

# AUDIO-VISUAL BASED RECOGNITION OF AUSCULTATORY HEART SOUNDS WITH FOURIER AND WAVELET ANALYSES

Fumio Nogata<sup>1</sup>, Yasunai Yokota<sup>2</sup>, Yoko Kawanura<sup>2</sup>, Hiroyuki Morita<sup>3</sup>, Yoshiyuki Uno<sup>3</sup>, W. R. Walsh<sup>4</sup>

<sup>1</sup>Department of Human and Information Systems, <sup>2</sup>Department of Information Science, <sup>3</sup>Graduate School of Medicine, Gifu University, Gifu, Japan

<sup>4</sup>Department of Surgery, Prince of Wales Clinical School, University of New South of Wales, Austraira  
E-mail: [nogata@gifu-u.ac.jp](mailto:nogata@gifu-u.ac.jp)

Received December 2010, Revised December 2011, Accepted January 2012

## Abstract

Since there are various difficulties associated with auscultation techniques (e.g., the detection/recognition of murmurs and sound tone changes within approximately one second), we have proposed an audio-visual based technique to examine and visualize heart sounds for both physicians and patients. To overcome auscultation difficulties, the technique can be used to assist in the understanding of the heartbeat, the detection of heart disease, and the digital database management of the auscultation examination. Auscultatory sounds are visualized by both an FFT image and a wavelet image to detect any abnormal heart sounds. A simple technique of pattern classification has been established using short-time Fourier transform and wavelet analyses to detect abnormal sounds. This new technique is expected to be simple and practical in this era of computer-managed clinical data. The result indicates that there is a possibility of developing a fully automatic detection system based on a map of standard sounds in the near future. A simple auscultation method is also expected to be developed for in-home use.

**Keywords:** Image recognition, auscultatory sound, heart murmur, frequency analysis

## 1. Introduction

With the invention of the stethoscope by René Laennec in 1816, "auscultation" became possible, introducing an exciting and practical new method of bedside examination [1]. Auscultation is performed to examine the circulatory, respiratory, and gastrointestinal (for bowel sounds) systems. It requires substantial clinical experience and good listening skills. When many medical students first attempt to use a stethoscope in clinical settings to detect/diagnose diseases of the heart and lungs, they often have difficulties hearing the characteristic sounds. As a result, there is concern about potential missed diagnoses. In fact, the shared bands (frequency and sound pressure) between audible sounds and heart sounds are very narrow (Fig. 1); thus, special techniques are required for auscultation. One solution is the amplification of heart and breathing sounds (to shift shared bands upward). However, most doctors do not use a stethoscope with an amplification unit due to changes in tone.

In the era of computer-managed clinical data, it is important to establish a new technique to analyze heart and breath sounds digitally, which will allow for better understanding by both physicians and patients. Some reports on the visualization, wavelet-based imaging, mapping, and computer-aided diagnosis of cardiac disorders [2-9] have been published. However, simple and practical techniques that use low-cost devices have not been established.

Therefore, we propose a technique for visualizing heart sounds to assist in the evaluation of heartbeats by both physicians and patients that is based on pattern recognition to detect abnormal sounds using Fourier transform and wavelet analyses. Additionally, a feasibility assessment is performed using thirteen types of heart diseases that are recorded on the training CD [10] for auscultation techniques.

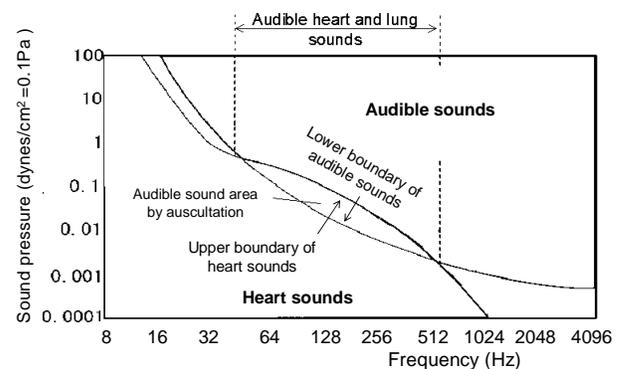


Fig.1: Diagram showing frequency bands of audible and heart sounds. Note that the audible sound area (40-600 Hz) by auscultation is small.

