International Journal of Advances in Bio-Informatics and Bio-Technology– IJABB
Volume 1 : Issue 2 [ISSN 2475-2258]
Publication Date : 30 September, 2014

A Timbre Model Based on Multidimensional Features for Analyzing Heart Sound

HaiYang WANG, GuangPei LI, BinBin FU and MingChui DONG

***Abstract*—Cardiovascular diseases (CVDs) are currently the leading cause of deaths worldwide. The traditional auscultation is cost-effective and time-saving for the public to diagnose CVDs early. While many approaches in analysis of the heart sound (HS) signal from auscultation have been utilized successfully, few studies are focused on acoustic perspective to interpret the HS signal. This paper creatively proposes a timbre model to interpret HS with much more attentions on feature extractions combined with medical knowledge. The timbre model based on multidimensional features has three dimensions, which are spectral centroid (*SC*), log attack time (*LAT*) and temporal centroid (*TC*). The simulation experiments indicate that the proposed model is of use in HS feature extraction and its following CVD diagnosis.**

***Keywords*—acoustics, feature extraction, Hilbert transform, spectral centroid, temporal centroid**

I. Introduction

Cardiovascular diseases (CVDs) have become a great threat to human' lives aggressively. A good way to prevent the death caused by CVDs is early discovery and interposition. Over the years, in spite of the advent of echocardiogram (ECHO), electrocardiogram (ECG), photoplethysmogram (PPG), sphygmogram (SPG), etc., the heart auscultation signal is still one of the most primary physiological signals. Physicians primarily analyze it thus to find signs of pathologic conditions, as it can provide clues to the diagnosis of many cardiac abnormalities including valvular heart disease, congestive heart failure and congenital heart lesions before requesting for ECHO, ECG, PPG, SPG, etc. Particularly, in remote areas or less developed regions, auscultation may be the only means available. This leads the research on heart sound (HS) signal analysis and interpretation in order to provide a cost-effective and time-saving prognostic approach for the victims of CVDs.

As for the analysis of HS, many approaches have been proposed in feature exaction in the literatures such as wavelet decomposition and reconstruction method [1-3], short time Fourier transform (STFT) method [4, 5], and S-transform method [6, 7]. Most solely analyze the time frequency domain for feature extractions [6,7]. This leads to a conclusion that feature extractions are less aligned with medical knowledge.

A doctor diagnoses the CVDs generally according to a stethoscope from acoustic perspective. Common descriptive terms about what it sounds like in auscultation include rumbling, blowing, machinery, scratchy, harsh, gallop, ejection, click, drum, cymbal, etc. From acoustic perspective to analyze bio-signal, its main advantage is that engineering is aligned with medicine. For instance, continuous machinery sound is heard on the left sternal border between the second and third ribs in auscultation indicating the patent ductus arteriosus. From acoustic perspective, the mel-frequency cepstral coefficient (MFCC) method has been utilized for HS feature extractions [8-10]. MFCC is based on the theory that human audition spaces linearly at low frequency band and logarithmically at high frequency. As there exist two inverse-transforms in MFCC, it encounters computation complexity. There exists another approach named timbre from acoustic perspective, which has not been utilized and can avoid computation complexity. In acoustics，timbre is a significant attribute of three acoustic attributes, which embodies the texture of acoustic source. CVDs, as the pathological changes in heart and blood vessels, provide different timbre information. As a result, the timbre could be well suitable for HS feature extractions. However, the current timbre analysis algorithms are all aimed at different music instrument recognition. Psychoacoustic literatures show the dimensions of timbre are fundamental frequency (*FF*), log attack time (*LAT*), temporal centroid (*TC*), spectral centroid (*SC*), harmonics, etc. [11-14]. Upon request, different dimensions are selected to extract various HS features. Compared with MFCC, timbre extraction can save plenty of computation.

A timbre model based on multidimensional features is creatively proposed to interpret HS in this paper. It is observed from simulations that the proposed model is of convenient usage in HS analysis with diagnostic significance from HS.

II. A Timbre Model Based on Multidimensional Features

A. Timbre Model for Music

Timbre expresses normally the quality of sound which distinguishes one music instrument from another, while there are a wide variety of instrument families and individual categories. Timbre is defined by ANSI as the attribute of auditory sensation, based on which a listener can judge quality of two sounds presented similarly and having the same loudness and pitch. Therefore, timbre is subjective and must be parameterized very carefully thus to apply it on automatic timbre recognition. In recent years, researchers have

