Journal of Systemics, Cybernetics and Informatics (2023) 21(2), 55-63
https://doi.org/10.54808/JSCI.21.02.55

Basic Research on the Development of an Automatic Heart Sound Diagnosis System - Analysis of Heart Sounds for Learning Policy and Experiment for the Prototype of the Auscultation Part -

Hirotoshi HISHIDA
Faculty of Engineering, Kogakuin University
1-24-2 Nishi-Shinjuku Shinjuku, Tokyo 163-8677, Japan

Koichi TOKUUYE
Radiology, Tokyo Medical University Hospital
6-7-1 Nishi-Shinjuku Shinjuku, Tokyo 160-0023, Japan

Keiko HISHIDA
Headquarter, Keiko’s Music Room, and Kamakura Player’s Association
2-37-7 Fueda Kamakura, Kanagawa 248-0027, Japan

Hayato TOJO
Postgraduate Student, Kogakuin University
1-24-2 Nishi-Shinjuku Shinjuku, Tokyo 163-8677, Japan

Yasuhiro HISHIDA
Faculty of Science and Technology, Keio University
4-1-1 Hiyoshi Kouhoku Yokohama, Kanagawa 223-8521, Japan

Tomomi KOIDE
Postgraduate Student, Kogakuin University
1-24-2 Nishi-Shinjuku Shinjuku, Tokyo 163-8677, Japan

ABSTRACT

A design policy was established for a specific data flow and learning method for the automatic heart-sound diagnosis system under development. The production of each part becomes possible, and auscultation and learning begin. It can be used over clothes as long as it is applied well to the skin surface of the chest. It would be nice to be able to set multiple auscultation positions, but there is a limit to what ordinary people can be asked to do, so this should be considered while having AI learn.

We analyzed normal heart sounds to explore learning strategies. Sounds I and II are considered to be important anchor information sources for identifying other heart sounds. Abnormal heart sounds may not be heard at every beat and the rhythm may be abnormal. AI refers to the multiple beats of heart sounds during auscultation. Heartbeat analysis is a multidimensional information analysis related to time and space, and heart sounds are factored if normal and abnormal heart sounds can be organized based on the score. For pitch that tends to depend on individuals and devices, a relative discussion would be more appropriate.

Keywords: FT analysis, Neural-network, Learning data, Sound information, Image information.

Authors thank to Editage (www.editage.com) for English language editing.

1. INTRODUCTION

The authors are currently developing an automatic heart-sound diagnosis system. The overall policy of the system is described in a previous study [1]. In the present study, based on the progress since then, a specific system configuration was determined. The history of heart sounds and stethoscopes [1] is reviewed, the stethoscope problem in the proposed system is extracted, and experiments are conducted. In addition, how artificial intelligence (AI) [4] employed in this system learns, is described in another paper [5].

2. HEARTBEAT AND HEART SOUNDS

Definitions

Primitively, a sound, particularly a series of striking sounds, is called a beat. The term “beat” sometimes refers to the pressure or impact caused by continuous sound. Surprisingly, the term appears to be flexibly applied across fields. For example, in the music field, "beat" basically means a regular sound or sequence of sounds. However, it is sometimes used as 3-beat, 4-beat, etc. In the medical field, "heartbeat" seems to be defined as the regular physical movement of the heart to create blood circulation. On the other hand, the speed of the heartbeat, for example, 70bpm, is called the heart rate, HR.

In this paper, the terms "heartbeat" and "heart sound" would be assigned intermediate definitions, because engineering is applied in the medical field with a musical perspective. All sounds originating from the heart are called “heart sounds.” In normal cases, Sounds I and II, which are the two heart sounds, cyclically create (play) a heartbeat. Auscultation reveals the heartbeat, and the diagnosis is conducted by analyzing its rhythm and components, such as heart sounds.

Relationship between Humans and Heartbeat
A heartbeat is probably the first sound heard by a person after birth. The hearing organs and heart of the fetus are both formed approximately six weeks after fertilization. The fetus's ears in the womb hear the mother's heartbeat along with the sounds of her circulatory, respiratory, and digestive organs. A fetus spends almost 40 weeks in the mother's womb listening to its mother's and own heartbeats. Even after birth, until the day people die, they continue to listen to the sounds of their heartbeats . Because people have been listening to heartbeats for a long time, it would not be strange if they were unconsciously familiar with heartbeats.

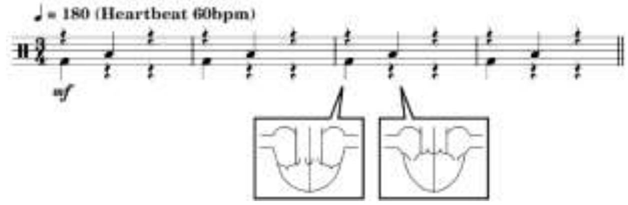
The heartbeat expresses the physical condition, particularly the state of the autonomic nervous system. As a result, heart sounds may reflect not only diseases around the heart but also the physical condition of the entire body. This is what we would like to highlight. In the womb, the fetus must be sensing the mother's physical and mental state by listening to the mother's heartbeat. It may have synchronized its heartbeat with that of the mother. The authors believe that this is the origin of emotional education. We hope that further research on the heart will lead to current studies on music therapy [2,3]. Our hearts and minds are significantly similar.