## 2. Heart Sound Visualization Procedures Using Short-time Fourier Transform and Wavelet Analyses

A heart that is functioning normally produces two basic sounds: the first sound (S1) and the second sound (S2). Extra heart sounds are also heard in both normal and abnormal situations and can include heart murmurs, adventitious sounds, and gallop sounds as the third sound (S3) and the fourth sound (S4). To help understand the visualization technique proposed, a summary of heart sounds is provided below:

S1 is produced by the closure of the mitral and tricuspid valves at the beginning of ventricular contraction, systole, reverse blood flow, and vibration of the ventricle walls caused by increasing pressure.

S2 is produced by the closure of the aortic and pulmonary valves at the end of systole (i.e., the beginning of ventricular diastole) when ventricular pressure falls rapidly, causing slight backflow of blood from the aorta and the pulmonary artery. The pressure drop due to temporary backflow of blood and the recoiling events cause the aortic and pulmonary valves to close. S3 is produced at the beginning of diastole after S2 and is lower in pitch than S1 and S2. It is thought to be caused by the oscillation of blood between the ventricular walls initiated by the inrushing of blood from the atria. S3 is known to be benign in children, youth, healthy people younger than age twenty, and some trained athletes.

S4 is produced when the ventricles fill and stretch, just after arterial contraction at the end of diastole and immediately before S1. The combined presence of S3 and S4 is a quadruple gallop at rapid heart rates.

Figure 2 is a schematic graph showing these sounds within the time frame of one heartbeat (0.8-1.0 s), which includes the four kinds of sounds. The time between S1 and S2 is about 0.6 s in systole, and the time between S3 and S4 is about 0.4 s in diastole. As shown in Fig. 2(b), there are murmurs during systole and diastole as well as four kinds of periodic changes in sound tone: crescendo, decrescendo, crescendo-decrescendo, and continuous. A physician is required to recognize and diagnosis murmur sounds, which are defined as periodic sounds of short duration of cardiac or vascular origin, within about one second. Therefore, both stethoscope sounds and visual-based evidence are needed. To evaluate auscultatory sounds, including sound volume, intonation, and tone, time-frequency analysis is needed. Frequency range of the heart sound can roughly divide four bands as under 100, 200-300, 300-400 and over 400Hz. From a viewpoint of visual based recognition technique for detecting heart murmur, it needs to show overall and specific frequency bands. Therefore, two techniques were used: short-time Fourier transform (FFT) with a Gauss window and continuous wavelet transform (CWT) with a complex Morlet function; these are powerful tools to analyze spectrum and time-frequency in nonlinear sequence.

### FFT Analysis

The Fourier transform is computed as a fixed-length window slides along the time axis, resulting in a two-dimensional representation of the signal. One of the weak points of short-time Fourier transform (FFT) with a Gauss window is that it has a fixed resolution; the width of the windowing function relates to how the signal is represented. A wide window gives good frequency resolution but poor time resolution, while a narrow window gives good time resolution but poor frequency resolution. The analysis condition is shown in Table 1.

### CWT Analysis

Wavelet transformation has the ability to simultaneously clarify the spectral and temporal information within the signal [11]. This method overcomes the basic shortcoming of Fourier analysis, which is that the Fourier spectrum contains only globally averaged information, which leads to location-specific features in the signal being lost. If the scale parameter of the mother wavelet function is constant, it gives the same analysis

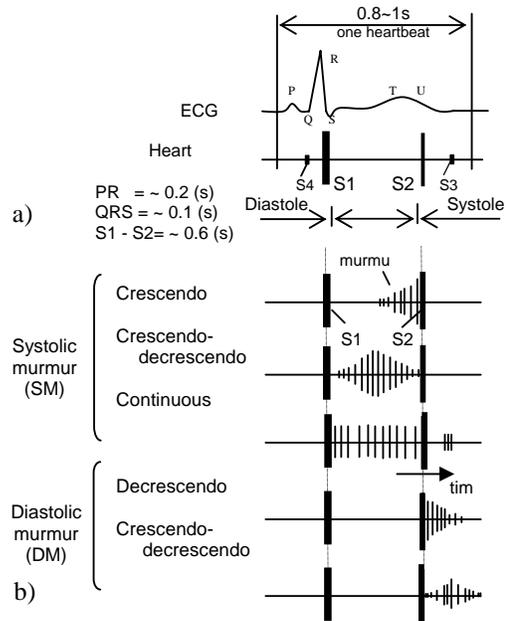


Fig.2: a) Relative timing of ECG graph and heart sounds at the apex and b) types of periodic murmur sounds (SM and DM), where S1, S2, S3, and S4 refer to the 1st, 2nd, 3rd, and 4th heart sounds, respectively (as seen in [10]).

