No Need to Hurry! Taking a Roundabout Way Means Time to Gain Experience. The Essence of a Researcher Is "Finding Something Interesting in Everything."

Hiroshi Sawada
Senior Distinguished Researcher, NTT Communication Science Laboratories

Overview

Nonnegative matrix factorization (NMF) enables big data such as Internet of Things sensor data to be represented in the form of matrices having only nonnegative values and to be analyzed by using simple mathematical expressions. NMF has found application in many areas including the analysis of audio, image, and text data. Dr. Hiroshi Sawada is a Senior Distinguished Researcher at NTT Communication Science Laboratories known for his research and development efforts in NMF and signal separation technology. We asked him about his current research projects and the mind frame needed by a researcher.

Keywords: nonnegative matrix factorization, independent component analysis, audio source separation

Analyzing complex data in a simple manner—development of groundbreaking technology for predicting the near future

—Dr. Sawada, please tell us about your current line of research.

I am now working on the integration of nonnegative matrix factorization (NMF) technology, which excels in uncovering the structure and features of an information source such as data or signals, and independent component analysis (ICA) technology, which estimates how data or signals were observed by sensors through an observation system (**Fig. 1**). Of course, I am also focusing on enhancing each of these technologies prior to integration and applying them effectively to social issues.

NMF is an algorithm for extracting frequency patterns. It leverages the fact that much data in the real world can be represented in matrix form and that individual matrix elements can be assumed as 0 or positive values, and it analyzes that data by matrix factorization using simple mathematical expressions (**Fig. 2**). Application to a wide variety of fields can be
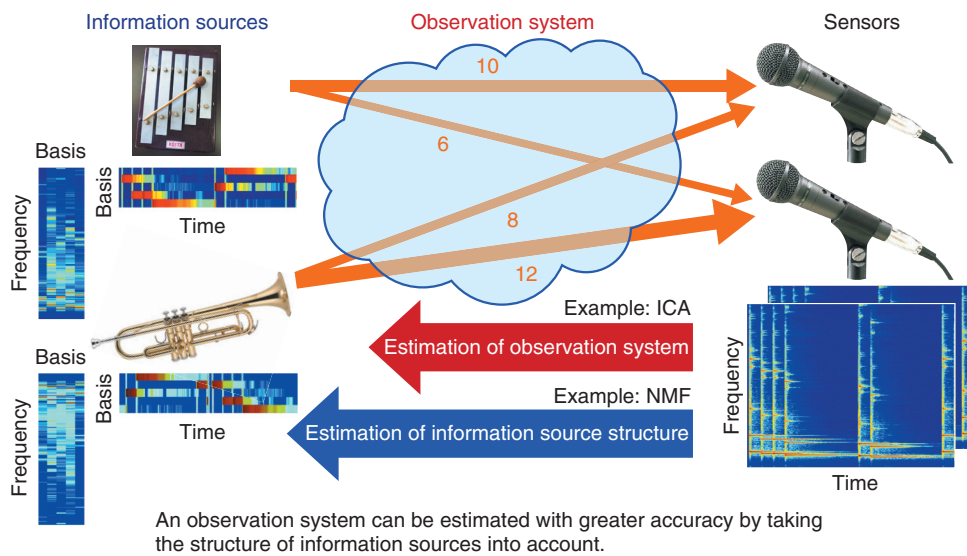
Fig. 1.   Simultaneous estimation of information source structure and observation system.