Sound of the Valve
Heartbeats were mainly composed of the first heart sound (sound I, S_1) and the second (sound II, S_2). Sound I is the sound of the valves closing when the heart pumps blood to the entire body, and sound II is the sound when the heart collects blood from the entire body. Sound I is produced by the tricuspid and mitral valves, while Sound II is produced by the pulmonary and aortic valves. The pulmonary and aortic valves are round and smaller than the tricuspid and mitral valves, respectively. It seems that this fact would make a slight difference in these sounds: sound I was slightly longer and lacked sharpness, whereas sound II was slightly higher in pitch.

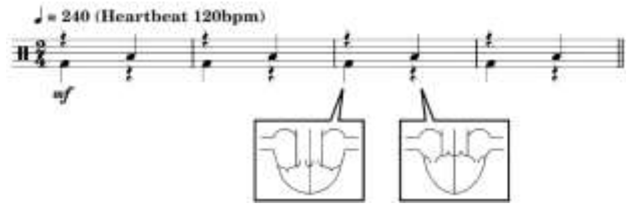
Normal Heartbeat
In a normal heartbeat, there is a rest for one beat after sound II, sounds I and II, and the rest make the overall triple beat (3-beat), as shown in Fig. 1(a). One cycle was composed of sounds I and II, and the rest as another cycle.

The number of cycles in one minute was defined as the heart rate. The heart rate of healthy adults is defined to be approximately 60-75bpm. However, there is no need to be over-concerned about numerical values, as the range varies depending on the reference and doctor. The heart rate varies significantly from person to person because it depends on how much nutrients and oxygen the body needs, which is its metabolism. Because metabolism slows down with age, heart rate also decreases with age. Infants possess high heart rates.

Exercise Heartbeat
However, heart begins to pound at the time of running or when we become nervous, or excited. The beat is a double beat (2-beat) as shown in Fig. 1(b), to the extent that one can feel their heartbeat. At this time, the rest next to the second sound at normal times disappeared. The heart works hard to pump blood throughout the body; therefore, it cannot rest. The heart-rate in double time was also situational. The beat rate can easily exceed 120 bpm when an athlete is exercising.



(a) Normal heartbeat.



(b) Excrcise heartbeat.

Fig. 1: Examples of heart sounds expressed in musical notation

3. PROPOSED SYSTEM

Overview
Fig. 2 shows an overview of the automatic heart sound diagnostic system that our laboratory aims to put into practical use. The authors are aiming for a system in which the auscultation part is connected to the patient's local device such as a PC or smartphone, and the patient's auscultated heart sound is sent to the remote AI [4] via a communication system such as a telephone or internet line. The AI receives it for diagnosis, and the diagnosis result is returned to the patient. While the basic policy [1] was being studied, the specific structure of the system, method of use, and development procedure were understood to be adopted.

Auscultation Part

Positioning of chest piece: The first important step is to decide where to place the stethoscope. Because a patient who is not a doctor holds the stethoscope on their chest, and it requires some time to adopt to it. It would be better to have a sample diagram showing the auscultation position, and it would be beneficial to have a service asking the doctor about this position over the phone. It would be beneficial if a system for receiving training in auscultation at an early stage, similar to that for injecting insulin, could be created. If a recorded heartbeat which is not sufficiently good in quality is sent, the AI must be able to judge it as "undecipherable." The heartbeat of the patient was then re-recorded. In any case, it is essential that the patient hears their own auscultatory sounds during auscultation.

Collecting air vibrations: There are generally good stethoscopes in the market; therefore, one that is light and easy to handle should be chosen. Alternatively, a stethoscope that is more basic and easier to use may be better because it can be used by patients. A stereo stethoscope may be difficult to use because of its anisotropic application, whereas a monaural stethoscope may be suitable for patients who are uncertain of its position. Rather than recording a good sound, it may be better to take other

