

Contemplate the Essence of a Problem from Multiple Perspectives

Kunio Kashino

**Senior Distinguished Researcher, NTT
Communication Science Laboratories
and Bio-Medical Informatics
Research Center, NTT Basic Research
Laboratories**



Overview

Auscultation (auditory diagnosis) has been used for centuries in the fields of health screening and medical care. It has the advantage of being repeatable and non-invasive while providing immediate results; however, because it is conducted in close proximity to the patient, it is challenging to perform on patients suspected of carrying infectious diseases. We interviewed Kunio Kashino, a senior distinguished researcher, who is working on *teleauscultation* with an eye toward addressing the above challenge and the future of health management, about the progress of his research activities and his attitude as a researcher.

Keywords: auscultation, telestethoscope, AI auscultation

Development of a telestethoscope that can collect and analyze biological sounds remotely

—Could you tell us about your current research?

When I was interviewed about six years ago, we talked about my research on scene analysis and media exploration based on media information, which I described as “creating a media dictionary.” Since then, I have also started research on *teleauscultation* and *artificial intelligence (AI) auscultation* in July 2019 while concurrently working at the Bio-Medical Informatics Research Center.

We often see stethoscopes around doctors’ necks when we are examined at medical institutions. According to medical professionals, sounds coming from inside the body (biological sounds) via stetho-

scopes contain a wealth of information. Last year, we developed a system called the “telestethoscope,” which enables the collection and analysis of such sounds from a remote location (**Figs. 1 and 2**). The system consists of a wearable device worn by the patient and a receiving terminal (e.g. the tablet in the right photo in Fig. 1) operated remotely.

During auscultation, it is necessary to hear sounds at various places on the body. Accordingly, 18 acoustic sensors are installed in the wearable device of the current prototype telestethoscope. The 18-channel acoustic signals captured by these sensors are simultaneously collected with a 1-channel electrocardiographic waveform. On the screen of the receiving tablet, an outline of the torso is shown, and the part of the torso to be listened to can be specified by touching it on the screen or clicking on it with a mouse. The

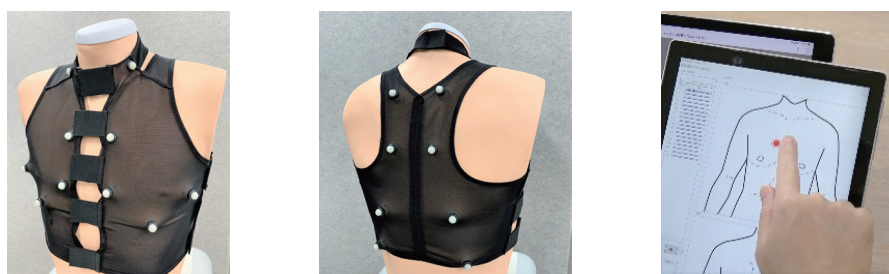


Fig. 1. Telestethoscope.

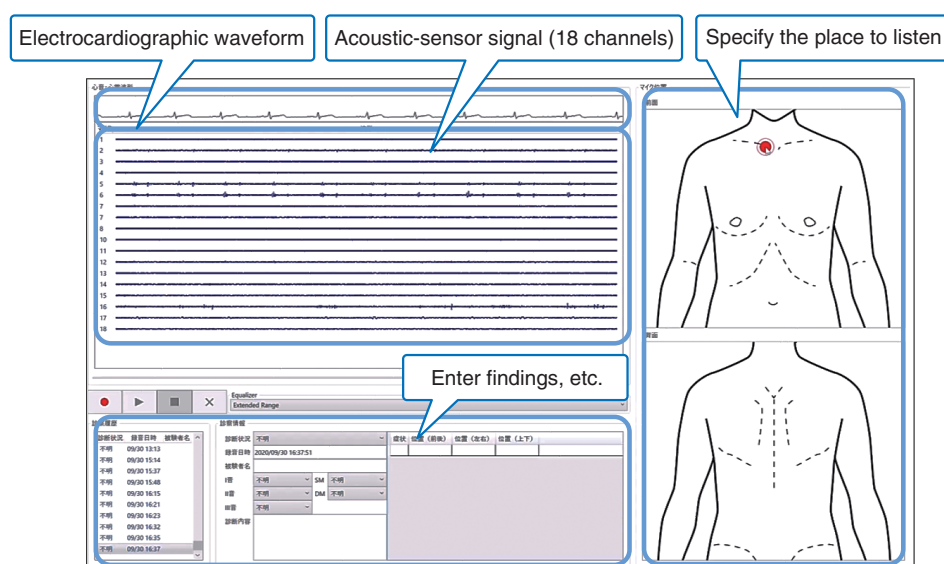


Fig. 2. Screenshot of receiving tablet of telestethoscope.

sound emitted at that part (an estimated sound synthesized from signals from the multiple sensors) can then be heard. The acoustic signal waveforms captured by each sensor are also displayed on the screen, and it is possible to listen directly to those acquired raw waveforms. Taking computer analysis into consideration, the system is designed to obtain more information spatially and in frequency domains compared with ordinary stethoscopes, namely, not only the main frequency bands used in conventional auscultation but also other bands. All received signals are recorded so that they can be listened to at any time.