window width for all frequencies. For analyzing heart sounds, a small window width is needed in higher frequency bands. Then, the complex Morlet function [11, 12] and a scaling parameter are applied. The relationship between the frequency ( $F$ ) and the center frequency ( $F_c$ ) is given by

$$F = \Delta \cdot F_c / a_n, \quad \text{---- (1)}$$

where  $\Delta$  is the sampling frequency,  $F_c$  is the center frequency, and the scaling parameter  $a_n$  is given by a numerical sequence

$$a_n = a_1 + k \cdot n(n-1), \quad \text{---- (2)}$$

$k$  = arbitrary constant,  $n$  = time of scaling.

Figure 3 shows a relation between frequency and scale,  $a_1 = 5$ ,  $k = 0.001$ ,  $\Delta = 5\text{kHz}$  and  $n=121$ . Thus, the technique of wavelet transform allows analysis of sounds in detail at lower frequency bands with larger value of  $a_n$ . In this report, all wavelet images were shown for the range of 30-500Hz by nonlinear scale as in Fig. 3 with  $a_n = 10 - 150$ .

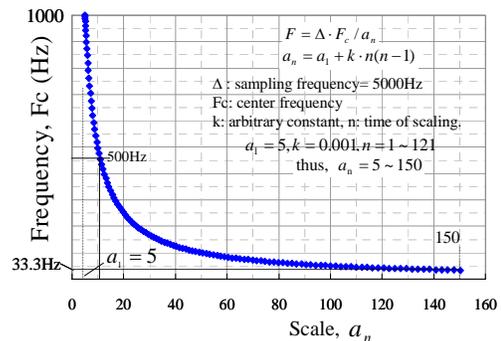


Fig.3: The relationship between frequency ( $F$ ) and scale  $a_n$  in Eq. 1:  $a_1 = 5$ ,  $k = 0.001$ , and  $n=121$ .

**3. Specimens**

There are many types of heart diseases, including coronary artery disease, irregular heart rhythm, heart valve disease, congenital heart disease, heart muscle disease (cardiomyopathy), dilated cardiomyopathy, and hypertrophic cardiomyopathy. In order to demonstrate the audio-visual based recognition technique proposed, thirteen heart sound examples recorded in the auscultation training textbook (with CD) [10] were used (Table 1). The analysis conditions are shown in Table 2.

**4. Results and Discussion**

The thirteen examples are discussed as follows with respect to the frequency, contour, and distribution of the two types of images:

**1) Healthy heart sounds at the apex and at the left sternal border (LSB) of the 4th rib**

In general, physicians auscultate at five locations related to common diseases as the first step. Two types of healthy heart sounds were analyzed at the apex and at the LSB of the 4th rib, and these results are shown in Figs. 4(a) and (b). Clear signal images for S1 and S2 can be seen; and the amplitude of S1 is larger than that of S2 at the apex, while S1 and S2 are approximately the same amplitude at the LSB of the 4th rib. This is a typical heart sound described as a "lub-dup," and there are no heart murmurs. The frequency, amplitude, and time interval between S1 and S2 and between S2 and the next S1, which relate to heart function, can be estimated. From an STFT image, it is easy to understand that S1 has four frequency bands (approximately 80, 250, 450, and 600 Hz) and S2 has two frequency bands (80 and 200 Hz); also, the sounds below 100 Hz have higher intensities than the high-frequency sounds. In the wavelet image, the relationship between scale and frequency is nonlinear, as shown in Fig. 3, where the lower frequency band is shown in detail. Thus, a wavelet image makes distributive analysis possible as a pattern which focused frequency using appropriate value of the  $a_n$  in Eq. 1. Comparing the two sounds at the apex and the LSB, there is a clear difference in the frequency distribution pattern. Physicians can confirm their diagnoses using both the auscultatory sounds and the audio-visual results, and they can re-examine the recorded heart sounds on a personal computer for educational purposes.

