AUTOMATIC CLASSIFICATION OF SYSTOLIC HEART MURMURS

M. Markaki[⋆], *I. Germanakis*,[†] *Y. Stylianou*[⋆]

[⋆] Computer Science Department, University of Crete, Greece
[†] Dept of Pediatrics, Faculty of Medicine, University of Crete, Greece

ABSTRACT

This paper describes a system for discriminating innocent from pathologic systolic heart murmurs in children based on auscultation recordings. For sound signal analysis the use of reassigned spectrogram is suggested. Both dimensions and noise of the time-frequency representation were significantly reduced using higher order singular value decomposition. Optimal dimensions were selected through cross-validation experiments on a database of auscultation recordings with systolic murmurs from the University Hospital of Heraklion. The database only consisted with recordings of high misclassification rate by general practitioners. Using support vector machines for classification, the suggested approach achieved an Equal Error Rate of $6.71 \pm 1.18\%$ and an Area Under the Curve score of 0.9758 ± 0.0053 (95% confidence intervals). The performance of the suggested classification system is comparable to the reported accuracy of experienced pediatric cardiologists on the same database, while it outperforms alternative signal representations based on simple STFT schemes.

Index Terms— Auscultation, heart murmurs, reassigned spectrogram, higher order SVD.

1. INTRODUCTION

Classic heart auscultation using a conventional stethoscope to detect abnormal heart sounds is the most common and widely recommended method to screen for structural abnormalities of the cardiovascular system [1]. Normal blood flow within the heart and vessels is mainly laminar and therefore silent. Passing through abnormal communications or narrowed valves, the blood flow becomes turbulent; vibration generated to surrounding tissues, gives rise to audible additional sounds referred to as murmurs [1]. Heart murmurs are non-stationary signals, characterized by sudden changes in intensity and spectral content, varying temporal duration and position in the cardiac cycle (systolic or diastolic phase) [2–4]. Except for pathological murmurs which are related to underlying heart disease, a significant percentage of children exhibits *innocent* functional murmurs [5]. Moreover, murmurs in the same category (innocent or pathological) differ among individuals due to the variation in the physiology of the heart and other body parts (e.g. chest) [4]..

Detecting relevant symptoms and forming a diagnosis based on the sounds heard through a stethoscope, is a skill that can take years to acquire and refine [6]. It has been proven that general practitioners and pediatricians cannot accurately distinguish pathological heart sounds from innocent functional murmurs [5]. Their diagnostic accuracy is internationally reported as low to moderate. Suboptimal clinical skills of physicians and pediatricians result in inappropriate referrals and misuse of expensive diagnostic methods such as the echocardiogram, which is the predominant noninvasive diagnostic method in cardiology nowadays [5, 6].

An interesting alternative is the remote, on-line or off-line, auscultation in children based on high quality digital recording of heart sounds (phonocardiography) [7–9]. Depending on the digital stethoscope, a sensitivity of 87-100% and a specificity of 82-88% in the discrimination of innocent and abnormal murmurs has been reported when *experienced pediatric cardiologists* evaluate digital recordings referred to as phonocardiograms (PCG) [8].

There have been a few attempts towards *automatic* discrimination between functional and abnormal murmurs based on auscultation recordings. Earlier approaches mostly relied on wavelet analysis for feature extraction [10, 11]. More recently, quadratic energy distributions have been tested and compared to magnitude of Short Time Fourier Transform (STFT) and wavelet transform representations for detection - not classification - of cardiac murmurs [12]; various nonparametric techniques such as Mel frequency cepstral coefficients (MFFC) were used for estimating the spectral power contours which were evaluated as features for classification [12]. It was shown that the quadratic energy distributions combined with MFCC contours achieved the best classification performance [12].