If this telestethoscope is put into practical use for medical purposes, it will enable medical professionals to conduct auscultation from a separate room without directly coming into contact with the patient or during online medical treatment (teleauscultation).

Once we establish a technology for automatically analyzing the collected sounds, the telestethoscope will not only enable a medical professional to listen to the sounds from a sick person remotely but also enable a healthy person to use it on a daily basis for self-healthcare by using AI to analyze the sounds.

In fact, we are currently researching AI auscultation which analyzes internal bodily conditions on the basis of biological sounds (**Fig. 3**). An example of such research is translating the meaning of the captured biological sounds into words. We have developed a unique technology that directly translates bio-sound signals into a series of words. Unless you are an expert, you will not be able to clearly understand the meaning of your own biological sounds when you listen to them; however, by going through a “translator,” so to speak, you will benefit from being able to

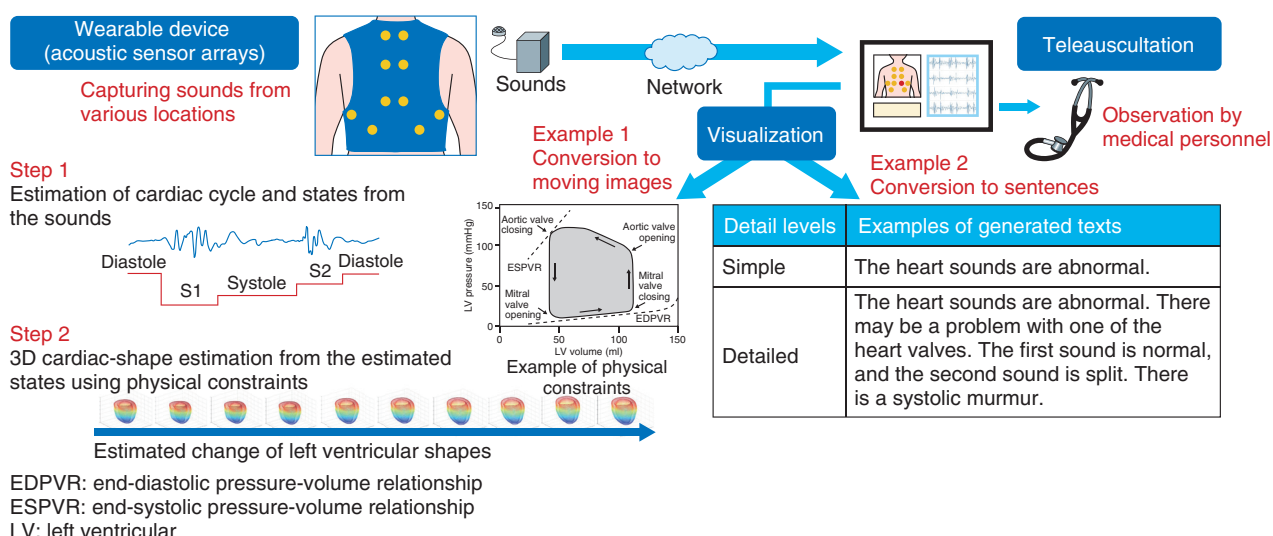


Fig. 3. AI auscultation.

understand the state of your own body. In particular, by taking advantage of using sentences rather than just indicating words such as the name of a disease, this technology is designed to convert biological sounds to sentences with the desired level of detail on the spot, for example, the presence and degree of abnormalities and changes over time since the last checkup.

—Your research on telestethoscope and AI auscultation is generating high expectations. What difficulties are there in making them practical?

We have faced a few difficulties; for example, diagnostic-imaging technology is making remarkable progress year after year, and automated diagnosis using AI is being actively researched. In contrast, auscultation data are rarely collected and usually discarded without being recorded. Therefore, we had no choice but to start our research by collecting data and building the equipment to do so ourselves.

It has also been a very interesting challenge to reproduce the skills of medical professionals on a computer. There are aspects in common with auditory-scene analysis, which I have been working on for the past 30 years. When a person hears a sound, vibration generated somewhere by a physical cause is transmitted through a medium and shakes the eardrum. Auscultation is the problem of estimating the causes of the (mixtures of) vibrations transmitted. To solve this problem, it is necessary to use as much

information as possible in addition to the vibration waveform. Just as medical professionals make judgments on the basis of their anatomical knowledge and medical experience, computers will need to fully use physical models of the body's organs and statistical models learned from a huge amount of case data. Although how to achieve this is a point of difficulty, it is also a point to which we can apply our ingenuity.