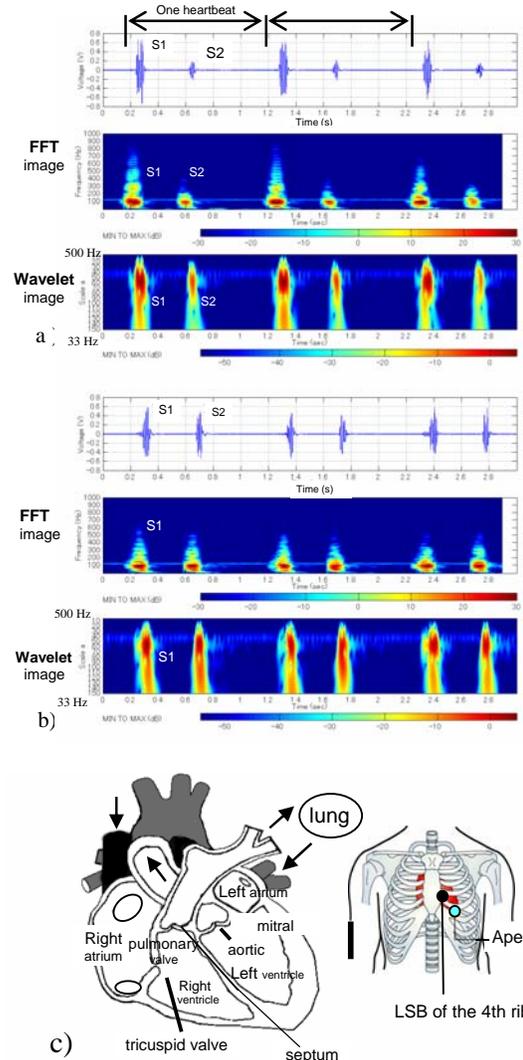


Fig.4: Audio-visual based recognition images analyzed for healthy heart sounds; a) at the apex and b) at the LSB of the 4th rib, c) an illustrative figure of the heart and the locations of auscultation.

Table 2 Conditions of FFT and wavelet analyses (sampling frequency:  $\Delta=5$  KHz)

FFT analysis	Wavelet analysis
Gauss window	Complex Molet Wavelets
Size= 512 points	Time of Scaling (n)=121, a1=5
%overlap rate=98	Coefficient, k=0.01
	Analyzed frequency=33~1000 Hz

**2) Mitral regurgitation (MR) sound at the apex: 1st degree**  
 Insufficiency of the mitral valve makes a systolic murmur (SM) due to reverse flow between S1 and S2, but only a small amount (i.e., mild case). The center frequency of the SM is 200-350 Hz from the FFT image.

**3) Mitral regurgitation (MR) sound at the apex: 2nd degree**  
 In addition to the SM in the 1st degree case, the slight sound of S3 appears after S2. The STFT and wavelet images show a clear pattern, and the frequency of S3 is about 50 Hz in the FFT image.

**4) Mitral regurgitation (MR) sound at the apex: 3rd degree (serious case)**

In this disease, the mitral valve does not close properly when the heart pumps out blood. This causes abnormal leaking of blood through the mitral valve and into the left atrium when the left ventricle contracts [12]. Two types of murmurs occur after S1: an SM and a diastolic murmur (DM); almost noise sounds are existed in a heartbeat. The frequencies from the FFT image are as follows: S1 is 50-100 Hz, SM is 200-300 Hz, S3 is 80 Hz, and DM is 100 Hz. Clear differences in frequency are seen between S3 and DM from the wavelet image as well.

**5) Mitral stenosis (MS) sound with DM and OS (opening snap of the atrioventricular valve) at the apex**

The mitral valve becomes thickened and immobile due to the narrowing of the valve and, as a result, does not completely open. When this occurs, blood tends to backup in the left atrium, leading to increasing pressure in the chamber [12]. Over long periods of time, significant problems can result. Many types of murmurs are mixed and overlap continuously at

frequencies from 100-200 Hz, which, unfortunately, is the same frequency range as S1 and S2.

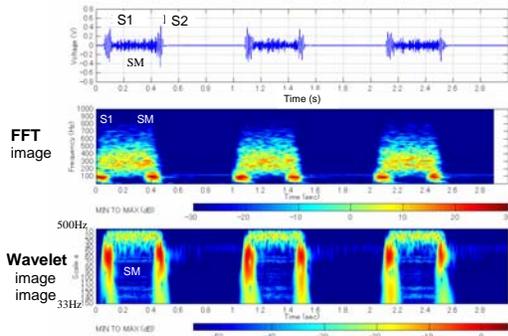


Fig.5: Audio-visual based recognition images analyzed for MR: 1st degree.