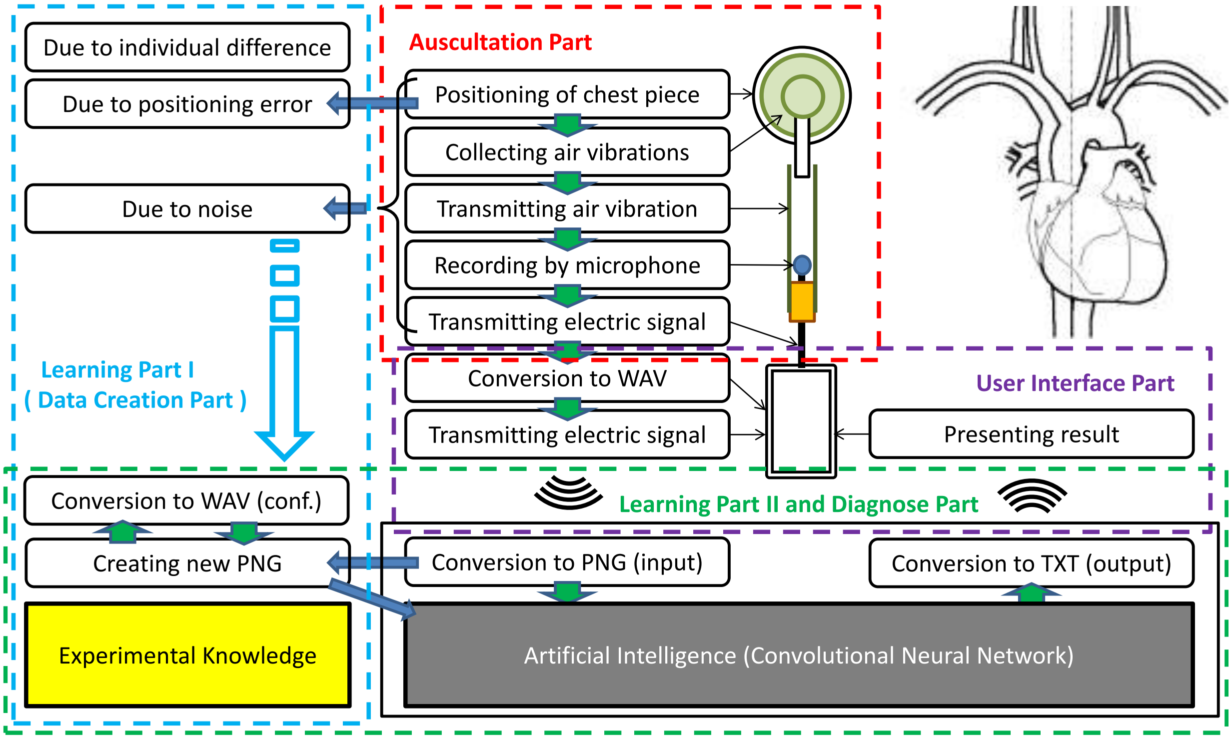


Fig. 2: Overview of research and development of automatic heart sound diagnosis system

measures, such as holding one's breath while recording, to ensure that the sound does not deteriorate.

Transmitting air vibrations: The distance from the sound collector to the electronic device that converts the air vibrations into electrical signals should be as short as possible. Considering ease of handling and sound insulation, a commercially available stethoscope tube is considered sufficient.

Recording by microphone: A microphone was attached to the end of the tube to transmit the air vibrations. A special jig was fabricated and fixed to the tube. Small microphones that fit into commercial stethoscope tubes already exist, such as those for broadcasting.

Transmitting electric signal: A method in which sound waves are transmitted to the interface by a wire should be adopted. If a wireless system is adopted, there is a possibility that noise will be introduced.

Use Interface Part

Conversion to WAV: It is necessary to convert the electrical sound data into WAVE data [n] without compression. This requires software and few of them are available free of cost. While an older PC requires an amplifier and a sound board or USB port, a modern PC requires only a microphone jack. Dedicated applications are required for smartphones A smartphone is essentially a phone with a recording function; therefore, it should be easy to create sound data. Determining whether to save and hold the sound data created thus far will be the next theme after this system is completed.

Transmitting electric signal: WAVE data must be transported to the workstation where the AI is waiting via telephone or Internet functions. Wired communication is the preferred noise countermeasure. However, considering the damage to infrastructure in the event of a disaster, radio may be tougher.

Delivering result: It is necessary to receive automatic diagnosis results sent from the AI via telephone and Internet functions. This can be done via email or an automated voice call.

Diagnose Part

Conversion to PNG: The workstation creates the input data for the AI by Fourier transforming (FT) WAVE data. Various conversion programs are available, some of which are free of charge. The PNG format was used as the input data. It was decided that a high-resolution tiff format was unnecessary. An uncompressed BMP format is ideal, but its size should be as small as possible, considering processing by AI. The JPG format is common, but it has been lossy compressed, and it is not recommended for use assuming some contingencies. Whether to deal with color images is a matter of consideration. In this case, the GIF format was avoided.

Conversion to TXT: It is desirable for the automatic diagnosis results of AI to be reported in simple sentences to patients. Numerical values are effective in showing the morbidity probability for each disease. It may be better to use words than numbers when describing health levels because the output result of each unit of AI is a numerical value from 0 to 1, and a program is necessary to convert it into a sentence or another number, such as a probability. Communicating data in TXT format is easy.

Convolutional Neural Network, CNN: CNN [8,9] is considered suitable for image diagnosis. Because good image-processing AIs are being proposed one after another, a structure in which the input and output layers are fixed and only the intermediate layer can be exchanged would be effective. The input data were heart sound images. The design of the output data was one of the research objectives.

Learning Part

Experimental Knowledge: Learning (training) is performed in three stages, as shown in Fig. 2 [1]. First, a heart-sound sample from a medical education textbook was used. It will be possible to check whether AI can easily learn the data for each lesion. Next, new artificial learning data are created by considering various lesions based on conventional phonocardiograms [1] (called additional data). As the amount of learning data increased, AI became educated. Only after completing this stage, AI can be used in clinical applications. Finally, the patient data were added to the learning data, and the AI became increasingly precise.

Creating new PNG and Conversion to WAV: This research work is currently in the second stage. Additional data corresponding to one lesion were created and added to the learning data. The image data were converted into audio data and the work proceeded while checking for sound.

Inaccuracy due to three factors: After clinical application, various factors, such as patient individuality, errors due to different listening positions, and physical noise, hinder the accuracy of AI estimation [1]. Anything that needs to be considered should be done in advance.