---

HaiYang WANG
University of Macau
Macau, SAR, China

Binbin FU
University of Macau
Macau, SAR, China

GuangPei LI
University of Macau
Macau, SAR, China

MingChui DONG
University of Macau
Macau, SAR, China

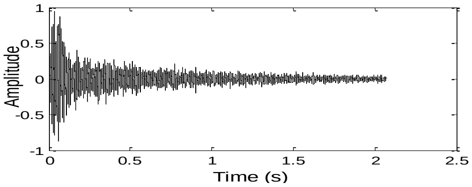

Figure 1. A drum sound signal

extensively utilized various acoustical features to build computational models for automatic music timbre estimation [15-17]. As the timbre is complex, it is difficult to describe through only one parameter. Consequently, timbre models based on multidimensional features are utilized, which reveal continuous perceptual dimensions correlated with acoustic parameters corresponding to spectral, temporal and temporal-spectral properties of the sound events. In some literatures these acoustic parameters are called as parameters of timbre space [18]. Using such models, judging the degree of similarity and dissimilarity between timbres of instruments becomes a relatively simple task. The criticism to these models is that user can get only what put in. Such models were applied to measure musical timbre by Plomp (1970), subsequently by Wessel (1973), Miller and Carterette (1975), and Lakatos (2000).

B. *Features Selected for Constructing Timbre Model for HS*

Extracting HS features is not as easy in distinguishing common descriptive terms as what it sounds like in auscultation. The first research task is to explore depending on what features and how to construct timbre model to distinguish the sound of drum or cymbal. The cymbal is plate type and drum is membrane type. Both belong to percussive instruments. Unfortunately, due to the indefinite pitch and the signal complexity shown in Fig.1, most proposed timbre models for music give much fewer attentions to percussion instruments, though they could deal with string, wind harmonic instruments properly [12-16]. In this paper, *SC*, *LAT* and *TC* are selected to construct the timbre model for HS.

III. Features of the Timbre Model

Actually, *SC* reflects the signal power distribution in frequency domain. *LAT*, which is also called rising time in other literatures, is of great use in feature extraction on morphology. *TC* reflects the strength of the signal distribution in time domain, which is the geometric centroid of the signal in temporal domain.

A. *SC*

SC is computed as power spectrum weighted average of the frequency in power spectrum. It reflects the signal power distribution in frequency.

$$S(k)=\sqrt{\sum_{i=1}^{M}P_i(k)/M} \tag{1}$$

$$SC=\frac{\sum_{K=1}^{NFFT} f(k)\cdot S(k)}{\sum_{K=1}^{NFFT} S(k)} \tag{2}$$

where M is the total number of frames in a sound segment, $P_i(k)$ is the k^{th} power spectrum coefficient in the i^{th} frame, $f(K)$ is the k^{th} frequency bin.

B. *LAT*

LAT is the very short period of time that elapses before the sound has formed its spectrum. How instruments are identified usually depends on *LAT*.

It is defined as the logarithm of duration between the time where signal starts and the time where it reaches its stable part.

$$LAT=\log_{10}(T_1-T_0) \tag{3}$$

where T_0 is the time when signal starts. Here we use the time while the amplitude is 0.001 dB. T_1 is the time when signal reaches its sustained part of maximum part.

C. *TC*

TC is defined as time average over the energy envelope of signal.

$$TC=\frac{\sum_{n=1}^{length(SEnv)} n/s_r\cdot SEnv(n)}{\sum_{n=1}^{length(SEnv)} SEnv(n)} \tag{4}$$

where s_r *is* the sample rate. The envelope is computed through Hilbert transform. *TC* reflects the energy distribution in the temporal domain.

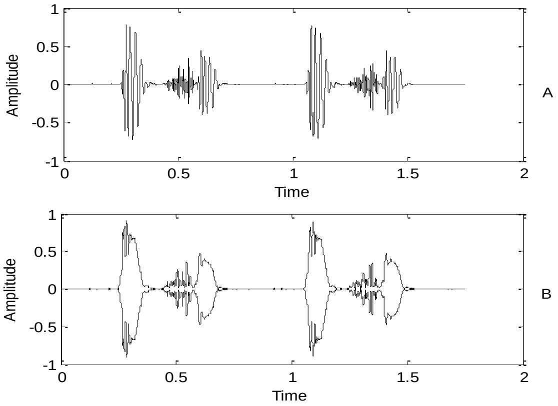

Figure 2. HS signal and envelop computed by Hilbert transform: A is the original HS signal; B is its envelope computed by Hilbert transform

IV. Analysis of Simulation Results

A. *Setting of Environment*

The data used in tests are a set of different eight cymbals and seven drums, comprised of five tom drums and two snare drums. The recorded samples have divergent loudness levels. Moreover, the signal recorded only reflects one stroke on the percussive instruments.

The sounds are digitized using a sampling frequency of 44100 Hz. The spectrograms are computed with Fast Fourier Transform (FFT). All processing of the signals are performed in MATLAB (The Math Works, Inc., Natick, MA, USA).

