011; Yuenyong et al., 2011]. However, these algorithms suffer in the presence of noise and due to the variation of cardiac cycle duration, which varies due to the heart rate variability and due to arrhythmia. In another approach, a dynamic clustering based method has been proposed to segment the PCG signal using a density function and that makes the method robust to noise [Tang et al., 2012]. The main problem with this method is the requirement of prediction of the cardiac cycle, which is difficult in most cases of pathologies. Therefore, the performance of the method is unsatisfactory in such cases.

The most commonly used approach for the PCG signal segmentation is based on enve-

lope method. Envelope of the PCG signal has been obtained using Shannon entropy [Yadolahi and Moussavi, 2006], NASE [Gavrovska et al., 2014; Ramos et al., 2013], HT [Atbi et al., 2013], Homomorphic filtering [Yuenyong et al., 2011; Gupta et al., 2007], Short-term log energy [Zia et al., 2011], Rényi entropy [Boutana et al., 2011], and Blanket fractal dimension [Paskaš et al., 2014]. In [Jiang and Choi, 2006], a new method has been proposed to obtain the envelope of the PCG signal called as CSCW, which is based on the single-degree-of-freedom model of a spring-mass system. The performance of CSCW, Shannon energy, and Hilbert transform was analysed to obtain the envelope of the PCG signal and found that the CSCW relatively better emphasise the FHS [Choi and Jiang, 2008]. Some researchers obtained the envelope using combinations of multiple time domain features [Ramos et al., 2013; Kumar et al., 2006] and morphological features [Safara et al., 2012]. Viola integral method and short-time Hilbert transform were integrated for this purpose in [Sun et al., 2014]. However, the envelope obtained by time-domain methods gets significantly affected due to the presence of noise and due to the presence of murmur [Ramos et al., 2013].

To address challenges faced in the time-domain analysis, as stated above, researchers first transformed the signal into a domain where S1 and S2 are emphasised [Ahlström, 2008] and then the envelope is obtained. Several choices of the transformation have been presented over the years, such as S-transform [Moukadem et al., 2013], pseudo-affine Wigner-Ville Distribution (WVD) [Gavrovska et al., 2014; Kudriavtsev et al., 2007], STFT [Boutana et al., 2011], and empirical mode decomposition [Bajelani et al., 2013]. The envelope has also been obtained by correlating the Morlet wavelet to the signal [Yuenyong et al., 2011; Rajan et al., 2006]. In another approach, a singular spectrum analysis based method has been proposed to separate murmurs from the FHS [Sanei et al., 2011]. WT has been used extensively to emphasise the FHS by retaining only the low-frequency components related to the FHS [Yuenyong et al., 2011; Kumar et al., 2006; Vaisman et al., 2012; Song et al., 2012; Kumar et al., 2007]. WT based method is able to suppress the out-of-band noise as well as in-band noise from the FHS. However, it provides little ability to tune the quality factor (ratio of centre frequency to bandwidth) of the wavelet according to the oscillations of FHS [Selesnick, 2011]. It imposes limitations on WT based methods in pathological cases where murmur overlaps the FHS severely in the frequency domain.

Recently, Patidar and Pachori [Patidar and Pachori, 2013] proposed a Constrained Tunable Quality Wavelet Transform (C-TQWT) based method to suppress the murmur sounds. TQWT has the ability to tune its parameters including the quality factor of the wavelet, according to oscillations of the FHS. Authors in [Patidar and Pachori, 2013] optimised the parameters of TQWT using the GA such that the kurtosis of the approximation level (last sub-band) is maximised. Such approach provides the adaptability of the parameters according to the signal, however, needs high computational time as the signal is decomposed multiple times at each intermediate generation of GA. kurtosis will be relatively larger for the FHS as compared to the murmur sounds. It is because the distribution of FHS is nearly super-Gaussian, while murmur sounds have Gaussian or sub-Gaussian like distribution [Sanei et al., 2011]. The obtained results for various pathological cases show the effectiveness of the method to segment the PCG signal in most of the cases. However, in some pathological cases where murmurs have sharp peaks, the kurtosis is unable to distinguish between the murmur and FHS, consequentially the segmentation is adversely affected. Moreover, selection of approximation level makes the algorithm prone to real-life noises as these noises are dominant in the low-frequency band [Gradolewski and Redlarski, 2014].

In view of the above-discussed issues with C-TQWT related to the computational cost and vulnerability to real-life noises, we propose a TQWT based improved method. The proposed method first denoises the signal using DWT and then decomposes the denoised signal up to twenty levels using the TQWT. Then a particular level is selected adaptively from the

detailed levels only, instead of selecting the approximation level as in [Patidar and Pachori, 2013]. Due to this approach, the proposed method provides two advantages: 1) it discards the approximation level and hence discards most of the real-life noise components, and 2) the computational time of the algorithm is reduced significantly because the signal decomposition is performed only one time as against multiple times required in [Patidar and Pachori, 2013]. Another contribution of this work is the exploration of a new quality parameter, called as Fano factor, to overcome the issues related to kurtosis being unable to distinguish between the murmur and FHS, as stated above. Fano factor is the ratio between the variance and mean of the signal and provides a measure of relative variability [Fano, 1947]. The variance will be higher for the clean signal at the same time mean value will be relatively low. Thus, the Fano factor will be relatively larger for the signal with FHS as compared to the signal with a murmur. In the case of sharp murmur sounds, due to the presence of extra components, the mean value will be relatively high and consequentially the value of Fano factor will be relatively low. Thus, the Fano factor is expected to select a level with emphasised FHS. To further improve the robustness of the method, the signal of the selected level is thresholded. For this purpose, a threshold value is obtained adaptively based on the statistical parameters of the signal under consideration. A new parameter Med_{75} is introduced, which is 75th percentile value of the sorted absolute value of the coefficient vector. 75th percentile value is selected, instead of 50th percentile value (median), because typically the sum of the time duration of both FHS, S1 and S2, remains less than the 25% of the time duration of cardiac cycle [Singh and Anand, 2007]. After thresholding of the signal, the envelope is obtained using the NASE method [Gavrovska et al., 2014] and peaks are identified using a threshold value, obtained by Otsu's method [Otsu, 1979]. Finally, peak conditioning and identification are performed based on the domain knowledge about the duration of the systole and diastole periods.

The rest of the chapter is organised as follows: Section 5.2 provides the theoretical background of the tunable quality wavelet transform and the proposed method is described in Section 5.3. Results of the proposed method are presented and discussed in Section 5.4. Section 5.5 provides the concluding remarks of the work.