4. DESIGN OF LEARNING DATA AND AI CONSTRUCTION

Diagnosis of lesions by heart sounds

Phonocardiogram, PCG [1] is an established therapeutic technique [1]. However, it is difficult to pass down the technology of heart sound diagnosis in recent medicine, which makes extensive use of equipment. The authors initially planned to use sound data in this study; however, it could be thought that images would be more effective if unknown lesions remain. Visually, the sound data representing waveforms are extremely difficult to analyze. Sound data can only be understood as various pieces of information after they have been heard, which requires competence. Fig. 3 shows the audio-digital data of heart sounds composed only of normal heart sounds obtained from the medical education textbook [1]. It was originally in the m4a format and was converted to the WAVE format. A set of graphs was used for the stereo data, and each graph had its horizontal axis as time and its vertical axis as sound pressure.

The Fourier transform (FT) is the only physical means of visually confirming the properties of sound data. Figs. 4(a) and 4(b) show the spectral envelopes of sounds I and II, respectively. Each envelope is labeled as a through d, where a, b, c, and d correspond to the 1st, 2nd, 3rd, and 4th beat. Although the entire envelope curve of the spectrum differed depending on the beat, the features below the peak frequency were the same. There were also fine peaks below the peak frequency and the fundamental tone of each sound was lower than that of the peak frequency. Various overtones and, in some cases, noise were unstably mixed in the frequency region higher than the peak frequency. The basic object of this investigation was the region below the peak frequency, and various pieces of information may be hidden in the high-frequency region.

It is believed that visual information facilitates the discovery of new information. The proposed system has the potential to identify and discover unknown lesions.

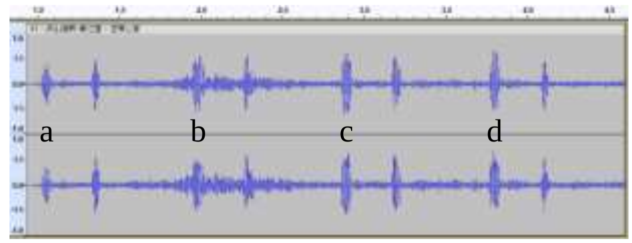
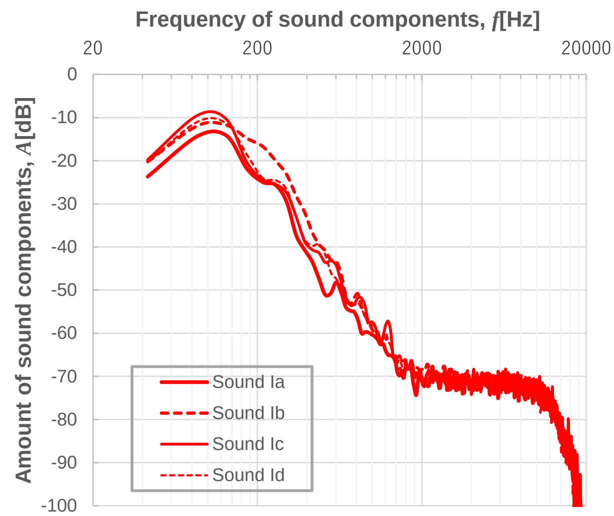
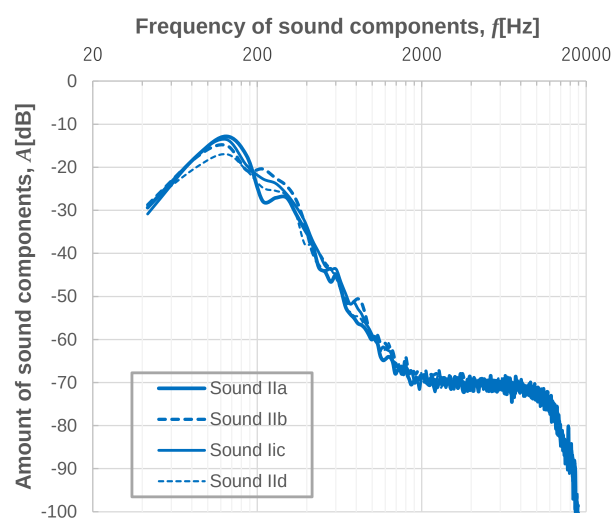


Fig. 3: Audio digital data of heartbeat shown as WAVE format in the medical education textbook [1].



(a) Sounds I.



(b) Sounds II.

Fig. 4: Spectral envelope of sounds I and II in each beat cycle a to d.

Sounds I and II

A normal heartbeat comprises sounds I and II. As the sail valve that constitutes the atrioventricular valve is oblong, the sound I produced by its closure is low-frequency, long, and dull. On the other hand, the semilunar valve, which constitutes the arterial valve, is fan-shaped with a central angle of 120°; therefore, sound II, due to its closure, is high frequency, short, and sharp. These pitches are said to be 400 Hz or less than 1000 Hz; however, according to the measurements carried out in our laboratory, the pitch is considered to be quite low. The FT results showed that the frequency was not high, up to 400 Hz. Heart sounds contain high-frequency sounds and ultrasonic waves as overtones. Partial components below 20 Hz are also present. The sounds outside the audible frequency range of 20 Hz–20000 Hz cannot be auscultated and might not be discussed.

Table 1: Comparison of peak values of sounds I and II.