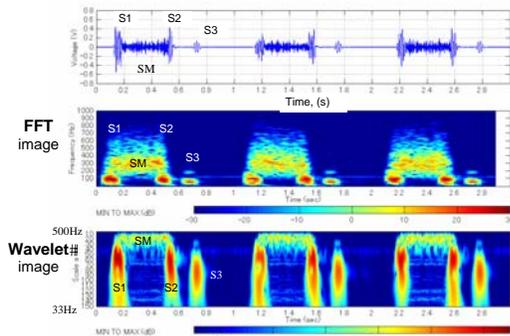


Fig.6: Audio-visual based recognition images analyzed for MR: 2nd degree.

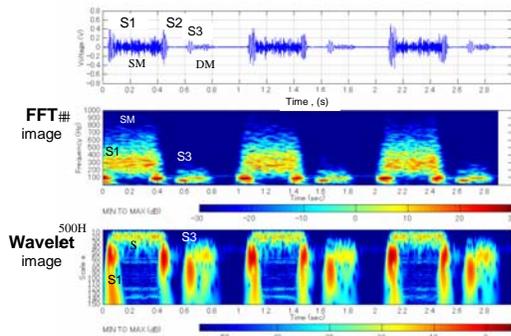


Figure 7: Audio-visual based recognition images analyzed for MR: 3rd degree.

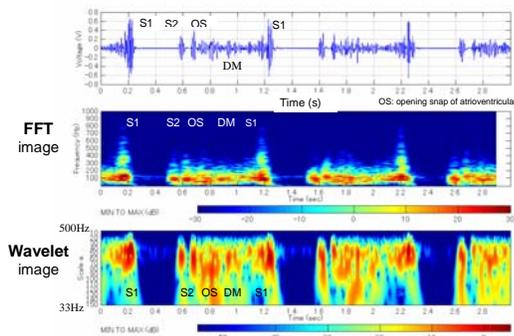


Fig.8: Audio-visual based recognition images analyzed for MS with DM and OS.

**6) Mitral stenosis and insufficiency (MSI) sound at the apex**

In general, all cases of mitral stenosis are due to heart disease secondary to rheumatic fever and the consequent rheumatic heart disease [13]. The sound includes SM, OS, and DM. The frequency bands of all sounds (S1, S2, OS, SM, and DM) are about 100 Hz but are easily distinguishable in these results.

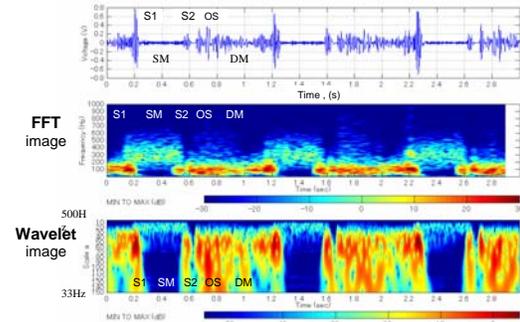


Fig.9: Audio-visual based recognition images analyzed for MSI.

**7) Ventricular septal defect (VSD) sound at the apex: 3rd degree (serious case)**

A VSD is a defect in the ventricular septum. The murmur depends on the abnormal flow of blood from the left ventricle, through the septal defect, and to the right ventricle. If there is not much difference in pressure between the left and right ventricles, it may be silent [13]. In the 3rd degree case, there are three types of murmur sounds: S3, SM, and DM

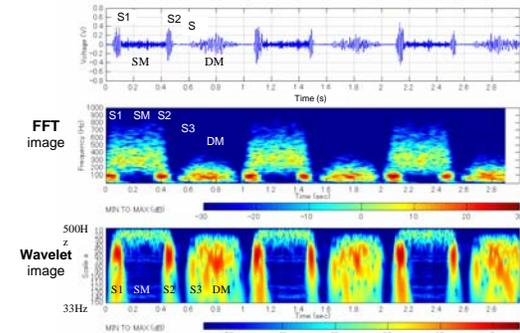


Fig.10: Audio-visual based recognition images analyzed for VSD: 3rd degree.

**8) Hypertrophic obstructive cardiomyopathy (HOCM)**

HOCM is an obstructive variant of hypertrophic cardiomyopathy (HCM). Due to the distortion of the normal heart anatomy, blood outflow from the left ventricle of the heart is obstructed. This is a systolic murmur, and its sound level depends on the position of the patient (i.e., standing, sitting, lying down, or crouching down) [10]. Low sound levels of S4 and SM are also included.