In this study, a system is developed for automatically discriminating innocent from pathologic murmurs based on the reassigned spectrogram [13] of the PCG. Since diastolic murmurs are usually regarded as pathological [5], only systolic murmurs were considered. The short -time Fourier transform (STFT) is used as the input time-frequency representation for the reassignment algorithm. The database consists of PCG

*Thanks to John Latsis Public Benefit Foundation for funding.

recordings of high misclassification rate by general practitioners, from the Pediatric Cardiology Unit of University Hospital of Heraklion, Crete. For the dimensionality reduction of the time - frequency representations higher order singular value decomposition (HOSVD) was employed [14]. The singular value decomposition (SVD) of a time-frequency distribution was first proposed in [15] for the Wigner distribution. By truncating the series after the first few terms of the decomposition, a significant noise reduction is accomplished while retaining most of the signal [16]. Comparing to SVD, reduction in dimensionality of feature space through HOSVD can be performed in every subspace separately. Cross-validation experiments - using Support Vector Machines as classifier - are performed then in order to determine the number of singular vectors to be retained in each subspace.

The paper is organized as follows: Section 2 briefly describes the reassignment method for the time-frequency representation of heart sounds. In Section 3 the HOSVD of the reassigned spectrogram is presented. The pattern classification algorithm and the performance analysis measures are presented next, followed by a discussion of the results and future directions.

2. TIME-FREQUENCY REPRESENTATION OF HEART SOUNDS

To follow the time varying characteristics of the heart sounds, a time-frequency representation (TFR) is required. The most common TFR is that obtained by the short time Fourier Transform (STFT). STFT contains just a few cross-terms arising from interference phenomena; however this is obtained at the expense of poor resolution of the individual components of multi-component signals [13]. The reassignment method was selected to be used since it is specifically designed to overcome this trade-off between resolution in time and frequency [13, 17]. Reassignment aims at sharpening time-frequency and time-scale representations in order to improve their readability [17]. The reassignment method used in this work, initially employs a short-time Fourier transform (STFT), denoted as $\mathcal{F}_x(t,\nu;h)$; t, ν refer to time and frequency, respectively, and h is the frequency analysis window. The length of the analysis window $h(n)$ - as well as the degree of overlap between successive windows - controls the trade-off between resolutions in the time and frequency axes. The reassigned time-frequency distribution uses information from the phase spectrum to sharpen the amplitude estimates for each bin of the STFT [13].

3. MULTILINEAR ANALYSIS OF TIME-FREQUENCY FEATURES

Every PCG segment is represented in the time-frequency plane as a two-dimensional matrix $\mathbf{B} \in \mathbb{R}^{I_1 \times I_2}$, where I_1 and I_2 correspond to the frequency and time dimensions, respectively. Let I_3 denote the number of signal segments contained in the training set. The mean value is computed over I_3, and it is subtracted from all the spectra in the training set. The zero-mean spectra are then stacked, creating the data tensor $\mathcal{D} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$. A generalization of Singular Value Decomposition (SVD) algorithm to tensors referred to as Higher Order SVD (HOSVD) [14] enables the decomposition of tensor $\mathcal{D}$ to its mode$-n$ singular vectors:

$$\mathcal{D} = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)} \tag{1}$$

where $\mathcal{S}$ is the core tensor with the same dimensions as $\mathcal{D}$; $\mathcal{S} \times_n \mathbf{U}^{(n)}$, $n = 1, 2, 3$, denotes the $n-$mode product of $\mathcal{S}$ by matrix $\mathbf{U}^{(n)}$; $\mathbf{U}^{(1)} \equiv \mathbf{U}_\nu \in \mathbb{R}^{I_1 \times I_1}$, $\mathbf{U}^{(2)} \equiv \mathbf{U}_t \in \mathbb{R}^{I_2 \times I_2}$ and $\mathbf{U}^{(3)} \equiv \mathbf{U}_s \in \mathbb{R}^{I_3 \times I_3}$ are the unitary matrices of the corresponding subspaces of frequency, time and samples. Each matrix $\mathbf{U}^{(n)}$ contains the n-mode singular vectors (SVs):

$$\mathbf{U}^{(n)} = \left[U_1^{(n)} \; U_2^{(n)} \; \ldots \; U_{I_n}^{(n)} \right]. \tag{2}$$