B. Simulation Results and Discussions

The simulation results of cymbals are shown in Table I, where *LAT* values are from 2.6053s to 3.9668s. The simulation results of drums are shown in Table II, where *LAT* values are from 0.6554s to 2.6352s. Their ranges are different with small ranges overlapped. And, the average *LAT* of drums is smaller than cymbals'. It is observed from simulations that the smaller *LAT* is, the stronger the sense of stroke a listener heard is. In Table I, both No.4 and No.5 sound like a slight rub on the cymbal surfaces by a stick rather than a stroke on them by a stick. Therefore *LAT* could be utilized in distinguishing ejection, gallop, rumble, etc., as their dashing degrees are various in acoustic perceptions. In Table I, *SC* values of cymbals are all around 10^4 Hz. Instead, in Table II, *SC* values of drums are all around 10^3Hz. *TC* values of cymbal are one order bigger than drum's. Finally come to *TC*. *TC value* ranges of cymbals and drums have large intersections shown clearly in Table I and Table II. Consequently *TC* performs badly in distinguishing cymbal and drum comparing with previous two dimensional features. However, even *TC* is not effective in extracting timbre of cymbal and drum, it is rather crucial in extracting features of HS. For instance, for normal and different abnormal HSs shown in Fig. 3, all of them have obvious different geometric centers. So *TC* provides significant information in diagnosing aortic stenosis with ejection murmur, mitral regurgitation with pan systolic murmur, aortic regurgitation with early diastolic murmur, mitral stenosis with open snap and patent ductus arteriosus, etc. The geometric center of the click rhythms HS is also divergent observed from *TC* values of sounds A to F shown in Fig. 3. Therefore *TC* is a promising clue in CVDs diagnosis.

TABLE I. SIMUALTIOAN RESULTS OF CYMBALS

No.	Dimensions		
	***TC* (s)**	***LAT* (s)**	***SC*(×10^4 Hz)**
1	0.2888	2.6053	2.11
2	0.2918	2.8686	1.88
3	0.2025	3.2462	2.50
4	0.8863	3.8400	3.47
5	1.7590	3.9668	2.62
6	0.2135	3.1274	3.12
7	0.0918	3.1337	1.76
8	0.2043	3.0459	2.25
Mean Value	0.49223	3.2292	2.46

TABLE II. SIMUALTIOAN RESULTS OF DRUMS

No.	Dimensions		
	***TC* (s)**	***LT* (s)**	***SC* (×10^3 Hz)**
1	0.2175	2.6352	2.37
2	0.0760	0.8132	1.91
3	0.1288	1.8316	2.61
4	0.1526	0.6554	8.21
5	0.4057	1.6066	2.07
6	0.3632	2.2642	1.80
7	0.2943	2.5789	3.88
Mean Value	0.2340	1.7693	3.26

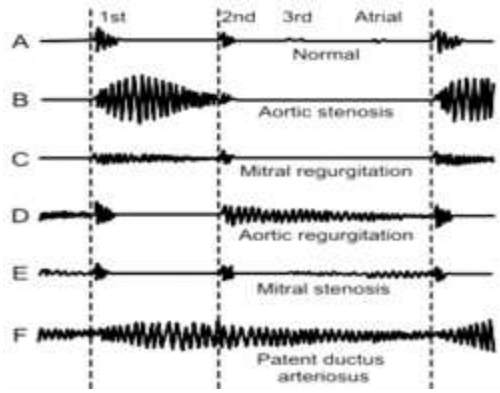

Figure 3. Normal HS and pathological HS categories

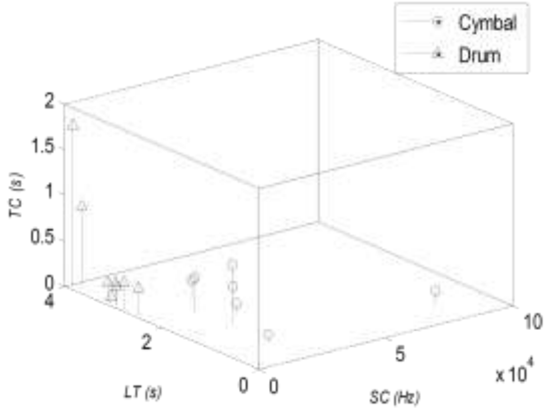

Figure 4. Multidimensional features of drums and cymbals represented in 3-D diagram

The multidimensional features of drums and cymbals are described in 3-D diagram in Fig. 4. The timbre model based on multidimensional features distinguishes them undoubtedly.

V. Conclusion

A timbre model based on multidimensional features for analyzing HS from acoustic perspective is creatively proposed. The very encouraging results were observed in this research. The model presented in this paper has potential talent in extracting timbre features of HS and provide theoretical basis of a novel CVDs diagnosis method.

The future research work will be concentrated on employing such a timbre model to analyze HS signals and realize CVDs diagnosis.