■ NMF: Factorizes an I×J matrix into the product of an I×K matrix and K×J matrix
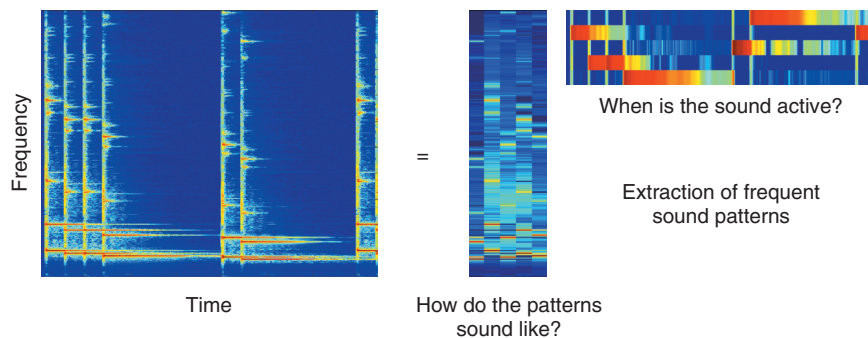

■ Application to musical instrument sounds

Fig. 2.   Overview of NMF.

considered, including document data, purchase histories, sounds, images, biological signals, and genetic data.

In contrast, ICA has progressed as an audio source separation technology (**Fig. 3**). If we take, for example, an audio recording of more than one person talking at one time, ICA is a technology that can cleanly

extract the voice of each person from that mixture. It is said that Prince Shotoku (an early regent of Japan) could listen to ten people at once and understand what each was saying. ICA accomplishes this feat by computer. As human beings, we hear sounds with both ears, and it is easier for us to make out a voice and to shut out superfluous information with two ears rather
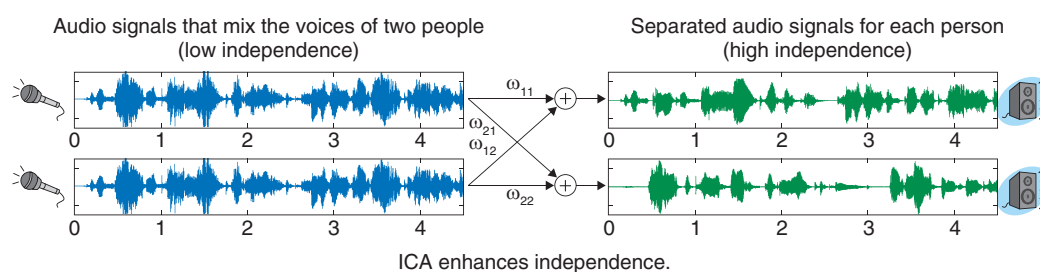
Fig. 3. ICA and audio source separation.

than one. In addition, we tend to catch sound coming from a certain direction with the ear closest to that sound, while catching the same sound arriving at the farthest ear slightly later. Consequently, by compensating for that time difference and subtracting the time-compensated sound from the original sound, it becomes possible to delete that sound. This may be easier to understand if we imagine replacing our two ears with two microphones and performing compensation and subtraction processing of audio signals (waveforms) by computer.

As in the case of a stereo mixer, the simple mixing of sound is called instantaneous mixing. The simple result of multiplying the waveform of one sound by three and that of the other by two and mixing the two sounds with a mixer can be easily analyzed. In real space, however, reverberation from desks, the ceiling, and other objects would make it difficult to cleanly extract those sounds when such echoes were present, or in other words, to delete those echoes. However, processing those sounds on a frequency-by-frequency basis would make such extraction possible. In actuality, though, sounds targeted for such extraction are not always clear, so it would be necessary to skillfully estimate a variety of conditions such as the time difference of the sound targeted for deletion and the environment in which the sounds including echoes are being recorded.

As an alternative, attempts at processing sounds using blind source separation first began in 2000. Here, *blind*, as the word implies, refers to processing using technology that separates audio sources in an environment in which "your eyes are closed" to the conditions under which the audio is being recorded. At that time, researchers throughout the world began to compete fiercely to develop this technology, but progress was slow at first. Starting in 2003–2004, however, various techniques embodying the concept of blind source separation were announced, and our

research as well came to contribute to this effort.

Support from senior colleagues and a robust human network reflecting NTT's strength in the research of sound

—Research activities have a mutually beneficial effect on other research work

In our research on audio source separation, it became possible to separate audio sources in the case of human voices, but when we applied the technique to music, good results were not forthcoming. We therefore thought that perhaps the use of NMF would enable music data as well to be separated, but it was not that simple. However, after engaging in beneficial discussions with researchers having specialized knowledge in this field and determining the direction that this research should take, we completed a number of studies culminating in a paper that we presented in 2013. Researchers in various areas have since cited this paper, so I feel that it marked a successful application of NMF to audio source separation.