	Ia	Ib	Ic	Id
Peak frequency, f_{peak} [Hz]	107.82	107.83	106.80	107.34
Peak amplitude, A_{peak} [dB]	-11.72	-10.45	-9.24	-10.01

	IIa	IIb	IIc	IId
Peak frequency, f_{peak} [Hz]	129.20	127.95	128.38	128.92
Peak amplitude, A_{peak} [dB]	-12.86	-14.90	-13.54	-16.94

The average peak frequency of sound I is 107.82 Hz (No.23.65 tone) with a standard deviation of 0.427, and that of sound II is 129.20 Hz (No.26.79 tone) with a standard deviation of 0.4846. The standard deviation is sufficiently small, and each average value can be considered the true peak pitch of sounds I and II for the person. The spectral envelope is believed to depend on the valve size. If the heart is normal, it is natural for every beat to have the same pitch because it comes out of the same valve. However, due to valve pain, the pitch may fluctuate each time, and the heartbeat to be checked may require several beats, not just one beat. It should be discussed in the future whether to give the AI a cut image for each beat at the same time or to give the whole image without cutting (without knowing the number of beats).

Fig. 5 compares the FT results for sounds I and II, the blank between sounds I and II, and the entire sound for this beat. The spectral difference between sounds I and II was clear but small. Moreover, fine peaks were observed and the spectral curve was more complicated. Therefore, it is expected that the spectral characteristics will be easier to analyze by converting them into a color map, where the horizontal axis represents time, the vertical axis represents frequency, and the color or brightness is adjusted according to the amount of each sound component.

Excessive Heart Sounds

The third heart sounds (sounds III and S_3), fourth heart sounds (sounds IV and S_4), and ejection sound (ES), which cannot be heard normally, are collectively called excessive heart sounds. Owing to the failure of compliance of the ventricular wall, the outflow of blood increases, and the blood is ejected against the ventricular wall. It is best heard at the apex of the heart. A bell-type stethoscope is suitable for auscultation owing to its low frequency.

Sound III

Sound III is generally called "gallop." An excessive heart sound

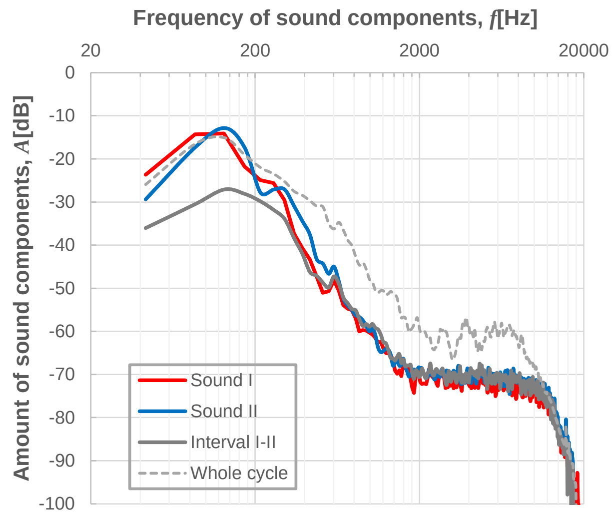


Fig. 5: Spectral envelope of normal heartbeat in a first beat cycle a.

occurs when a large amount of blood flows from the atrium to the ventricle during the rapid filling period of early diastole. It occurs after sound II with a peculiar gallop rhythm. Occasionally, sound III is heard when auscultating a healthy person under the age of 30 years or a pregnant woman. However, when a middle-aged or older person breathes or stands up, it is diagnosed as abnormal if sound iii does not disappear. Sound III indicates suspected conditions, such as mitral regurgitation, interarterial patency, left heart failure, tricuspid regurgitation, pulmonary regurgitation, and atrial septal defect.

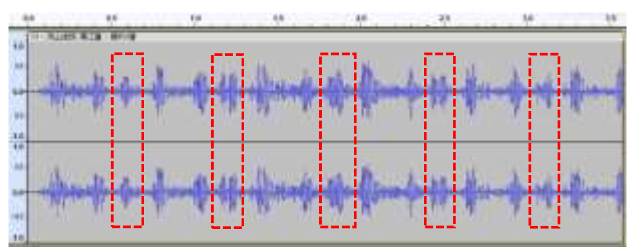


Fig. 6: Audio digital data of heartbeat with sound III shown as WAVE format in the medical education textbook [1].

In some examples of heart sounds with sound III in the medical education textbook, the waveform graphs of sounds I and II were also disturbed, as shown in Fig. 6; however, as far as the authors could hear the sounds, they were not disturbed. In addition, it is difficult to specify the timing of sound III as the 3rd beat, and it occurs once or multiple times between the 2.5th beat and the 3.2nd beat. The FT results showed that the peak frequency of sound III was lower than that of sound I; however, it was rather high for the listener. When discussing whether automatic heart-sound diagnosis should be performed by sound or image, it was pointed out that more advanced learning might be required if sound is used.

Sound IV

Sound IV appeared before sound I. It is also called the atrial squeeze sound or atrial constriction sound. When the right ventricle or left ventricle is stressed, oscillations may occur in accordance with contraction of the atria if the extensibility of the ventricular wall is reduced. The strength of the vibration was proportional to the degree of the load. It is normal for Sound III to appear during an infant's heartbeat. However, when it is heard during the heartbeat of an adult, it is diagnosed as abnormal. Adults are suspected of having hypertension, left heart failure, aortic stenosis, hypertrophic cardiomyopathy, ischemic heart disease, pulmonary hypertension, and pulmonary artery stenosis, which are not observed in atrial fibrillation.

