A SURVEY OF AUDIO PROCESSING ALGORITHMS FOR DIGITAL STETHOSCOPES

Fabio de Lima Hedayioglu, Miguel Tavares Coimbra
Instituto de Telecomunicações, Faculdade de Ciências da Universidade do Porto
Rua do Campo Alegre, 1021/1055, 4169 - 007 Porto, Portugal

Sandra da Silva Mattos
Fetal and Pediatric Cardiology Unit (UCMF) at Real Hospital Português de Benecifiencia em Pernambuco
Av. Portugal, 163 - Recife, PE Brazil

Keywords: Digital stethoscope, Audio processing.

Abstract: Digital stethoscopes have been drawing the attention of the biomedical engineering community for some time now, as seen from patent applications and scientific publications. In the future, we expect 'intelligent stethoscopes' to assist the clinician in cardiac exam analysis and diagnostic, potentiating functionalities such as the teaching of auscultation, telemedicine, and personalized healthcare. In this paper we review the most recent heart sound processing publications, discussing their adequacy for implementation in digital stethoscopes. Our results show a body of interesting and promising work, although we identify three important limitations of this research field: lack of a set of universally accepted heart-sound features, badly described experimental methodologies and absence of a clinical validation step. Correcting these flaws is vital for creating convincing next-generation 'intelligent' digital stethoscopes that the medical community can use and trust.

1 INTRODUCTION

Auscultation is one of the oldest, cheapest and most useful techniques for the diagnosis of heart disease. Since their invention in 1816, stethoscopes have been used as part of the initial evaluation of all patients with suspected heart or lung problems. An experienced physician can diagnose a large number of clinical conditions just from the initial auscultation of the patient's chest (Tilkian and Conover, 1984). There have been several attempts to create electronically enhanced stethoscopes, with better sound amplification and frequency response. However, and according to Durand (Durand and Pibarot, 1995), their introduction into clinical practice has been hindered by factors such as their background noise, unfamiliar sounds to clinicians due to filtering or fragility and bad ergonomic design. Recent advances in electronics and digital circuits allow us to not only overcome these problems but also to exploit the benefits of digital signal processing for signal analysis and visualization. In this paper we will embrace this novel perspective and analyze the state-of-the-art in audio processing of heart sounds that might be adequate for integrating into this next generation of stethoscopes. A deeper explanation of digital stethoscopes is given in Section 2, a review of audio processing methods is described in Section 3, followed by a discussion (Section 4) on the future of digital stethoscopes and how biomedical engineers can contribute to the success of this technology.

2 DIGITAL STETHOSCOPE

It is essential that we define digital stethoscope since there can be various interpretations from the name alone. Traditional stethoscopes depend solely on acoustics to amplify and transmit the heart sounds to the clinician. The concept of electronic stethoscope arrives when electronic components were first used to amplify, filter and transmit the sound (Fig.1) (Durand and Pibarot, 1995).

There are various examples in literature regard-

de Lima Hedayioglu F., Tavares Coimbra M. and da Silva Mattos S. (2009).
A SURVEY OF AUDIO PROCESSING ALGORITHMS FOR DIGITAL STETHOSCOPES.
In *Proceedings of the International Conference on Health Informatics*, pages 425-429
DOI: 10.5220/0001512104250429

Figure 1: Lab prototype of an electronically enhanced stethoscope.

ing the development of digital and electronic stethoscopes. Bredesen and Schmerler (Bredesen and Schmerler, 1991) have patented an “intelligent stethoscope” designed for performing auscultation and for automatically diagnosing abnormalities by comparing digitized sounds to reference templates using a signature analysis technique. Several other electronically enhanced and digital stethoscopes have been developed and described in literature (F. L. Hedayioglu, 2007; M.E. Tavel and Shander, 1994; Durand and Pibarot, 1995; Brusco and Nazeran, 2005).