When we focus on sounds generated inside the body, however, we come across something new that we do not in typical computational auditory-scene analysis research. For example, research on acoustic signal processing has mainly focused on the scenario in which the medium for transmitting sound is a uniform substance such as air or water. However, in the body, sound is transmitted through multiple media with significantly different physical properties, such as bone, muscle, air in the lungs, fat, and liquid, are mixed in a small space. I think this mixed media makes teleauscultation or AI auscultation an even more challenging problem than conventional auditory-scene analysis.

The key to advancing research is to have an interest inherent in the research and maintain that interest

—What are some of the things that you keep in mind when you set research themes and problems?

It seems to me that the more complex the problem,

the more important it is to think straightforwardly and simply. I also think that to find a problem, set a research theme, and solve the problem, it is important to contemplate what the real problem is. What is the essence of the problem? I also think that a researcher should take a balanced attitude among three perspectives, namely, a first-person perspective focused on yourself such as what you want to do and how you can increase your productivity as a researcher, a second-person perspective focused on other people that your research results directly target, and a third-person perspective other than these two perspectives. In particular, if you do not pay careful attention to the third-person perspective, you will tend to be dragged back by the first- or second-person perspective.

In a sense, what I said could be similar to a marketing strategy. Modern marketing emphasizes the importance of efforts to find latent needs that have not yet emerged from multiple perspectives, breaking away from conventional approaches such as product-centric perspective, namely, “We want to sell this product,” as well as an existing-market perspective, namely, “This type of product is selling well.” This way of thinking can also be applied to basic research. It is also about asking what kind of research can bring what kind of value to whom at what time from an objective perspective. Having said that, it is difficult to predict research value. Sometimes value in one’s research may be discovered in a field where you do not expect. Therefore, I try to have both conviction and an eye for the unpredictable.

—So, what does it take to have strengths as a researcher?

From my viewpoint, it seems that all the researchers around me have unique strengths; however, they do not necessarily recognize their strengths as such. They just say they like what they are doing. Things that are not too tough for a person to do are often seen as that person’s unique strength in the eyes of others. With that in mind, I think the key to advancing research is to have an interest inherent in the research itself. It may be important for researchers to be able to have the interest inherent in the problem and the way of solving it and maintain that interest.

As exemplified by teleauscultation and AI auscultation, it seems that auscultation is going against the times as diagnostic-imaging equipment is being developed and used in the medical field and its importance is increasing. However, I believe that by infusing new technology into auscultation, its usefulness

will be reevaluated in the context of an aging society, new medical treatments, and self-healthcare. What’s more, the extraordinary skill of estimating what is happening inside the body, which is invisible to the eye, from a one-dimensional waveform, has actually been used in the medical field for 200 years, so just thinking about it seems quite interesting to me.

The importance of teamwork

—Please give some words of advice to our junior researchers.

Through my recent research activities, I realized that I have a lot to learn from people with diverse experiences, not only members of my own research team but also non-researchers. Communicating with these people is invaluable. The Bio-Medical Informatics Research Center assembles people with various backgrounds from multiple NTT laboratories. Our team at the Center has been actively collaborating with external research institutes and medical institutions, so we are having more opportunities to meet with doctors and other medical professionals. These meetings are beneficial, and remind me of the importance of teamwork.

For teamwork to function properly, it is necessary to build relationships between members and enhance the strength of each member. It is important to create an atmosphere in which junior and senior people can respect each other’s strengths and contributions, regardless of age or position. I want to ask all our employees, especially those in a higher position in the workplace in relation to younger employees, to always keep that in mind.

I think it is beneficial for researchers to not only develop their existing specialties but also diversify their specialties. For example, a person who studied sound processing at university then joined the company may also take up language processing, and if an opportunity to work on tasks other than research comes up, it would definitely be worth the experience. In my case, prior to joining the company, in addition to pursuing my doctoral research on acoustic-information processing, I was involved in agile development of commercial systems in an emerging industry at the time. After joining the company, I had the opportunity to learn about image processing and Bayesian inference theory. I also had many opportunities to work with people from NTT operating companies. Looking back on those experiences, I realize that they have all been useful in regard to my current

work. I hope that you all will take advantage of every opportunity.

■ Interviewee profile

Kunio Kashino

Senior Distinguished Researcher at NTT Communication Science Laboratories and Bio-Medical Informatics Research Center, NTT Basic Research Laboratories, and Visiting Professor at National Institute of Informatics (NII).

He received a Ph.D. from the University of Tokyo in 1995. His research interest includes audio and video analysis, synthesis, search, and recognition algorithms and their implementation. He is a fellow of the Institute of Electronics, Information and Communication Engineers (IEICE), and a member of the Institute of Electrical and Electronics Engineers (IEEE), the Association for Computing Machinery, Information Processing Society of Japan, the Japanese Society for Artificial Intelligence, and the Acoustic Society of Japan.