The contribution $\alpha_{n,j}$ of the j^{th} n-mode singular vector $U_j^{(n)}$ is defined as a function of its singular value $\lambda_{n,j}$:

$$\alpha_{n,j} = \lambda_{n,j} \; / \sum_{j=1}^{I_n} \lambda_{n,j} \tag{3}$$

By setting a threshold, the R_n, $n = 1, 2$ singular vectors (SVs) whose contribution exceeds that threshold are retained. The truncated matrices $\hat{\mathbf{U}}^{(1)} \equiv \hat{\mathbf{U}}_\nu \in \mathbb{R}^{I_1 \times R_1}$ and $\hat{\mathbf{U}}^{(2)} \equiv \hat{\mathbf{U}}_t \in \mathbb{R}^{I_2 \times R_2}$ are obtained then. Time-frequency representations $\mathbf{B} \in \mathbb{R}^{I_1 \times I_2}$ extracted from phonocardiographic signals are projected on $\hat{\mathbf{U}}_\nu$ and $\hat{\mathbf{U}}_t$:

$$\mathbf{Z} = \mathbf{B} \times_1 \hat{\mathbf{U}}_\nu^T \times_2 \hat{\mathbf{U}}_t^T = \hat{\mathbf{U}}_\nu^T . \mathbf{B} . \hat{\mathbf{U}}_t \tag{4}$$

where $\mathbf{Z} \in \mathbb{R}^{R_1 \times R_2}$, $R_1, R_2 \ll I_1, I_2$ [14].

4. EXPERIMENTS AND RESULTS

4.1. Database

Pediatric cardiology outpatients referred for murmur evaluation or followed-up for heart disease at the University Hospital of Heraklion, Crete, which had undergone a complete echocardiographic study, were eligible for enrollment in the study. Therefore, the database is accordingly labelled (ground truth is known). Informed consent was obtained from the parents.

A sensor based electronic stethoscope with incorporated 3 lead electrocardiograph (ECG) was used (TheStethoscope, Welch Allyn-Meditron, AS). Five recordings of 6 sec duration each were performed, corresponding to the apical, lower left sternal border (LLSB), upper left (ULSB) and upper right sternal border (URSB) as well as a jugular fossae recording

(JGLR). PCG and ECG signals were synchronously recorded using the designated software (Meditron Analyzer 4) with 44,100 KHz sampling rate and 16 bit dynamic resolution. Recordings were evaluated by a pediatric cardiologist (2nd co-author); if three or more recordings in a given case could not be accurately interpreted due to low quality, this case was excluded from the original database.

The final database includes 25 cases with innocent murmurs and a normal echocardiographic study, and 25 patients with abnormal systolic murmurs and an echocardiographic diagnosis of a congenital heart disease.

4.2. Preprocessing of PCG recordings

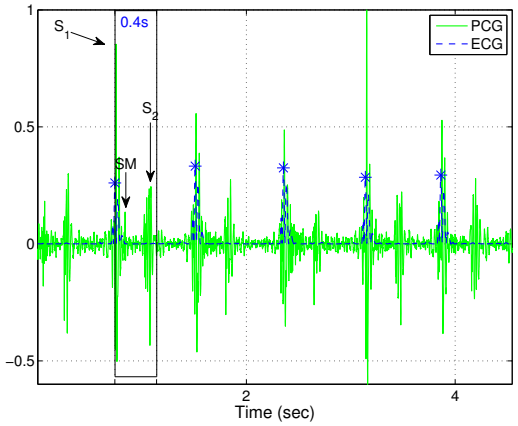

Fig. 1. PCG signal (solid) with innocent murmur (SM) and envelope of ECG signal (dash). Stars at maxima of ECG envelope (R-peaks), point to the beginning of heart cycles.

The PCG-synchronous ECG signals were used as reference for detecting heart sounds since the R-peaks in ECG coincide with the beginning and end points of heart cycles. Both ECG and PCG signals were down sampled to the same sampling frequency of 2205 Hz. The ECG signal was band-pass filtered with cut-off frequencies of 10 and 25 Hz and the R-peaks were detected using an envelope based detection algorithm [18]. The R-R intervals were used then as a reference for heart cycle detection and segmentation of the PCG signals [19]. Heart cycles have different lengths depending on the cardiac rhythm. Heart rates in children are quite variable depending on their age and physiological state (fever, fear, activity level etc). The length of the systole however is quite stable as increased heart rate is achieved physiologically at the cost of reduced diastolic (relaxing) interval [19]. For example, if the heart rate varies between 60 and 120 beats per minute (with corresponding heart cycles between 1 sec and 500 ms), then the systole length will vary between 291 and 231 ms respectively [19]. We have chosen then to analyze a fixed-length segment comprising of the first 400 ms of the heart cycle. In our database, this 400ms segment always included the first heart sound, S_1, the whole systolic phase and the second heart sound, S_2 (Figure 1).