**9) Hypertrophic cardiomyopathy (HCM) – asymmetric septal hypertrophy sound at the apex**

HCM is a disease that affects the heart muscle. It causes thickening of the heart muscle (e.g., the ventricles or lower heart chambers), left ventricular stiffness, mitral valve changes, and cellular changes. The auscultatory sound includes a weak S4, which is a typical sound for HCM. However, there is an

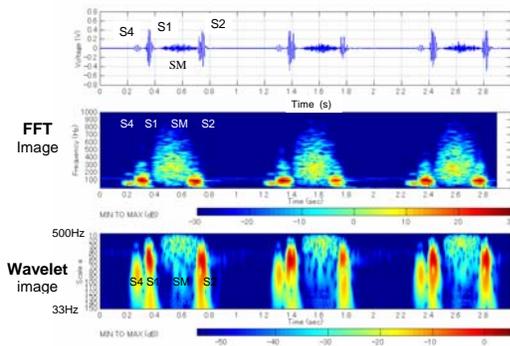


Fig.11: Audio-visual based recognition images analyzed for HOCM.

inaudible case due to a very small S4 [13]. Figure 12 shows typical HCM with an asymmetric septal hypertrophy sound with a slight S4 signal and an S3 noise-contour with a frequency of approximately 50 Hz. Physicians may recognize S4 in the STFT and wavelet images. This is one example showing the benefit of this audio-visual based technique

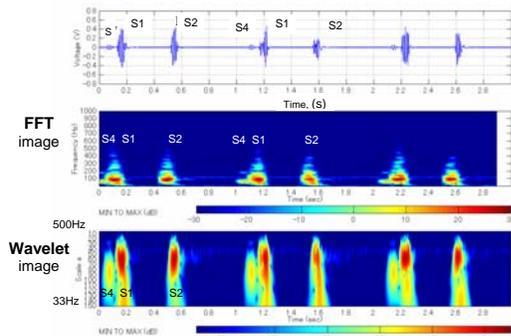


Fig. 12: Audio-visual based recognition images analyzed for HCM – asymmetric septal hypertrophy.

**10) Dilated cardiomyopathy with S3 sound at the apex**

Dilated cardiomyopathy was previously called congestive cardiomyopathy, and its auscultatory sound includes S3 after S2. However, for shorter heartbeats, it can become difficult to distinguish between S3 and S4, thus producing a single sound called a summation gallop. In this case, the audio-visual based recognition technique becomes extremely beneficial, as shown in Fig. 13. S3 may be normal in people under 40 years of age and in some trained athletes but should disappear before middle age.

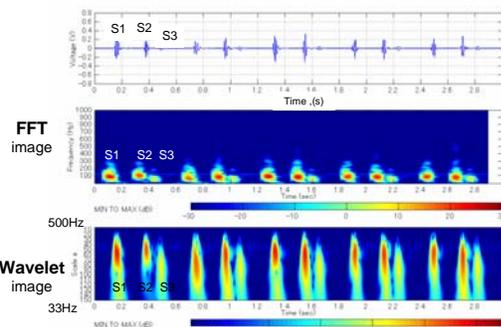


Fig.13: Audio-visual based recognition images analyzed for dilated cardiomyopathy with S3.

**11) Acute anterior myocardial infarction with quadruple rhythm sound at the apex (very serious case)**

This disease includes sinus beat sounds with frequent gallop rhythms. S3 gallop and S4 gallop may sometimes occur together, forming a quadruple gallop [10], which is a very serious case. There may be difficulties distinguishing between S1 and S4 and between S2 and S3 in this sound, but the two images clearly show each frequency band. [10], which is a very serious case. There may be difficulties distinguishing between S1 and S4 and between S2 and S3 in this sound, but the two images clearly show each frequency band.

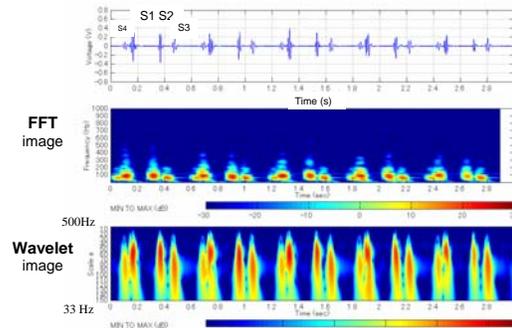


Fig. 14: Audio-visual based recognition images analyzed for acute myocardial infarction with quadruple rhythm (very serious case).