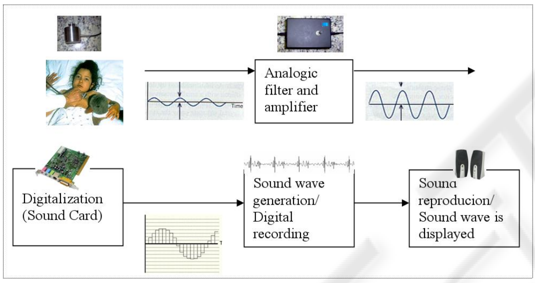

Figure 2: Block diagram of a digital stethoscope prototype developed by our group and field-tested at Real Hospital Português de Beneficiência em Pernambuco in Recife, Brazil. Over 100 auscultations were performed during the clinical validation stage.

Fig. 2 shows a block diagram of a digital stethoscope prototype developed by our group. The auscultation quality was considered satisfactory in clinical trials when compared to auscultation using acoustic stethoscopes, but clinicians still perceived differences in audio pitch although this did not affect their ability to diagnose heart conditions. Our field experience confirms Durand’s (Durand and Pibarot, 1995) opinion that audio enhancement alone is not enough for the clinical community to adopt this new technology. In order to make digital stethoscopes attractive to clinical cardiologists, we clearly need to address the numerous potential improvements provided by a fully functional, robust digital stethoscope: real-time acquisition, analysis, display and reproduction of heart sounds and murmurs. Digital stethoscopes must also open the doors for digital audio archives, simplifying the acquisition, storage and transmission process of cardiac exams and murmurs, potentiating functionalities such as the teaching of auscultation, telemedicine, and personalized healthcare.

3 AUDIO PROCESSING

For the analysis of the state-of-the-art on audio processing in cardiology, we have very loosely adopted some concepts of clinical systematic reviews. A rigorous systematic review of such a multi-disciplinary vast field is quite difficult to implement in practice due to the large number of papers retrieved by analysis of both engineering and medical scientific databases. Our review methodology was as follows:

- Considered that Durand’s (Durand and Pibarot, 1995) excellent review paper fully covers this topic up to 1995.
- Consulted the IEEE Xplore (ieeexplore.ieee.org) database with the following query: “(((feature extraction)<in>metadata) <and> ((cardiology)<in> metadata))”, obtaining 159 results after 1995.
- By title and abstract inspection, we kept only papers dealing with phonocardiogram data analysis, reducing this number to 19.
- We analyzed the references from all these papers, and selected all papers published after 1995 and with more than 10 citations, obtaining 20 results. This enabled us to cover additional articles besides the ones published in IEEE journals and conferences, artificially expanding the scope of our review to other scientific databases.
- The total number of papers covered by this review is thus 39 (19+20).

Although we are certain that it is possible to miss some papers using this methodology, we feel that we have covered a sufficiently vast and interesting sample to draw some important conclusions, as described in Section 4.

3.1 Heart Sound Analysis and Feature Extraction

The main constituents of a cardiac cycle are the first heart sound (typically referred to as S1), the systolic period, the second heart sound (S2) and the diastolic period. Whenever a clinician is performing an auscultation, he tries to identify these individual components, and is trained to analyze related features such as rhythm, timing instants, intensity of heart sound components, splitting of S2, etc (H. Liang and Hartimo, 1997b). This analysis allows him to search for murmurs and sound abnormalities that might correspond to specific cardiac pathologies. From a signal processing perspective, Heart Sound Analysis (HSA) is not only interesting by itself (allowing quantitative measures to be displayed automatically in a digital stethoscope), but is also an essential first step for the subsequent task of automatic pathology classification. In this paper, we will distinguish two sub-tasks of HSA: Heart Sound Segmentation (HSS) and Aortic Pulmonary Signal Decomposition (APSD).

3.1.1 Heart Sound Segmentation

In HSS we expect to identify and segment the four main constituents of a cardiac cycle. This is typically accomplished by identifying the position and duration of S1 and S2, using some sort of peak-picking methodology on a pre-processed signal. Liang (H. Liang and Hartimo, 1997a) has used discrete wavelet decomposition and reconstructed the signal using only the most relevant frequency bands. Peak-picking was performed by thresholding the normalized average Shannon energy, and discarding extra peaks via analysis of the mean and variance of peak intervals. Finally, they distinguish between S1 and S2 peaks (assuming that the diastolic period is longer than the systolic one, and that the later is more constant), and estimate their durations. A classification accuracy of 93% was obtained on 515 periods of PCG signal recordings from 37 digital phonocardiographic recordings. The same authors further improved the statistical significance of their results by obtaining the same accuracy using 1165 cardiac periods from 77 recordings (H. Liang and Hartimo, 1997b), and later attempted murmur classification based on these features and neural network classifiers, obtaining 74% accuracy (Liang and Hartimo, 1998b). Omran (Sherif Omran, 2003) has also studied this problem using normalized Shannon entropy after wavelet decomposition of the audio signal, but their experimental methodology is not so convincing.