With NMF, we can break down various kinds of data such as music data, purchase histories, and data on the activities of inbound tourists using a matrix format, and we can analyze such data using simple mathematical formulae. In this simplification process, a "researcher sense" based on an understanding of past research plays a major role in classifying such data. Incidentally, I would say that in our research up to now, as to what techniques to apply to what types of data, the sense of not only our research team but also that of the entire NTT research planning team as reflected by their viewpoints and activities have been beneficial.

Partly due to being a longtime provider of telephone services, NTT is recognized the world over as

being strong in the area of research involving sound. It has a diverse history in this area, such as giving birth to technology for digitizing and compressing audio/speech signals for efficient transmission and for making mobile phones a reality. In addition, my senior colleagues have a long history in the research of speech and acoustics that I just mentioned, and their support has been immense. I joined the speech research team in 2000, and I intuitively felt from the start that this team was a little different from other teams and was highly competent. Consequently, when I started to give presentations at international conferences, I was often approached in the manner of "Ah, Dr. Sawada of NTT" by people who had an interest in my work, so I was able to find fellow researchers from around the world early on. Of course, these researchers were good rivals in a sense, but I was able to make a connection with the world thanks to my seniors at NTT. In this way, I was able to feel NTT's strength in research. This NTT research team had certainly built up a diverse and robust network of prominent researchers from around the world.

—You seem to have formed reassuring relationships with your colleagues through research activities.

Apart from research, there is not much that I have done for any length of time. I have tried a variety of things, but I tend to lose interest in the end. My research activities, however, have endured for a long time. I believe that producing something new is one reason for being a researcher. To present something new to the world, I put down my ideas or findings in a journal paper. I am pleased if other researchers read my paper and expand upon it, and I am happy to see my paper explained in easy-to-understand terms to people on the business side and to then have my findings used.

By the way, there are all kinds of people involved in research, and I myself am one that likes to try experimenting, prototyping, and programming on my own. After all, there are many things that you will not understand unless you conduct tests on your own. You can find that something examined in a journal paper or using a mathematical formula differs from reality when you do tests on your own.

In recent years, the work I do in areas outside of my own research, for example, in organization management, has been increasing, so opportunities for doing my own work have been decreasing, but I try to make as much time as possible for myself. On returning home, I spend time with my family and help out with the housework, but there is always a corner of my mind that is thinking about research. Something may pop into my head before I go to bed or while I'm soaking in a hot bath, and I may think about giving it a try when I get back to work the next day. I have not thought much as to why I became a researcher. The truth is, I have come this far without any misgivings about the path I have chosen.

I was originally involved in research on LSIs (large-scale integrated circuits), but in 2000, that research group was dissolved. I think my group leader at the time took my aspirations into account, as I then spent half a year studying overseas. On returning to Japan, I shifted to my present line of research. Then, for about two years starting in 2007, I worked in a joint department for organizational purposes, and my research work was suspended. However, this applied only to my work as a professional researcher, as I continued to keep in touch with research activities. For example, I would provide consultation to young researchers as part of human resource development, and I would look over the research plans unfolding at various research laboratories.

At first, I was irritated at being separated from my research, but by advising college students worried about their future path solely from the standpoint of a researcher, I was able to look back at my own history. Additionally, by touching on research outside of the fields that I had worked in, I was able to broaden my horizon and my personal values. Furthermore, though I was able to present my research on audio source separation at an international conference in 2007, I had to step away from my research just when I was planning to expand on that paper and present it again two years later. As a result, that presentation was delayed by another two years. However, I was eventually able to make that presentation despite that blank period in my research, so in hindsight, it was just a two-year gap, and in a sense, no more than a part of my ten-year period of research.

