Annual Report

NHK
Science & Technology
Research Laboratories

2018

Nippon Hoso Kyokai
[Japan Broadcasting Corporation]

Table of Contents

Greetings

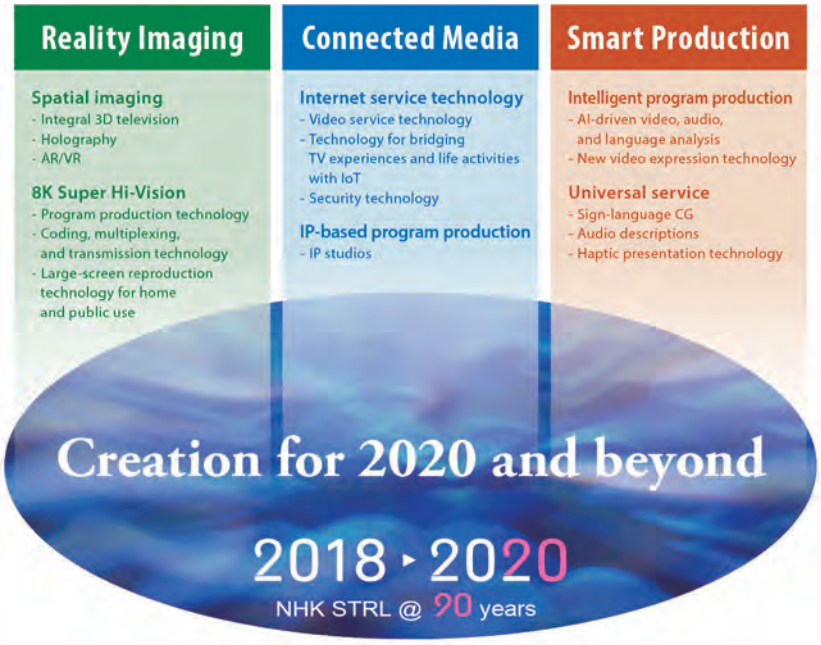
Kohji MITANI
Director of NHK Science & Technology Research Laboratories

NHK Science & Technology Research Laboratories (STRL), the sole research facility in Japan specializing in broadcasting technology and part of the public broadcaster NHK, is working to create a rich broadcasting culture through its world-leading R&D on broadcasting technologies.

Fiscal year 2018 saw the dawn of a new era of broadcasting with the launch of 4K and 8K satellite broadcasting services. In this milestone year, we are in the midst of implementation of our NHK STRL 3-Year R&D Plan (FY 2018-2020), which aims at creating new broadcasting technologies and services for 2020 and beyond, into 2030 and 2040. In accordance with this plan, we have been driving our R&D under the three pillar concepts of "Reality Imaging" to deliver video and audio with a higher sense of presence and reality, "Connected Media" to achieve more convenient broadcasting and services by the use of the internet, and "Smart Production" to deliver information of value to viewers in a timely manner by applying artificial intelligence (AI) technology.

This annual report summarizes our research results in FY 2018, the first year of our three-year plan. It is my hope that this report will serve as an impetus for you to better understand NHK STRL's research and development activities. I also hope it will help us build collaborative relationships that promote research and development and places of co-creation utilizing the results of our efforts.

Finally, I would like to express my sincere gratitude for your support and look forward to your continued cooperation in the future.

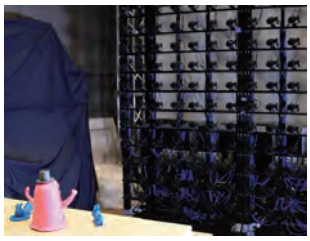


Accomplishments in FY 2018

Reality Imaging - Spatial imaging

NHK STRL is researching 3D imaging technology that offers more natural 3D images with the goal of providing a future broadcasting service that goes beyond 2D image presentation. We made progress in our research on a high-definition 3D imaging technology and a 3D imaging technology for portable devices and studied elemental technologies for 3D display devices. We also began investigating system requirements for enjoying 3D images in diverse viewing environments. In addition, we began studying new viewing experiences and services that utilize augmented reality (AR) and virtual reality (VR).

→See p. 4 for details.



High-definition 3D capture technology

Reality Imaging - 8K Super Hi-Vision

We continued with our R&D on program production in full-featured 8K, which is the ultimate format of 8K Super Hi-Vision (SHV). For video technologies, we developed a 1.25-inch 8K three-chip camera, a slow-motion recorder and player, and a full-featured 8K liquid crystal display (LCD). For audio technologies, we developed an adaptive downmixer that automatically downmixes 22.2 multichannel sound (22.2 ch sound) of a program to stereo or 5.1 multichannel sound for efficient sound production. For transmission technologies, we researched video coding standards, technologies for increasing the transmission capacity of satellite broadcasting systems and ways of enhancing media transport technologies. We also made progress in our development of a lightweight, thin and portable sheet-type display and in our research on a 22.2 ch sound reproduction system for home use for the easy viewing of 8K SHV at home. To further promote 4K/8K broadcasts, which were launched in December 2018, we progressed with our R&D on a broadcasting system for terrestrial 4K/8K broadcasting and developed a method for transmitting 4K/8K programs from cable TV stations by using coaxial cables installed in apartment buildings. We also developed a wireless camera using a millimeter-wave band to produce a wide variety of 4K/8K programs and researched a Radio over Ethernet (Radio over Fiber) technology that could increase the mobility of wireless cameras.

→See p. 9 for details.

1.25-inch 8K three-chip camera head

Connected Media

To make broadcasting more user-friendly and convenient by making use of the internet, we conducted R&D on a media-unifying platform that links TV with smartphones and IoT-enabled devices to offer more engaging broadcasting services. We also researched a high-response delivery technology to stabilize the viewing quality of video delivered on the internet and achieve smooth viewing operation. In our research on cryptography and information security, which are essential to ensure high security and reliability of these services, we progressed with our development of cryptography algorithms including one that can be used for post-quantum computer measures in the future. We also researched a TV-watching robot that makes TV viewing more enjoyable. With the aim of establishing an efficient workflow of IP-based program production, we developed IP-based transmission equipment and a device to monitor the status of IP packets. We demonstrated the effectiveness of each device through a verification experiment of IP live production connected with existing program production equipment.

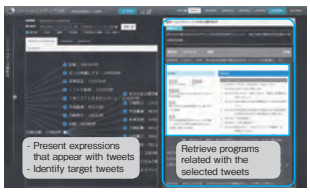
→See p. 22 for details.

Example of linkage between TV and smartphone

Smart Production - Intelligent program production

We are researching intelligent program production using AI technologies to achieve an efficient program production environment. We progressed with our research on a social media analysis technology aimed at extracting newsworthy information by using text big data such as social media. We also studied a technology for analyzing viewers' opinions about programs after broadcasts. For image analysis technologies, we researched a technology that can reduce the amount of work required for the colorization of monochrome video significantly and a technology for automatically producing digest videos of programs. In our work on speech recognition technologies, we investigated a recognition technology that uses the information of video as well as audio as a technology for supporting the transcription of speech in video footage. For more advanced program production, we researched a new image representation technique for sports and other live programs. For the effective use of AI technologies for program production, we also enhanced the development structure of Smart Production Lab, which we built in NHK STRL, and supported the practical application of our technologies such as an AI announcer using a speech synthesis technology.

→See p. 28 for details.



Retrieval of related programs on the basis of social media posts

Smart Production - Universal service

We are conducting research on universal broadcasting services that all viewers can enjoy. We studied an automated closed-captioning technology to convey TV program speech in text to those with hearing difficulties and the elderly and a sign language CG generation technology to automatically generate sign language CGs for explaining the status of sports events. In our work on services for visually impaired people, we researched an automated audio description technology, which automatically inserts commentary audio for providing supplementary explanation of a program, and a robot commentary technology, which automatically creates and reads out manuscripts describing the status of sports events. In response to the need for prompt provision of information to non-native speakers due to the increasing number of tourists visiting Japan, we studied a machine translation technology to translate news scripts from Japanese to English. We are also researching the use of sensation other than sight and hearing for conveying program information. As a haptic presentation technology, we investigated haptic devices that convey the moving direction of a ball and the timing a ball hits the racket. We also began studying the possibility of conveying information by smell.

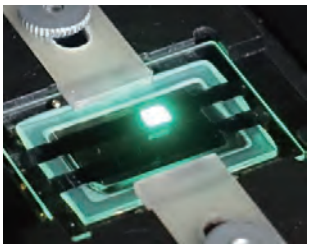
→See p. 33 for details.

Example of the screen layout of sports sign language CG service

Devices and materials for next-generation broadcasting

We continued researching elemental technologies of imaging, recording and display. For imaging technologies, we conducted research on 3D integrated imaging devices that could be applied to advanced imaging devices in the future, RGB-stack-type image sensors for compact, lightweight single-chip color cameras, and solid-state image sensors overlaid with multiplier films using a new structure for an 8K camera with higher sensitivity. For recording technologies, we conducted R&D on multi-level holographic memory to achieve a recording system with a very large capacity and high transfer rate for the long-term storage of 8K video. We also researched a high-speed magnetic recording device with no moving parts to achieve a high reliability and began investigating a recording method using novel magnetic materials. For display technologies, we worked to identify the principle of longer lifetime and higher color purity to realize a flexible OLED display. We also researched ways to improve the quality of a semiconductor film by a solution method to produce a large flexible display. We also progressed with our R&D on thin-film transistors (TFTs), which are driving elements of pixels, to increase the image quality and lower the power consumption of sheet-type displays.

→See p. 37 for details.



Quantum dot light-emitting diode

Research-related work

We promoted our research on 8K SHV and other technologies in various ways, including through the NHK STRL Open House, various exhibitions, and reports. We also actively collaborated with other organizations and program producers. The theme of the FY 2018 NHK STRL Open House was “Connecting to you smartly and realistically.” It featured a presentation of NHK STRL 3-Year R&D Plan (FY 2018-2020), which was announced in April, and exhibits on our latest research results. We contributed to the development of technical standards by participating in activities at international and domestic standardization organizations. We also cooperated with outside organizations through collaborative research and commissioned research efforts. We hosted visiting researchers from home and abroad and dispatched NHK STRL researchers overseas. We collaborated in the production of NHK sports programs. Our technologies were used in golf and ski jumping competition programs and also for the production of a figure skating program for the BS8K channel. In addition, our research results on AI technologies were utilized for the program production of NHK Special. In recognition of our research achievements, we received external awards including the Maejima Award.

→See p. 41 for details.

STRL Open House 2018

1 Reality Imaging - Spatial imaging

1.1 3D imaging technology

With the goal of developing a new form of broadcasting, NHK STRL is researching a spatial imaging three-dimensional (3D) television that shows more natural 3D images to viewers without special glasses. We conducted research on capture, display and coding technologies for high-resolution 3D images and on a 3D imaging technology for portable terminals for personal viewing. For the practical use of these technologies, we also worked to identify the characteristics of 3D images and the requirements for 3D imaging systems that are suitable for diverse viewing environments.

■ High-resolution 3D imaging technology

To capture high-resolution 3D images, we developed a technology for generating 3D images from images captured with a camera array[1]. This technology performs viewpoint interpolation process on the 154-viewpoint images captured with a camera array consisting of 154 HD cameras (Figure 1-1 (a)) to generate images of as many viewpoints as needed for 3D displays (Figure 1-1 (b)). This viewpoint interpolation process estimates the depth of an object from multi-view images and determines the pixels of interpolated images from the color information of the pixels of camera images. This technology enabled the generation of 3D images that have about 330,000 pixels. Using a prototype system, we captured persons and clay animations and exhibited the generated 3D images at the NHK STRL Open House 2018.

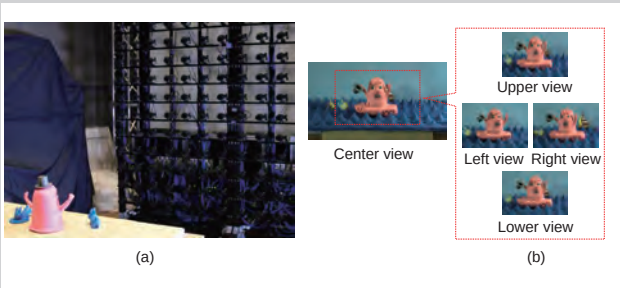


Figure 1-1. High-resolution 3D capture technology (a) Capture using a camera array (b) Examples of images generated from camera images by viewpoint interpolation

As a technology for displaying high-resolution 3D images, we developed a system called Aktina Vision, which uses multiple projectors and a special diffusion screen[2]. This system reproduces high-density multi-view images by arranging multiple projectors and superimposes the images on the display screen through an optical system. This enables the display of high-resolution 3D images (Figure 1-2 (a)). It can display 3D images with reduced crosstalk of light between light rays by adapting the diffusion characteristics of the display screen to the angle between projected light rays. Since the resolution of displayed 3D images agrees with the resolution of images from each viewpoint in this system, the resolution of 3D images can be increased more easily than the lens array system by increasing the resolution of multi-view images. Our prototype of this system using fourteen 4K projectors achieved the display of 3D images having 330,000 pixels, about three times that of a prototype lens-array system that we fabricated in FY 2017 (Figure 1-2 (b)).

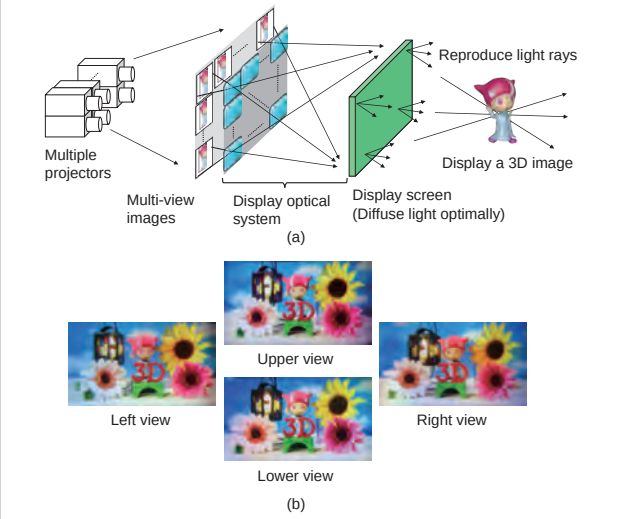


Figure 1-2. Aktina Vision (a) Configuration diagram (b) Displayed images from various viewpoints

Since high-resolution 3D images contain a huge amount of information, it is necessary to develop a high-efficiency coding technology to realize a 3D television. We continued to attend MPEG meetings and promoted standardization activities for 3D video coding standards. We conducted experiments on applying existing coding schemes to the test sequences of elemental images provided to the MPEG meeting and submitted the results to the meeting as input to contribute to the promotion of standardization.