3.1.2 Aortic Pulmonary Signal Decomposition

Besides the four main components of the cardiac cycle, there is a clinical interest in the analysis of some of its associated sub-components (JingPing Xu, 2000). It has been recognized that S1 may be composed of up to four components produced during ventricular contraction (Durand and Pibarot, 1995), although the complexity of this task has been a very difficult hurdle for the signal processing community. The S2 sound is more well known, being composed of an aortic component (A2), which is produced first during the closure and vibration of the aortic valve and surrounding tissues, followed by the pulmonary component (P2) produced by a similar process associated with the pulmonary valve (JingPing Xu, 2000). Durand (JingPing Xu, 2000) demonstrated that it is possible to model each component of S2 by a narrow-band nonlinear chirp signal. Later (JingPing Xu, 2001) he adapted and validated this approach for the analysis and synthesis of overlapping A2 and P2 components of S2. To do so, the time-frequency representation of the signal is generated and then estimated and reconstructed using the instantaneous phase and amplitude of each component (A2 and P2). In this paper the accuracy evaluation was made by a simulated A2 and P2 components having different overlapping factors. The reported error was between 1% and 6%, proportional to the duration of the overlapping interval. Nigam (Nigam and Priemer, 2006) also presented a method for extracting A2 and P2 components by assuming them as statistically independent. To do so, four simultaneous auscultations are analyzed using blind source separation. The main advantage of this method is the lower dependence on the A2-P2 time interval, although it needs a non-conventional 4-sensor stethoscope. Leung (T. S. Leung, 1998) also analyzed the splitting of S2 using time-frequency decomposition.

3.2 Automatic Pathology Classification

The vast majority of papers we have found regarding audio processing algorithms, adequate for the integration into a digital stethoscope, concern the detection of specific heart pathologies. This highlights the interest of the scientific community on this topic but, as our analysis shows, there are still some major flaws in most of them such as the absence of a clinical validation step and unconvincing experimental methodologies. Most papers use the well-established pattern recognition approach of feature extraction followed by a classifier. Due to space limitations, we will describe the most interesting papers, leaving a more detailed discussion on this topic to Section IV. Bentley (P. M. Bentley and Grant, 1995) uses Choi-Williams Distribution (CWD) as features, working with 45 normal/abnormal valve subjects. Some features were determined via visual inspection, others automatically from the CWD by simple rule-based classification. Latter (P. M. Bentley and McDonnell, 1998), the authors show that CWD is a better method to represent the frequencies in PCG and to get heart sound descriptors, than other time-frequency (T-F) representations. According to them, a simple description of the T-F distribution allows an analysis of the heart valve's condition. However, they highlight the need of a more comprehensive evaluation using a larger population of test patients. Wang (P. Wang and Soh, 2005) proposes a representation of heart sounds that is robust to noise levels of 20dB, using mel-scaled wavelet features. However, details regarding the used dataset are not clear enough for robust conclusions. Liang (Liang and Hartimo, 1998a) developed an interesting feature vector extraction algorithm where the systolic signal is decomposed by wavelets into subbands. Then, the best basis set is selected, and the average feature vector of each heart sound recording is calculated. Neural Networks (NN) are used for classifying 20 samples after being trained with 65, obtaining an accuracy of 85%. NNs are also used by Abdel-Alim (Onsy Abdel-Alim and El-Hanjouri, 2002) for the automatic diagnostics of heart valves using wavelets feature vectors and stethoscope location information. They use two NNs: one for systolic diseases and the other for diastolic diseases. A total of 1200 cases were used: 970 cases for training and 300 for testing. The recognition rate was 95%. Turkoglu (Turkoglu and Arslan, 2001), Ozgur (Ozgur Say and Olmez, 2002) and El-Hanjouri (M. El-Hanjouri and Alim, 2002) also used wavelets as feature vectors for classification, although they provide too few details regarding the used data sets. Trimmed mean spectrograms are used by Leung (T.S. Leung and Salmon, 2000) to extract features of phonocardiograms. Together with the acoustic intensities in systole and diastole, the authors quantified the distinctive characteristics of different types of murmurs using NNs. One of the few papers that is conscious about the important clinical validation step is from Kail (E Kail and Balázs, 2004). The authors propose a novel sound representation (2D and 3D) and feature extraction algorithm using Morlet wavelet scalograms. After manual classification of the resulting graphs performed by two cardiologists on 773 subjects, they clinically validated the features as useful for sound and murmur extraction. Sharif (Zaiton Sharif and Salleh, 2000) also proposes other features for classification systems based on central finite difference and zero crossing frequency estimation.