For every recording, five consequent heart cycles were considered. PCG signals were band-pass filtered with cut-off frequencies of 40 and 1100 Hz [19], since the maximum of frequency spectrum of heart sounds reaches up to 850 Hz [1]. The 3rd order Butterworth type high-pass and low-pass filters were used. The amplitude of each signal was scaled then by the absolute maximum of each PCG recording [20] (see Figure 1). Reassignment spectra were computed using the Time - Frequency Toolbox [21]. Figure 2 (a) depicts the reassigned spectrogram amplitude of the PCG of a 10-years old, with an innocent systolic murmur of intensity equal to 3 in a 3-degree scale. Figure 2 (b) depicts a pathologic systolic murmur in the PCG of an 11-years old child with atrial septal defect (ASD). The auscultation area was the same in both cases (LLSB). Due to the variability of different heart pathologies, the case of pathological murmur cannot be considered "typical": frequencies in the systolic phase are below the 200 Hz limit, resembling an innocent sound according to the common screening criteria [20]. Still, they are sustained for the whole systolic duration. Also the second heart sound spectro-temporal content is a characteristic of ASD [20].

For the computation of the singular matrices of HOSVD, a random subset of 15 subjects with innocent and 15 subjects with pathological murmurs was selected once; three recordings - with five consequent heart cycles each - were considered for each subject. For every contribution threshold, the resulting matrices $\mathbf{Z} \in \mathbb{R}^{R_1 \times R_2}$ (according to eq (4)) were subsequently reshaped into vectors before classification. Classification was carried out using Support Vector Machine (SVM) classifiers. SVM find the optimal boundary that separates two classes maximizing the margin between separating boundary and closest samples to it (support vectors) [22]. In this work, SVMlight [22] with a Radial-Basis-Functions kernel was used. A 5-fold stratified cross-validation was used, repeated 25 times. In every run, the classifier was trained on the 80% of subjects of both classes, then tested using the remaining 20%. Training and testing was based on 400ms segments from a single heart cycle. Classification of PCG recordings was based on the median of the decisions over segments from five consequent heart cycles of this recording.

The system performance was evaluated using the Receiver Operating Characteristic (ROC) curve as well as the detection error trade-off (DET) curve between false rejection rate (or miss probability P_{miss}, equal to one minus sensitivity) and false acceptance rate (or false alarm probability P_{false}, equal to one minus specificity) [23]. The optimal detection accuracy (DCF_{opt}) occurs when the threshold is set such that the total number of errors is minimized. The Equal Error Rate (EER) refers to the point at the DET curve where the false-alarm probability equals the miss probability. Stratified

Table 1. Classification scores with 95% confidence intervals based on RSP and STFT for one or five heart cycles (one recording)

	AUC	$DCF_{opt}(\%)$	$P_{miss}(\%)$	$P_{false}(\%)$	EER (%)
one heart cycle - RSP	0.8938±0.0066	85.49±0.80	13.14	16.02	17.34±0.90
one heart cycle - STFT	0.8534±0.0062	80.11 ±0.57	19.05	20.49	22.62±0.79
one recording - RSP	0.9758±0.0053	95.87±0.72	4.36	3.93	6.71±1.18
one recording - STFT	0.9274 ±0.0076	90.78± 0.67	5.53	12.95	13.78 ±1.07