As shown in Fig. 7, sound IV may be continuous with sound I or distinctly separate from it. It was difficult to distinguish between sounds I and II until the auscultator became accustomed to auscultation. Sounds I and II are often regular in creating the base heartbeat, and the higher pitch sound is sound II. The spectral characteristics tended to be lower than those of sound I and did not exhibit a clear peak. Moreover, Sound IV, which continues to sound, occasionally hides the peak of Sound I. This may explain why the pitch is heard to be slightly higher than the following sound I. The FT results suggest that the pitch of sound IV is often lower than that of sound I. Because judging from only one beat is extremely difficult, the data for several beats should be used as input data for the AI.

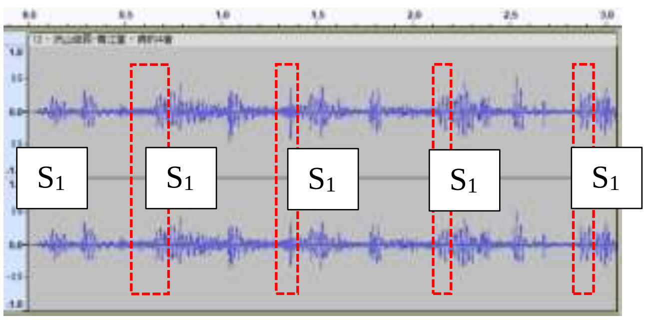


Fig. 7 : Audio digital data of heartbeat with sound IV shown as WAVE format in the medical education textbook [1].

Ejection Sound
Immediately after sound I, a high-pitched sound may be heard owing to the ejection of blood into the aorta or pulmonary artery. It is believed to be caused by pulmonary hypertension or high blood pressure (BP). Although it is difficult to distinguish the ejection sound from sound I, comparing the case of listening on the open bell side and on the membrane side, where high-pitched sounds are dominant, the volume balance between the ejection sound and sound I is different. Whether the patient was actually auscultated in both planes must be discussed. On auscultation, the pitch heard was almost the same as or higher than sound II. The FT results agree with the auditory sensibility results. The pitch varied frequently over a wide range.

Murmur
Other sounds heard between the hearts that are not normally heard are collectively referred to as murmurs. It occurs when blood flow is abnormal due to stenosis, regurgitation, shunt, or outflow tract obstruction. They are classified into systolic, diastolic, continuous, and to-and-fro murmurs according to the cardiac cycle. Its pitch varies. It is possible to estimate which valve was damaged based on the timing of pronunciation and which part was heard prominently when auscultated. However, in the proposed system, the patient places a stethoscope on the patient; therefore, the auscultation points must be in one place. It may be difficult for AI to guess the parts. The AI may also have to estimate where the patient will place the stethoscope.

Heartbeat rhythm is disturbed under medical conditions such as arrhythmia, atrial fibrillation, atrial flutter, ventricular fibrillation, ventricular flutter, and premature contraction. Simultaneous observation of the carotid artery pulse may enable a detailed diagnosis. The present system does not consider placing a stethoscope on the carotid artery. However, it will hopefully be easy to judge the irregularities in the rhythm in detail from the image data.

Fixed Splitting of the Sound II
Sound II is pronounced in both the aortic and pulmonary valves; however, their timing may not match perfectly. The closing sound of the aortic valve, called sound IIA (S_{A2}), is synchronized with the heartbeat. However, during inspiration, the rib cage expands, the pressure in the thoracic cavity decreases, and the pressure in the right atrium also decreases, thus promoting blood flow from the vena cava to the right atrium and resulting in more blood being sent to the blood vessels. During inspiration, the pulmonary vessels are dilated and stretched, and the closing sound of the pulmonary valve, called sound IIP (S_{P2}), generated by the peak pressure applied to the pulmonary valve, is slightly delayed. As the aortic valve is slightly larger than the pulmonary valve, the pitch of sound IIA is slightly lower than that of sound IIP.

Analysis of Multidimensional Information
Heart sounds can be considered multidimensional information of valve location, that is, spatial and temporal information. If AI can completely decompose these factors, it will be possible to identify the lesion. This idea may be similar to that when conducting an ensemble score. The heart sounds explained thus far were expressed as scores based on these peak frequencies, as shown in Fig. 8. The time histories of sounds I and II are described in the 1st and 2nd bars on the 1st and 2nd staffs, respectively. The peak pitch of each heart sound did not exactly match the frequency of the equal-temperament tone, so it was approximated to the nearest pitch.

Fig. 8: Examples of musical notation of heartbeats in the medical education textbook [1] *based on their peak frequency.*

In a previous study [1], musical scores were used to determine which musical tones would be substituted for other heart sounds based on their fundamental tones, not by FT, but by human auditory sensibility. The difference of 2 octaves is thought to be due to the fact that the overtones were presently referred to instead of the fundamental tone. Finally, the fundamental tone showed a peak frequency; however, this time, the overtone component also showed a peak frequency. In fact, it is thought that the fundamental tone of the heartbeat is the same as that of the last one, that is two octaves lower. (Specifically, there may be a fundamental tone with a frequency lower than that of the audible tone.) The sounds created by percussion that have no pitch with various overtones are inherently difficult to recognize. In addition, it is sometimes difficult to compare the pitches of sounds at different times auditorily. Considering the experimental results described later, it is possible that high-frequency sounds are emphasized to make the sounds in medical education textbooks easier to understand.