4 DISCUSSION

By covering the most interesting papers on audio-processing from a digital stethoscope perspective, we can make some observations regarding the state-of-the-art on this field. Section 3.1 has shown us that there are already important results regarding audio feature extraction. The S1 and S2 sounds can be robustly segmented and there is promising work regarding the extraction of secondary sounds such as A2 and P2.

The scenario is not so bright for automatic pathology classification (Section 3.2). Reviewing some of the papers and simply observing the disparity in the number of publications when compared with the other challenges, we conclude that there is a strong interesting in this topic. However, in our opinion, there is still a long way to go before we can have robust automatic classification systems that can be introduced in the clinical routine of hospitals. We have identified three major problems that afflict most of the papers reviewed:

- Absence of a set of well-accepted features - We rarely found papers that selected the same features for pathology classification. Most acknowledge that the presence of S1 and S2 is important but there is no consensus of the scientific community on how these should be used. We have collected more than 25 different features with minimum overlap between papers. We clearly need more studies on the statistical significance and clinical importance of heart sound features, from an automatic pattern recognition perspective.
- Badly descript data-sets - It is not enough for authors to mention that they have worked with 300 cardiac cycles. Where were these obtained? From how many patients? In what conditions? Using which equipment? All these factors are vital in the analysis of a system's performance and robustness. Studies need to be much more rigorous on this topic so their results can be reasonably convincing.
- Absence of clinical validation - Almost no papers bothered to handle this vital task of all assisted-diagnostic systems. No medical specialist will trust any kind of automatic system without it proving to be robust and accurate in real field testing. These conditions are very different from a typical biomedical engineering research lab, which can drastically affect results.

As a final conclusion, we can say that working towards next-generation 'intelligent' digital stethoscopes is highly desirable judging from the significant number of scientific publications on this topic but also examining the undeniable benefits that such systems can provide. There is already solid work regarding audio feature extraction and many unsolved challenges in this field such as the complex analysis of the sub-components of S1. Automatic pathology classification is still too undeveloped to be of any practical usage and we hope that the valuable lessons learned from this study can correct previous mistakes and provide a precious boost to the challenging field of audio processing for digital stethoscopes.

ACKNOWLEDGEMENTS

This work was supported by the Programme Alban, the European Union Programme of High Level Scholarships for Latin America, scholarship no. E07M402298BR.

REFERENCES

Bredesen, M. and Schmerler, E. (1991). Us patent no. 5,010,889: Intelligent stethoscope.

Brusco, M. and Nazeran, H. (2005). Development of an intelligent pda-based wearable digital phonocardiograph. In *Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*.

Durand, L.-G. and Pibarot, P. (1995). Digital signal processing of the phonocardiogram: Review of the most recent advancements. *Critical Reviews in Biomedical Engineering*.

E Kail, S Khoór, B. K. K. F. and Balázs, F. (2004). Internet digital phonocardiography in clinical settings and in population screening. *Computers in Cardiology*.

F. L. Hedayioglu, S. S. Mattos, L. M. M. E. d. L. (2007). Development of a tele-stethoscope and it's application in pediatric cardiology. *Indian Journal of Experimental Biology*, 45.