**12) High blood pressure (HBP) sound at the apex**

In the case of HBP, S2 becomes a sthenic sound [10]. Compared to the healthy sound in Fig. 4, S2 has a wide frequency band from 100-500 Hz, as estimated from the FFT image result.

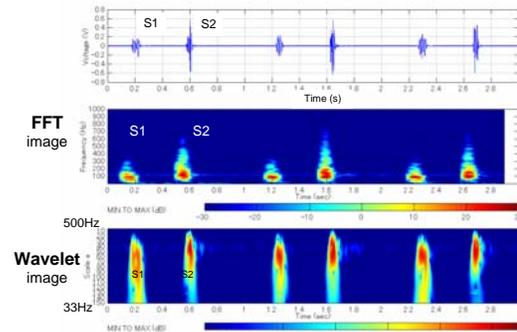


Fig.15: Audio-visual based recognition images analyzed for HBP.

**13) High blood pressure (HBP) with systolic murmur (SM) sound at the apex**

The sound of advanced HBP includes, in many cases, an SM and a sthenic S2 [10]. The SM may be due to calcification of the artery. In Fig. 16, slight SM and sthenic S2 are clearly visualized for detecting the disease. Early detection of arteriosclerosis is the key to good health management; thus, the audio-visual based recognition technique for auscultatory sounds would be an invaluable tool.

In summary, thirteen major types of heart diseases are characterized based on both STFT and wavelet images in Fig. 17. These indicate that almost all heart diseases can be detected by pattern classification; thus, the technology can overcome the

difficulties associated with hearing abnormal sounds by comparing with a typical healthy heart sound. The result indicates that there is a possibility of developing a fully automatic detection system based on a map of standard sounds in the near future

**5. Conclusion**

Since there are various difficulties associated with auscultation, we have introduced an audio-visual based recognition technique to assist in the evaluation of heartbeats by both doctors and patients. A feasibility assessment was performed using thirteen types of heart diseases recorded on the CD for the auscultation techniques training textbook [10].

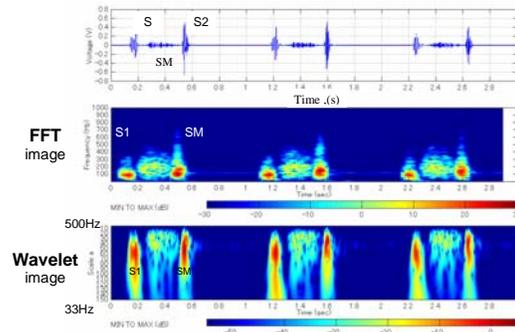


Fig. 16: Audio-visual based recognition images analyzed for HBP with SM.

Table 1 Heart sounds referred from a textbook [10] for auscultation training.

No.	Name of heart disease	Remarks
1	Healthy heart sound at two locations; the apex and at the LSB of the 4th rib	Fig. 4 a), b)
2	Mitral regurgitation (MR) sound at the apex: level of the 1st degree	Fig. 5
3	Mitral regurgitation (MR) sound at the apex: level of the 2nd degree	Fig. 6
4	Mitral regurgitation (MR) sound at the apex: level of the 3rd degree	Fig. 7
5	Mitral stenosis (MS) sound at the apex	Fig. 8
6	Mitral stenosis and insufficiency (MSI) sound at the apex	Fig. 9
7	Ventricular septal defect (VSD) sound at the apex: level of the 3rd degree	Fig. 10
8	Hypertrophic obstructive cardiomyopathy (HOCM) sound at the apex	Fig. 11
9	Hypertrophic cardiomyopathy (HCM), asymmetric septal hypertrophy sound at the apex	Fig. 12
10	Dilated cardiomyopathy with S3 sound at the apex	Fig. 13
11	Acute anterior myocardial infarction with quadruple rhythm sound at the apex, very serious case	Fig. 14
12	High-blood pressure (HBP) , sound at the apex,	Fig. 15
13	High-blood pressure (HBP) with systole murmur (SM), sound at the apex	Fig. 16

