Spotlight on NTT Laboratories

Toward Next-generation High-quality, High-reality Video Surpassing HDTV

Akira Nambu
General Manager, NTT Cyber Space Laboratories
(Cyber Communications Laboratory Group)

In this issue, we turn the spotlight on NTT Cyber Space Laboratories, which is known in Japan and overseas for its development of SARA, the world's first H.264 codec LSI (large-scale integrated circuit) for professional broadcast use. As an integral part of the NTT Cyber Communications Laboratory Group, Cyber Space Laboratories researches and develops media-related technologies to support the expansion of ubiquitous broadband services. We talked to General Manager Akira Nambu to learn more about research and development at Cyber Space Laboratories.

Role and mission of Cyber Space Laboratories

—What is the role and mission of Cyber Space Laboratories, which is known for its ongoing development and release of forward-looking technologies?

In a few words, we are involved in the research and development of application-oriented base technologies and upper-layer services. In contrast to the research laboratories in the NTT Information Sharing Laboratory Group that are focused on network-related technologies, Cyber Space Laboratories is researching and developing upper-layer applications and Web-2.0 next-generation Web services that will run on top of those network technologies. Our laboratories are also playing an important role in the convergence of telecommunications and broadcasting through the development and implementation of video compression technologies as broadcasting migrates toward an all-digital format. We work on R&D of speech and natural language media, video and graphics media, and open source software with the mission of developing technologies that will become a driving force behind NTT Group business.

—What are your core technologies?

In the field of speech processing, we have technology for high-quality, efficient encoding of speech, the effect of which can be seen in cell phones and IP (Internet protocol) telephony. We also have acoustics-related technologies such as our stereo echo canceller as well as speech-recognition and speech-synthesis technologies for converting the spoken word to text and converting text to fluent speech. In the field of natural language processing, we have text-analysis technology for extracting information from freely written text. And in the field of video, I can point out technology for high-quality, efficient encoding of video for use in DVDs and digital broadcasts, digital watermarking technology for adding invisible ID (identification) information to video, video monitoring technology for recognizing the video output from monitoring cameras, and 3D model-construction technology. Finally, in the field of open source software, we are researching and developing Linux systems and other types of operating systems and database management technologies such as PostgreSQL.

Photo 1. NILFS file system capable of restoring past states.

Making the ubiquitous broadband society a more pleasant and familiar experience through research strategy and developed technologies

—Cyber Space Laboratories develops a variety of technologies in support of a more pleasant information society. In this regard, what are your R&D targets making use of core technologies and what achievements have you made so far?

We have been involved in the digitization of speech and video for many years, but to enhance realtime communication in telephones and videophones, our target is high-reality communications through high-quality video and audio to enable *high-presence* conversations without transmission interruptions. At the cell-phone level, we will need a bit more time to achieve such high-reality communication, but I think we are quite close in the case of IP videophones (such as NTT's FLET'S Phone) and landline telephones.

In particular, in the R&D of technology for digitizing and compressing video, as cultivated through the development of the videophone since the days of the Nippon Telegraph and Telephone Public Corporation, we came to develop technology that became "MPEG-2", the key standard in the video world in this era of analog-to-digital conversion. The MPEG-2 HDTV (high-definition television) codec LSI (large-scale integrated circuit) called VASA that we developed was the world's first practical single-chip LSI for MPEG-2 encoding. This LSI evolved into SARA, a high-quality, high-compression H.264/MPEG-4 audio/video codec LSI that we announced in April 2007.

In our open-source computing project, we developed a file system called NILFS (**Photo 1**) that is capable of restoring past states. This system is being provided to the whole world as open source software. For the future, we would like to see NILFS evolve into an ultralarge-capacity grid-storage system that stores a variety of information on the network with the change history appended.

—In what fields are you applying research results?

We developed VASA for video transmission equipment used for relaying digital TV. At present, VASA technology can be found in various types of equipment in broadcast stations. We have also applied this technology in the development of MPEG-2 and codec chips for consumer devices, and these chips are now being used in home appliances such as video cameras, DVD players, and hard-disk recorders. At the same time, we are making progress in research on large-screen, high-quality video that surpasses HDTV. Last year, on New Year's Eve, we performed a UHDV (ultrahigh-definition video) experiment together with NHK. We successfully performed a live UHDV broadcast of NHK's popular New Year's Eve music show using our developed technology. This was the world's first attempt to transmit UHDV pictures, which have 16 times the resolution of HDTV, to a remote location over an IP network.

As for other achievements, we were the first to implement Internet searching based on natural Japanese-language statements. This service is now being provided on the popular "goo" portal. In addition to providing answers to questions entered in spoken-Japanese style, an outstanding feature of this service is that candidates for the words and expressions used in the reply are selected from Web pages themselves. Also, as an example of the application of our speech synthesis technology, we have developed singing-voice synthesis software called "WonderHorn" that has come to be used in the "Yoshimoto Chaku Messe" ring-tune creation service.

Power of researchers who pursue their targets persistently gives birth to new technologies

—What is the atmosphere like at Cyber Space Laboratories?

Cyber Space Laboratories is located in the Yokosuka R&D Center, which is surrounded by green mountains and the ocean at Miura. This is an outstanding location from which Mt. Fuji can be seen in the distance. This advantageous environment, which is depicted by the keywords "cheerful and enjoyable," promotes free and energetic R&D amongst laboratory personnel.

At present, with targets like high picture quality, high compression, and high reality fixed clearly in our minds, I feel that everyone is on a straight heading to meet their targets. We have about 200 researchers in all. Of these, 20 are women, which is a considerable percentage compared with other research laboratories. In fact, in the development of the H.264 video-encoding LSI and remote monitoring technology, for example, women played hard-working, leadership roles. We also have two foreign researchers on our staff as well as visiting overseas researchers who come from all over the world every year. Thirty members of our staff conduct their research at the Musashino R&D Center, and we communicate with them by holding laboratory meetings at Musashino and by using videoconferencing systems.

—What would be your message to researchers at Cyber Space Laboratories?

On assuming my current office, I said "Please increase your individual potential." If each individual researcher increases his or her potential, the potential of Cyber Space Laboratories will likewise increase, and research achievements will multiply. I also want researchers to go beyond their own little world and find out what the world at large really needs and to bring together individual technical expertise in a great team effort to research and develop things that will find widespread use in society.

—What issues do you expect to face in the future and what is your medium-term outlook?

One byproduct of the ubiquitous broadband era is information overload. To solve this problem, our aim through R&D is to provide useful knowledge in a timely fashion and to promote a vision of knowledge sharing. This calls for innovative technology development on a breakthrough level. A major problem is how to come up with revolutionary technology in each project that we take on.

I also think that services toward the commercialization of the next-generation network are important and that innovation in raising the quality of video is something that we, whose work is close to commercially oriented development, must achieve. Things that can be usefully applied, like the VASA LSI, can also expand as a business, and it is this aspect of R&D that I think we should focus on.

Short profile of Cyber Space Laboratories

VASA/SARA

VASA is the world's first MPEG-2 codec LSI for professional broadcasting. It enables the encoding and decoding of HDTV video to be performed on a single chip (**Photo 2**). It is being used, for example, in the digital relay network (trunk network) that supports digital terrestrial broadcasting. This operation began in December 2003. Multiple VASA chips can be used to achieve high-reality, large-screen video codec processing that surpasses HDTV. VASA stands for versatile and advanced signal processing architecture.

SARA is the world's first H.264 realtime codec LSI for professional broadcasting (**Photo 3**). It supports the high 4:2:2 profile of color sampling that is prescribed for high-quality video and that can be applied to the transmission of raw broadcast materials. It is equipped with a transcoding function that can handle the multi-stage links essential to broadcasting facilities. It also achieves a high compression rate on par with true H.264 performance and achieves HDTV ($1920 \times 1080i$) encoding/decoding processing in a package the size of a postcard. SARA stands for super advanced realtime codec.

Many outside commendations

Cyber Space Laboratories has had the honor of receiving many commendations from outside NTT. In fiscal year 2006, it received 19 commendations including a Science and Technology

Award from the Minister of Education, Culture, Sports, Science and Technology. In fiscal year 2005, it received 14 commendations. Recognition from the outside helps to motivate researchers, and research themes that could lead to commendations are actively encouraged.

Presentation scene

General Manager Akira Nambu's motto is "work can be enjoyable." At Cyber Space Laboratories, researchers make their presentations in a relaxed and sociable atmosphere (**Photo 4**).
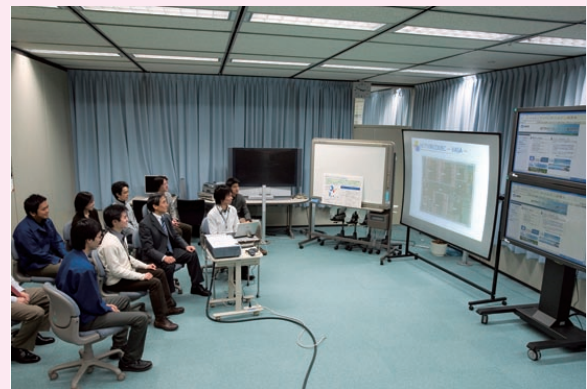

Photo 3.   SARA.

Photo 2.   VASA.

Photo 4.   Discussion with General Manager Nambu.

Notable places at Cyber Space Laboratories

Anechoic chamber and echoic chamber

The acoustics-experiments building at the Musashino R&D Center includes an anechoic chamber and an echoic chamber (**Photos 5** and **6**). Both of these chambers are vital to acoustic research, especially in the areas of speech-quality enhancement, generation of high-reality acoustic fields, and creation of rich acoustic environments. These chambers are known in Japan for their high-performance characteristics. The anechoic chamber is constructed so that the floor, walls, and ceiling absorb sound to achieve a space with no reflected sound. The echoic chamber, in contrast, is constructed so that the floor, walls, and ceiling reflect sound to contain it in the room and make the sound volume nearly constant throughout. It is also known as a reverberant chamber.

3D-object-restoration experimental room

This experimental room is used for 3D object restoration (**Photo 7**).
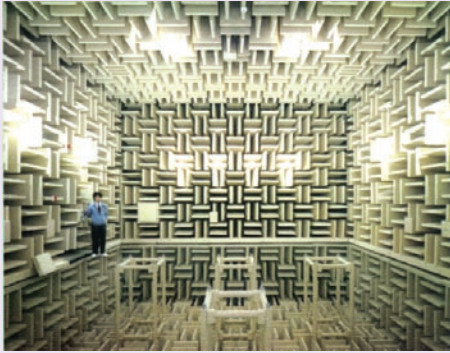
Photo 5.   Anechoic chamber.

Photo 6.   Echoic chamber.

Photo 7.   3D-object-restoration experimental room.