5.2 THEORETICAL BACKGROUND ABOUT TQWT

I.W. Selesnick [Selesnick, 2011] developed a wavelet transform for discrete-time signals for which the Q-factor is easily tunable and denoted it as TQWT. Q-factor of the wavelet is the ratio of its centre frequency to its bandwidth. Tuning of the Q-factor makes the transform tunable according to the oscillatory behaviour of the signal under consideration. The TQWT was implemented using the filter banks consisting of low pass and high pass filters, same as in the case of DWT. However, in TQWT, filters are oversampled and thus have redundancy that provides benefits of near shift invariance and more flexibility to design the required filter bank. TQWT provides facility to parameterize the redundancy factor along with the Q-factor. For continuity, a brief description of the decomposition and reconstruction using the TQWT, as given in [Selesnick, 2011], is given as follows.

5.2.1 Decomposition of the Signal using TQWT

As shown in Figure 5.1, decomposition of the signal is achieved by applying high pass and low pass filters to the signal and then the obtained coefficients are decimated. The process of applying filters and scaling is described in the following steps.

STEP 1: Initialization of scaling parameters, α and β , as follow:

$$\beta = \frac{2}{Q+1} ; \alpha = 1 - \frac{\beta}{r} \quad (5.1)$$

where Q is the quality factor and r is the redundancy factor. Although redundancy is ben-

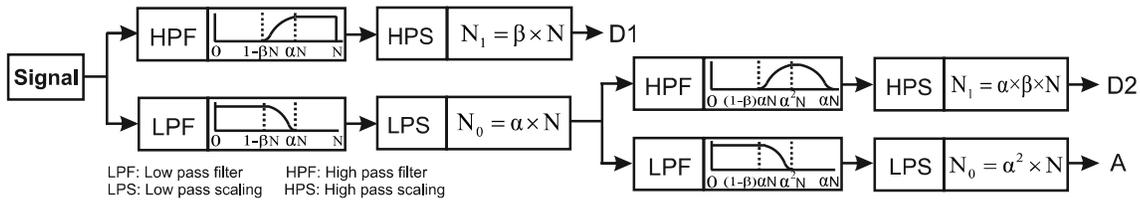


Figure 5.1 : Decomposition of the signal using TQWT

eficial for shift invariance, to avoid high redundancy, the values of the individual scaling parameters i.e. α and β should be less than one and, for the perfect reconstruction, $\alpha + \beta > 1$.

STEP 2: Calculate the unitary Discrete Fourier Transform (uDFT) coefficients of the N -length signal $x(n)$ as follows [Selesnick, 2011]:

$$X(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x(n) \exp\left(-j \frac{2\pi}{N} nk\right) \quad (5.2)$$

for $0 \leq k \leq N - 1$.

STEP 3: Obtain three frequency bands: pass band (P), transition band (T), and stop band (S) for both the low-pass and high-pass filters. As shown in Figure 5.2, pass band of the low-pass filter is $0 - N_1$ and stop band is $N_0 - N$. For the high-pass filter it is the reverse of the low-pass filter, i.e. pass band is $N_0 - N$ and stop band is $0 - N_1$. Transition band, $N_0 - N_1$, is same for both the filters. N_0 and N_1 are obtained as follows:

$$N_0 = \alpha \times N ; N_1 = \beta \times N \quad (5.3)$$

Thus, the frequency band of each level depends on the values of α and β . The sum of the output responses of both the filters in the transition band should be one. Such type of transition band can be constructed using any 2π -periodic power-complementary function [Selesnick, 2011]. In [Selesnick, 2011], I.W. Selesnivk used the Daubechies function with two vanishing moments, expressed as follows.

$$\theta(n) = \frac{1}{2} \left(1 + \cos\left(\frac{n\pi}{N+1}\right) \right) \sqrt{2 - \cos\left(\frac{n\pi}{N+1}\right)} \quad 0 \leq n < N \quad (5.4)$$

where $0 \leq n < N$. The centre frequency (F_c) of the level j can be obtained as:

$$F_c = \frac{1}{4} \alpha^{j-1} (2 - \beta) F_s \quad (5.5)$$

where, F_s is the sampling frequency of the signal. The bandwidth of the level j can be obtained as:

$$BW = \frac{1}{4} \alpha^{j-1} \beta F_s \quad (5.6)$$

STEP 4: The output responses of both the filters, high pass and low pass, are decimated with factor α and β , respectively, as shown in Figure 5.1. A detailed description of the scaling operation can be obtained from [Selesnick, 2011].

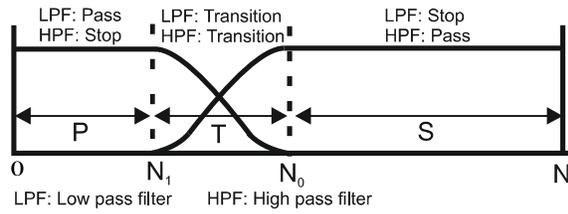


Figure 5.2 : Characteristics of the filters used in TQWT decomposition

5.2.2 Reconstruction using TQWT

To reconstruct the signal, the last level of approximation and detailed coefficients are rescaled by the factor $1/\alpha$ and $1/\beta$, respectively [Selesnick, 2011]. Then the rescaled coefficients are passed through the synthesis filters, low pass filter and high pass filter. The summation of the output responses of these two filters works as the approximation coefficient for the next level reconstruction. In this way, successively, the signal can be reconstructed.

5.3 METHODOLOGY

In the proposed method, the signal is decomposed using TQWT and then particular level with emphasised FHS, based on Fano factor, is selected from detailed levels. The block diagram of the proposed method is shown in Figure 5.3 and various steps are described in the following subsections.

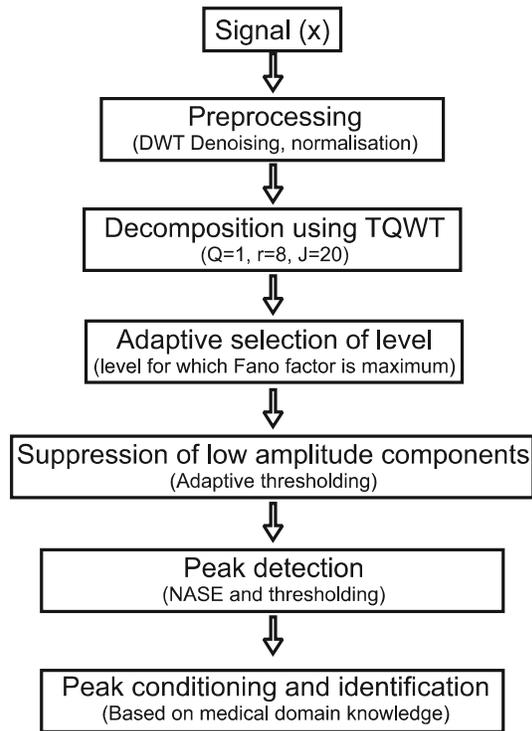


Figure 5.3 : Block diagram of the proposed method

5.3.1 Preprocessing

The signal acquisition systems available, off-the-shelves, generally sample the signal at a much higher rate than the required rate for heart sound signals with its frequency, typ-