In addition to the octave difference, there was a small pitch difference, which was thought to be due to individual differences. It is surprising that different heart sounds match pitch. There are larger pitch differences in medical settings. The third bar describes the main pitch ranges of heart sounds given in medical education textbooks. However, because heart sounds are individual, the spectrum and pitch are determined based on the overall relative relationship. We hope that this score will become more precise as research progresses.

Because sound propagates along the bloodstream, it is best heard where the bloodstream is closest to the chest wall. It is recommended that the right sternal valve sounds at the right sternal margin of the fifth intercostal space (5R), the left velocular valve sounds at the apex of the left fifth costal space (APEX), the pulmonary valve sounds at the left thoracic margin in the 2nd or 3rd intercostal space (2 L or 3 L), and the aortic valve sounds at the right sternal margin in the 2nd intercostal space (2R). The patient should not be required to specify where the stethoscope should be placed. It would be interesting to create an AI that can guess which valve has an abnormality regardless of the auscultation position.

5. EXPERIMENTS FOR PROTOTYPE OF AUSCULTATION PART

Microphone Selection

A microphone was used as the ausculatory component. COS-11D manufactured by SANKEN, which is sufficiently small, was selected as the first candidate among several candidates. Fig. 9 shows the results of the comparison of the frequency response of the microphone with that of a calibrated standard microphone. The normal heartbeat shown in Fig. 3 was reproduced by the calibrated speaker, and two microphones recorded it at the same location. The 3 times frequency response recorded by the COS-11D matched well with that of a standard microphone. The COS-11D's overall sensitivity is 1/3 that of a standard microphone, but its frequency response is undisturbed, indicating that it is recording sound correctly.

Accuracy of Self-Auscultation

The major difference between auscultating yourself and someone else is how the chest piece of the stethoscope fits against the target chest surface. There was a lot of noise in the sound, and I auscultated myself without being a good fit. However, it seemed that important sounds, such as sounds I and II, were also included. This may require some training, but eventually anyone may be able to auscultate.

Selection of Chest Piece

The operability and sound pickup ability of various stethoscopes were examined. The two-sided type requires ingenuity in how to hold it; therefore, it is difficult for amateurs to handle. The larger the auscultatory surface, the easier it is to fix the auscultatory position. The two-sided type tends to have a small auscultatory surface and no planar shape.

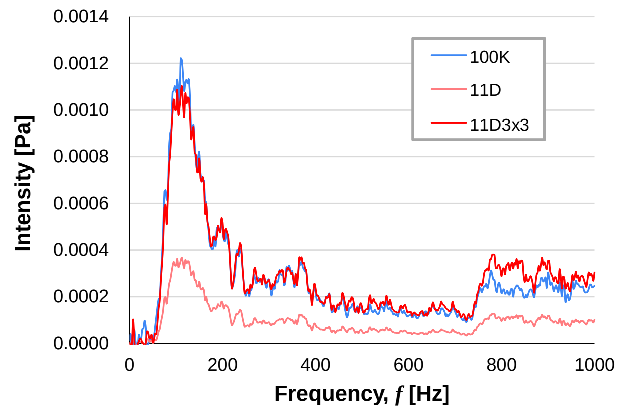


Fig. 9: Recorded frequency characteristics of a normal heart sound [1] *reproduced from a speaker*

A chest piece with good operability was selected from each of the stainless steel, hard plastic, and brass chest pieces. The three devices and electronic stethoscope were compared in detail. COS-11D was connected to the auscultation part of the three types of chest pieces. This is an analogous recording. It was an analog that was close to the doctor's hearing. However, it was digital, which made it easier for amateurs to listen. Strangely, the digital stethoscope seemed to emphasize the high frequencies. It was suggested that the heartbeat in the medical education textbook was either digitally recorded or its high frequencies were emphasized to make it easier to hear. It is unclear whether this was a deliberate process, a natural phenomenon caused by an electronic circuit, or because the device was as small as an earphone. Finally, it was concluded that the 3M-Littmann's Master Cardiology (made of stainless steel, 190 g), which has the best operability, looks good for general use, and that Welch Allyn's HARVEY DLX-Double Head (made of brass, 235 g) seemed to be useful for research purposes.

Fig. 10: Example of the auscultation part.

Consequently, the auscultation shown in Fig. 10 was completed. The audio amplifier was US-322 (TASCAM).

Auscultation Through Clothes

Auscultation was conducted in two ways: over clothes and directly on the skin, and the recorded frequency characteristics were compared. Just to be sure, all three stethoscopes which were listed as the final candidate above were used. Regardless of the stethoscope used, sounds from above the clothes were extraneous at various frequencies. This seems to be the sound of clothes rubbing against the stethoscope or hand. By taking the difference between these sounds, it was found that there were main noises below 20 Hz and around 120 Hz, which did not interfere with the heart-sound diagnosis.

6. CONCLUSIONS

A specific design policy for an automatic heart-sound diagnostic system was developed. Various examinations were conducted to create an auscultation unit prototype. What became clear in this process is summarized below:

1) A specific data flow was established. Thus, it is possible to manufacture each part. In the future, while creating a prototype, the data flow and each component will be improved, if necessary.

2) A prototype of auscultation was completed. Two types of auscultation units were proposed: one that was considered to be suitable for research purposes, and the other that was considered to be good for wide use by the general public.