H. Liang, S. L. and Hartimo, I. (1997a). Heart sound segmentation algorithm based on heart sound envelogram. *Computers in Cardiology*, 24.

H. Liang, S. L. and Hartimo, I. (1997b). A heart sound segmentation algorithm using wavelet decomposition and reconstruction. In *19th International Conference - IEEE/EMBS*, Chicago, IL, USA.

JingPing Xu, L.G. Durand, P. P. (2000). Nonlinear transient chirp signal modeling of the aortic and pulmonary components of the secound heart sound. *IEEE Transactions on Biomedical Engineering*, 47(7).

JingPing Xu, L.G. Durand, P. P. (2001). Extraction of the aortic and pulmonary components of the secound heart sound using nonlinear transient chirp signal model. *IEEE Transactions on Biomedical Engineering*, 48(3).

Liang, H. and Hartimo, I. (1998a). A feature extraction algorithm based on wavelet packet decomposition for heart sound signals. In *Proceedings of the IEEE-SP International Symposium*.

Liang, H. and Hartimo, I. (1998b). A heart sound feature extraction algorithm based on wavelet decomposition and reconstruction. In *Proc. IEEE EMBS*, volume 20.

M. El-Hanjouri, W. Alkhaldi, N. H. and Alim, A. (2002). Heart diseases diagnosis using hmm. In *Electrotechnical Conference -MELECON*.

M.E. Tavel, D. B. and Shander, D. (1994). Enhanced auscultation with a new graphic display system. In *Arch. Intern. Med.*, volume 154, page 893.

Nigam, V. and Priemer, R. (2006). A procedure to extract the aortic and the pulmonary sounds from the phonocardiogram. In *Proceedings of the 28th IEEE EMBS Annual International Conference*, New York City, USA.

Onsy Abdel-Alim, N. H. and El-Hanjouri, M. A. (2002). Heart diseases diagnosis using heart sounds. In *Radio Science Conference*.

Ozgur Say, Z. D. and Olmez, T. (2002). Classification of heart sounds by using wavelet transform. In *Proceedings of the Second Joint EMBS/BMES Conference*, volume 1.

P. M. Bentley, J. T. E. M. and Grant, P. M. (1995). Classification of native heart valve sounds using the choi-williams time-frequency distribution. In *IEEE-EMBC and CMBEC*.

P. M. Bentley, P. M. G. and McDonnell, J. T. E. (1998). Time-frequency and time-scale techniques for the classification of native and bioprosthetic heart valve sounds. *IEEE Transactions on Biomedical Engineering*, 45(1).

P. Wang, Y. K. and Soh, C. B. (2005). Feature extraction based on mel-scaled wavelet transform for heart sound analysis. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference*.

Sherif Omran, M. T. (2003). A heart sound segmentation and feature extraction algorithm using wavelet. In *Proc. of IEEE MWSCAS '03*, volume 1, pages 27–30.

T. S. Leung, P. R. White, J. C. W. B. C. E. B. A. P. S. (1998). Analysis of the secound heart sound for diagnosis of paediatric heart disease. In *IEE Proceedings - Sci. Meas. Technol.*, volume 145.

Tilkian, A. and Conover, M. (1984). *Understanding heart sounds and murmurs with an introduction to lung sounds*. W.B. Saunders Company.

T.S. Leung, P.R. White, W. B. C. E. B. and Salmon, A. P. (2000). Classification of heart sounds using time-frequency method and artificial neural networks. In *Proceedings of the 22nd Annual International Conference of the IEEE*, volume 2.

Turkoglu, I. and Arslan, A. (2001). An intelligent pattern recognition system based on neural network and wavelet decomposition for interpretation of heart sounds. In *Proceedings of the 23rd Annual International Conference of the IEEE*, volume 2, pages 25–28.

Zaiton Sharif, Mohd Shamian Zainal, A. Z. S. and Salleh, S. H. S. (2000). Analysis and classification of heart sounds and murmurs based on the instantaneous energy and frequency estimations. In *Proceedings of TENCON*, volume 2.