Acknowledgment

This work was supported in part by the Research Committee of University of Macau under Grant MYRG184 (Y2-L3)-FST11-DMC, and Science and Technology Development Fund (FDCT) of Macau S.A.R under Grant 018/2009/A1, respectively.

References

[1] Z. Dokur, T. Ölmez, Digital Signal Processing, vol.19, 2009, pp.521-531

[2] H. Liang, “feature extraction algorithm based on wavelet packet decomposition for heart sound signals,” Time-Frequency and Time-Scale Analysis, 1998 [Proceedings of the IEEE-SP International Symposium 1998]

[3] P.Vikhe , S. Hamde, N. Nehe,“Wavelet Transform Based Abnormality Analysis of Heart Sound,” Advances in Computing, Control, & Telecommunication Technologies, 2009 [ACT 9th. International Conference on, pp.367-371, 2009.]

[4] W. Kao, C. Wei, J. Liu, P. Hsiao. Automatic heart sound analysis with short-time Fourier transforms and support vector machines. Circuits and Systems, 2009. MWSCAS 9th. 52nd IEEE International Midwest Symposium, pp.188-191, 2-5 Aug. 2009.

[5] A. DJEBBARI F. BEREKSI REGUIG, “Short-time Fourier transform analysis of the phonocardiogram signal,” the 7th IEEE International Conference on Electronics, Circuits and Systems, 2000. pp. 844- 847 vol.2

[6] Hadi, H.M., Mashor, M.Y., Suboh, M.Z., Mohamed, M.S. Classification of heart sound based on s-transform and neural network. Information Sciences Signal Processing and their Applications (ISSPA), 2010 [10th International Conference, pp.189-192, 10-13 May 2010]

[7] M. Varanini, G. De Paolis, M . Emdin, Macerata, Pola S, Cipriani M,Marchesi C. “Spectral Analysis of Cardiovascular Tme Series by theS-Transform,”. Computers in Cardiology 1997; 24:383-86.

[8] C. Sunita, et al. “A computer-aided MFCC-based HMM system for automatic auscultation.” Computers in Biology and Medicine 38.2 (2008): 221-233.

[9] F. Wenjie, X. Yang Y. Wang. “Heart sound diagnosis based on DTW and MFCC.” Image and Signal Processing (CISP), 2010 3rd International Congress on. Vol. 6. IEEE, 2010.

[10] I. Kamarulafiza, et al. “Heart sound analysis using MFCC and time frequency distribution.” World Congress on Medical Physics and Biomedical Engineering 2006. Springer Berlin Heidelberg, 2007

[11] J. Grey,. (1975). An exploration of musical timbre. Ph.D. dissertation (Dept. of Music Report No. STAN-M-2, Stanford University, CA) (unpublished)

[12] Stephen lakatos, “ A common perceptual space for harmonic and percussive timbres,” Perception & Psychophysics 2000,62 (7) 1426-1439

[13] Stephen McAdams, Computer Music Journal, vol.23. MIT Press Cambridge,1999, pp. 85-102

[14] H. V. Gianpaolo Evangelista,“Separation of harmonic instruments with overlapping partialsin muti-channel mixtures,” 2003 IEEE Workshop an Applications of Signal Prwessinp to Audio and Acoustics.

[15] K. Jensen ,“Timbre Models of Musical Sounds” Nr. 99/7 ISSN 0107-8283

[16] S. McAdams, “Perspectives on the Contribution of Timbre to Musical Structure”computer Music journal, Vol.23, No.3, Recent Research at IRCAAAM (Autumn, 1999).

[17] Anne Caclin, “Acoustic correlates of timbre space dimensions: A confirmatory” 18 April 2005

About Authors:

HaiYang WANG received the B.Sc degree in automation from Beijing Union University in 2012. He is pursing a Master’s degree in the University of Macau. His research interest is biomedical processing and applications etc. His postal address is Faculty of Science and Technology, University of Macau, Macau

GuangPei LI is now pursuing for his bachlor’s degree of electrical and computer engineering in University of Macau. His interest is in biomedical engineering and microprocessors. His postal address is Faculty of Science and Technology, University of Macau, Macau

BinBin FU is currently perusing PhD degree in ECE department of University of Macau. Her research interests focus on embedded system, biomedical signal processing and artificial intelligence, towards the goal of synthesizing them for wearable/portable monitoring with computerized risk stratification..

MingChui DONG is a Full Professor and PhD supervisor of ECE Department of University of Macau, Automation Department of Tsinghua University, China and Professor of YanTai University, China. He received MSc degree in EEE in 1975 at Tsinghua University, China, Visiting Scholar in EEE in 1981 at Rome University, Italy.

His main research interests are AI and its application in biomedical engineering, CIMS, fault diagnosis, AI in vision-to-text, voice-to-text machine translation etc.