ically, below 700 Hz [Gupta et al., 2007]. Therefore, the acquired signal is decimated such that the effective sampling frequency of the signal is set to 1600 Hz, adequately higher than the Nyquist sampling rate. Decimation reduces unnecessary large data processing and hence reduces the execution time of the algorithm. The decimated signal is denoised using DWT based method [Gradolewski and Redlarski, 2014]. According to it, the signal is decomposed up to five levels using the ‘coif-5’ as a mother wavelet. Decomposition of the signal with 1600 Hz sampling frequency will result in five detailed levels (400-800, 200-400, 100-200, 50-100, and 25-50 Hz) and one approximation level (0-25 Hz) [Singh and Anand, 2007]. Thus, the 3rd, 4th, and 5th detailed levels and approximation level cover the frequency range 0-200 Hz and hence cover the frequency range of the FHS, which is approximately 20-120 Hz [Singh and Anand, 2007]. Therefore, denoising is achieved by reconstructing the signal using approximation coefficients and detailed coefficients at 3rd, 4th, and 5th levels. The filtered signal is normalised to suppress the amplitude variation of FHS, as follows:

$$x_{norm}(n) = \frac{x(n)}{\max(|x|)} \quad (5.7)$$

5.3.2 Decomposition using TQWT

As discussed in the previous section, the quality parameter Q and redundancy parameter r control the frequency band of each level. The values of these parameters should be selected such that the murmur sounds get separated out, as much as possible, from the FHS. However, it is a challenging task due to the varying frequency range of the FHS across various pathological cases [Meziani et al., 2012]. In [Patidar and Pachori, 2013], values of the quality and the redundancy parameters were adaptively obtained by optimising for the kurtosis of the approximation level signal using the GA [Man et al., 1996]. Thus, the decomposition is performed multiple times for each intermediate generation of the scaling parameters created by GA, which requires high computational time. Moreover, selection of approximation level may degrade the performance of the algorithm in the presence of real-life noises as these noises are more dominant in the low frequencies [Gradolewski and Redlarski, 2014].

In view of these issues, the proposed method decomposes the signal up to twenty levels using TQWT for only one time as against approximately 55 times, according to the mentioned parameter values of the GA in [Patidar and Pachori, 2013]. Hence, the computational time is reduced significantly. The value of quality parameter Q is set to one because Q close to unity provides effectively separates FHS from murmur, as observed in [Patidar and Pachori, 2013]. The redundancy parameter r controls the undesired excessive ringing in order to localise the wavelet in the time domain [Selesnick, 2011]. For $r < 3$, time domain response will not be well localised [Selesnick, 2011]. However, at a large value of r , a large number of decomposition levels will be required to separate the FHS and low-frequency noise. In view of this trade-off, a moderate value of the redundancy parameter $r = 8$ is set.

Values of the parameters $Q = 1$ and $r = 8$ result values of the β and α as 1 and 0.875, respectively, according to Eq.5.1. Then the signal is decomposed up to twenty ($J = 20$) levels. Using these values of the scaling parameters and sampling frequency of 1600 Hz, the centre frequency of the detailed level at $J = 20$ will be 31.63 Hz and the bandwidth will also be 31.63 Hz, according to Eq.5.5 and Eq.5.6. Thus, the detailed level ($J = 20$) covers the frequency band 15.82 Hz ($F_c - BW/2$) to 47.44 Hz ($F_c + BW/2$) and the approximation level covers the lower frequency band. As discussed above, the approximation level is prone to real-life noises and therefore, it is discarded from the consideration. The level with emphasised FHS is selected from the detailed levels adaptively based on a quality index, called as Fano factor, described as follows.

5.3.3 Adaptive Selection of the Level Based on Fano Factor

As discussed above, a quality index is required to select a level from the decomposed levels, such that the selected level has emphasised FHS. Authors in [Sanei et al., 2011] and [Patidar and Pachori, 2013] used the kurtosis [Pearson, 1905] for this purpose. kurtosis is a measure of whether the data are heavy-tailed or light-tailed relative to a normal distribution [Westfall, 2014]. It is used generally in the statistical field to describe the distribution of observed data around the mean. The kurtosis (K) can be expressed as follows.

$$K = \frac{\frac{1}{N} \sum_{n=0}^N (x(n) - \text{mean}(x))^4}{\left(\sum_{n=0}^N (x(n) - \text{mean}(x))^2 \right)^2} \quad (5.8)$$

Generally, the magnitude of the FHS is larger and shorter in time duration as compared to the murmurs [Sanei et al., 2011]. Hence, the FHSs have super-Gaussian distribution and sharper peaks than the murmur sounds, which is nearly Gaussian or sub-Gaussian [Sanei et al., 2011]. Thus, the kurtosis for the signal having prominent FHS will be larger as compared to the signal with murmur sounds and noise. This can be observed from the Figure 5.4 (case 1), which shows the signal of the aortic regurgitation pathological case at two different levels ($l_1 = 5$ and $l_2 = 19$), decomposed using TQWT. In this case, the murmur sound is present at level l_1 and has the sub-Gaussian distribution, and the FHS is present at level l_2 and has nearly Gaussian distribution. Therefore, the kurtosis is higher for the signal at level l_2 . However, when the murmur sounds have sharp peaks and overlap the FHS predominantly as in the case of aortic valve ejection, shown in Figure 5.4 (case 2), this assumption does not remain valid. In this case, the signals at two different levels $l_1 = 5$ and $l_2 = 18$ (decomposed using TQWT) are shown. The signal at level l_1 has murmur sound, which has sharp peaks and the signal at level l_2 has the FHS. The kurtosis is larger for the signal at level l_1 than at the level l_2 . Thus, using the kurtosis as a quality parameter to select the level with emphasised FHS, the signal at level l_1 will get selected where the murmur sounds have dominance over FHS.

Due to the limitation of the kurtosis, as mentioned above, a new quality index called as Fano factor [Fano, 1947] is explored. It is also known as the index of dispersion and variance-to-mean ratio [Cox and Lewis, 1966]. In statistics, it is generally used to quantify whether a set of observed occurrences is clustered or dispersed compared to a standard statistical model. The Fano factor (F) can also be viewed as a kind of noise-to-signal ratio and it is obtained as follows [Fano, 1947].

$$F = \frac{\sigma_w^2}{\mu_w} \quad (5.9)$$

where σ_w^2 and μ_w are the variance and the mean of the signal over a window with finite length w . For the signal with noise, the mean value of the absolute signal will be relatively larger than the clean signal and hence the value of Fano factor will be lower. Therefore, in case 1 (Figure 5.4) the value of Fano factor is lower for the signal at level l_1 than the value at level l_2 . In the second case (Figure 5.4), where the murmur has sharp peak, the variance will be higher for the signal with the murmur at level l_1 as compared to the signal with FHS at level l_2 . However, the mean value increases relatively more compared to the variance, which results in smaller value of Fano factor for the signal at level l_1 as compared to the level l_2 . Thus, using the Fano factor as a quality factor, the signal at the level l_2 will get selected which have more emphasised FHS as compared to the signal at level l_1 .

To show the efficacy of the quality index, the characteristics of the denominator and numerator terms of both quality indices, kurtosis and Fano factor, are analysed for the signal contaminated with AWGN. Figure 5.5 shows the relationship between denominator and

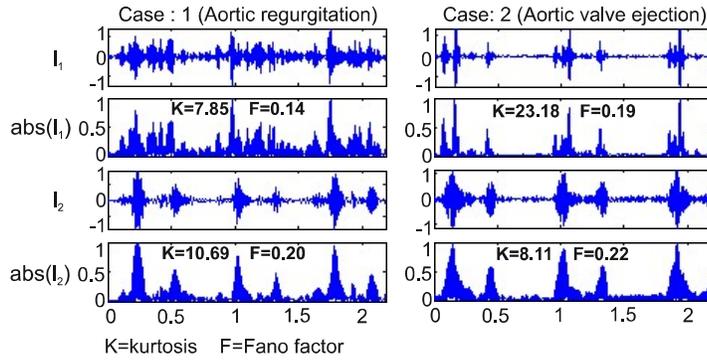


Figure 5.4 : Analysis of quality indices in two different cases of pathological signals at two different levels, decomposed using TQWT

numerator terms of both quality indices with varying SNR of the signal. In the case of Fano factor, as the SNR of the signal decreases, the denominator term (μ) increases exponentially and the numerator term (σ^2) decreases only a little. On the other hand, in the case of kurtosis, as the SNR value of the signal decreases the numerator term decreases a little and the denominator term remains almost constant. Thus, as the SNR of the signal decreases, the Fano factor decrease exponentially as compared to kurtosis. Therefore, the Fano factor may be a useful parameter to select a relatively clean level among various decomposed levels using TQWT.

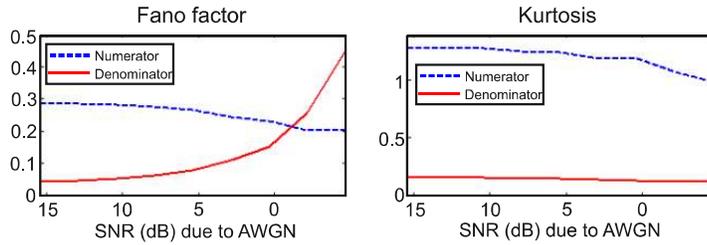


Figure 5.5 : Analysis of denominator and numerator terms of Fano factor and kurtosis for signal with various SNR, due to AWGN

5.3.4 Suppression of Low Amplitude Components

The previous step provides a signal with the emphasised FHS. Although, it may contain unwanted components with lower amplitude than FHS, due to the murmurs and noise. These components may affect the peak detection algorithm and, therefore, it is important to remove them. For this purpose, the signal is thresholded using a threshold value, which is obtained adaptively based on the statistical parameters of the signal. A new parameter, med_{75} , is proposed based on the domain knowledge about the heart sound signal that the sum of the time duration of S1 and S2 remains less than 25% of the time duration of a cardiac cycle [Singh and Anand, 2007; Naseri and Homaeinezhad, 2013; Atbi et al., 2013]. med_{75} represents 75th percentile value of sorted absolute values of detailed coefficients in ascending order. The threshold value is estimated using variance (v) and mean (m) of the absolute of the signal, as follows:

$$T = \begin{cases} med_{75} \times [1 - (v - med_{75})] & \text{if}(med_{75} < v) & \text{Case:1} \\ med_{75} & \text{if}((med_{75} > v) \&\&(med_{75} < m)) & \text{Case:2} \\ med_{75} + (med_{75} - m) & \text{if}(med_{75} > m) & \text{Case:3} \end{cases} \quad (5.10)$$

The threshold is obtained according to the level of noise present in the signal. In the first case, when the med_{75} is lower than the variance (v) and mean (m), the noise level is very low and therefore the threshold value should be low. The second case represents a signal contaminated with a moderate level of noise. The third case represents a high level of noise.

The obtained threshold is applied to the signal according to the soft threshold function [Luo and Zhang, 2012]. In this function, if the absolute value of the signal, $|x[n]|$, is less than the threshold value (T), then it is forced to zero, otherwise, it is shrunk by the threshold value (T) as follow.

$$x_s^T[n] = \begin{cases} \text{sign}(x[n]) (|x[n]| - T) & \text{if } |x[n]| > T \\ 0 & \text{otherwise} \end{cases} \quad (5.11)$$

5.3.5 Peak Detection and Conditioning

Envelope extraction methods emphasise FHS by attenuating the noise and averaging operation provides smoothness to the envelope [Choi and Jiang, 2008]. In the proposed method, the envelope of the signal is obtained using the NASE [Choi and Jiang, 2008]. Shannon energy emphasises the medium amplitude components more effectively as compared to the high amplitude components. It also attenuates the low amplitude components. Thus, Shannon energy based envelope method helps to identify the FHS with medium amplitude, especially in pathological cases where one of the FHS may have a lower amplitude than the other one. NASE is obtained as follows [Choi and Jiang, 2008].

$$NASE(n) = \frac{ASE(n) - \text{mean}(ASE)}{\text{standard deviation}(ASE)} \quad (5.12)$$

ASE is average Shannon energy, over the window length N , obtained as follows:

$$ASE(n) = -\frac{1}{N} \sum_{i=n-N/2}^{n+N/2} x^2(i) \times \log(x^2(i)) \quad (5.13)$$

In the proposed method, window length N is set to 32 that corresponds to signal duration of 0.02 seconds with the sampling rate of 1600 samples/second. The averaging operation smooths the envelope, which helps in near correct peak detection [Messer et al., 2001].

Peak detection is performed using a threshold value obtained by Otsu's threshold method [Otsu, 1979]. Otsu's method calculates a threshold value to classify data into two classes such that the interclass variance will be maximized and intra-class variance will be minimized. This matches with the requirement for the heart sound signal as the amplitude of the FHS will lie in one class with high amplitude, while noise components, mostly, will lie in another class with low amplitude. However, few false peaks may be identified due to high amplitude of noise and therefore, peak conditioning is performed to remove these false peaks. False peaks are removed based on the duration between detected peaks, as described in [Liang et al., 1997]. If two peaks appear within 50ms, which is the largest split normal sound interval, the peak with higher energy is retained and another one is discarded. After the peak detection and conditioning, peaks are identified based on the biological information i.e. the systole duration remains lower than the diastole and the systolic period is fairly constant than the diastole period.

5.4 RESULTS AND DISCUSSION

Experiments are performed using MATLAB® (MathWorks, Inc.) on the dataset obtained from the heart sound series produced by the Texas Heart Institute at St. Luke's Episcopal Hospital [Institute]. The dataset contains the PCG signals of normal cases and various pathological cases. Furthermore, the performance of the proposed method is tested on

these PCG signals contaminated with simulated noise. For the simulation, white Gaussian noise, pink noise, and red noise models are considered. Pink and red noise models are considered because the Gradolewski and Redlarski [Gradolewski and Redlarski, 2014] observed that the characteristics of a real-life noise match to these noises, which are generated through MATLAB® code created by Hristo Zhivomirov [Zhivomirov].

The performance of the proposed method is compared with recently proposed method C-TQWT [Patidar and Pachori, 2013] and another WT based popular method used to emphasise the FHS [Yuenyong et al., 2011; Vaisman et al., 2012], as discussed in Section 1. The performance of WT based method significantly depends on the selection of its parameters such as mother wavelet, the number of levels, and thresholding rules. In order to determine the optimum parameters, Gradolewski and Redlarski [Gradolewski and Redlarski, 2014] performed experiments on the PCG signals contaminated with various types of noises. The determined optimum parameters in [Gradolewski and Redlarski, 2014] are used to obtain the results for comparison.

The results are compared subjectively using plots and objectively in term of the segmentation rate (%). Segmentation rate is obtained as:

$$SR(\%) = 100 \times \frac{\text{Number of segmented cycles}}{\text{Number of total cycles}} \quad (5.14)$$

The cycle is considered to be segmented if both S1 and S2 are correctly identified. The performance of the proposed method is also compared with other methods with respect to computational efficiency.

5.4.1 Impact of Quality Indices: Kurtosis Vs Fano Factor

The proposed method uses Fano factor, as the quality index, to select the level with emphasised FHS, instead of the kurtosis used in [Patidar and Pachori, 2013; Sanei et al., 2011]. To show the efficacy of the quality indices, the performance of the proposed method is analysed using both the indices, separately. White Gaussian noise, pink noise, and red noise were added to the PCG signal of a normal case (shown in Figure 5.6) and the obtained results are presented in Figure 5.7. From the figure, it can be observed that the use of Fano factor results in emphasised FHS as compared to the one obtained using kurtosis.

Figure 5.8 shows the obtained results using both indices for various types of noise at various SNR. At each SNR, the results of multiple instances (10 times) were averaged out to find the performance at the given SNR. The Figure 5.8 depicts that the segmentation rate is higher using the Fano factor as compared to the one obtained using kurtosis in all cases of noise. In the case of AWGN, the results using the kurtosis degrades significantly mainly due to the Gaussian nature of the noise. In the case of pink noise, segmentation rate is slightly reduced using both the indices, although it is higher using the Fano Factor. Results using both the indices are more satisfactory in the case of red noise as compared to the pink noise. It is because red noise is mainly dominant in lower frequency [Gradolewski and Redlarski, 2014], as compared to pink noise, which is discarded by not considering the approximation level.

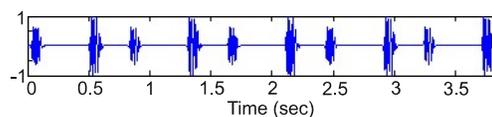


Figure 5.6 : PCG signal of a normal case

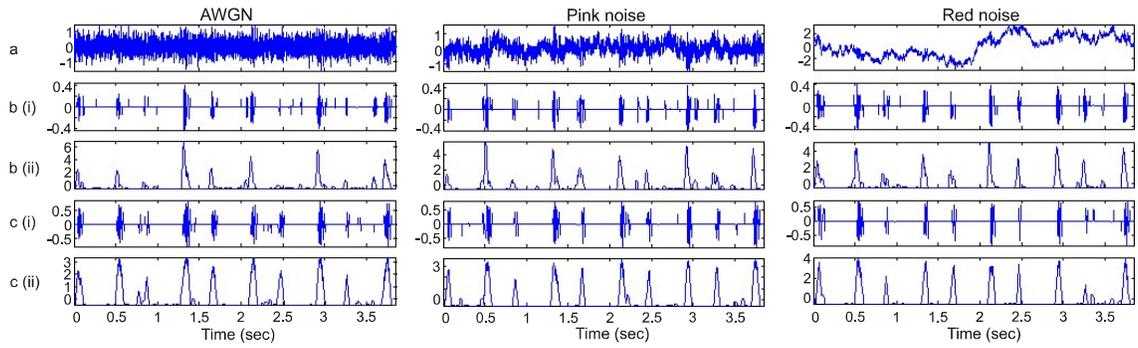


Figure 5.7 : Obtained results using the proposed method with different quality indices for the PCG signal contaminated with various types of noise: (a) noisy signal, (b) Results using kurtosis, and (c) Results using Fano factor. The Roman numbers (i) and (ii) represent reconstructed signal at a selected basis level and its envelope using NASE, respectively

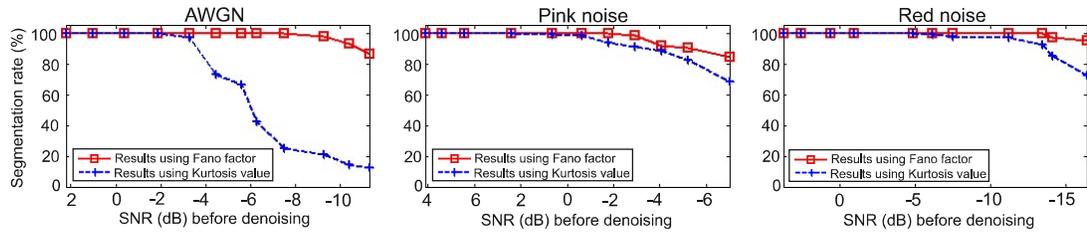


Figure 5.8 : Obtained segmentation rate (%) using the proposed method with kurtosis and Fano factor, for the signal contaminated with various types of noise at various levels.

5.4.2 Result for the Normal Heart Sound Signal with Simulated Noise

The proposed method and two other methods, WT and C-TQWT, are applied to the PCG signal contaminated with the simulated noise, as before. The obtained results are presented in Figure 5.9. From the figure, it is manifested that in all three types of noises, the proposed method suppresses the noise significantly, which shows the effectiveness of the use of the Fano factor to select the level with low noise. The figure also depicts the effectiveness of the proposed adaptive thresholding method for the suppression of the low amplitude noise. WT method also suppressed the noise significantly, however, it is comparatively less effective. On the other hand, the results of the C-TQWT method leaves a lot of scope for the noise removal. It is because C-TQWT method selects the approximation level where these noises have dominant energy.