3) The user interface is a subject for future research. From a technical perspective, it is possible to create a prototype without any problems.

4) Normal heart sounds were analyzed. Although the spectral envelope varied from beat to beat, there were no significant differences between features. In other words, sounds I and II are likely to be important anchor information sources for heart-sound diagnosis.

5) The pitch and volume of the first and second notes differed slightly. However, it is thought that there are individual differences in pitch and volume. Other abnormal heart sounds did not mask sounds I and II, but were likely to ride on them. The AI must be educated so that it can reliably recognize sounds I and II.

6) Abnormal heart sounds do not necessarily occur at every beat, and the rhythm of the beats may be abnormal. Ultimately, the heart-sound input into the system should be performed over multiple beats. However, because AI is easy to learn, it may be better to do so beat by beat first. The specific method is discussed later in this paper.

7) Heart-sound analysis can be understood as a multidimensional information analysis of time and space. If each part of the heartbeat could be written as an ensemble score, it might be worthwhile to arrange normal and abnormal heart sounds on the score.

8) Whether the auscultation position should be determined should be considered when learning about the AI. It was not expected that the patient would undergo auscultation at multiple sites.

9) Auscultation through clothes should not be problematic. Instead, it is important to apply it skillfully to the skin surface of the chest.

10) The heart sounds in the medical education textbook referred to here may have emphasized high tones. Regardless of the results of this FT analysis, it would be better to discuss the pitch and spectrum. The actual envelope peak may have been lower than that in the analytical results.

11) The learning component is another core component. Although the heart sounds corresponding to various lesions are organized, AI learning is currently running in parallel [5].

ACKNOWLEDGMENTS

The authors would like to express their gratitude to Yoshitaka MORISAKI, Shiori HIGUCHI, and other university students for their cooperation in this paper.

In addition, this research was supported by the Grants-in-Aid for Scientific Research for three years.

REFERENCES

[1] Hirotoshi HISHIDA, Yasuhiro HISHIDA, Hayato TOJO, Koichi TOKUUE, and Keiko HISHIDA: "Development of an Automatic Heartbeat Diagnosis System - Development Policy -", *Proceedings of the 26th World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 2022)*, Volume II, ISSN: 2771-0947 (Print), ISBN - Volume II: 978-1-950492-65-7 (Print), DOI: 10.54808/WMSCI2022.02, pp. 22-25.

[2] Hirotoshi HISHIDA, Riku KASAHARA, Keiko HISHIDA, Yamato FUJII, Yasuhiro HISHIDA, Hina ETOH, and Mitsuhiro HISHIDA: "Experiment of Music Therapy Conducted at a Classical Music Recital - Measurement of Pulse Wave, Blood Pressure and Mind Orientation -", *J. SYSTEMICS, CYBERNETICS AND INFORMATICS*, **19**(3), 2021, pp.58-65, ISSN: 1690-4524.

[3] Hirotoshi HISHIDA, Yoshihiro KOMATSU, Keiko HISHIDA, Dai YAMAMOTO, and Yasuhiro HISHIDA: "Construction of a Music Database for Earphone Hearing Loss Prevention and Music Therapy - Discussions on the Relationship between Beethoven's Music and His Deafness -", *J. SYSTEMICS, CYBERNETICS AND INFORMATICS*, **19**(6), 2021, pp.1-8, ISSN: 1690-4524.

[4] Hirotoshi Hishida: "Study on Viscoplastic Constitutive Equation and Its Application to Some Engineering Problems", *Doctor Thesis*, Tokyo University, Graduate School of Engineering (1992-3) pp.1-160.

[5] Hirotoshi HISHIDA, Yuhiro MUROYA, Hayato TOJO, and Tomomi KOIDE: "Basic Research on the Development of an Automatic Heart Sound Diagnosis System - Adding Another Lesion Data in AI Learning -", *Proceedings of the 27th World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 2023)*, ISSN: 2771-0947 (Print), ISBN (Volume):

78-1-950492-73-2 (Print), DOI: 10.54808/WMSCI2023.01, pp.32-35.

[6] Saad Albawi, Tareq Abed Mohammed, and Saad AI-Zawi: ”Understanding of a convolutional neural network”, *Proceedings of the 2017 International Conference on Engineering and Technology* (2017), DOI: 10.1109/ICEng Technol.2017.8308186, https://www.researchgate.net/publication/319253577_Understanding_of_a_Convolutional_Neural_Network.

[7] Matthew A. Reyna, Yashar Kiarashi, Andoni Elola, Jorge Oliveira, Francesco Renna, Annie Gu, Erick A. Perez Alday, Nadi Sadr, Ashish Sharma, Jacques Kpodonu, Sandra Mattos, Miguel T. Coimbra, Reza Sameni, Ali Bahrami Rad, and Gari D. Clifford: “Heart murmur detection from phonocardiogram recordings: The George B. Moody PhysioNet Challenge 2022”, *PLOS DIGITAL HEALTH*, https://doi.org/10.1371/journal.pdig.0000324 September 11, 2023.

[8] Keiron O’Shea, and Ryan Nash: “An Introduction to Convolutional Neural Networks”, *arXiv*, Cornell University, 1511.08458 (2015), https://arxiv.org/abs/1511.08458.

[9] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton : “ImageNet Classification with Deep Convolutional Neural Networks”, *Proceedings of the Neurips*, https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf