33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Noise Detection in Heart Sound Recordings

Mohammad K. Zia,Benjamin Griffel, Vladimir Fridman, Cesare Saponieri, and John L. Semmlow

Abstract— **Coronary artery disease (CAD) is the leading cause of death in the United States. Although progression of CAD can be controlled using drugs and diet, it is usually detected in advanced stages when invasive treatment is required. Current methods to detect CAD are invasive and/or costly, hence not suitable as a regular screening tool to detect CAD in early stages. Currently, we are developing a noninvasive and cost-effective system to detect CAD using the acoustic approach. This method identifies sounds generated by turbulent flow through partially narrowed coronary arteries to detect CAD. The limiting factor of this method is sensitivity to noises commonly encountered in the clinical setting. Because the CAD sounds are faint, these noises can easily obscure the CAD sounds and make detection impossible. In this paper, we propose a method to detect and eliminate noise encountered in the clinical setting using a reference channel. We show that our method is effective in detecting noise, which is essential to the success of the acoustic approach.**

I. INTRODUCTION

Coronary artery disease (CAD) is the leading cause of death in the United States. It causes one in every six deaths and the associated cost burden exceeds $177 billion [1]. CAD results when the coronary arteries, which supply blood to the heart tissues, narrow due to plaque deposition. This narrowing restricts blood flow and hence limits transportation of vital oxygen and nutrients to heart tissues, which can lead to a heart attack. Although methods exists to detect CAD, they are invasive and/or costly, which limits their use as a regular screening tool to detect CAD in early stages, when its progression can be controlled using drugs and diet. Because CAD often goes undetected and untreated, it the leading cause of death in the United States. A noninvasive and cost-effective screening tool that can detect CAD in early stages would greatly reduce the number of deaths and the associated cost burden.

In 1983, Semmlow et al. first proposed the acoustic approach to detect CAD, which is the most promising, noninvasive and cost-effective method to detect CAD [8]. Based on the analogy of how blood flow through partially narrowed carotid arteries generates sounds, termed *bruits*, Semmlow and colleagues hypothesized that coronary artery narrowing should also produce sounds. Although reports of audible CAD sounds are rare [4],[6], there are occasional reports of diastolic murmurs whose timing closely coincides with maximum blood flow through the coronary arteries during diastole, and these sounds have been attributed to CAD [7], [5], [2]. Based on these reports, Semmlow and colleagues argued that CAD sounds are likely present in most patients, but are too faint to be audible under normal circumstances.

To detect CAD sounds, Semmlow and colleagues applied advanced signal enhancement techniques to acoustic measurements from the chest. They observed clear differences in the frequency distribution of normal and CAD patients, where the latter had more high frequency energy above 120 Hz [8]. Based on this observation, they concluded that signal enhancement techniques could be used to detect the faint CAD sounds, which would allow for noninvasive and cost-effective detection of CAD. For a review of past work on CAD detection using the acoustic approach, see the review article by Semmlow et al [9].

Currently, we are working with an industrial partner, SonoMedica, Inc. (McLean, Virginia), to develop a device to detect CAD using the acoustic approach introduced by Semmlow and colleagues. The limiting factor to the acoustic approach is sensitivity to clinical noise. Noise in the clinical setting can obscure the faint CAD sounds, making CAD detection difficult. There are two sources of noise in the clinical setting: external and internal. External noise includes talking in nearby rooms, intercom activity, machine and ventilation noise and door slams. Internal noise is primarily from stomach growls. In their original study, Semmlow and colleagues acquired their data in a soundproof booth. Such a requirement is impractical and would raise the cost to develop a device to detect CAD using the acoustic approach. Other researchers manually edit their data for noise. This is also impractical and imposes an operator bias. Therefore, for the successful detection of CAD sounds, it is essential to automatically detect and eliminate noise before any CAD signal detection efforts. In this paper, we propose an algorithm that uses the signal from a reference microphone to detect noise, verify that this noise is also present in the active channel as some stomach growls are low-level and do not corrupt the heart sound recordings, and eliminate the corrupting noise.

II. DESCRIPTION OF DATA ACQUISITION

Since we expect noise to be the limiting factor in detection of CAD using the acoustic approach, we use a reference microphone placed on the stomach to capture both external and internal noise. We also acquire heart sound data from microphones placed on the chest. Recordings are taken using a 4-channel data acquisition system developed by SonoMedica, Inc.. Channel one and two record heart sounds from

Mohammad K. Zia and Benjamin Griffel and are with the University of Medicine and Dentistry of New Jersey – Graduate School of Biomedical Sciences and Rutgers, the State University of New Jersey, Piscataway, NJ, USA. E-mail: mkhawarzia@gmail.com.

Vladimir Fridman and Cesare Saponieri are with the Department of Cardiology, Long Island College Hospital, Long Island, NY, USA.

John L. Semmlow is with the Department of Surgery, University of Medicine and Dentistry of New Jersey, Piscataway, NJ, USA.

978-1-4244-4122-8/11/$26.00 ©2011 IEEE

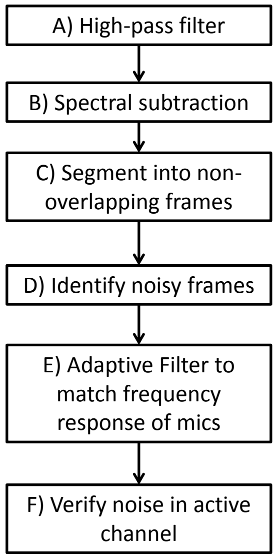

Fig. 1. Steps for detecting noise using signal from the reference channel.

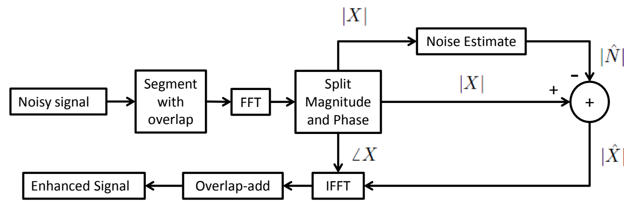

Fig. 2. Steps of spectral subtraction.

the chest. Channel three records the reference microphone for noise. Channel four records the electrocardiogram. Data were sampled at 22099 Hz using 16 bits. Data were acquired under the Institutional Review Board (IRB) approval of the University of Medicine and Dentistry of New Jersey and with the consent of the patients. Recordings were made in cardiology clinics under the direction of two of our clinical co-authors. From our database, we selected a noisy record to demonstrate the effectiveness of our method.

III. METHOD

For each record we have active channels (channel one and two) that contain heart sounds and a reference channel (channel three) that captures external and internal noise. Our objective is to detect noise activity in the reference channel and verify that it is also present in the active channels; some stomach growls are faint, so they do not corrupt the active channels and need not be eliminated. To detect noise, we perform a series of steps. A) We begin by applying a high-pass filter to the signal from the reference channel. B) We enhance the reference signal using spectral subtraction. C) We segment the enhanced signal into non-overlapping frames. D) We identify noisy frames using an eigenvalue-based method. E) We use an adaptive filter to match the frequency response of the active and reference microphone. F) We verify that frames identified as noisy in the reference channel have similar noise activity in the active channels. See Fig. 1 for a summary of the method.

A. Bandpass Filter

From previous work, we know that CAD sounds are above 120 Hz. Therefore, we first apply a fifth-order Butterworth high pass filter with a cutoff frequency of 90 Hz to focus on detecting noise above 120 Hz. Also, although we purposefully place the reference channel on the right side of the stomach to reduce possible interference from loud heart sounds, applying a high-pass filter further ensures that these sounds do not trigger false noise detection.

B. Spectral Subtraction

We used Berouti's [3] method for spectral subtraction, which is the simplest method for narrowband signal enhancement and broadband signal attenuation. Assuming the clean signal is corrupted by noise that is additive and uncorrelated, we can obtain an estimate of the clean signal by subtracting an estimate of the noise spectrum from the spectrum of the noisy signal. The preserved phase of the noisy signal is then combined with the estimate of the clean signal spectrum to re-synthesize the enhanced signal.