To demonstrate the robustness of the proposed method, experiments are carried out at various levels of additive noises (white Gaussian, red, and pink). Results obtained at multiple instances were averaged out to find the performance, same as before. These results are plotted in Figure 5.10, which shows that the proposed method outperforms the other two methods, WT and C-TQWT. In the case of AWGN, where noise has a uniform distribution over all frequencies, the main reason behind the robustness of the proposed method is the use of Fano factor, which efficiently selects the level with low noise. In the case of pink and red noise, the proposed method efficiently segments the PCG signal because the approximation level was discarded, where these noises have dominance. WT method produced better results as compared to the C-TQWT method. The performance of the C-TQWT method is significantly inferior in the case of pink and red noise. The reason behind it is the selection of approximation

level, as discussed above.

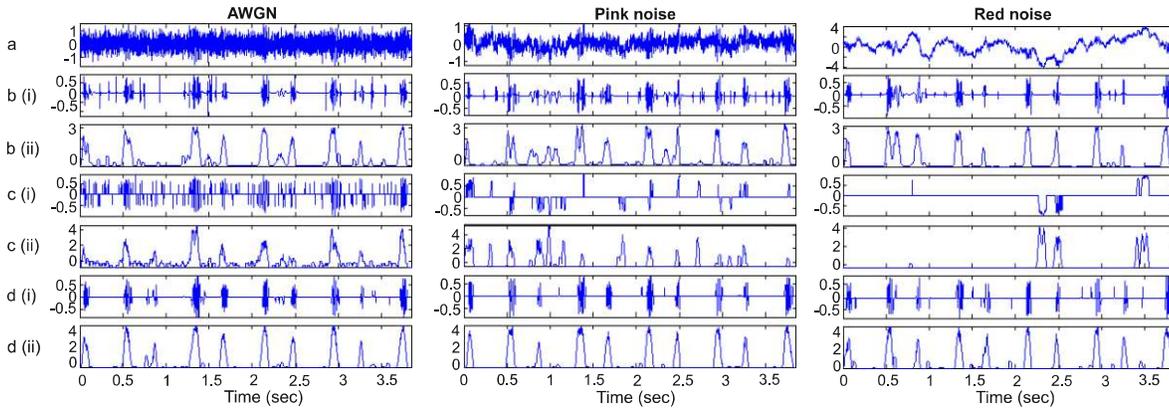


Figure 5.9 : Obtained results for the PCG signal contaminated with various types of noise: (a) noisy signal, (b) Results using WT method, (c) Results using C-TQWT method, and (d) Results using proposed method. The Roman numerals (i) and (ii) represent reconstructed signal at selected basis level and its envelope using NASE, respectively

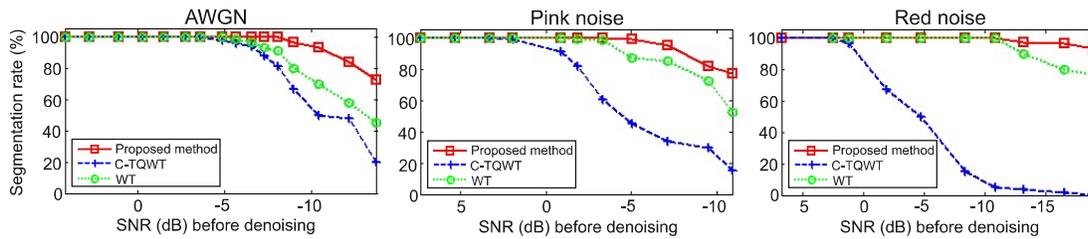


Figure 5.10 : Obtained segmentation rate (%) using various methods for the signal contaminated with various types of noise at various levels

5.4.3 Result for Pathological Cases Without and With Simulated Noise

Experiments are also performed on the PCG signals of various pathological cases obtained from the dataset [Institute], without and with simulated noise. Results obtained using the proposed method and other two methods, WT and C-TQWT, for the PCG signal of few pathological cases without simulated noise are presented in Figure 5.11. The figure depicts that the proposed method efficiently suppresses the murmur sounds and hence emphasised the FHS. Furthermore, the performance of the proposed method is superior in comparison to WT and C-TQWT methods. The main reason behind this efficiency is due to the use of Fano factor, which selected the signal with emphasised FHS. It was observed that, in most of the cases, Fano factor selected the best level in terms of emphasised FHS.

Obtained segmentation rates for various pathological cases using the proposed method, WT, and C-TQWT methods are given in Table 5.2. The table contains the result for the signals without additive noise and with additive noise simulated using various noise models. In the case of without additive noise, the table depicts that the segmentation rate obtained using the proposed method is higher in most cases compared to other methods. However, in a few cases, the performance of the proposed method is inferior to the C-TQWT method. The reason for the same is as follows: In the case of third and fourth heart sound, the Fano factor selected the lower frequency subband, which was relatively clean as compared to the higher frequency subbands. However, the selected level contains third or fourth heart sound because these sound components have low-frequency range and overlap the frequency range

of FHS, dominantly [Kumar et al., 2007]. The presence of these extra components affected the segmentation accuracy of the proposed method. The same problem also occurred in one case of mitral opening snap diseases. In this case, the subband selected using the Fano factor was cleaner as compared to other subbands. However, few components were missing in the selected band.

The performance of the C-TQWT method is satisfactory for most of the cases, due to the adaptive decomposition of the signal using TQWT. The reason behind the low performance of the WT method is its limitation to tune its parameters, which does not allow decomposition of the signal into fine-tuned frequency bands.

The results obtained using various methods for the PCG signal of pathological cases with additive noise are also obtained and provided in Table 5.2. The results presented in the table manifested that the proposed method outperforms the other two state-of-art methods. It shows the effectiveness of the quality index, Fano factor, to select the subband with a low level of noise. However, in a few pathological cases such as aortic stenosis and pulmonary valve stenosis, where murmur sounds affected the FHS severely, the performance of the proposed method also degraded significantly. In the case of the signal with third heart sound, the WT based method produced better segmentation rate as compared to the proposed and the C-TQWT methods.

On the other hand, the performance of the C-TQWT degrades significantly in the presence of additive noise, especially in case of pink and red noises. It is mainly due to the selection of approximation level, which is prone to these types of noises.