Working to make social activities more efficient through experience, intuition, and courage

—The period in which you were not directly involved in research became an experience that brought more depth to your later research activities. What is your current research objective?

I would like to contribute to society by making good use of large volumes of data. For example, I
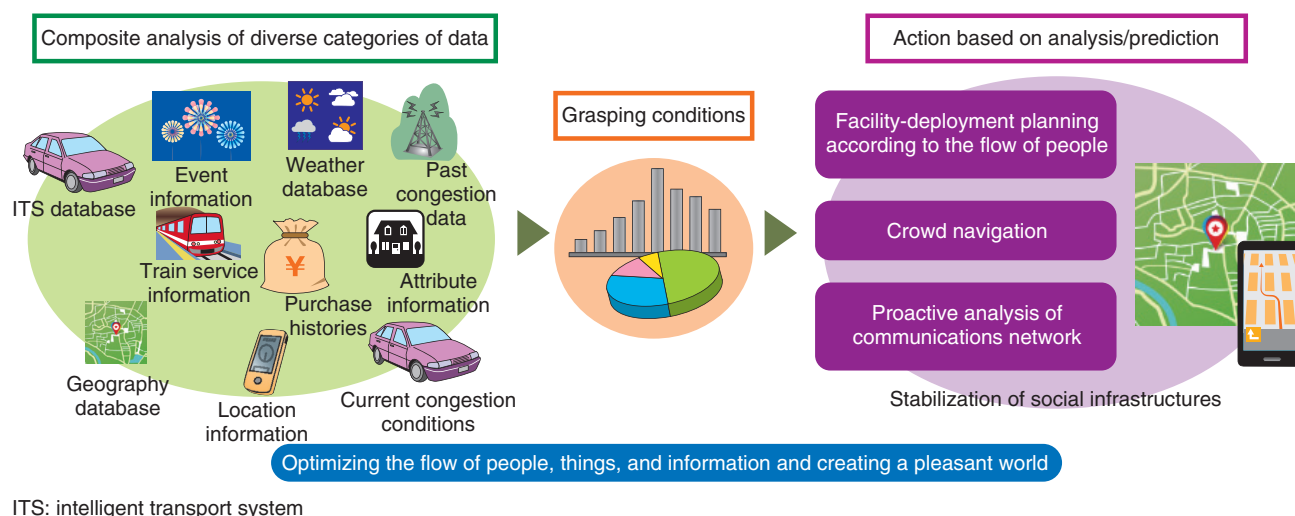
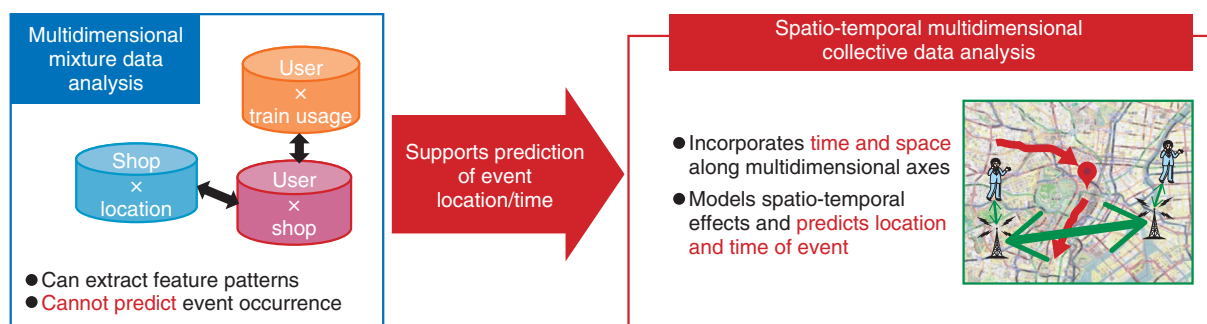Fig. 4.   Multidimensional mixture data analysis technology.