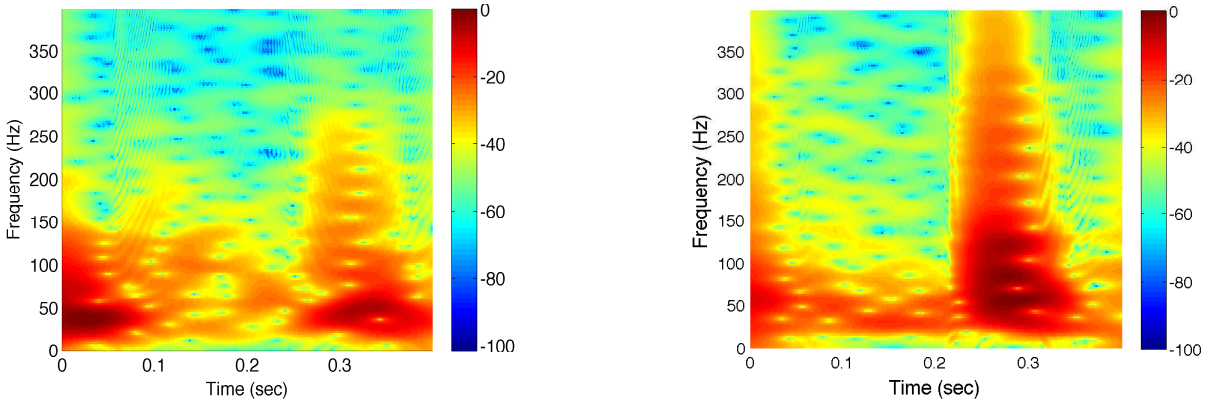

Fig. 2. Energy (in dB) of the reassigned spectrogram of the PCG (a) of a 10-years old with loud innocent murmur, (b) of an 11-years old with atrial septal defect and pathological murmur.

cross-validation was performed while varying the contribution threshold $\alpha_{n,j}$ (eq. 3) and retaining the R_1, R_2 singular vectors whose contribution exceeded that threshold. The area under the ROC curve (AUC) was the criterion for selecting the optimal dimensions R_1, R_2 for the classification of the TFR projections (eq. 4). The best system for reassigned spectrogram (RSP) used $[20 \times 19]$ dimensions (380 features vs $[440 \times 294]$ initially) which corresponded to an $\alpha_{n,j} = 0.02\%$ and an AUC of 0.9758 ± 0.0053 (95% confidence interval, CI). For STFT the best system used $[14 \times 12]$ dimensions (168 features vs $[513 \times 12]$ initially) which corresponded to an $\alpha_{n,j} = 0.6\%$ and an AUC of 0.9274±0.0076 (95% CI). Table 1 provides the average of classification score both for RSP and STFT, in terms of the AUC, DCF_{opt} (with corresponding P_{miss} and P_{false}), EER and the corresponding 95% CI as these were estimated from the 25 runs. The scores per heart cycle and per recording (five consequent heart cycles) are provided. The more pronounced effect of increased time resolution of reassigned TFR compared to STFT spectrogram is the significant reduction of false alarm rate when taking the median of decisions over five consequent heart cycles for each recording.

5. DISCUSSION

In this study we aimed to develop and validate the clinical efficacy of an automatic detection and classification system of pediatric heart sounds based on reassigned spectrogram. Selected cases were in the gray area of easily misclassified systolic murmurs by general practitioners, being either loud innocent murmurs radiating to all sites, or abnormal murmurs associated with mild to moderate heart defects. Normal subjects without any heart murmur were not considered.

Our results were comparable to the accuracy achieved by experienced pediatric cardiologists on the same database [9]. More specifically, 96% of moderate to severe congenital heart disease and more than 92% of children with functional murmurs were accurately detected in blind, off-line auscultation by two experienced pediatric cardiologists. The corresponding scores of our system were 95.87% and 96.07%, i.e., system specificity was better. System sensitivity could be further improved by employing additional information such as the variability of the wide split of the second heart tone with respiration, or the variable intensity of the first heart sound between consequent heart cycles [19, 24].

6. REFERENCES

[1] A. Tlkian and M. Conover, *Understanding Heart Sounds and Murmurs: With an Introduction to Lung Sounds*, W. B. Saunders Co., Philadelphia, 2001.

[2] S.M. Debbal and F. Bereksi-Reguig, "Time-frequency analysis of the first and the second heartbeat sounds," *Applied Mathematics Computing*, vol. 2, no. 184, pp. 1041–1052, 2007.

[3] D. Kumar, P. Carvalho, M. Antunes, and J. Henriques, "Noise detection during heart sound recording," in *IEEE EMBC09*, 2009, pp. 3119–3123.

[4] C. Mahnke, "Automated heartsound analysis/computer-aided ausculation: A cardiologist's perspective and suggestions for future development," in *IEEE EMBC09*, 2009.