Major seven heart diseases based on FFT and Wavelet analyses

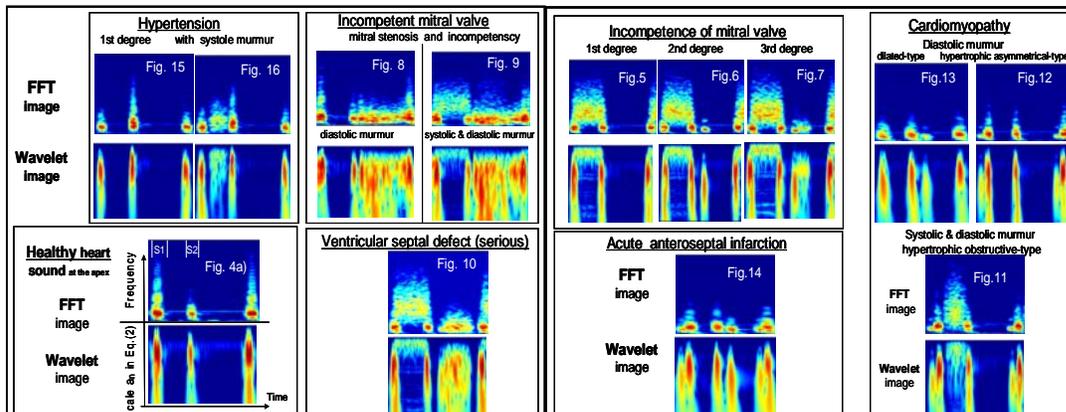


Fig.17: An image-based recognition map of thirteen heart diseases with FFT and Wavelet analyses

This technique can be an invaluable tool for detecting heart disease when combined with an auscultator, and is summarized as follows:

1) It is a strong tool that applies both Fourier and wavelet transform analyses for analyzing auscultatory sounds, and its frequencies and intensity scales clarify the causes of unknown sounds due to any heart disease.

2) In addition to auscultation, these sounds can be visualized to assist in the detection of any abnormalities in heart function. In particular, since weak sounds can be visualized using this technique, we recommend that it be used in combination with auscultation.

3) Since the auscultatory sounds are digitalized by the technology, we can isolate the abnormal sounds by

subtracting the healthy sounds. It can be shown that almost all heart diseases can be detected by pattern classification. The result indicates that there is a possibility of developing a fully automatic detection system based a map of standard sounds in the near future. A simple auscultation method is also expected to be developed in-home use.

#### Acknowledgements

The authors wish to thanks Mr. Shigeru Hayashi, who joined this project as a graduate student in 2005-2006.

#### References

- [1] I.R. Hanna and M.E. Silverman, "A history of cardiac auscultation and some of its contributors," *American J. of Cardiology*, 90-3, 259-267 (2002).
- [2] H. Makino, Y. Sanjo, T. Katoh, and S. Sato, "Description of a heart condition by three-dimensional visualization of cardiac murmur using a 'visual stethoscope'," *American Society of Anesthesiologists, Anesthesiology*, 101: A598 (2004).
- [3] S. Karki, M. Kaariainen, and J. Lekkala, "Measurement of heart sounds with EMFi transducer," *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 1683-1686 (2007)
- [4] S.A. Taplidou, "Nonlinear analysis of heart murmurs using wavelet-based higher-order spectral parameters," *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 1:4502-4505.
- [5] M. Cozic, L.G. Durand, and R. Guardo, "Development of a cardiac acoustic mapping system," *Med. & Biol. Eng. & Comput.*, 36, 431-437 (1998)
- [6] S. Chauhan, "A computer-aided MFCC-based HMM system for automatic auscultation," *Singapore Comput. Biol. Med.*, 38:221-233.
- [7] Z. Syed, D. Leeds, and J. Gutttag, "Audio-visual tools for computer-assisted diagnosis of cardiac disorders," *Proc. 19th IEEE Symposium on Computer-Based Medical Systems*, 207-212 (2006).
- [8] M. Okada, "Chest wall maps of heart sounds and murmurs," *Computers and Biomedical Research*, 15, 281-294 (1982)
- [9] M. Cozic, L.G. Durand, and R. Guardo, "Development of a cardiac acoustic mapping system," *Med. & Biol. Eng. & Comput.*, 36, 431-437 (1998)
- [10] T. Sawayama, "Consummator of auscultation - the heart sound recorded in CD," *Nippon Igaku Shupan*, 2nd ed. (1998).
- [11] P.S. Addison, "Illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance," *Institute of Physics Publications* (2002).
- [12] D. Kumar D, Carvalho P, Antunes M, et al. "Third heart sound detection using wavelet transform-simplicity filter," *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 1277-1281 (2007).
- [13] E.A. Derby and S.S. Agabegi, "Step-up to medicine (step-up series), Ch. 1: Diseases of the Cardiovascular system," *Lippincott Williams & Wilkins* (2008).