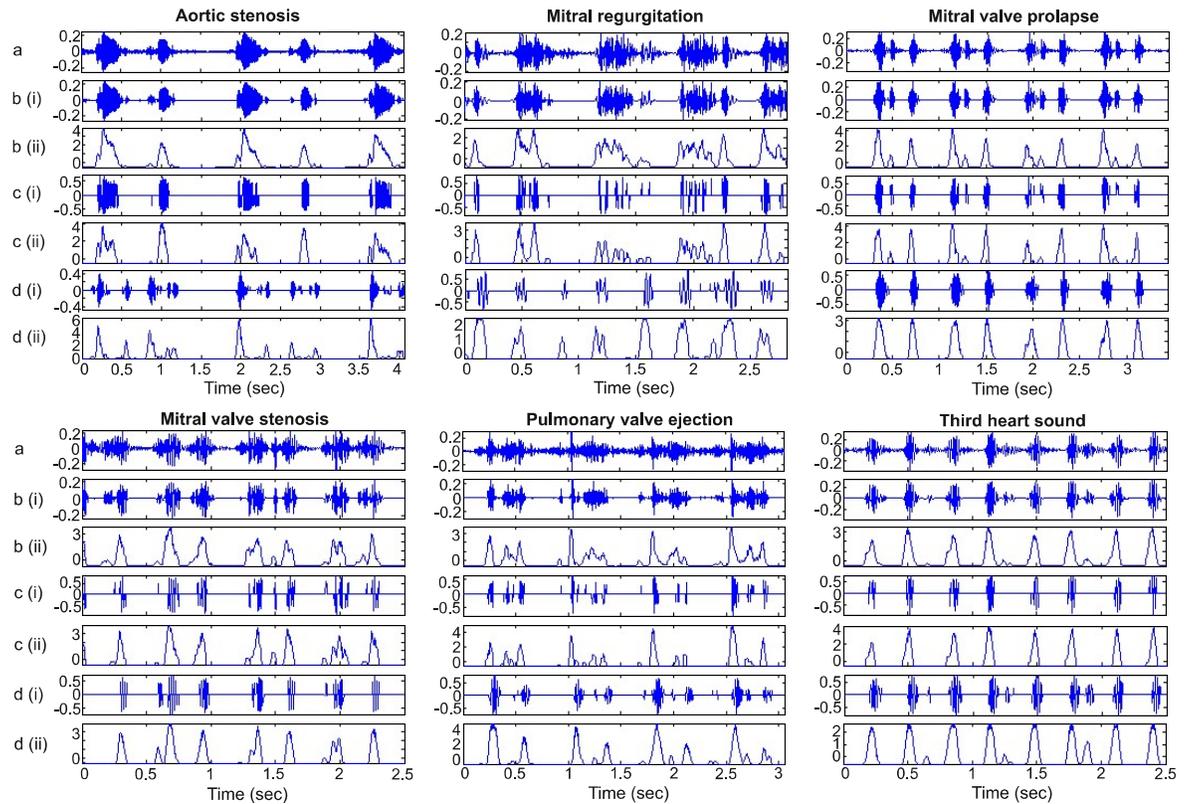


Figure 5.11 : Results obtained for the PCG signal with murmur sounds: (a) PCG signal with murmur, (b) Results using WT method, (c) Results using C-TQWT method, and (d) results using proposed method. The Roman numbers (i) and (ii) represent reconstructed signal at selected basis level and its envelope using NASE, respectively

5.4.4 Time Complexity

The computational complexity of an automatic analysis algorithm becomes a crucial factor when such system has to be used for long-term and when the heart is being monitored at the Intensive Care Unit (ICU). Therefore, the proposed method has been compared with the WT and C-TQWT methods with respect to computational time. The methods have been implemented in MATLAB® (MathWorks, Inc.) version R2009b. The experiments have been performed on a Dell personal computer consisting of the processor: Intel® core i5 CPU 760 @2.8 GHz, installed memory (RAM): 8.00 GB, and operating system: 64-bit Windows 7. The computational time of all three methods for the signal with various lengths is provided in Table 5.1. It is obvious that the WT method has a lowest computational time due to the low number of decomposition level, which is six in this case. On the other hand, the proposed method performs the decomposition up to twenty levels. However, as discussed in the previous section, WT method is not efficient in pathological cases, which are the most important as they lead to a diagnosis of the health status of the heart. The proposed method is approximately 25 times faster than the C-TQWT method. The reason behind this lower computational time is that the proposed method decomposes the signal up to twenty levels, only once as against the 55 times in C-TQWT method, as discussed in Section 5.1.

For the segmentation of heart sound signal, various methods based on envelope extraction and transformation techniques have been discussed in Section 1. The time complexity of the transformation-based methods is higher as compared to the time-complexity of the envelope based methods. It is because, in addition to the computational cost of envelope extraction, transformation-based methods include the computational cost required to transfer the time-domain signal into another domain. The computational complexity of the DWT is $O(N)$, STFT is $O(N^2)$, S-transform is $O(N^2 \log N)$, WVD is $O(N^2 \log 2N)$ [Toole et al., 2005], and TQWT is $O(rN \log_2 N)$ where N is the length of the input signal, and r is the redundancy factor [Selesnick, 2011]. Moreover, to perform the transformation, these methods require the buffering of samples of a certain time period. This requirement limits their use for the real-time application. However, transformation provides a better way to separate the FHS from the murmur and noise and therefore most of the recently reported works are focused on transformation-based methods.

Table 5.1 : Execution time (seconds) taken by the proposed method, WT method, and C-TQWT method for aortic regurgitation case at various signal length

Signal length (N)	TQWT	WT	proposed
20000	5.24	0.02	0.25
40000	11.27	0.03	0.52
60000	15.61	0.04	0.79
80000	20.48	0.05	1.02
100000	24.22	0.07	1.16
120000	33.69	0.09	1.47
140000	42.14	0.11	2.09
160000	61.37	0.14	2.57

5.5 CONCLUSIONS

In this chapter, a TQWT based segmentation method, robust to noise and murmur sound, is proposed for segmentation of PCG signal. There are three reasons that contributed towards the enhancement of robustness of the method. First one is the discarding of the approximation level coefficients by which most of the real-life noise components get discarded. The second reason is the use of Fano factor that effectively selects a level with the low level of noise. The third reason is the adaptive thresholding of the selected level and that suppresses the low amplitude noise components efficiently. The obtained results show that the proposed

Table 5.2 : Obtained segmentation rate (%) using the proposed method, WT method, and C-TQWT method for various pathological cases without additive noise and with added noise simulated using white Gaussian , red, and pink noise models