Fig. 5.   Spatio-temporal multidimensional collective data analysis technology.

would like to analyze data that can provide people with good hints as grounds for making important decisions. With this need in mind, we have come to establish and apply multidimensional mixture data analysis technology (**Fig. 4**). Moreover, with innovative analysis technology capable of predicting the future, our aim now is to foresee and obtain insight from things that will happen in the somewhat near future taking into account the four data aspects of time, space, multidimensionality, and collectivity (**Fig. 5**).

With a view to 2020, we envision using this technology to predict the occurrence of congestion at large event venues so that countermeasures can be taken and a stable communications infrastructure can be ensured. Furthermore, using such activities as a foundation, I would like to make social activities

more efficient. For example, it would be good if a governmental or administrative body could use this technology to make objective decisions such as on constructing a new road. I would also be pleased if the technology could be applied to the shared assets of society. I would like to join up with the NTT laboratories and partner companies to propose such applications to the real world and to create a variety of case studies.

In the past, when I was involved in analyzing the data of a company performing user surveys, I was very happy to hear comments from the company side such as "These results are the same as what we had thought to be true, so they corroborate our beliefs." This was gratifying because I was able to corroborate using mathematical data what is usually judged on the basis of "experience, intuition, and courage" in an

ordinary society. There is an aspect of my own research that is based on experience, intuition, and courage. This is why I have to try things on my own. It is also useful to corroborate what one has done on the basis of intuition using none other than data analysis. By the way, I don't think my intuition is that good, and as for courage, I wouldn't say I have more than anyone else! However, I have been able to gain much experience by being involved in research over many years, and this is very gratifying to me.

Actually, I have been able to secure a little more time for research than before, so I have been writing up patent applications based on ideas and creating programs and beginning experiments. I am excited about my work and look forward to the results. This is because as a researcher, I want to produce results that are good enough to have my basic research or papers cited by other researchers. In addition, as I look to the future, I feel it would be great if I could team up with other researchers or even students in these research endeavors.

—*Dr. Sawada, could you leave us with a word of advice for young researchers?*

Finding a research theme that is always on your mind, even while you're soaking in a hot bath as I mentioned before, is happiness for a researcher. Furthermore, on hearing about other research, it's important that you find it interesting as well. I believe that a researcher, in essence, is a person that finds something interesting in everything. Whatever you approach may be general in nature or stem from research in another specialized field, but it should not be something that exists solely in your world. Being motivated by external stimuli is good.

Nowadays, however, papers on the level of international conferences can be retrieved off the Internet, and new information is always being uploaded, so you may discover that your research is already over the instant you begin your search! However, even if your search results reveal that someone else has beat you to it, I think a good response would be: "It can't

be helped, but what matters is that I have chosen my direction." It is not rare for two or more groups in the world to be thinking about the same type of research at the same time. In such a case, I think it would be good that all groups concerned are acknowledged. Of course, if patents are involved, being the first to apply is important, but research also involves an element of having one's work understood, so while speed is important, please proceed without worrying too much. I am behind you in your work.

■ **Interviewee profile**
Hiroshi Sawada

Senior Research Engineer (Senior Distinguished Researcher), Head of Innovative Communication Laboratory, NTT Communication Science Laboratories.

He received a B.E., M.E., and Ph.D. in information science from Kyoto University in 1991, 1993, and 2001. He joined NTT in 1993. His research interests include statistical signal processing, audio source separation, array signal processing, machine learning, latent variable models, graph-based data structures, and computer architecture. From 2006 to 2009, he served as an associate editor of the Institute of Electrical and Electronics Engineers (IEEE) Transactions on Audio, Speech & Language Processing. He received the Best Paper Award of the IEEE Circuit and System Society in 2000, the SPIE ICA Unsupervised Learning Pioneer Award in 2013, and the Best Paper Award of the IEEE Signal Processing Society in 2014. He is an associate member of the Audio and Acoustic Signal Processing Technical Committee of the IEEE Signal Processing Society and a member of IEEE, the Institute of Electronics, Information and Communication Engineers, and the Acoustical Society of Japan.