Spectral subtraction consists of several steps, see Fig. 2. a) First, we segment a signal into overlapping frames. b) We calculate the fast Fourier transform (FFT) of each frame. c) We separate the magnitude and phase spectrum of each frame. d)We subtract an estimate of the noise magnitude spectrum from the noisy signal magnitude spectrum. Subtraction is performed using Eq. 1,

$$\hat{X}(\omega) = \max([|X(\omega)| - \eta|\hat{N}(\omega)|], \gamma|\hat{N}(\omega)|), \qquad (1)$$

where $\hat{X}(\omega)$ is the estimate of the magnitude spectrum of the enhanced signal, $X(\omega)$ is the magnitude spectrum of the noisy signal, and $\hat{N}(\omega)$ is an estimate of the magnitude spectrum of noise. The η parameter is the over-subtraction parameter which controls how much of the noise is subtracted from the noisy signal. The γ parameter ensures that negative values resulting from over-subtraction are corrected by imposing a baseline spectral floor. These two parameters control the trade-off between broadband signal attenuation and distortion introduced by spectral subtraction. e) We combine the enhanced signal spectrum with the phase of the original noisy signal. f) We calculate the inverse Fourier transform (IFFT) of each frame. g) Finally, we re-synthesize the signal in the time domain using overlap-add method.

For estimating the noise spectrum used in step d, we employed the method proposed by Stahl et al. [10]. Based on their observation that noise is usually contained in the top 0.8-0.9 to 1.0 quantile of a given frequency band, we can estimate the noise spectrum $\hat{N}(\omega)$ from the noisy signal spectrum $X(\omega, t)$ by first sorting the magnitudes $X(\omega)$ in each frequency band ω in ascending order. Then, we define the q-quantile noise estimation as

$$\hat{N}(\omega) = X(\omega, t_{[qT]}). \quad (2)$$

That is, we estimate the magnitude of the noise spectrum $N(\omega)$ as the magnitude of each frequency band ω at the q-th quantile. For example, at $q = 0.5$, the median, we are assuming that all the frequency bands contain noise at least half the time of the duration of the whole signal. For spectral subtraction, we segmented the frames into 80 ms length with 20 ms using hamming window. We used 1024 point FFT and used $q = 0.45$ for estimating noise.

C. Segment Enhanced Signal

We segment the enhanced signal into 100 ms non-overlapping frames using a rectangular window. These frames are then checked for the presence of noise in subsequents steps.

D. Identification of Noisy Frames

We identify noisy frames using an eigenvalue-based method that exploits the difference in the number of significant eigenvalues of noisy and noise-free frames. To calculate the eigenvalues, we construct a trajectory matrix of each frame. A trajectory matrix has lagged vectors of the enhanced signal as its column. We used a lag of 100 to construct the trajectory matrix. We apply singular value decomposition (SVD) on the trajectory matrix and extract the eigenvalues. Next, we determine the number of significant eigenvalues in each frame. Significant eigenvalues are those that account for 99% of the total eigenvalues. Afterward, we calculate the variance in each frame and determine the 10 frames with the least variance. The eigenvalue distribution of these frames are representative of noise-free frames. Then, we determine the mean and standard deviation (STD) of the significant eigenvalues of the 10 frames with the least variance. Finally, we identify noisy frames to be those whose number of significant eigenvalues is less than the mean minus three STD's of the significant eigenvalues in the 10 frames with the lowest variance.

E. Adaptive Filter to Match Microphone Frequency Response

We have observed that the frequency response of the microphones to the same signal is not the same between the active and reference channel. This imposes a problem because we cannot directly do cross-correlation to verify that the signal detected in the reference channel is also present in the active channels. To resolve this problem, we apply an adaptive filter to match the frequency response of the two microphones. We use the least mean square adaptive (LMS) filter. To select the optimum filter length and adaptation coefficient, we optimized the filter length and adaptation coefficient to give the best cross-correlation between the active and reference channel. The inputs into the LMS filter were the original high-passed signals from step A instead of the enhanced signals from step B.

F. Verify Presence of Signal in Active Channel

In our last step, we verify the presence of the signal in the reference channel by doing a simple cross-correlation. Only very high values, more than 0.95, of normalized cross-correlation imply signal match.

IV. Results

Fig. 3 shows a typical noisy recording. Marked in the top panel and second from top panel of the figure are the corresponding stomach sounds in both active and reference channel, respectively. The bottom two plots in this figure show details of the noisy segments from the active and reference channel. We clearly see that the frequency content is not an exact match. This is verified from Table I which shows the normalized cross-correlation to be 0.58 before adaptive filtering. After adaptive filtering the correlation increases to 0.99.

Fig. 4 shows the short-term variance before (top panel) and after (lower panel) spectral subtraction of the reference signal shown in Fig. 3. Also shown in lower panel of Fig. 4 are the frames that were identified as noisy using the eigenvalue-based method. Before spectral subtraction, the noise-free periods are not distinguishable from noisy periods based short-term variance. After spectral subtraction, variance of noise-free periods are noticeably smaller. This change in variance is reflected in the distribution of the eigenvalues and we exploit this to identify noisy frames.

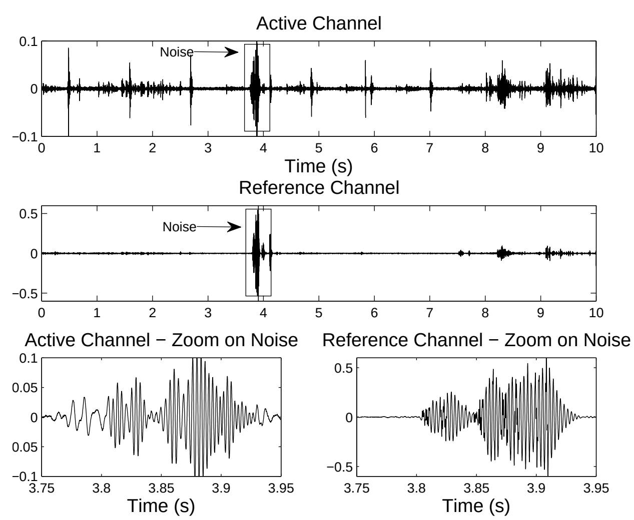

Fig. 3. An example noisy record. Top panel: Active channel with a noisy period marked. Second panel from top: Reference channel with a noisy period marked. Bottom left panel: An enlargement of the active channel noise period. Bottom right panel: An enlargement of the reference channel noise period. Since the marked noise period does not have the same frequency response in the active and reference channel, we cannot do direct cross-correlation to match the noisy period. We must first apply adaptive filter to match the frequency response of the two channels before applying cross-correlation to verify whether the detected noise in the reference channel is also present in the active channel.

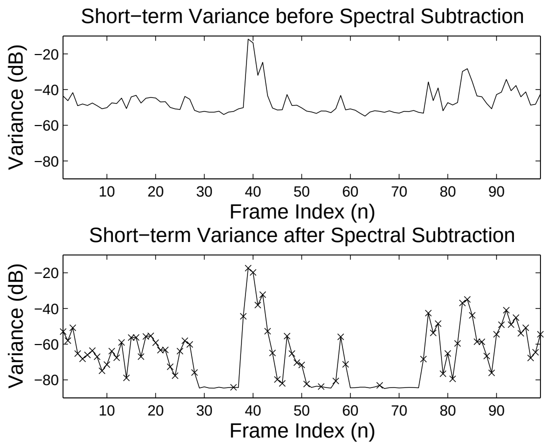

Fig. 4. Short-term variance (STV) of the reference signal shown in Fig. 3 second panel from top. Top panel: STV before spectral subtraction. Lower panel: STV after spectral subtraction. Spectral subtraction accentuates the difference between the STV of noisy and noise-free frames. This is reflected in the eigenvalue distribution of the frames and we exploit this difference to identify noisy frames, marked in 'x'.

TABLE I

CHANGE IN NORMALIZED CROSS-CORRELATION OF NOISE AFTER APPLYING ADAPTIVE FILTER TO THE MATCH THE FREQUENCY RESPONSE OF THE ACTIVE AND REFERENCE CHANNEL. CROSS-CORRELATION IS HIGH FOR NOISE THAT IS DETECTED IN THE REFERENCE CHANNEL AND IS ALSO PRESENT IN THE ACTIVE CHANNEL.

Before LMS	0.58
After LMS	0.99

V. DISCUSSION

Application of spectral subtraction is important because it enhances the noise commonly encountered in the clinical setting. Many times this noise has comparable amplitude to the broadband noise in the room. By applying spectral subtraction, we enhance the narrowband signal typical of noise in the clinical setting relative to the broadband noise in the room. This narrowband enhancement accentuates the difference in short-term variance of noisy and noise-free frames, compare Fig. 4 top panel and second from top panel, which makes subsequent detection of noisy frames using eigenvalues easier.

After spectral subtraction, noisy frames require relatively fewer number of eigenvalues to represent most of the variance compared to noise-free frames. This makes intuitive sense because a narrowband signal, characteristic of clinical noise, can be succinctly captured using fewer eigenvalues while broadband signal, characteristic of noise-free frames, require more eigenvalues to represent the signal. We exploit this difference to identify noisy frames. This method for noise detection is robust since it does not require explicit thresholding of the STV to identify noisy frames and is indifferent to gain setting, which can vary greatly in the clinical setting.

Finally, using adaptive filtering to match the frequency response of the active and reference channel was an essential step. Without this step, there would be no way to verify that noise that is detected in the reference channel is also in the active channel. This is required primarily for internal noise because not all stomach growls corrupt the heart sounds. Also, for noise detected in the reference channel that is not present in the active channel, the cross-correlation is low even after applying adaptive filter. We only accept noisy frames that have cross-correlation greater than 0.95 after matching the frequency response of the active and reference channel.

VI. CONCLUSION

In this paper, we have proposed a method to detect noise in the clinical setting that can corrupt heart sound recordings. This method can verify the presence of ambient and internal noise in the active channel using a reference channel. We anticipate that such a noise detection and elimination method will be key to the success of the acoustic approach to CAD detection.

VII. ACKNOWLEDGMENTS

This work was supported by the NIH-HIHLB Grant number: HL079672-02. John L. Semmlow has financial interest in SonoMedica, Inc..

REFERENCES

[1] "Heart Disease and Stroke Statistics–2010 Update: A Report From the American Heart Association," *Circulation*, vol. 121, no. 7, pp. e46–215, 2010.
[2] K. A. Berg, J. R.and Weaver, T. Russell, and D. G. Kassenbaum, "Disappearance of coronary artery stenosis murmur after aortwcoronary by-pass," *Chest*, vol. 63, p. 440, 1973.
[3] M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of speech corrupted by acoustic noise," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '79.*, vol. 4, apr 1979, pp. 208 – 211.
[4] T. O. Cheng, "Diastolic murmur caused by coronary artery stenosis," *Ann. Int. Med.*, vol. 72, p. 543, 1970.
[5] ——, "Physical diagnosis of coronary artery disease," *Amer. Heart J.*, vol. 80, p. 716, 1970.
[6] G. W. Duncan, J. O. Gruber, C. F. Dervey, G. S. Myers, and R. S. Lees, "Evaluation of carotid stenosis by phonangiography," *New Eng. J. Med.*, vol. 27, p. 1124, 1975.
[7] J. F. Sangster and C. M. Oakley, "Diastolic murmur of coronary artery stenosis," *Brit. Heart J.*, vol. 35, p. 840, 1973.
[8] J. Semmlow, W. Welkowitz, J. Kostis, and J. W. Mackenzie, "Coronary artery disease - correlates between diastolic auditory characteristics and coronary artery stenoses," *Biomedical Engineering, IEEE Transactions on*, vol. BME-30, no. 2, pp. 136 –139, Feb. 1983.
[9] J. Semmlow and K. Rahalkar, "Acoustic detection of coronary artery disease," *Annual Review of Biomedical Engineering*, vol. 9, no. 1, pp. 449–469, 2007.
[10] V. Stahl, A. Fischer, and R. Bippus, "Quantile based noise estimation for spectral subtraction and wiener filtering," in *Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on*, vol. 3, 2000, pp. 1875 –1878 vol.3.