■ 3D imaging technology for portable terminals

We continued with our R&D on integral 3D display with eye-tracking system with the aim of realizing a 3D image viewing service using portable devices for personal viewing. In FY 2018, we developed a method for the real-time generation of 8K elemental images in accordance with the viewer's eye

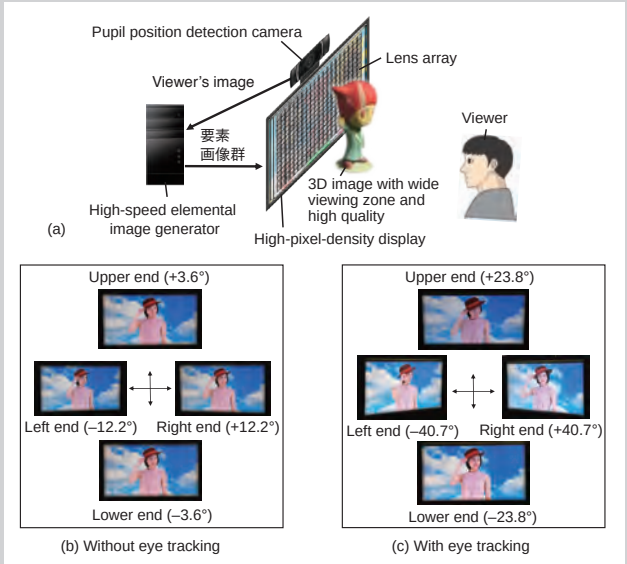


Figure 1-3. Integral 3D display with eye-tracking system (a) System configuration and viewing angle (b) Without eye tracking (c) With eye tracking

position and a viewing-zone formation method suitable for eye-tracking display[3]. Our prototype equipment using a high-pixel-density 8K organic light-emitting diode (OLED) display with a pixel density of 1058 ppi (pixel per inch) produced by Semiconductor Energy Laboratory Co., Ltd. achieved about 3.3 times the horizontal viewing zone and about 6.6 times the vertical viewing zone of a conventional device (Figure 1-3). The use of a lens array with a long focal length also almost doubled the light density for reproducing 3D images in both the horizontal and vertical directions over a prototype that we fabricated in FY 2017. This led to the improvement in image quality.

The integral method displays 3D images by adhering a lens array to elemental images shown on a direct-view display. This method, however, causes color moiré because the display's subpixel structure of red (R), green (G) and blue (B) is observed through a lens array. To reduce the color moiré, in FY 2018, we developed a method for optically shifting the pixels of elemental images in a time-sharing manner and multiplexing them. We conducted experiments to verify the operating principle of a system combining an OLED display (produced by Semiconductor Energy Laboratory Co., Ltd.), a lens array, and an optical wobbling device consisting of a double refraction element and a polarization control element (Figure 1-4). The results demonstrated that the system can reduce the color moiré to 66% that of a conventional device.

As a wide-viewing-zone imaging method suitable for 3D images on portable terminals, we studied a method for generating 3D models of an object from multi-view images and converting them into elemental images. In FY 2018, we produced high-quality 3D image content using a 3D model generation technology that supports 4K multi-view robotic cameras and a nonlinear depth compression technology. We

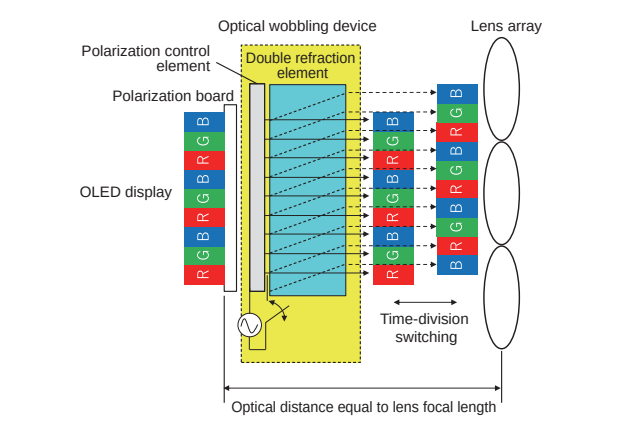


Figure 1-4. Structure of color moiré reduction method using optical wobbling device

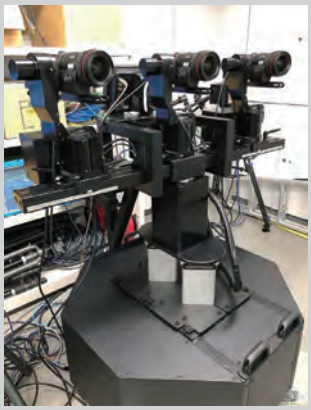


Figure 1-5. Stereo robotic cameras system

Figure 1-6. Interaction between 2D TV and table-type 3D display

also prototyped a stereo robotic cameras system to generate integral 3D images having a wider viewing zone with a smaller number of cameras (Figure 1-5). This system can control the directions of its three robotic cameras simultaneously, enabling camera work such as panning, tilting and zooming while capturing multi-view images of a 3D image reproduction area in a real space.

We investigated applications using the 3D imaging technology for portable terminals. As examples of such applications, we prototyped integral 3D displays (a table-shaped one and a smartphone-shaped one) that show 3D images in the air by using an optical system. We also prototyped an application that delivers and displays 3D images linked with 2D TV images on a 3D display and a system that enables 3D image viewing with an interactive operation. We exhibited them at the NHK STRL Open House 2018 (Figure 1-6).

3D image characteristics and 3D imaging system requirements suitable for the viewing environment

We are engaged in research to identify 3D image characteristics and 3D imaging system requirements that are suitable for diverse viewing environments. In FY 2018, we developed a new depth-compression method that could increase the quality of integral 3D display, which is a 3D imaging technology, and evaluated the influence of depth compression in a portable 3D display environment through psychological experiments.

In the integral method, an image tends to appear blurry when it is reproduced at a distance in the depth direction from the lens array on the display surface. To reduce the blurring, we use depth-compression expression that compresses the entire reproduced scene into a narrow depth while appropriately deforming the shape of the object to ensure the naturalness of its appearance. In the conventional (unidirectional) depth-compression technology, the shape of an object was compressed in the depth direction (the normal direction to the display surface). This caused a larger sense of unnaturalness when the viewer viewed the reproduced image obliquely from a different position from the original viewpoint because the deformation of the object was recognized. We therefore investigated an omnidirectional depth-compression method that compresses the object shape and space centering on the viewpoint. This method expresses a scene in a space of 3D polar coordinates whose origin point is the viewpoint and compresses the objects in the scene in the radial direction. This is expected to reduce the degree of recognizable deformation because the angular difference between the line of sight and the compression direction becomes smaller than that of the conventional method.

We analyzed and quantitatively evaluated the extent to which objects are deformed for display (the amount of distortion) when applying the unidirectional and omnidirectional depth-compression methods (Figure 1-7). We defined the amount of distortion on the basis of the ratio of the

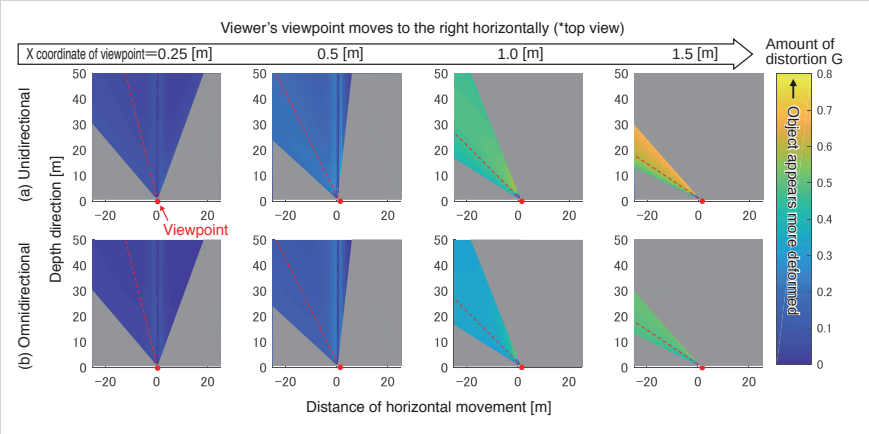


Figure 1-7. Comparison of the distortion amount at viewpoint movement between unidirectional and omnidirectional depth compressions

Figure 1-8. Experimental equipment for 3D imaging evaluation simulating a viewing environment with portable terminals

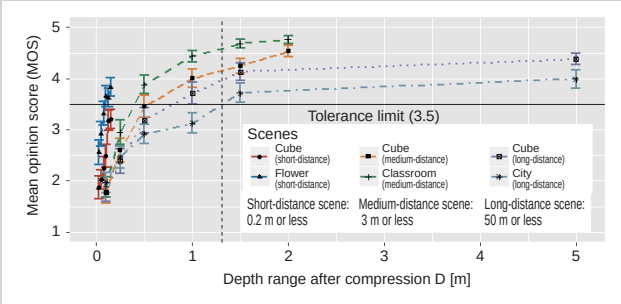


Figure 1-9. Relationship between the depth range after compression and the evaluation scores of naturalness

distance between two points on a 2D image that the viewer sees between before and after depth compression and used a scale with 0 for no deformation and a larger value for larger deformation. The results showed that the proposed method was more effective for reducing the amount of distortion and producing more natural expression than the conventional method particularly when the viewpoint position moved greatly in the horizontal direction (far-right upper and lower charts in Figure 1-7)[4].

Assuming a 3D image viewing environment using portable terminals, we conducted subjective evaluations in terms of the naturalness of depth-compressed images. For the experiments, it is necessary to evaluate the distortion of space and shape regardless of the depth reconstruction characteristics of the display equipment, but currently available integral 3D displays cannot provide a sufficient depth reconstruction range. We therefore developed a new experimental equipment for evaluation (Figure 1-8).

This equipment, which consists of a active shutter 3D system and a device for measuring the viewpoint position, displays images drawn in accordance with the viewpoint position and the display attitude. This enables the presentation of 3D images having motion parallax, which is a feature of the integral method. It is also possible to handle, with a small force, the position and direction of the display, which is suspended from the ceiling by a constant force spring and a wire.

The participants evaluated the unnaturalness of depth-compressed images (scenes with original depth ranges of 0.2 - 50 m compressed into 0.1 - 5 m). Figure 1-9 shows the relationship between the depth range after compression and the evaluation scores of naturalness. The results demonstrated that the minimum depth range for causing unnaturalness equal to or higher than the acceptable threshold (MOS = 3.5) was about 1.3 m. This value slightly exceeds the result (1 m) previously obtained in a viewing environment with a fixed display (standard TV in living room). The unnaturalness after depth compression tends to be conspicuous when the viewpoint moves largely. This is considered to be the reason why the range after necessary depth compression slightly increased for a viewing environment using portable terminals, on which relative viewpoint movement tends to be larger.

[References]

(1) M. Kano, H. Watanabe, M. Kawakita and T. Mishina: "3D Imaging with 330,000 Pixels Using Camera Array," ITE Winter Annual Convention 2018, 23D-1 (2018) (in Japanese)

(2) H. Watanabe, N. Okaichi, T. Oomura, M. Kano, H. Sasaki and M. Kawakita: "Light field display system with resolution of 330,000 pixels using top-hat diffusing screen," ITE Winter Annual Convention 2018, 23D-2 (2018) (in Japanese)

(3) N. Okaichi, H. Sasaki, H. Watanabe, K. Hisatomi and M. Kawakita: "Integral 3D display with eye-tracking system using 8K display," ITE Winter Annual Convention 2018, 23D-3 (2018) (in Japanese)

(4) Y. Miyashita, Y. Sawahata, T. Morita, K. Komine: "Design of depth compression method for natural 3D visualization," ITE Annual Convention 2018, 12A-1 (2018) (in Japanese)

1.2 AR/VR

We launched new research on augmented reality (AR) and virtual reality (VR) and studied the concept of services that will offer new user experiences to viewers by using AR and VR technologies. In FY 2018, we researched "By AR/VR" services, which provide new viewing experiences by combining existing technologies, and "For AR/VR" services, which provide new viewing experiences by implementing technologies that have yet to be introduced in the AR/VR field and newly developed technologies.

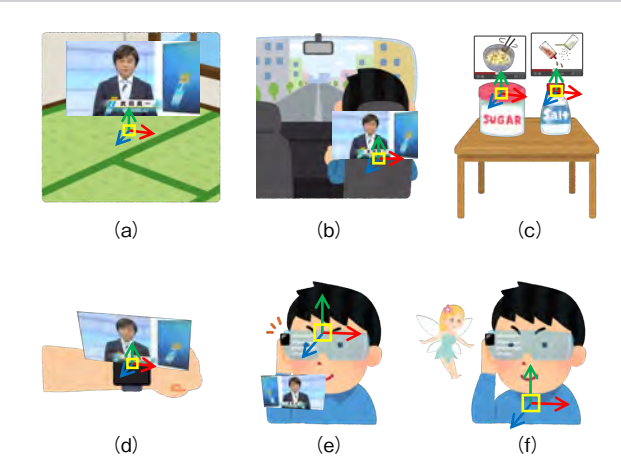


Figure 1-10. Coordinate systems suitable for each usage style

Table 1-1. Types of usage styles and ideal presentation size and coordinate systems

Examples of viewing experiences	Type of usage style	Expected ideal coordinate systems	Size*
Program use, Game	Focus on images for entertainment	World/Ambient environment/Head	Small - Large - 360
Information retrieval, Reference to manuals/recipes	Refer to something	World/Ambient environment/Object/ Wrist/Head	Small
SNS, Email	Handle private information	Ambient environ-ment/Wrist/Head	Small
Video chat, VR communication	See facial expressions and avatars	World/Ambient environment/Wrist/ Head	Small - Large - 360
Signboard, Ad	Signage	World/Ambient environment/Object	Small - Large
Pet, Partner	Virtual creature	Head/Body (Loose tracking)	Small

* (Size guideline) Small: Smartphone/tablet screen, Large: TV screen, 360: 360-degree image

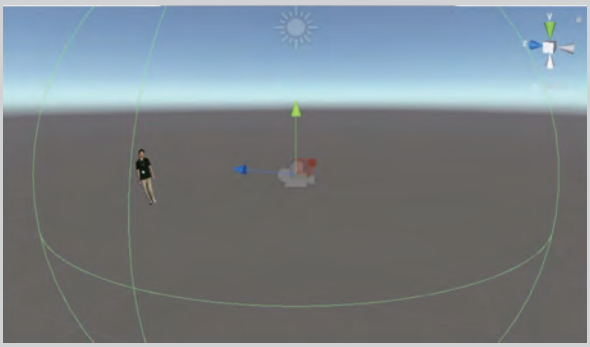


Figure 1-11. Presentation by 360-degree spherical images with an alpha channel

"By AR/VR" services

We investigated how 2D images should be presented in a 3D virtual space of AR and VR. We organized the styles of using 2D images, the size of presented images and coordinate systems for the use of AR glasses. We divided the coordinate systems into six categories of (a) world coordinate system, (b) ambient environment coordinate system, (c) object coordinate system, (d) wrist coordinate system, (e) head coordinate system and (f) body coordinate system (Figure 1-10), and proposed appropriate coordinate systems for each style of use (Table 1-1).

Focusing on 360-degree spherical images with an alpha (opacity) channel, we developed a method for cutting out an object to be presented from a 2D image and presenting it framelessly in a 3D space. We also devised a method that presents the images of an object with parallax by changing the radius of 360-degree spherical images with an alpha channel according to the distance between the camera and the object (Figure 1-11).

"For AR/VR" services

We investigated services that utilize high-resolution VR images to provide a new viewing experience. In FY 2018, we produced high-resolution VR images necessary for our investigation by combining ("stitching") images captured with an array of multiple 8K broadcast cameras equipped with a fisheye lens. This approach succeeded in producing highly immersive images that can be used for head-mounted displays and other displays. We conducted capture experiments in multiple scenes using several production methods with different shooting distances and camera arrangements. For the shooting distance, we captured images at different distances from short to long. For the camera arrangement, we tried three different methods. We tested and investigated a method in which two 8K cameras are arranged facing opposite directions to capture 360-degree images and a method in which two 8K cameras are arranged side by side to obtain 3D VR images having binocular parallax in a forward direction. We also devised a new method for capturing VR images covering a viewing angle of 180 degrees or more by arranging three 8K cameras in the forward horizontal direction at different angles to increase the resolution to the maximum and allow flexible operation of production effects during program production. This method produced VR images that exceed 8K (Figure 1-12). We plan to examine the concept of services using high-resolution VR images by applying these methods.

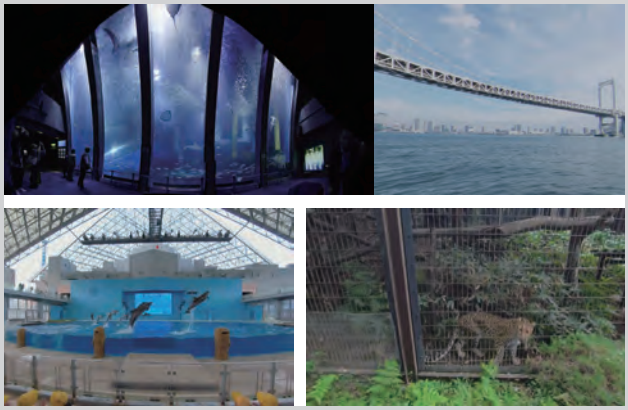


Figure 1-12. High-resolution VR images produced

[References]
(1) Kawakita, Yoshino, Koide, and Hisatomi: "Ideal form of 2D moving picture in 3D space," HCG Symposium 2018, HCG2018-I-1-1 (2018)
(2) Koide, Yoshino, Kawakita, and Hisatomi: "A study on high-immersive sense moving picture production for a high-definition virtual reality system using 8K-resolution cameras," ITE Technical Report, Vol.43, No.10, BCT2019-47, pp.45-48 (2019)

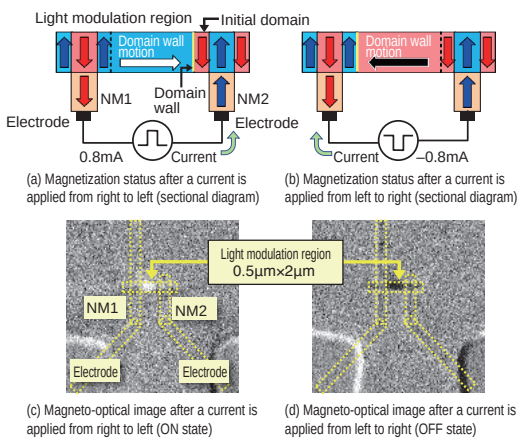
1.3 3D imaging devices

Ultra-high-density spatial light modulator

We are engaged in research on electro-holography that shows natural three-dimensional (3D) motion images. Displaying 3D images in a wide viewing zone requires the development of a spatial light modulator (SLM) having a very small pixel pitch.

We are researching a magneto-optical (MO) SLM driven by spin transfer switching (spin-SLM) that uses magnetic materials as pixels. The spin-SLM can modulate light by using the MO effect, in which the polarization plane of reflected light rotates according to the magnetization direction of the magnetic materials. We previously prototyped a device that can switch the magnetization direction by magnetic domain wall motion induced by pulse currents applied to the magnetic materials (a light modulation device driven by current-induced domain wall motion) and successfully verified its basic operating principle.

In FY 2018, we optimized the composition of the light modulation layer made of a gadolinium-iron alloy, and succeeded in driving at a low-current of 0.8 mA. The current can be supplied by a cell-selection micro transistor with a size of 1 μm×2 μm, which size is required for wide viewing 3D holographic displays. Using this composition, we prototyped an element with a size of 0.5 μm×2 μm and evaluated its light modulation operation (Figure 1-13). Each end of the light modulation region has a nano magnet having an antiparallel configuration of magnetization direction (NM1, NM2), through which currents flow into the light modulation region. An initial domain is formed at an end of the light modulation region by local magnetic fields from the nano magnet. We confirmed that applying a current from right (left) to left (right) in this state expanded the area of the initial magnetic domain to the entire area of the light modulation region by domain wall motion, which turned on (white) or off (black) the reflected light. This successfully demonstrated the light modulation operation using a micro device[1].



(a) Magnetization status after a current is applied from right to left (sectional diagram)
(b) Magnetization status after a current is applied from left to right (sectional diagram)
(c) Magneto-optical image after a current is applied from right to left (ON state)
(d) Magneto-optical image after a current is applied from left to right (OFF state)

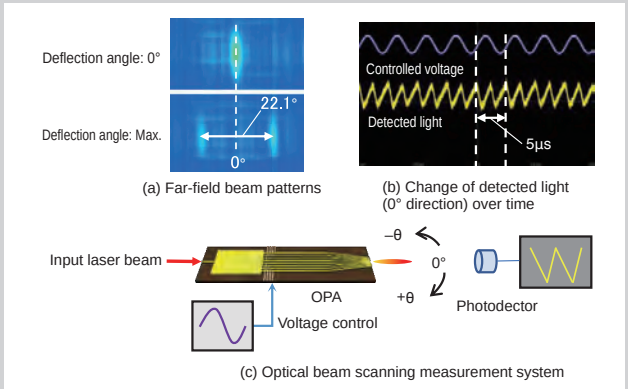
Figure 1-13. Light modulation operation of a spin light modulation element driven by current-induced domain wall motion

Elemental technologies for optical phased array

For a future integral 3D display with much higher performance than current displays, we are conducting research on a new beam-steering device that can control the direction of optical beams from each pixel at a high speed without using a lens array. Focusing on an optical phased array (OPA) consisting of multiple optical waveguides (channels) as a beam-steering device, we designed, fabricated and evaluated an OPA using an electro-optic (EO) polymer that can change the refractive index at a high speed on each channel by applying an external voltage.

The prototype OPA using an EO polymer can flexibly change the direction of output beam by applying a voltage via the channels to change the refractive index of the EO polymer and control the optical phase. We previously designed and prototyped an OPA consisting of eight channels and demonstrated an optical beam deflection of ±3.2 degrees.

In FY 2018, we developed a technology for tightly confining light in a waveguide by using a material having a large refractive index difference for an optical waveguide, which reduced the crosstalk between channels. This technology made it possible to narrow the beam output channel waveguide pitch of an OPA to 4 μm and achieved an optical beam deflection of 22.1 degrees. We also made an optimum layout design of the device, which decreased light propagation loss at the curve and bifurcation of an optical waveguide and significantly reduced unnecessary stray light components of far-field beam patterns. This increased the peak value of beam intensity of the prototype OPA and also succeeded in optical beam scanning at a high speed of 200 kHz[2] (Figure 1-14).



(a) Far-field beam patterns
(b) Change of detected light (0° direction) over time
(c) Optical beam scanning measurement system

Figure 1-14. Experiment result of optical phased array

[References]
(1) R. Higashida, N. Funabashi, K. Aoshima, K. Machida: "W-spin domain wall type optical modulation device for electronic holography," 2018 ITE Annual Convention Program, 32C-3 (2018) (in Japanese)
(2) Y. Hirano, Y. Miyamoto, Y. Motoyama, K. Machida, K. Tanaka, T. Yamada, A. Otomo and H. Kikuchi: "Beam Deflection of the Optical Phased Array using Electro-Optic Polymer Waveguide Arrays of 4 μm pitch," The journal of the ITE Vol. 73, No. 2, pp. 392-396 (2019) (in Japanese)

2 Reality Imaging - 8K Super Hi-Vision

2.1 Video systems

We are conducting R&D on a program production system for full-featured 8K, which is the ultimate format of Super Hi-Vision (SHV), and research on program production methods for high-dynamic-range television (HDR-TV) and its synergy effects.

■ Full-featured 8K program production system

We continued with our R&D on program production equipment and systems that support a 120-Hz frame frequency with the goal of realizing full-featured 8K program production. To enable live program production and transmission experiments outside the broadcast stations, we installed our previously developed full-featured 8K production equipment and transmission equipment using dense wavelength division multiplexing[1] in an experimental production van[2]. Using the experimental production van, we conducted operation tests and confirmed that there was no major issue with the interconnectivity and operability of the installed equipment (Figure 2-1).

As new full-featured 8K production equipment, we developed a video editing system that can input/output 8K/120-Hz video and 22.2 ch sound in real time to enable the creation of highlight video during live program production. We also began a study on full-featured 8K production using Internet Protocol (IP) interfaces and prototyped a device for comparative evaluation of the image quality of multiple lightweight compression codecs.

We researched measurement methods for the sensitivity and signal-to-noise ratio of TV cameras, which are useful for full-featured 8K program production, and compiled our findings into ARIB Technical Report TR-B45. In addition, we analyzed the relationship between the edge direction of an image and the measurement accuracy and precision of the modulation transfer function (MTF), which indicates the spatial resolution characteristics of a TV camera[3]. This led to the identification of practical measurement conditions and parameters[4].

■ High-dynamic-range television

For the program production of high-dynamic-range television (HDR-TV), which can reproduce a wider range of brightness than conventional televisions, we developed a conversion method between high dynamic range (HDR) content and standard dynamic range (SDR) content. We adopted a dynamic range conversion function in which a linear function and log function are combined for converting HDR content to SDR content and demonstrated that the function can produce converted SDR content with a quality comparable to that of directly produced SDR content [5]. We contributed these results to the International Telecommunication Union, Radiocommunication Sector (ITU-R) and our contribution were incorporated into ITU-R Report BT.2446.

Since HDR video expands the range of brightness that can be reproduced, it is necessary to set an objective metric for ensuring consistency in brightness to prevent unpleasantness given by excessively bright images to viewers. We began a study to develop the metric. We statistically demonstrated through subjective evaluation tests that the average of displayed pixel luminance level is not sufficient for the metric and found that correction using the angle at which each pixel is viewed and the luminance distribution is effective[6]. We also studied the perceived diffuse white level when viewing HDR video. The results of subjective evaluation tests demonstrated that the perceived diffuse white level is not the same as the peak luminance of a display but lower than the peak luminance[7].

A display color gamut renders a complex shape in a 3D uniform color space (a space in which a perceived color difference corresponds to a distance in the 3D space represented by hue, chroma and lightness) that can be reproduced by displays supporting HDR and a wide color gamut, and it is therefore difficult to represent it in two dimensions. To address this problem, we devised a "Gamut Rings" that develops a color gamut in 2D while saving the color volume per unit lightness and unit hue angle[8]. This method enabled the 2D evaluation of the volume and shape of a color gamut represented in 3D.

Figure 2-1. Experimental full-featured 8K production van (left) and operation test (right)

[References]

(1) T. Hayashida, D. Koide, J. Yonai, A. Arai, S. Aoki, K. Miura, Y. Takiguchi, Y. Nishida: "Development of compact dense wavelength division multiplexing transceivers for full-featured 8K," ITE Annual Convention 2018, 13D-2 (2018) (in Japanese)

(2) A. Arai, T. Hayashida, D. Koide, J. Yonai, S. Aoki, K. Miura, Y. Kikuchi, Y. Takiguchi, Y. Nishida: "Development of Full-Featured 8K Production system Outside Broadcasting Station," ITE Annual Convention 2018, 13D-3 (2018) (in Japanese)

(3) K. Masaoka: "Accuracy and precision of edge-based modulation transfer function measurement for sampled imaging systems," IEEE Access, Vol.6, pp.410479-41086 (2018)

(4) K. Masaoka: "Practical edge-based modulation transfer function measurement," Optics Express, Vol.27, No.2, pp.1345-1352 (2019)

(5) K. Nomura, Y. Kusakabe, Y. Ikeda, and Y. Nishida: "Method for HDR to SDR conversion considering HDR Reference White," Proc. IDW/AD'18 (2018)

(6) Y. Ikeda, and Y. Kusakabe: "Subjective Evaluations on Perceptual Image Brightness in High Dynamic Range Television," Proc. IS&T Electronic Imaging (2019)

(7) F. Jiang, K. Masaoka, and M. Fairchild: "Subjective Verification for Color Gamut Volume of HDR WCG Display," Proc. IDW/ AD'18 (2018)

(8) K. Masaoka, F. Jiang, M. Fairchild, and R. Heckaman: "2D Representation of Display Color Gamut," SID International Symposium (2018)

2.2 Cameras

1.25-inch full-featured 8K three-chip camera

With the aim of making a full-featured 8K SHV camera more compact and practical, we prototyped a full-featured 8K three-chip camera using a 1.25-inch optical format[1] (Figure 2-2). As image sensors, we used three 1.25-inch 33-megapixel CMOS image sensors supporting a 240-Hz frame frequency (that we developed in FY 2017) for RGB signals. We employed a system that performs sensor driving and signal processing all at 240 Hz and outputs full-featured 8K signals with a 120-Hz frame frequency by applying two-frame averaging at the last stage of signal processing in the camera control unit (CCU). This enabled the camera to be also used as an 8K 4× high-speed camera. To transmit image sensor signals driven at 240 Hz from the camera head to the CCU, we used the newly-developed optical transmission interface that can transmit signals of 448 Gbps in total over four optical fibers. Using a bidirectional wavelength division multiplexing technology, the interface can transmit 40-Gbps signals from the CCU to the camera head, which enables the transmission of return video signals with higher resolution to the viewfinder. The camera head and the CCU are connected by a cable including four optical fibers. The prototype camera head weighs 18 kg and the CCU weighs 37 kg, in the 8-RU size, both of which are less than half the size and weight of those of a conventional 8K three-chip camera.

We verified the effect of a noise reduction technology that we implemented on the 1.25-inch 33-megapixel image sensor[2]. The image sensor that we developed contains a folding-integration analog-to-digital converter (ADC) with a three-stage pipelined ADC architecture, and multiple sampling operations in the folding-integration ADC reduced noise. There is a limit to the number of samplings that can be performed during driving at a 120-Hz frame frequency. We therefore built a simulation model to investigate the relationship between the number of samplings and noise and estimated the noise reduction effect of this method. The calculated results closely agreed with measured values and indicated that six sampling operations, which can be performed in 120-Hz driving, are sufficient for noise reduction. This research was conducted in cooperation with Shizuoka University.

To realize autofocus (AF) function for 8K cameras, we developed a method for estimating the distance to the object precisely and with a small amount of calculation by using on-chip phase detection[3]. The on-chip phase detection system uses the amount of shift in images that is obtained from phase-difference detection pixels implemented on the image sensor. Applying complex analysis using rectanglar basis functions for phase calculation reduced the computation volume significantly compared with conventional distance estimation methods and achieved high-precision distance estimation.

8K slow-motion system

To enable 8K slow-motion imaging in the production of sports and other programs, we developed an 8K slow-motion system that consists of an 8K high-speed camera and an 8K slow-motion recorder/player.

We added high-speed capture modes to our 1.25-inch full-featured 8K camera to achieve the 8K high-speed camera that supports a 4× capture mode (4:4:4 color sampling, capture at a 240-Hz frame frequency) and an 8× capture mode (4:2:0 color sampling, capture at a 480-Hz frame frequency). We also upgraded a slow-motion recorder/player supporting 8K 4× speed[4], which we prototyped in FY 2017, to make it able to support signal input of up to 8× speed (480 Hz) in agreement with the output of the high-speed camera (Figure 2-3). We also added a frame interpolation mode using two-frame averaging to the signal output function to enable smooth slow-motion video reproduction. We used this slow-motion system for the 8K live program production of international sport events and exhibited it at international exhibitions such as the NAB Show 2018 and IBC 2018.

Figure 2-2. 1.25-inch 8K three-chip camera head

Figure 2-3. 8K slow-motion recorder/player and controller

8K solid-state image sensor overlaid with multiplier film

The sensitivity of 8K SHV cameras decreases as they become more compact and support higher frame rates because the amount of light incident on each pixel of the imaging device decreases. As a drastic solution to this problem, we are developing a high-sensitivity solid-state image sensor overlaid with a photoconductive film (multiplier film) on a CMOS circuit for reading signal charges (Figure 2-4). The multiplier film can obtain the effect of electric charge multiplication by only applying a low voltage. In FY 2018, we worked to improve the image quality of the 8K solid-state image sensor overlaid with a multiplier film that we prototyped in FY 2017 and investigated ways to reduce the noise of an 8K CMOS circuit on which a multiplier film is overlaid.

Previously, crystalline selenium films that constitute multiplier films were formed by immediately heating amorphous selenium, in which atoms are arranged irregularly, at high temperature. However, the degradation of surface flatness and the internal stress of the film those result from

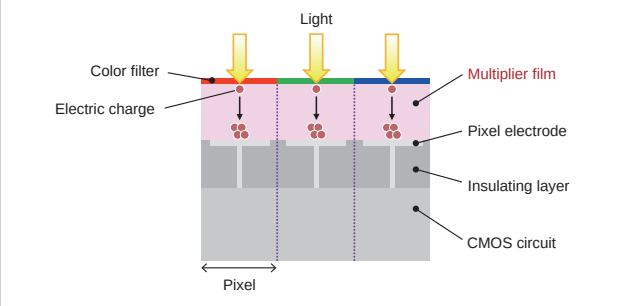


Figure 2-4. Structure of solid-state image sensor overlaid with multiplier film

sudden grain growth caused the image quality degradation of the sensor such as increased dark current and pixel defects. To address this problem, we introduced a two-step annealing method that heats selenium at low temperature to crystallize it briefly and then heats it again at high temperature for recrystallization after depositing an electron blocking layer and others on the crystallized selenium. This method successfully suppressed the degradation of image quality (Figure 2-5)[(5)(6)].

To increase the sensitivity of image sensors overlaid with multiplier films, it is important not only to multiply electric charges in the film but also to reduce the noise of the CMOS circuit. We therefore clarified the mechanism of dark current generation in the floating diffusion layer (i.e., a capacity for accumulating electric charges generated in the film) in the CMOS circuit and adopted a new structure that prevents the injection of dark current into the floating diffusion layer from the outside. We confirmed that this can reduce the shot noise of dark current to 1/8 that of the conventional structure[(7)]. We also began prototyping test circuits with the aim of reducing noise that occurs when electric charges accumulated in the floating diffusion layer are reset after signal readings (reset noise) and improving the charge-voltage conversion gain.

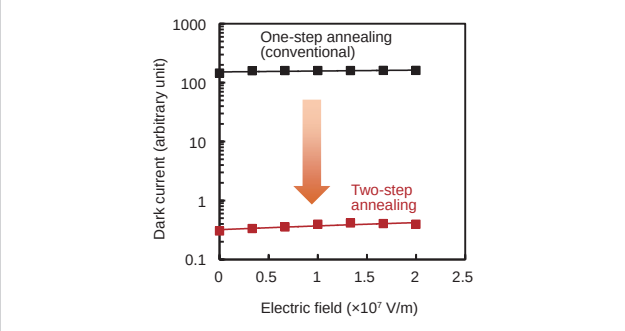


Figure 2-5. Dark current characteristics of device

[References]

(1) R. Funatsu, T. Kajiyama, T. Yasue, K. Kikuchi, K. Tomioka, T. Nakamura, H. Okamoto, E. Miyashita and H. Shimamoto: "8K 240-Hz full-resolution high-speed camera and slow-motion replay server systems," IBC 2018 Conference (2018)

(2) K. Tomioka, T. Yasue, R. Funatsu, T. Matsubara, T. Kosugi, S. W. Jun, T. Watanabe, M. Nagase, T. Kitajima, S. Aoyama and S. Kawahito: "Noise suppression effect of folding-integration applied to a column-parallel 3-stage pipeline ADC in a 2.1μm 33-megapixel CMOS image sensor," IS&T International Symposium on Electronic Imaging, IMSE-367 (2019)

(3) H. Okamoto, R. Funatsu and H. Shimamoto: "A Proposal of Distance Estimation Method from Images," ITE Annual Convention 2018, 14C-4 (2018) (in Japanese)

(4) T. Kajiyama, K. Kikuchi, R. Funatsu, E. Miyashita and H. Shimamoto: "8K240fps slow-motion replay server," ITE Annual Convention 2018, 31C-3 (2018) (in Japanese)

(5) S. Imura, K. Mineo, K. Miyakawa, M. Nanba, H. Ohtake and M. Kubota: "Fabrication of polycrystalline selenium with high crystallinity and surface flatness using lateral grain growth by the effect of surface capping," Ext. Abstr. of the 65th JSAP Spring Meet., 17p-P3-1 (2018) (in Japanese)

(6) S. Imura, K. Mineo, Y. Honda, T. Arai, T. Watabe, K. Miyakawa, M. Nanba, H. Ohtake and M. Kubota: "Reduction of pixel defects in stacked 8K image sensors using crystalline selenium films grown by two-step annealing," ITE Annual Convention 2018, 12C-4 (2018) (in Japanese)

(7) T. Arai and H. Ohtake: "Evaluation of the dark current of three-transistor pixel circuits for 8K image sensors with a photoconductive layer," ITE Annual Convention 2018, 12C-5 (2018) (in Japanese)

2.3 Displays

We have made progress in our research on large sheet-type displays and continued with our development of various displays that can handle full-featured 8K video.

Sheet-type display technologies

We are developing a lightweight, and thin sheet-type organic light-emitting diode (OLED) display for the easy viewing of immersive 8K video at home. In FY 2018, we developed an 88-inch 8K sheet-type display that uses a sheet glass substrate only several millimeters thick and exhibited it at various domestic and international events such as the NHK STRL Open House 2018 and the IBC 2018 (Figure 2-6). This display was demonstrated in cooperation with LG Display and ASTRODESIGN, Inc.

Figure 2-6. 88-inch 8K sheet-type OLED display

Full-featured 8K liquid crystal display

We are developing a full-featured 8K liquid crystal display that can be mounted on an outside broadcast (OB) van. In FY 2018, we developed a 31.5-inch display that supports HDR, a wide color gamut and a 120-Hz frame frequency (Figure 2-7) in cooperation with Sharp Corporation. The display uses the U-SDI interface and employs the local dimming technology, which individually controls the amount of light of the LED backlight for each area of the screen to increase the contrast.

Figure 2-7. Full-featured 8K liquid crystal display

By increasing the number of dividing areas for local dimming, the display achieved higher HDR image quality than with conventional displays.

Full-featured 8K projector

For the large-screen viewing of full-featured 8K video, we are developing an 8K projector that adopts red, green and blue laser diodes as light sources to achieve a wide color gamut and supports HDR, and a 120-Hz frame frequency (Figure 2-8). In FY 2018, we reduced the speckle noise caused by coherency of laser light sources (i.e., local brightness variations on a screen due to the interference of laser light) to improve the image quality. A speckle noise pattern depends on the polarization direction. We therefore used an optical system that can temporally change the polarization direction of the output light of the projector to change the noise pattern. This reduced the perceived speckle noise to about 70% that of a conventional projector.

Figure 2-8. Full-featured 8K projector

2.4 Sound systems

In our work on sound systems, we are researching a 22.2 multichannel sound (22.2 ch sound) system for 8K SHV and next-generation audio services and working on their domestic and international standardization.

Adaptive downmixer

We are studying adaptive downmix technologies for 22.2 ch sound to produce efficiently and simultaneously high-quality stereo and 5.1 surround. In FY 2018, we developed an adaptive downmixer capable of real-time processing by combining coherence control[1] to compensate the tone deterioration due to downmixing and loudness chasing[2] to compensate the integrated loudness levels of downmixed signals.

Reproduction of 22.2 ch sound at home

Figure 2-9. Loudness chaser

We are researching technologies for the easy reproduction of 22.2 ch sound at home. We continued with our study on binaural reproduction[3] using line array loudspeakers. In FY 2018, we developed an algorithm for separately controlling each channel and a design method for a reproduction controller optimized for it. We also implemented the design method into our signal processing device. The device was developed in cooperation with Sharp Corporation.

Figure 2-10. 22.2 ch sound reproduction system for home use

Next-generation audio services

We are studying next-generation audio services for advanced terrestrial TV broadcasting. In FY 2018, we added a function to synchronize an audio definition model (S-ADM: Serial-Audio Definition Model, ITU-R BS.2125-0) with audio signals and output them to an MPEG-H 3D Audio[5] encoder using a digital audio interface. The ADM is audio-related metadata that describes the configuration of sound materials and the location information during playback in program audio. We also developed a real-time audio encoder/decoder using MPEG-H 3D Audio[6]. This led to the successful development of metadata control, object control and rendering functions using S-ADM.

This research was conducted in cooperation with the Fraunhofer Institute for Integrated Circuits, Germany.

Acoustic devices

We have been researching a shotgun microphone, which has sharper directivity than a conventional one. Aiming for better performance in a low frequency band and an environment with diffusive noise, in FY 2018, we developed a microphone that can record target sounds more clearly by using sub-arrays set in the microphone and signal processing in combination[7].

We also developed an 8 ch switching amplifier for thin loudspeakers using a piezoelectric bendable electro-acoustic transducer with the goal of realizing a practical 22.2 ch sound loudspeaker for home use (Figure 2-11). This research was conducted in cooperation with Fujifilm Corporation.

Standardization

We are engaged in domestic and international standardization activities to promote 22.2 ch sound systems and realize next-generation audio services.

Figure 2-11. 8 ch switching amplifier for thin loudspeakers

At ITU-R, we produced a New Recommendation on S-ADM, ITU-R BS.2125. We also added required bit rates for MPEG-4 AAC, which is used for transmitting 22.2 ch sound materials, to the Recommendations on audio coding schemes, ITU-R BS.1548-6, BS.1196-7 and BT.1872-2, on the basis of our evaluation results[8]. We produced new Reports ITU-R BS.2419-0, which describes room response adjustment for 3D sound systems including 22.2 ch sound, and ITU-R BS.2434-0, which describes the standards and examples for loudness levels used in the internet delivery of broadcast programs. We also prepared a Preliminary Draft New Recommendation on the subjective evaluation method of 22.2 ch sound systems with 8K video.

At the Society of Motion Picture and Television Engineers (SMPTE), we updated a preliminary draft of ST2116, a standard for S-ADM transmission using an existing digital audio interface.

At the European Broadcasting Union (EBU), we contributed to the development of Tech 3388, a Recommendation on a dedicated renderer for the ADM used for object-based audio and also participated in the preparation of Report TR043 by conducting subjective evaluations on the performance of multiple renderers including our proposed method. We also produced Tech 3392, a Recommendation on an audio definition model (ADM) profile for program production.

At ARIB, we compiled 22.2 ch sound production guidelines into Technical Report TR-B44. We also prepared a Preliminary Draft Standard for an IP-based production interface based on SMPTE ST2110.

At the Japan Electronics and Information Technology Association (JEITA) and the International Electrotechnical Commission (IEC), we continued with our works to revise a standard for the general channel assignment for multichannel sound systems (IEC 62574) to add channel labels for various systems including 22.2 ch sound system. Additionally, we produced a standard for transmitting 22.2 ch sound signal stream encoded by MPEG-4 AAC (IEC 61937-11). We also helped the Consumer Technology Association (CTA) revise a standard for transmitting 22.2 ch sound signal stream encoded by MPEG-4 AAC using HDMI (CTA-861.5). At the Audio Engineering Society (AES), we contributed to the standardization of technical guidelines which prescribe that each country's broadcasting rules (i.e., -24LKFS for Japan) should be followed in principle for the target loudness of over-the-top broadcast programs.

[References]
(1) T. Sugimoto and T. Komori: "Tone compensation method for downmixing of 22.2 ch sound," ITE Winter Annual Convention 2017, 12C-5 (2017) (in Japanese)
(2) T. Sugimoto and T. Komori: "Loudness chaser for 22.2 multichannel sound broadcasting," 2018 Autumn Meeting ASJ, 2-1-14 (2018) (in Japanese)
(3) K. Matsui, A. Ito, S. Mori, M. Inoue and A. Adachi, "A method to relief the binaural reproduction controller applying output tracking control," Journal of Acoustical Society of Japan, vol.73, no.5, 2017, p.281-290 (2017)(in Japanese)
(4) A. Ito, K. Matsui, K. Ono, H. Hattori, T. Suenaga, K. Iwauchi, and S. Adachi: "Design and implementation of binaural reproduction controller applying output tracking control," Audio Engineering Society Convention paper 10121 (2018)
(5) ISO/IEC 23008-3:2015/AMD3:2017 (2017)
(6) T. Sugimoto, S. Aoki, T. Hasegawa, and T. Komori: "Real time audio encoding/decoding system based on MPEG-H 3D Audio," 2018 Autumn Meeting ASJ, 2-1-13 (2018) (in Japanese)
(7) Y. Sasaki and K. Ono: "Shotgun microphone with high directivity by extra-long acoustic tube and digital noise reduction," Audio Engineering Society Convention paper 10097 (2018)
(8) S. Kitajima, T. Sugimoto, S. Oode, T. Komori and J. Urano: "Required Bit Rate of MPEG-4 AAC for 22.2 Multichannel Sound Contribution and Distribution," Audio Engineering Society Convention paper 10078 (2018)

2.5 Video coding

We are researching video coding techniques to transmit full-featured 8K SHV and to realize SHV terrestrial broadcasting.

Performance improvement of 8K/120-Hz HEVC encoder

We developed an encoder which compresses an 8K video with a frame frequency of 119.88-Hz (hereafter simplified to 120-Hz) using the high efficiency video coding (HEVC) scheme and exhibited it at the NHK STRL Open House 2018. In FY 2018, we improved its video encoding control method to achieve better image quality and added an interface using TS over IP (SMPTE ST 2022-2) to support higher bit rate than ever before. The addition of the TS over IP interface increased the bit rate to the maximum of 480 Mbps and also enabled the transmission of a high-quality 8K/120-Hz video through diverse channels (Figure 2-12).

For the encoder, we employed a quasi 2-pass encoding technology, which encodes a down-converted 4K/59.94-Hz (hereafter simplified to 60-Hz) video ahead of the corresponding 8K/120-Hz video and controls the 8K encoding using the compression results of the 4K video. This technology uses an encoding control method that adequately allocates the bit amount for an 8K/120-Hz video by detecting areas in a video frame, such as high complexity parts and slice boundaries, in which conspicuous deterioration caused by encoding can be seen. We evaluated the improvement effect of this method using peak signal-to-noise ratio (PSNR). The results demonstrated an improvement in the image quality at a bit rate of around 100 Mbps, including an improvement of about 2.5 dB in a test image[1]. This research was conducted in cooperation with FUJITSU LABORATORIES LTD.

Evaluation of backward-converted 120-Hz video

ARIB Standard STD-B32, which includes the specifications of the video coding scheme for digital broadcasting, employs a stream structure that supports both 120-Hz and 60-Hz. When 120-Hz broadcast is provided, a receiver supporting only 60-Hz decodes and displays video that is frame sub-sampled from 120-Hz video (backward compatibility). We conducted subjective evaluation experiments by non-specialists to verify

Figure 2-12. 8K/120-Hz HEVC encoder

image quality degradation by the stroboscopic effect caused by frame sub-sampling. The results showed no significant degradation in image quality and demonstrated even a possibility that the subjective quality of sub-sampled fast-moving images is better than that of images produced by a 60-Hz system[2].

Development and standardization of 8K file format

We are researching 8K file-based recording technology to enable file-based exchange of 8K content within the broadcast stations. In FY 2018, we studied coding parameters of a file format used for play-out, program exchange and archiving and developed a prototype HEVC decoder for verification (Figure 2-13).

To study the coding parameters, we investigated the PSNR values for the various combinations of the bit rate and the GOP (Group of Pictures) length through coding experiments with three repetitions of sequential encoding and decoding and verified the subjective image quality for each combination.

For the standardization of the file format, we participated in the newly launched ARIB JTG (Joint Task Group) on 4K/8K file format and formulated the requirements.

Development and standardization of next-generation video coding technologies

We are developing high-efficiency video coding technologies for next-generation terrestrial TV broadcasting. As coding tools for intra prediction, we developed a method for improving intra prediction by controlling the coding order, a method for increasing the prediction accuracy by changing the filter applied for prediction signals according to the distance from reference signals when generating prediction samples, and a method for controlling the transform adaptively according to the prediction mode of chroma samples[3][4]. As coding tools for inter prediction, we developed a method for smooth interpolation using the neighboring motion vectors and a method for extrapolation from a certain direction. We also developed a deblocking filter that changes the filter intensity according to the luminance level and a method for changing the transform when generating prediction images using both intra prediction and inter prediction[5]. We proposed these technologies as coding tools for Versatile Video Coding (VVC), a next-generation video coding scheme for which standardization efforts began at the JVET (Joint Video Experts Team) international standardization working group formed between ITU-T and ISO/IEC. The proposed method for improving the deblocking filter was adopted in a working draft JVET-L0414. We also contributed to the development of common test conditions for technical evaluation in standardization efforts and formulated the conditions for HDR video coding[6].

Additionally, we helped the JCT-VC international standardization working group prepare guidelines for the combinations of practical video formats and interfaces, which are industrially required for HEVC codec development[7].

Figure 2-13. Appearance of prototype decoder

Development of coding techniques using machine learning and super-resolution reconstruction and image quality assessment method

We studied the use of machine learning to increase the speed of the intra prediction mode decision for video coding. As an alternative to a conventional method for deciding the prediction mode using rate-distortion optimization with high computational complexity, we investigated a method for building a convolutional neural network using the information of pixels around the coding unit and the intra prediction mode applied to the neighboring blocks as inputs. Also, to reduce the computation load of neural networks, we used multiple types of neural networks with a small number of parameters by switching among them in multiple stages according to the frequency of the intra prediction mode. We confirmed that this method can reduce the computation volume while suppressing the deterioration in coding efficiency. This research was conducted in cooperation with Meiji University.

We developed an inter prediction technology that uses a super-resolution technology and a blurring technology (Figure 2-14). Conventional video coding methods perform inter prediction by comparing the locally decoded images of past and present. This means that the resolution of images captured by swivelling the camera (camera panning) tends to be higher when the image is still but lower when it contains large motion due to the influence of the charge storage effect of the camera sensor. These resolution variations between frames cause a resolution difference between reference signals and the signals to be coded in inter prediction, leading to a decline of prediction efficiency. Focusing on this phenomenon, we applied a registration super-resolution process between wavelet multi-scale components and a blurring process using wavelet decomposition. We demonstrated that coded images with a higher quality can be achieved in coding of moving images containing camera panning and the local motion of objects in the frame by using image signals applied with super-resolution reconstruction and blurring process as prediction reference

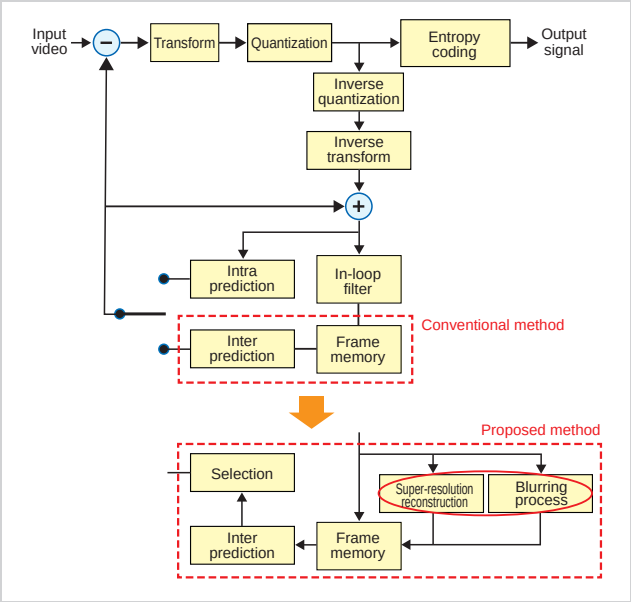


Figure 2-14. Process block diagram of video coding

candidates in addition to conventional prediction reference signals for inter prediction[8].

We studied objective quality metrics suited for HDR image coding using the Hybrid Log-Gamma (HLG) method. Previously, objective quality metrics for HDR image coding were studied for methods using a gamma curve based on the human vision system, such as the perceptual quantizer (PQ) method. Using compressed HLG images, we investigated correlations between the results of subjective evaluation experiments and the values derived from various objective quality metrics. The HLG gamma curve has largely different properties from those of PQ and other methods. The results of our investigation showed that objective quality metrics using the HLG curve were the best and that some objective metrics considered to have a good performance in previous studies are not suited for HLG image coding[9]. This research was conducted in cooperation with Universitat Pompeu Fabra.

Pre-coding processor

We developed a technology for automatically controlling the parameters of a video processor[10] that performs noise reduction and low-pass filtering as a pre-coding process to suppress possible image breakdown when a moving image subject to coding degradation is entered. We implemented the technology into our video processor for performance improvement. Focusing on the fact that the high-frequency-band components after wavelet-packet decomposition of input video are noise and strong edge components, this technology controls the amount of noise reduction and low-pass filtering from that level. We used the equipment that we developed for transmission experiments on advanced terrestrial TV broadcasting technology in Tokyo and Nagoya, and demonstrated that it was effective for suppressing image breakdown and improving comprehensive broadcast quality. This research was conducted as a government-commissioned project from the Ministry of Internal Affairs and Communications titled "R&D on Advanced Technologies for Terrestrial Television Broadcasting."

[References]
(1) K. Chida, X. Lei, S. Iwasaki, Y. Sugito, H. Miyoshi, Y. Uehara, K. Iguchi, K. Kanda: "Development of 8K120Hz Real-time Video Codec," ITE Annual Convention, 22C-1 (2018) (in Japanese)
(2) K. Miura, K. Chida, A. Ichigaya, K. Kanda, Y. Takiguchi and Y. Nishida: "A Study of Backward Compatibility in High Frame Rate Broadcasting," IEICE General Conference, D-11-1 (2019) (in Japanese)
(3) S. Iwamura, S. Nemoto, K. Iguchi, A. Ichigaya, et al.: "Description of SDR and HDR video coding technology proposal by NHK and Sharp," JVET-J0027 (2018)
(4) S. Iwamura, S. Nemoto, A. Ichigaya: "CE6-related: Implicit transform selection for Multi directional LM," JVET-M0480 (2019)
(5) S. Iwamura, S. Nemoto, A. Ichigaya: "CE6-related: Implicit transform selection for Multi-hypothesis inter-intra mode," JVET-M0482 (2019)
(6) A. Segall, E. François, S. Iwamura, D. Rusanovskyy: "JVET common test conditions and evaluation procedures for HDR/WCG video," JVET-L1011 (2018)
(7) Y. Syed, A. Ichigaya, C. Seeger: "Baseband Signalling Specifications Carriage of Usage of Video Signal Coding Types," JCTVC-AH0022 (2019)
(8) Y. Matsuo, A. Ichigaya, K. Kanda: "Super-Resolved and Blurred Decoded Pictures for Improving Coding Efficiency in Inter-Frame Prediction," Proceedings of IEEE ICFSP, pp.120-124 (2018)
(9) Y. Sugito and M. Bertalmío: "Performance Evaluation of Objective Quality Metrics on HLG-Based Image Coding," 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Anaheim, US, pp.96-100 (2018)
(10) Y. Matsuo, K. Iguchi and K. Kanda: "Pre-Processing Equipment of 8K Video Codec with Low-Pass Filtering and Noise Reduction Functions," Proceedings of IEEE ISSPIT, Smart and Sustainable Communities (2018)

2.6 Media transport technologies

We are conducting R&D on content multiplexing technologies on the basis of MPEG Media Transport (MMT) technology, which was adopted for the advanced satellite broadcasting for 4K/8K, with the aim of realizing services that are equally available irrespective of the channel, whether broadcasting or broadband networks, and creating new viewing experiences. Our work included research on a multiplexing transmission method for next-generation terrestrial TV broadcasting, demonstration experiments of IP multicast delivery and a delivery technology for 3D content.

Multiplexing transmission method for next-generation terrestrial TV broadcasting

Aiming for next-generation terrestrial TV broadcasting, we previously researched a multiplexing scheme for IP packets that conforms to the channel coding system for terrestrial broadcasting and an IP transmission system used over studio to transmitter links (STLs) to enable a single-frequency network (SFN) and prototyped a remux. In FY 2018, we conducted verification experiments using the remux and confirmed the effectiveness of the schemes we studied. The experiments were conducted using the Higashiyama and Nabeta stations, which we launched in Nagoya as new experimental transmission stations for next-generation terrestrial broadcasting. We connected a remux (in Higashiyama) and a modulator for next-generation terrestrial broadcasting (in Nabeta) over a commercial IP network and transmitted transmission signals for next-generation terrestrial broadcasting. The results demonstrated that an SFN can be built by using signaling information in the transmission signals for synchronization control. To improve reliability and operability for practical use of the remux, we also developed a network redundancy device that improves the reliability of transmission over STLs by using IP networks of multiple network service providers simultaneously and a network status monitoring device for improving the operability.

Part of this research was conducted as a government-commissioned project from the Ministry of Internal Affairs and Communications titled "R&D on Advanced Technologies for Terrestrial Television Broadcasting."

Demonstration experiment on IP multicast delivery

To promote 4K/8K content delivery, we verified MMT-based IP multicast delivery technology that could be used for the IP retransmission of broadcasting in closed networks of cable TV stations and other service providers and for the IP delivery of the relevant content linked with broadcasting. In the verification experiments, we up-linked multiple channels of live video from a stadium in Fukui Prefecture to an internet exchange (IX) service provider in Tokyo and delivered the content to nine cable TV service providers across the country simultaneously through IX equipment. The results of the experiments showed that packets observed in the reception point had no problem with quantitative transmission quality such as jitter, demonstrating the feasibility of simultaneous multichannel

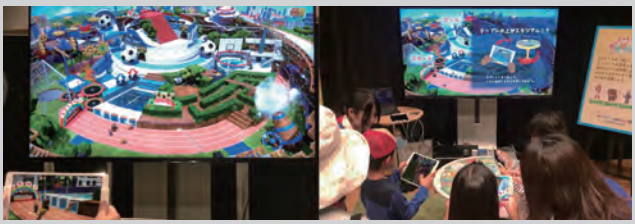


Figure 2-15. Domo's Slapstick Race

delivery of 4K/8K content on a nationwide scale using existing IX connection lines owned by cable TV service providers.

Delivery technology for 3D content

We began research on a delivery technology for 3D content in anticipation of future implementation of Diverse Vision. We developed a system that transmits the motion data of objects in a 3D space in real time over a broadband network using the same time base as that for video and sound of TV broadcasting and allows synchronized viewing of the objects on the screen of a tablet device using Augmented Reality (AR)[(1)(2)]. We exhibited interactive content "Domo's Slapstick Race," in which NHK character Domo and his friends compete in a race at a fantasy stadium, at the NHK STRL Open House 2018. This exhibit can present free-viewpoint images of the race using AR on a tablet device in synchronization with an overview of the race at the stadium displayed on the TV screen. Feeling like a cameraperson, each viewer can enjoy a new viewing experience offered by broadcasting and AR, in which they can take a peek at a place not displayed on the TV screen or follow their favorite character on their tablet device. We exhibited this system at various local events such as "Okhotsk Mirai Festival" held in July in Kitami City, Hokkaido, "Saga Saiko Festival" in October in Saga City, Saga, and "Seto City Digital Festival" in November in Seto City, Aichi.

[References]
(1) Y. Kawamura and K. Otsuki: "3D Free-viewpoint Augmented Reality on a Second Display Synchronized with a Broadcast Program," Proceedings of ICCIP (2018)
(2) Y. Kawamura and K. Otsuki: "Real-time Distribution for Free-Viewpoint Augmented Reality Content Synchronized with a Broadcast Program," ITE Technical Report, vol.42, no.41, BCT2018-93, 2018, p.49-54 (2018) (in Japanese)

2.7 Satellite broadcasting technology

We are researching 12-GHz-band satellite broadcasting system that provides BS digital broadcasting services and 4K/8K UHDTV BS broadcasting services to improve the satellite transmission performance, and researching next-generation satellite broadcasting systems such as 21-GHz-band satellite broadcasting, which is considered as a promising transmission channel for future new broadcasting services.

12-GHz-band satellite broadcasting

The 4K/8K UHDTV satellite broadcasting using ISDB-S3 (Integrated Services Digital Broadcasting for Satellite, 3rd generation), a transmission system researched and developed by NHK STRL, was officially started on December 1, 2018. Prior to the start, we evaluated the transmission performance of 16APSK (Amplitude Phase Shift Keying) by the BSAT-4a broadcasting satellite launched in 2017. The evaluation results demonstrated that adequate transmission performance with a required carrier-to-noise ratio (C/N) of 12.2 dB was obtained.

With the aim of increasing the capacity of 12-GHz-band satellite broadcasting by using multi-level modulation for more effective use of ISDB-S3, we are investigating a way of increasing the output power of the broadcasting satellite transmission. Since the side lobes of an on-board antenna, which cause radio wave interference to other countries, need to be suppressed when the output power is increased, we designed a reflector antenna with a wider aperture area by using a super-elliptical-shaped main reflector. We confirmed that the antenna, which takes an offset Gregorian type with 2.5 m aperture diameter considering the mountability onto the satellite body, reduced the side lobes by about 3 dB. Besides, for simplified reception of satellite broadcasting using a modulation scheme with a low required C/N, we also designed a compact flat receiving antenna for 12-GHz-band satellite broadcasting which does not require elevation adjustment when installed. The antenna uses patch antennas as elements and has an array structure that provides a phase difference in power feeding to each element, which achieved an elevation angle of 39 degrees, a maximum gain of 16 dBi and a half-power angle of 22 degrees for an aperture with sides of about 10 cm.

Next-generation satellite broadcasting

Using a 21-GHz-band transponder mounted on BSAT-4a, we conducted experiments on satellite transmission of 300-MHz-class wide-band signals. We evaluated the C/N-to-BER performance of wide-band signal transmission using the modulation scheme (QPSK, 8PSK) and LDPC (Low Density Parity Check) code rate as parameters for different loopback types of "intermediate frequency (IF)-loopback," which loops back the signals in our prototype modulator and demodulator, "earth-station-loopback," which loops back the signals by in the transmitting earth station, and "satellite-loopback," which goes through BSAT-4a (Figure 2-16). The required C/N of satellite-loopback degraded by 1.0 dB compared to IF-loopback when using the QPSK modulation while it degraded by 1.8 to

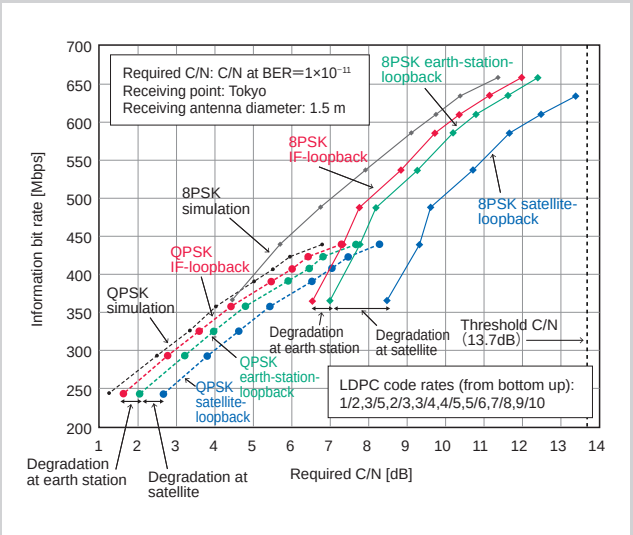


Figure 2-16. Required C/N vs. information bit rate during 21-GHz-band satellite transmission (performance measured with a prototype modulator and demodulator)

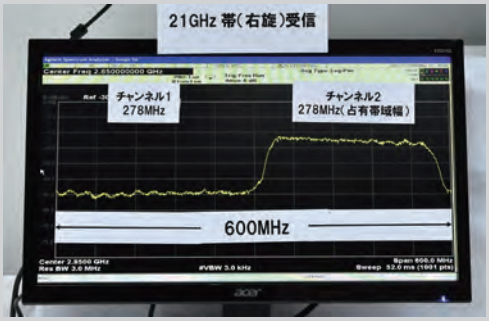


Figure 2-17. Received signal spectrum in 21-GHz-band satellite transmission experiment

2.2 dB when using 8PSK. These experiments clarified the relationship between the information bit rate that can be transmitted by BSAT-4a and the required C/N in the 21GHz band. We also conducted the world's first experiments on 4K and 8K multiplexing transmission using a 21-GHz-band broadcasting satellite (Figure 2-17)[1].

To evaluate rain attenuation characteristics in the 21 GHz band, in April 2018, we started measuring the receiving power of a left-hand circular polarization (LHCP) beacon signals constantly transmitted from BSAT-4a by using a 21-GHz-band receiving parabolic antenna with 1.5 m aperture diameter set up on the rooftop of our laboratory. We also installed a 21-GHz-band rain attenuation monitor station in NHK Fukuoka Broadcast station to evaluate rain attenuation characteristics in the 21 GHz band by region.

We prototyped interface devices for inputting and outputting IP signals to/from the wideband modulator and demodulator for 21-GHz band transmission using BSAT-4a. These devices enabled the transmission of MMT signals and other IP signals in the 21-GHz band broadcasting satellite.

We are studying a way to increase the output satellite transmission power by spatial synthesis using a 21-GHz-band array-fed reflector antenna. We designed a two-shaped-reflector antenna using a sequential array structure in which six-element horn antennas are arranged with rotational symmetry and confirmed that it can obtain a gain of about 37 dBi in major areas across the country when the cross-polarization discrimination is 27 dB or more.

With the aim of further increasing the capacity of satellite transmission, we prototyped a transmitter/receiver equipped with 64APSK coded modulation based on set partitioning and evaluated its performance in an environment with AWGN (Additive White Gaussian Noise). The prototype device was capable of up to 158.6-Mbps transmissions in a satellite transponder with a bandwidth of 34.5-MHz. Comparison with computer simulations showed that the required C/N degradation attributed to hardware was about 1.0 dB, and it was confirmed that the device had adequate transmission performance.

[References]
(1) Y. Suzuki et.al: "21-GHz Wide-band satellite transmission experiment with BSAT-4a broadcasting satellite," IEICE Technical report, Vol.118, No.136, SAT2018-14, p.19-23 (2018) (in Japanese)

2.8 Terrestrial broadcasting transmission technology

For the terrestrial broadcasting of 4K/8K, we made progress in our research on an advanced terrestrial broadcasting system by evaluating its characteristics, making improvements and conducting large-scale field experiments. We also conducted R&D on time-division multiplexing (TDM) and other systems that are different from the advanced system using frequency-division multiplexing (FDM), which we have been studying, and worked with international organizations in efforts such as investigation on the use of the fifth-generation mobile communications system (5G) for broadcasting. Part of this research was being performed under the auspices of the Ministry of Internal Affairs and Communications, Japan as part of its program titled "Research and Development for Advanced Digital Terrestrial Television Broadcasting System," in cooperation with Sony Corporation, Panasonic Corporation, Tokyo University of Science and NHK Integrated Technology Inc.

Evaluation of characteristics of advanced terrestrial broadcasting system and technology for their improvement

In FY 2018, we evaluated the characteristics of the advanced terrestrial broadcasting system through laboratory experiments using a modulator and demodulator that we prototyped in FY 2017. We demonstrated that the advanced system, which has more robustness by using Low-Density Parity-Check (LDPC) codes and a non-uniform constellation and a larger transmission capacity by having a new signal structure, can increase the transmission capacity by about 10 Mbps compared with ISDB-T, the current standard for digital terrestrial broadcasting, when the required C/N is the same while it can reduce the required C/N by about 7 dB when the transmission capacity is the same[1] (Figure 2-18). Moreover, the results of laboratory experiments simulating an actual SFN environment showed that the advanced system can reduce the degradation in the required C/N caused by an SFN by about 2.7 dB from that of ISDB-T[2].

We also worked to improve the advanced terrestrial broadcasting system that we previously formulated. In FY 2018, we devised a method for increasing the transmission

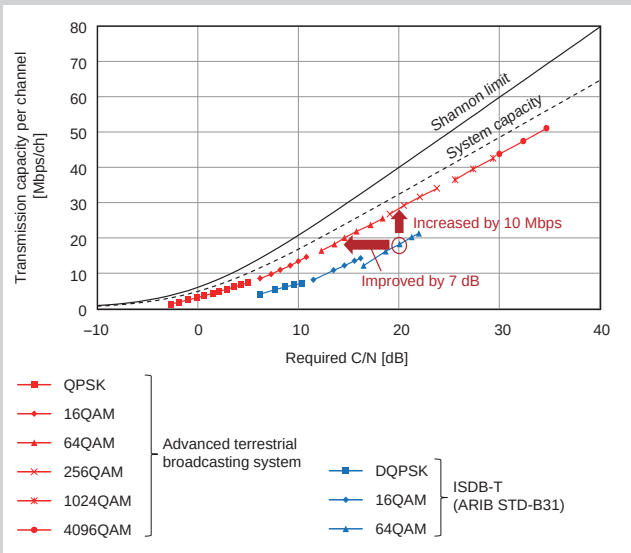


Figure 2-18. Relationship between required C/N and transmission capacity

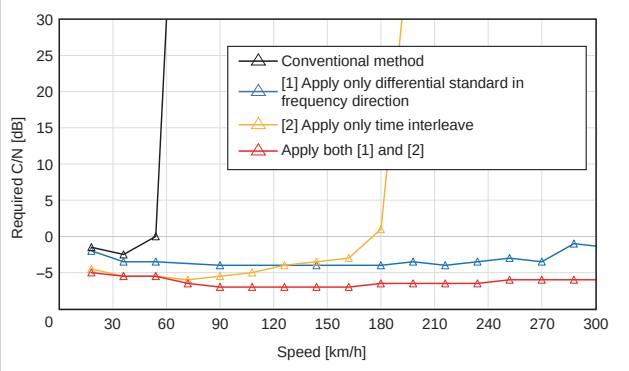


Figure 2-19. Comparison of speed tolerance of TMCC signal transmission

robustness of transmission and multiplexing configuration and control (TMCC) signals, which are transmission control signals, evaluated the method through computer simulations and implemented it into hardware. The conventional TMCC signal transmission used differential BPSK (Binary Phase Shift Keying) as the carrier modulation scheme and used signals adjoining in the time direction as differential reference signals. Our proposed method reduced the performance degradation caused by the time variations of the channel by using signals adjoining in the frequency direction as differential reference signals. In addition, we applied a cyclic shift to TMCC signals in the time direction only for the amount of shift determined based on the segment number to obtain the time interleave effect. We confirmed that the use of these two elemental technologies improved speed tolerance in a mobile reception environment[3] (Figure 2-19).

We also evaluated the effect of cyclic-shift frequency interleave, which we proposed as a technology for improving the performance of data signal transmission, through laboratory experiments. The conventional method needed to use two different types of frequency interleave depending on the scattered pilot (SP) signal configuration between layers, but the evaluation results demonstrated that our proposed method can replace the two types of frequency interleave with one type while ensuring the same level of reception characteristics.

Other systems than advanced terrestrial broadcasting system

We conducted research on hierarchical transmissions using TDM and layered division multiplexing (LDM), which are used for broadcasting systems in Europe and the U.S. For TDM, we studied a synchronization algorithm and a transmission method for preamble signals, which are not used by FDM. We also prototyped a TDM modulator and demodulator that has the same sampling frequency and signal bandwidth as those of an FDM device for performance comparison between TDM and FDM. For LDM, we conducted laboratory experiments using a modulator and demodulator that we developed in FY 2017 to evaluate the mobile reception characteristics of the use of LDM for the partial reception layer of FDM. The results demonstrated

Table 2-1. Transmission specifications of experimental stations

	Tokyo	Nagoya	
Station name	Shiba	Higashiyama (main station)	Nabeta (relay station)
Location	Minato Ward, Tokyo	Nagoya City, Aichi	Yatomi City, Aichi
Transmission channel	UHF ch28	UHF ch35	
Polarization	Horizontal and vertical (Dual-polarized MIMO)		
Transmission power	1kW per polarization	1kW per polarization	10W per polarization
Signal bandwidth	5.83MHz		

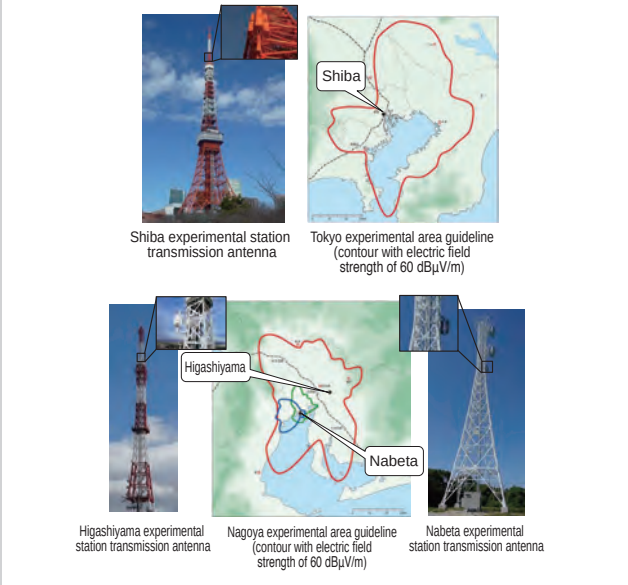


Figure 2-20. Transmission antennas of experimental stations and experimental area guidelines

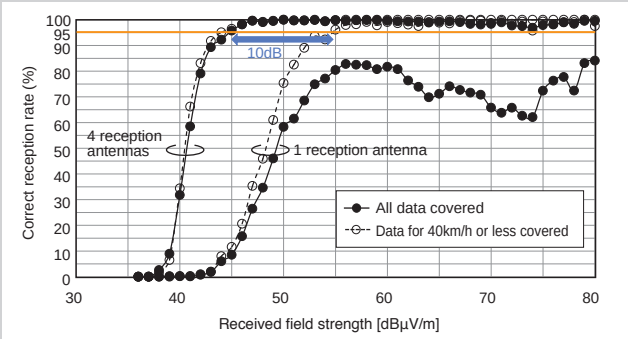


Figure 2-21. Effect of diversity reception

its effectiveness[4]. We also developed prototype modulator and demodulator that implement a gray mapped LDM which can improve characteristics when simultaneously demodulating multiple layers multiplexed by LDM.

We researched a wide-band terrestrial broadcasting system, which can improve interference tolerance by expanding the bandwidth and reducing the modulation level and the code rate of error correction code when we focus on ensuring a certain level of transmission capacity. We evaluated the tolerance to interference through computer simulations assuming a case that different programs are broadcast in neighboring broadcast areas. The results demonstrated that successful reception was achieved even if undesired waves with the same power as desired waves were received.

Preparation of experimental stations and field experiments for evaluating advanced terrestrial broadcasting system

As part of the government-commissioned project from the Ministry of Internal Affairs and Communications titled "R&D on Advanced Technologies for Terrestrial Television Broadcasting," in FY 2018, we constructed large-scale experimental environments in Tokyo and Nagoya (Table 2-1, Figure 2-20). We increased the transmission power in stages up to a specified level over a period of about one month while ensuring that radio waves from the experimental stations give no impact on existing wireless stations. Licenses were issued for the Higashiyama and Nabeta experimental stations in

November 2018 and for the Shiba experimental station in December 2018. Following the issuance of the licenses, we began field experiments using a modulator and demodulator for the advanced system that we developed in FY 2017.

In Tokyo, advanced system signals transmitted from the Shiba experimental station were received by a reception antenna 10 m high. At the measurement point, we collected channel responses as well as electric field strength and bit error rate characteristics using our signal analysis equipment[5]. We conducted measurements at 41 points and analyzed data such as the degradation of required received power due to multipath propagation and the received power ratio of horizontal polarization and vertical polarization. In Nagoya, we conducted SFN field experiments using the main station (Higashiyama) and the relay station (Nabeta). We measured the degradation of transmission characteristics, that is caused by SFN waves, by varying the desire and undesire received power ratio (D/U) and the delay time between the main station and relay station. At various measurement points, where the antenna direction between the main station and relay station are different, the results were compared with that of laboratory experiments. In addition, we conducted mobile reception experiments in Nagoya. We investigated electric field strength and evaluated reception performance by driving a measurement vehicle over a distance of about 350 km in the broadcast area of the Higashiyama experimental station[6]. In our evaluation of diversity reception with a larger number of reception antennas, we found that the use of four reception antennas can reduce the impact of the degradation caused at a fast driving speed. We also confirmed that it can obtain diversity gains of 10 dB or more compared with the use of a single reception antenna (Figure 2-21).

■ International collaboration

ITU-R WP6A (Terrestrial broadcasting delivery) is preparing a report on network planning and transmission methods for advanced digital terrestrial TV broadcasting. In FY 2018, we proposed adding information about dual-polarized multiple-input multiple-output (MIMO) for expanding transmission capacity, LDPC codes for enhancing transmission robustness and a non-uniform constellation.

Meetings of the Future of Broadcast Television (FOBTV), where broadcasters and standardization organizations around the world gather, were held at the venues of the NAB Show 2018 and IBC 2018. We attended the meetings to share the latest trends of next-generation broadcasting. We also investigated broadcast-broadband experiments over terrestrial broadcasting networks which the Electronics and Telecommunications Research Institute (ETRI) of South Korea conducted on Jeju Island in South Korea.

As part of activities at the Digital Broadcasting Experts Group (DiBEG) of ARIB, we exchanged opinions about next-generation terrestrial broadcasting with SBTVD-Forum, a standardization organization in Brazil.

As part of study toward next-generation terrestrial broadcasting, we are investigating the use of 5G for broadcasting in cooperation with EBU. In FY 2018, we attended the meetings of the 3rd Generation Partnership Project (3GPP), which is working on the standardization of 5G, to study the trends.

[References]
(1) Y. Yamakami, et al.: "Performance Evaluation of Modulator and Demodulator for the Advanced Terrestrial Broadcasting (1) - AWGN performance -," ITE Annual Convention 2018, 33D-1 (2018) (in Japanese)
(2) T. Ijiguchi, et al.: "Performance Evaluation of Modulator and Demodulator for the Proposed Specification of the Next Generation Terrestrial Broadcasting (2) - in SFN environment -," ITE Annual Convention 2018, 33D-2 (2018) (in Japanese)
(3) H. Miyasaka, et al.: "A Study on the Robust Transmission of TMCC for the Proposed Specification of the Next Generation Terrestrial Broadcasting," ITE Tech. Rep., Vol.42, No.36, BCT2018-84, pp.69-72 (2018) (in Japanese)
(4) A. Satou, et al.: "Laboratory experiments of hierarchal transmission system combinated FDM with LDM," ITE Annual Convention 2018, 21C-2 (2019) (in Japanese)
(5) T. Ijiguchi, et al.: "Development of Signal Analysis Equipment for Advanced Digital Terrestrial TV Broadcasting System," ITE Tech. Rep., BCT2019-36 (2019) (in Japanese)
(6) H. Miyasaka, et al.: "A Study on the Mobile reception for Advanced Digital Terrestrial TV Broadcasting System," IEICE Annual Convention 2019, B-5-117 (2019) (in Japanese)

2.9 Wireless transmission technology for program contributions

We are conducting R&D on a 1.2-GHz/2.3-GHz-band field pick-up unit (FPU) for transmitting video and sound materials and a millimeter-wave-band wireless camera with the goal of using them for program production in 4K/8K live broadcasting of emergency reports and sports coverage.

■ 1.2-GHz/2.3-GHz-band 4K/8K mobile relay FPU

To enable the mobile relay broadcasting of 4K/8K signals by using the 1.2-GHz/2.3-GHz-band, we conducted R&D on a multiple-input multiple-output (MIMO) system with adaptive transmission control using the time division duplex (TDD) scheme and contributed to the standardization[1][2].

In FY 2018, we improved the functions of the MIMO system with adaptive transmission control to enhance transmission performance. To overcome a problem of transmission performance degradation caused by long-delay multipath reflected waves in urban areas with many obstacles, we designed a transmission parameter to expand the guard interval length of orthogonal frequency-division multiplexing (OFDM) signals from the conventional value of 6 μs to 18 μs. Also for a further increase in the transmission capacity, we prototyped experimental equipment that increases the amount of information that can be transmitted per OFDM carrier symbol from the conventional value of 16 bits to 20 bits (Figure 2-22). We achieved a maximum transmission capacity of about

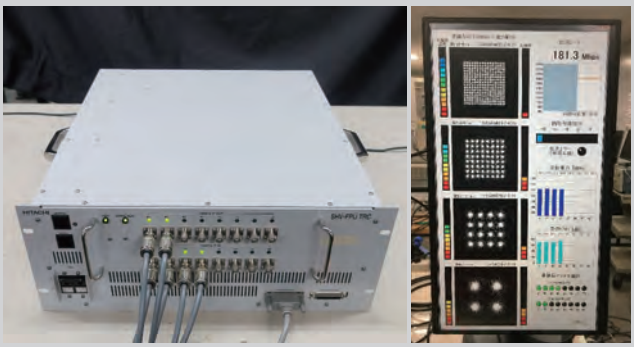


Figure 2-22. Prototype control unit of 4K/8K mobile relay FPU

Figure 2-23. Prototype compact transmitter for millimeter-wave-band 4K/8K wireless camera

180 Mbps by using a rate-matching technology[3], which controls the coding rate of error correction codes adaptively to prevent transmission errors even in a channel environment with instantaneous variations such as during mobile transmission.

We also contributed to the establishment of a national standard on advanced 1.2-GHz/2.3-GHz-band FPUs at the Information and Communications Council of the Ministry of Internal Affairs and Communications and the preparation of a standard at the Association of Radio Industries and Businesses (ARIB) so that these FPUs can be put into practical use in 2020.

Millimeter-wave-band 4K/8K wireless camera

We made progress in our research on wireless cameras that can transmit video for 4K/8K program production by using 42-GHz-band radio waves.

As a transmission system, we are studying a single-carrier-frequency domain equalization (SC-FDE) scheme, which is generally robust to the distortion of a power amplifier and has high power efficiency. In FY 2018, we developed a compact transmitter using the SC-FDE scheme (Figure 2-23)[4] and conducted wireless transmission experiments. The prototype compact transmitter, which supports a transmission frequency of 42 GHz, a channel bandwidth of 125 MHz and a transmission power of 200 mW, achieved a transmission capacity of 185 Mbps when using the 32APSK (Amplitude Phase Shift Keying) modulation scheme and an error correction coding rate of 1/2 using convolutional codes. We also studied ways to improve the transmission performance of the SC-FDE scheme. We identified an optimum power ratio of pilot signals used for channel estimation during demodulation and data signals[5], and we also improved the performance of frequency domain equalization[6].

Radio over Ethernet System

We are researching a Radio over Ethernet system to transmit video and audio signals transmitted by radio waves over Ethernet. In FY 2018, we developed Ethernet intermediate-frequency (IF) transmission equipment that would allow the easy connection of reception base stations for millimeter-wave-band 4K/8K wireless cameras (Figure 2-24).

An effective way to ensure stable reception from a moving wireless camera is to perform demodulation using signals received by multiple reception base stations. Conventional analog optical transmission equipment required the installation of optical fibers to each reception base station and the level adjustment of signals. Meanwhile, the Radio over Ethernet system can digitize received signals and transmit them over a local area network (LAN), which enables the easy installation just by connecting a cable to the hub without the need for signal level adjustment.

Figure 2-24. Ethernet IF transmission equipment

To digitize and transmit radio signals, it is necessary to accurately synchronize the sampling clock between the transmitting device and the receiving device. This equipment can modify errors of the sampling clock with high precision in a short time by using the departure and arrival timings of each packet that transmits digitized radio signals. This way, it can start transmission immediately after the transmitting side and the receiving side are connected with a cable.

To verify the effectiveness of the Ethernet IF transmission equipment, we conducted an experiment in which we constructed a wireless camera system using three reception base stations in a circuit-type underground parking lot with a circumference of about 100 m. The results demonstrated that camera images were successfully transmitted without interruption even when the wireless camera circled the underground parking lot.

[References]

(1) F. Ito, F. Uzawa, K. Mitsuyama and N. Iai: "Field Trial of 4×4 TDD-SVD-MIMO System with Rank Adaptation," IEEE Global Communications Conference 2018 (IEEE GLOBECOM 2018), Technical Session: WCS-I5-Wireless communication-5 (2018)

(2) F. Ito, F. Uzawa, K. Mitsuyama and N. Iai: "Mobile Transmission Trial of 8K Video using 4x4 TDD-SVD-MIMO System with Adaptive Transmission Control," ITE Technical Report, Vol.42, No.36, BCT2018-85, pp.73-76 (2018) (in Japanese)

(3) F. Uzawa, F. Ito, K. Mitsuyama and N. Iai: "Implementation of Rate-matching Function for Next Generation Mobile Relay FPU," IEICE Multiple Innovative Kenkyu-kai Association for wireless communications 2018 (MIKA 2018), P2-2 (2018) (in Japanese)

(4) Y. Matsusaki, F. Yamagishi, A. Yamasato, S. Okabe and N. Iai: "Development of Compact Transmitter for Millimeter-wave SHV Wireless Camera," ITE Technical Report, Vol. 43, No. 10, BCT2019-42, pp. 25-28 (2019) (in Japanese)

(5) F. Yamagishi, Y. Matsusaki, A. Yamasato, S. Okabe and N. Iai: "A Study on Boost Ratio for SC-FDE Millimeter-wave SHV Wireless Camera under Mobile Transmission Circumstances," ITE Technical Report, Vol. 43, No. 10, BCT2019-43, pp.29-32 (2019) (in Japanese)

(6) Y. Matsusaki, F. Yamagishi, A. Yamasato, S. Okabe and N. Iai: "A Study on Improving Transmission Performance for SC-FDE," Proceedings of the 2019 IEICE General Conference, B-5-68 (2019) (in Japanese)

(7) K. Aoki, K. Murase and K. Imamura: "A Study on Radio over Ethernet System for Mobile Wireless Camera," ITE Technical Report, Vol.42, No.23, pp.13-16 (2018) (in Japanese)

2.10 Wired transmission technology

We are studying the FTTH (Fiber to the Home) digital baseband transmission system for transmitting 4K/8K programs to homes over cable TV networks. This system divides multichannel streams of 2K/4K/8K broadcasting into IP packets, multiplexes them using time-division multiplexing and transmits them by optical signals. With its capability of large-capacity transmission of 10 Gbps, it can transmit about 100 channels of 4K/8K programs.

In FY 2018, we studied an intra-building transmission system for condominium buildings installed only with coaxial cables, without optical fibers[1]. Coaxial cables in condominium buildings cannot transmit baseband signals directly because there is a limit to the transmission bandwidth. We therefore devised a method that selects IP packets in response to a viewer request for a program and converts them to radio frequency (RF) signals. By transmitting signals only for the programs being viewed in the building, this method can transmit multiple channels of 4K/8K programs to individual households over existing coaxial cables. The Data Over Cable Service Interface Specifications (DOCSIS) standard used for internet communications on cable TV can be used as a way of transmitting IP packets in the form of RF signals. With existing DOCSIS, however, it is difficult to control in detail which RF signal transmits which program, preventing the effective use of transmission bandwidth. To address this problem and enable efficient delivery, we added an additional IP address for DOCSIS to an IP packet for a program so that transmission by a DOCSIS transmitter can be controlled by changing the IP address according to the viewer request for a program. We call this technology "IP encapsulation". We also studied measures against packet loss because the bit-error rate specified by the DOCSIS standard is too high to maintain video quality in intra-building transmission of programs. We prototyped intra-building transmission equipment having these functions and demonstrated the effectiveness of our new method.

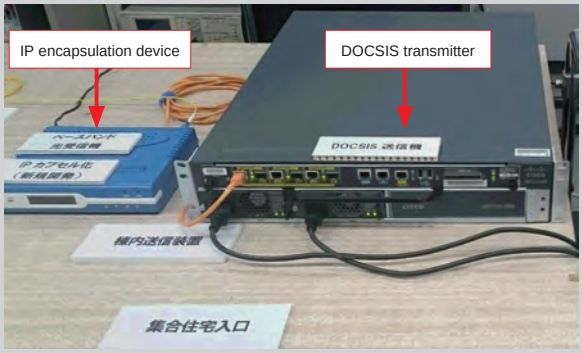


Figure 2-25. Intra-building transmission equipment using DOCSIS

[References]

(1) T. Kusunoki, T. Kurakake, K. Otsuki and K. Saito: "Improvement of transmission efficiency of 4K/8K multi-channel IP multicast using DOCSIS over in-building coaxial cable network," IEEE International Conference on Consumer Electronics 2019 (ICCE 2019), (2019)

3 Connected Media

3.1 Content provision platform

We are researching technologies for making use of the internet to provide more user-friendly and convenient broadcasting services. In our research on a media-unifying platform, in FY 2018, we developed a system that can link multiple presentation devices and a system that can present not only the main part of a program but also program-related information such as digest video and program website, and demonstrated their effectiveness. We also studied a technology for increasing responsiveness to achieve diverse viewing styles and comfortable viewing in video delivery. Moreover, we contributed to the domestic and international standardization of technical specifications for the advanced satellite broadcasting for 4K/8K launched in December 2018.

Media-unifying platform

Against the backdrop of the diversification of program-viewing methods including video delivery on the internet, we are researching a "media-unifying platform" that automatically selects appropriate media for viewing programs in accordance with the user situation. In FY 2018, we developed a system to realize the efficient selection of a program to be viewed according to the practical delivery environment and to meet viewers' diverse intentions for content viewing.

The media-unifying platform consists of servers that manage the content distribution status of each medium and media-unifying engines that automatically select appropriate media on the terminal side. For the content distribution status management server, we demonstrated that it is possible to select from programs of multiple operators and allow more flexible program selection by appropriately supporting content with different provision ranges such as home-use recorders and internet redistribution. We exhibited implementation examples at the NHK STRL Open House 2018 (Figure 3-1).