[5] K. Rajakumar, M. Weisse, A. Rosas, E. Gunel, L. Pyles, W.A. Neal, A. Balian, and S. Einzig, "Comparative study of clinical evaluation of heart murmurs by general pediatricians and pediatric cardiologists," *Clinical Pediatrics*, vol. 38, no. 9, pp. 511–518, 1999.

[6] S. Mangione, "Cardiac auscultatory skills of physicians-in-training: a comparison of three english-speaking countries," *Am J Med*, vol. 110, pp. 210–216, 2001.

[7] L.B. Dahl, P. Hasvold, E. Arild, and T. Hasvold, "Heart murmurs recorded by a sensor based electronic stethoscope and e-mailed for remote assessment," *Arch Disease Childhood*, vol. 87, no. 4, pp. 297–301, 2002.

[8] J.P. Finley, A.E. Warren, G.P. Sharratt, and M. Amit, "Assessing children's heart sounds at a distance with digital recordings," *Pediatrics*, vol. 118, no. 6, pp. 2322–2325, 2006.

[9] I. Germanakis, S. Dittrich, R. Perakaki, and M. Kalmanti, "Digital phonocardiography as a screening tool for heart disease in childhood," *Acta Pædiatrica*, vol. 51, no. 3, pp. 327–333, 2008.

[10] C.S. Hayek, W.R. Thompson, C. Tuchinda, R.A. Wojcik, and J.S. Lombardo, "Wavelet processing of systolic murmurs to assist with clinical diagnosis of heart disease," *Biomed Instrum Technol.*, vol. 37, no. 4, pp. 263–270, 2003.

[11] W.R Thomson, C.S. Hayek, C. Tuchinda, J.K. Telford, and J.S. Lombardo, "Automated cardiac auscultation for detection of pathologic heart murmurs," *Pediatric Cardiology*, vol. 22, pp. 373–379, 2001.

[12] A. F. Quiceno-Manrique, J.I. Godino-Llorente, M. Blanco-Velasco, and G. Castellanos-Dominguez, "Selection of dynamic features based on time-frequency representations for heart murmur detection from phonocardiographic signals," *Ann Biomed Eng*, vol. 38, no. 1, pp. 118–137, 2010.

[13] P. Flandrin, F. Auger, and E. Chassande-Mottin, *Applications in Time-Frequency Signal Processing*, chapter Time-Frequency Reassignment - from principles to algorithms, pp. 179–203, CRC Press, 2003.

[14] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, pp. 1253–1278, 2000.

[15] N.M. Marinovich and G. Eichmann, "An expansion of the Wigner distribution and its applications," in *Proc. ICASSP*. IEEE, 1985, vol. 3, pp. 1021–1024.

[16] Cohen L., *Time-Frequency Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1995.

[17] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1068–1089, 1995.

[18] M.-E. Nygards and L. Sörnmo, "Delineation of the qrs complex using the envelope of the e.c.g.," *Med. & Biol. Eng. & Comput.*, vol. 21, no. 5, pp. 538–547, 1983.

[19] M. El-Segaier, O. Lilja, S. Lukkarinen, L. Sörnmo, R. Sepponen, and E. Pesonen, "Computer-based detection and analysis of heart sound and murmur," *Ann. Biomed. Eng.*, vol. 33, no. 7, pp. 937–942, 2005.

[20] A.-L. Noponen, S. Lukkarinen, A. Angerla, and R. Sepponen, "Phono-spectrographic analysis of heart murmur in children," *BMC Pediatrics*, vol. 7, no. 23, pp. 1–10, 2007.

[21] F. Auger, P. Flandrin, P. Goncalves, and O. Lemoine, "Time-frequency toolbox for MATLAB, Users' Guide and Reference Guide," http://tftb.nongnu.org, 2005.

[22] T. Joachims, *Advances in Kernel Methods - Support Vector Learning*, chapter Making large-scale SVM Learning Practical, MIT-Press, 1999.

[23] A. Martin, G.R. Doddington, T. Kamm, M. Ordowski, and M. Przybocki, "The DET curve in assessment of detection task performance," in *Proc. Eurospeech '97*, 1997, vol. IV, pp. 1895–1898.

[24] Z. Syed, D. Leeds, D. Curtis, F. Nesta, R.A. Levine, and J. Guttag, "A framework for the analysis of acoustical cardiac signals," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 4, pp. 651–662, 2007.