Disease	Cycles	Without additive noise				Additive white Gaussian noise SNR (\approx 0dB)				Pink noise SNR (\approx 0dB)				Red noise SNR (\approx 0dB)			
		WT	TQWT	Proposed	J	WT	TQWT	Proposed	J	WT	TQWT	Proposed	J	WT	TQWT	Proposed	J
		SR(%)	SR(%)	J	SR(%)	SR(%)	SR(%)	J	SR(%)	SR(%)	SR(%)	J	SR(%)	SR(%)	SR(%)	J	SR(%)
Aortic regurgitation																	
Case 1	31	70.97	93.55	19	100	70.97	87.10	16	90.32	51.61	25.81	14	61.29	45.16	25.81	16	48.39
Case 2	32	71.88	84.38	19	93.75	68.75	81.25	16	93.75	43.75	18.75	18	43.75	34.38	21.88	16	56.25
Aortic valve ejection																	
Case 1	9	100	100	17	100	100	100	14	100	100	22.22	11	100	100	22.22	12	100
Case 2	9	100	100	14	100	100	55.56	14	100	100	11.11	14	100	100	33.33	14	100
Case 3	13	46.15	76.92	15	76.92	15.38	46.15	15	61.54	30.17	38.46	14	61.54	15.38	38.46	15	53.85
Case 4	16	81.25	62.50	12	93.75	62.50	56.25	12	93.75	62.50	25.00	12	75.0	75.0	25.0	12	87.50
Aortic Stenosis																	
Case 1	25	44.00	84.00	13	88.00	8.0	16.0	18	48.0	8.0	32.0	16	44.0	0.0	8.0	4	52.0
Case 2	46	34.78	86.96	13	91.30	15.22	23.91	18	52.17	0.0	13.04	2	26.09	0.0	4.35	18	26.09
Flail mitral regurgitation																	
Case 1	13	100	100	18	100	69.23	53.85	18	76.92	46.15	61.54	15	61.54	61.54	53.85	18	69.23
Case 2	12	75.00	100	16	100	75.0	83.33	16	83.33	66.67	88.33	17	83.33	58.33	58.33	16	75.0
Fixed splitting of S2																	
Case 1	31	90.32	87.10	13	93.55	80.65	77.42	13	80.65	83.87	87.10	13	90.32	83.87	25.81	13	83.87
Fourth heart sound																	
Case 1	29	72.41	100	17	93.10	72.41	68.97	16	86.21	55.17	75.86	15	82.76	68.97	62.07	17	75.86
Case 2	30	73.33	100	17	93.33	80.0	73.33	17	80.0	63.33	83.33	17	83.33	66.67	66.67	17	70.0
Hypertrophy obstructive cardiomyopathy																	
Case 1	10	100	100	14	100	100	100	14	100	90.0	30.0	14	100	90.0	0.0	15	100
Case 2	39	74.36	89.74	18	92.31	74.36	79.49	12	92.31	71.79	28.21	12	89.74	58.97	15.38	13	74.36
Case 3	46	69.57	86.96	12	91.30	60.87	76.09	10	82.61	67.39	30.43	10	84.78	63.04	28.26	11	73.91
Mitral opening snap																	
Case 1	11	81.82	100	18	90.91	72.73	36.36	15	81.82	63.64	27.27	14	63.64	81.82	27.27	12	81.82
Case 2	15	100	100	11	100	100	73.33	12	100	100	80.0	11	100	100	6.67	11	100
Case 3	14	100	100	16	100	100	100	12	100	78.57	57.14	14	92.86	100	57.14	14	100
Mitral valve regurgitation																	
Case 1	18	66.67	88.88	19	94.44	44.44	55.56	16	88.89	27.78	38.89	18	77.78	55.56	22.22	16	88.89
Case 2	9	22.22	88.88	19	100	0.0	11.11	18	66.67	0.0	33.33	19	55.56	0.0	0.0	17	66.67
Case 3	16	25.00	72.23	18	100	0.0	18.75	19	68.75	12.50	56.25	19	56.25	18.75	12.50	19	68.75
Mitral valve stenosis																	
Case 1	16	100	93.75	15	100	100	93.75	13	100	93.75	68.75	13	93.75	100	6.25	14	100
Case 2	14	100	100	19	100	100	92.86	12	100	71.43	57.14	5	71.43	100	14.29	12	100
Case 3	12	75.0	75.00	13	75.0	58.33	66.67	13	66.67	50.0	66.67	12	66.67	66.67	25.0	13	66.67
Case 4	12	100	91.67	16	100	91.67	91.67	15	91.67	83.33	83.33	15	83.33	100	33.33	14	100
Mid-systolic click																	
Case 1	17	100	100	14	100	100	100	15	100	100	47.06	14	100	82.35	58.81	14	100
Case 2	6	100	100	15	100	100	100	15	100	100	83.33	14	100	100	33.33	14	100
Case 3	68	95.59	100	14	100	97.06	97.06	13	100	76.47	55.88	13	91.18	64.71	38.24	14	79.14
Case 4	35	97.14	97.14	14	100	94.25	97.14	14	100	80.0	74.29	13	91.43	57.14	68.57	15	88.57
Case 5	43	95.35	100	13	100	95.35	100	15	100	76.74	72.09	13	88.37	74.42	25.58	13	90.70
Case 6	27	96.30	100	13	100	92.59	92.59	12	100	77.78	70.37	12	77.78	66.67	44.44	13	96.30
Case 7	19	100	100	14	100	100	100	14	100	89.47	68.42	12	94.74	84.21	42.11	14	100
Case 8	27	92.59	100	13	100	92.59	92.59	13	96.30	74.07	44.44	13	85.19	74.07	81.48	13	92.59
Case 9	20	85.0	80.0	13	95.0	80.0	75.0	13	85.0	50.0	30.0	13	60.0	60.0	65.0	13	75.0
Pericardial knock																	
Case 1	18	100	100	15	100	66.67	94.44	16	94.44	33.33	83.33	15	83.33	83.33	16.67	15	88.89
Physiologic split S2																	
Case 1	23	82.61	91.30	15	100	73.91	60.87	13	86.96	60.87	73.91	12	78.26	73.91	47.83	12	82.61
Case 2	23	95.65	91.30	19	100	86.96	91.30	12	91.30	69.57	26.09	12	82.61	78.26	26.09	12	86.96
Pulmonary valve ejection																	
Case 1	16	68.75	93.75	19	100	50.0	62.50	12	68.75	37.50	31.25	4	37.50	50.0	12.50	5	50.0
Case 2	28	71.43	92.86	19	100	46.43	57.14	11	64.29	67.86	39.29	9	67.86	64.29	14.29	6	64.29
Third heart sound																	
Case 1	11	90.91	90.90	18	72.73	90.91	90.91	13	90.91	81.82	72.73	13	81.82	90.91	18.18	11	72.73
Case 2	30	93.33	96.66	16	96.66	90.0	93.33	14	93.33	80.0	70.0	15	53.33	73.33	67.67	15	86.67
Tricuspid valve insufficiency																	
Case 1	27	74.07	96.29	14	96.29	37.04	22.22	5	59.26	37.04	66.67	14	81.48	59.26	22.22	14	74.07
Tumor plop																	
Case 1	20	100	100	14	100	75.0	70.0	15	75.0	80.0	40.0	14	80.0	90.0	20.0	14	95.0
Wide splitting S2																	
Case 1	21	100	100	13	100	95.24	95.24	13	100	85.71	66.67	12	95.24	95.24	23.81	13	100

method outperforms the state-of-art methods in terms of segmentation rate. The method is also compared with WT and C-TQWT based methods in terms of computational cost. Obtained results show that the proposed method is computationally efficient as compared to the C-TQWT, while, the WT based method is more efficient as compared to the proposed method. However, the performance of the WT method is limited in various pathological cases.

...