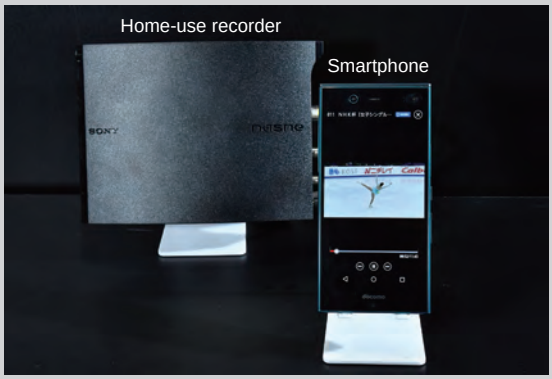


Figure 3-1. Example of the linkage between home-use recorder and smartphone

On the basis of the functions of the media-unifying platform, in FY 2018, we developed a system that realizes a new viewing style by linking multiple devices that can be used for program viewing, such as program viewing enabled by a linkage between a smartphone, TV and smart speaker[1]. We also extended the functions of the content distribution status management server, which previously allowed the selection of only the main parts of programs, to enable the selection of program-related information including digest video of programs, text content such as program websites and related SNS information. This achieved the provision of flexible content access according to the viewer situation including available viewing hours and degree of interest[2] (Figure 3-2).

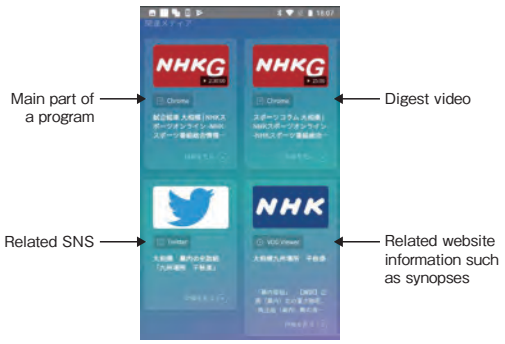


Figure 3-2. Example of the selection screen containing program-related information

We prototyped a system that links multiple viewing devices for program viewing and a system that allows the viewer to select program-related information in addition to the main part of a program and conducted evaluation experiments. Prior to the experiments, we asked 400 participants to respond to an online questionnaire for deciding evaluation items and collected data on the genres and related content of video they frequently view and their viewing behaviors. On the basis of the results, we decided the program genres to be used for evaluation and the scenarios of content viewing. The results of evaluation experiments participated by 50 general users showed that presenting program-related information increases viewers' appetite for and interest in program viewing[3] and that presentation using device linkage can offer a viewing style matching viewer preferences[4], demonstrating the effectiveness of the prototype system.

To promote the development and practical use of the media-unifying platform, we prepared a specification document describing cases of use, input/output specifications between functional elements and other information.

Adaptive-switching shaping technology

As a way of achieving both the stabilization of viewing quality in internet video delivery and high-speed response to viewing operations such as seeks, we developed a technology for adaptive switching of shaping (a technique that delays communication packets exceeding the speed limit to smooth the communication speed).

The technology that we developed performs bandwidth control using shaping for viewing terminal communications on a traffic control device to prevent packet loss, thus stabilizing the viewing quality. It also cancels shaping temporarily for communications of a viewing terminal on which a user performs a viewing operation to deliver video data packets preferentially to the terminal. This reduces the delay before the start of reproduction (response time) and increases the response speed.

We implemented this system into our traffic control device and video player and conducted delivery experiments. In the experiments, ten viewing terminals simultaneously accessed the distribution server connected with the traffic control device to view video and a viewer performed a seek operation during viewing on one of the terminals. We measured the viewing quality during simultaneous viewing by the multiple terminals and the response time to a seek operation for each of the

developed method (adaptive switching of shaping) and the conventional method (constant application of shaping). The results of the experiments using shaping parameters set according to conditional expressions based on video bit rate and the number of viewing terminals showed that the developed method reduced the response time to about 1/10 to 1/9 that of the conventional method without deteriorating the viewing quality[5].

Smooth switching of program sound according to user operation

We developed a video delivery technology that allows users to switch program sound smoothly according to their preferences for new user experiences while viewing video.

The technology produces a video stream and multiple types of audio streams with different combinations of audio sources or different level balances in MPEG-DASH (Dynamic Adaptive Streaming over HTTP) format on the distribution side, and efficiently delivers the streams only selected by user operation on the reception terminal. We prototyped a distribution system and video player equipped with this technology and verified their operation on a Hybridcast-enabled receiver (Figure 3-3). The results of experiments on the switching of program sound by remote controller operation using an audio segment length (a unit of stream distribution) of 0.5 seconds demonstrated that the switching took about 1.2 seconds. We also confirmed that the stable and smooth switching can be achieved without any sound interruption or noise[6].

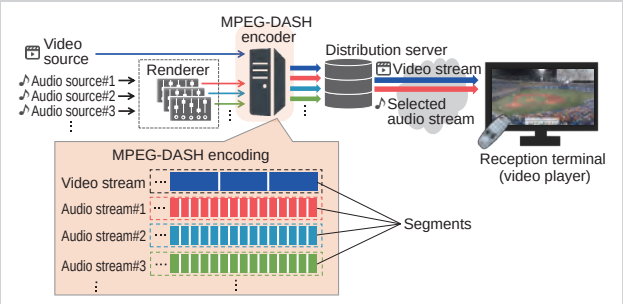


Figure 3-3. Structure of prototype program sound switching system

Server-side rendering technology

We made progress in our development of a server-side rendering technology that enables the viewing of program-linked content that involves advanced graphics processing such as 3D CGs and 360-degree images on terminals with a limited CPU and memory such as TVs. This technology enables interactive viewing by performing image processing in accordance with the user operation on a rendering server on the cloud and transmitting the rendered images to the terminal at low latency using web real-time communication (WebRTC) technology.

In FY 2018, we studied a function to automatically adjust the number of rendering servers to operate in accordance with the number of connected terminals with the aim of reducing the cost of content delivery servers using this technology. As a basic verification, we implemented an experimental system using this technology on a commercial cloud and conducted delivery experiments in which multiple terminals connected to a single rendering server simultaneously to view content including video. We equipped the rendering server with a graphics processing unit (GPU) and measured the usage rates of the CPU and the graphics processing part and decoding part in the GPU in accordance with the number of terminals. The results showed a tendency of the usage rates of the CPU and decoding part being proportional to the number of connected terminals, demonstrating that they are effective for estimating the upper limit of the number of terminals connectable to a single rendering server[7].

Promotion of video distribution for TV

We continued with our work to promote video distribution for TV by using Hybridcast and other means. We previously developed a video player that operates stably on a TV receiver and made it available to the members of the MPEG-DASH Interoperability Study Group at the IPTV Forum under the name of "dashNX." The player has been utilized in 4K delivery experiments and other efforts involving many broadcasters. We also contributed to the establishment of specifications for 2K-enabled "Hybridcast video" and 4K-enabled "Hybridcast 4K video," which are the video distribution services for TV specified at the Forum, and the development of verification test for determining the conformity of receivers with the specifications. This effort led to the announcement of logos for indicating receivers and programs compliant with these specifications in May 2018[8].

In FY 2018, we developed a low-cost MPEG-DASH-compatible 4K live encoder which uses a High Efficiency Video Coding (HEVC) codec board and is based on a general-purpose PC to promote video distribution service using 4K live delivery. This encoder is capable of flexibly adding and changing functions such as adaptive bit rate and digital rights management (DRM) in accordance with PC specifications and service requirements for video distribution.

Using this encoder, we prototyped a sports viewing application for 4K live delivery equipped with device linkage functionality (Figure 3-4). We also conducted 4K live delivery to a Hybridcast-compatible receiver by using this application at the NHK STRL Open House 2018 and demonstrated the feasibility of stable viewing with adaptive bit rate and smooth fast-reversing to a certain scene by operation from a remote controller and a smartphone linked with the receiver[9]. The development of this prototype and the 4K live delivery were conducted in cooperation with Asahi Television Broadcasting Corporation.

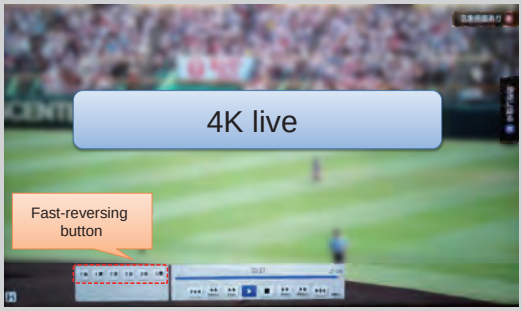


Figure 3-4. Viewing screen of sports viewing application (Screenshot provided by Asahi Television Broadcasting Corporation)

Standardization activities for 4K/8K multimedia broadcasting and hybrid systems

To expand and enhance the advanced satellite broadcasting for 4K/8K launched on December 1, 2018, we contributed to the revision of ARIB standards regarding device linkage functionality in data service and compression formats for the transmission of closed captions and superimposed characters[10].

In our international standardization effort, we prepared the additional description of device linkage functionality for ITU-R Recommendation BT.2075, which describes Integrated broadcast-broadband systems. We also contributed to the revision of ITU-R Report BT.2342, which describes closed-captioning schemes, to add descriptions on the use of ideographic variants and closed-caption file exchange formats. At the technical committee of the Asia-Pacific Broadcasting Union (ABU), we reported on domestic and international trends

of the multimedia schemes, CAS technology and security in broadcasting to promote Japan's broadcasting systems and Hybridcast.

[References]
(1) H. Endo, S. Taguchi, S. Takeuchi, K. Fujisawa and K. Kai: "Multi-Device Media-Unifying Viewing System," ITE Annual Convention, 14D-3 (2018) (in Japanese)
(2) S. Taguchi, H. Endo, S. Takeuchi, K. Fujisawa and K. Kai: "System Model of Media-Unifying Platform for Various Types of Content," ITE Annual Convention, 14D-4 (2018) (in Japanese)
(3) H. Endo, S. Taguchi, S. Takeuchi and K. Fujisawa: "Evaluation of Multi-Device Cooperation System for Content Viewing Independent of Distribution Media," IPSJ SIG Technical Report, Vol.2019-HCI-182, No.24 (2019) (in Japanese)
(4) S. Taguchi, H. Endo, S. Takeuchi and K. Fujisawa: "Evaluation of the System Providing Alternative or Related Content of TV Program," IEICE General Conference, D-21-7 (2019) (in Japanese)
(5) M. Kurozumi, S. Nishimura and M. Yamamoto: "A Study of Response Improvement in Video Delivery by Traffic Shaping Policy According to User Operation," ITE Annual Convention, 14D-2 (2018) (in Japanese)
(6) S. Mori, H. Kubo, S. Oode, S. Takeuchi and S. Nishimura: "A Study on a Smooth Switch Method for User-Selected Sound Sources with Video Streaming for TV Receivers," ITE Winter Annual Convention,13C-2 (2018) (in Japanese)
(7) M. Ohnishi, S. Nishimura and K. Fujisawa: "Evaluation of content delivery system for HTML5 compliant TV receiver using cloud rendering," ITE Winter Annual Convention, 13C-1 (2018) (in Japanese)
(8) S. Nishimura: "R&D Trends in Video Streaming for TV," NHK STRL R&D, No.171, pp.20-31 (2018) (in Japanese)
(9) T. Nakai, K. Hiramatsu, H. Terada, Y. Yamano, and S. Nishimura: "Development of a Prototype Application Using 4K Live Streaming," ITE Annual Convention, 32D-1 (2018) (in Japanese)
(10) S. Takeuchi and M. Ohnishi: "A Study on Compression Format of Closed-caption Data in UHDTV Satellite Broadcasting Service," ITE Winter Annual Convention, 13C-3 (2018) (in Japanese)

3.2 Device linkage services

We made progress in our research on device linkage services that will offer new TV experiences in various scenes of daily life by taking advantage of the internet and linking TV with various terminals such as IoT (Internet of Things)-enabled devices.

Hybridcast Connect X

To connect daily activities with broadcasting services more easily, we are developing Hybridcast Connect X, which enables viewers to start the interaction with Hybridcast applications on a TV from a smartphone or IoT-enabled devices that they use every day.

We developed a device linkage protocol to connect TV with IoT-enabled devices (Figure 3-5)[1], which was then released by the IPTV Forum Japan as a technical standard. As concrete use cases of the device linkage protocol, we demonstrated this protocol implemented into news, SNS and other smartphone applications (apps) and applications for interactive smart speakers at the NHK STRL Open House 2018 and the International Broadcasting Convention (IBC) 2018. The demonstration showed that the implementation of this protocol makes it possible to start the linkage with broadcasting service from smartphone applications and notification functions in daily use and thus enables services that offer new user experiences[2].

In addition, we developed a software module called "Hybridcast Connect Library"[3] that allows easy implementation of this protocol into smartphone applications and IoT-enabled devices and a test tool to verify the protocol operation. We exhibited them at various exhibitions such as the NHK STRL Open House 2018 and Inter BEE 2018 and also helped the IPTV Forum Japan to hold workshops to promote this technology.

Linkage of TV viewing with daily activities

With the aim of realizing more convenient broadcast-broadband services that connect TV viewing with internet services and daily activities outside of TV viewing, we are conducting R&D on a content-matching technology, which is a key technology for linking broadcast content with the data of various applications and IoT-enabled devices.

At the NHK STRL Open House 2018, we exhibited multiple use cases of Hybridcast Connect Library such as linkage between a smartphone application and broadcasting in cooperation with commercial broadcasters and manufacturers and demonstrated the feasibility of service enhancement and the improvement of convenience by data linkage with applications and IoT-enabled devices (Figure 3-6).

To enable viewing data to be utilized for various services under the management of individual viewers, we investigated the trends in personal data management methods and examined those methods. As an example of service linkage among different industries, we prototyped a system that allows the information of programs watched and places visited during a drive to be shared among various services such as

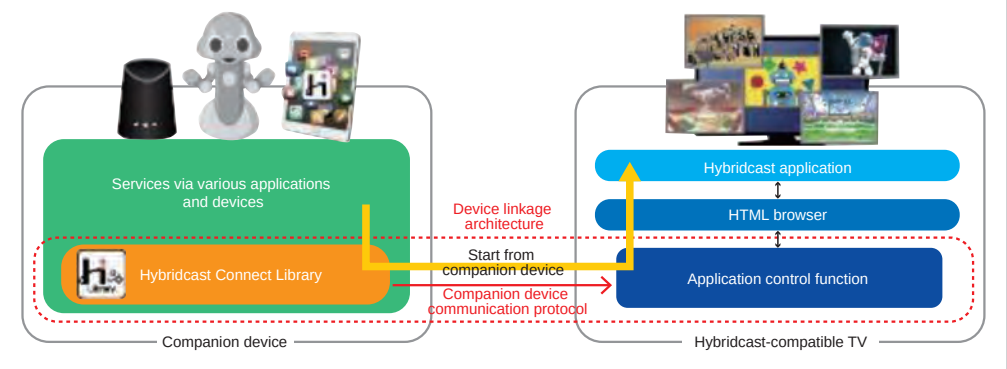


Figure 3-5. Device linkage protocol and Hybridcast Connect Library

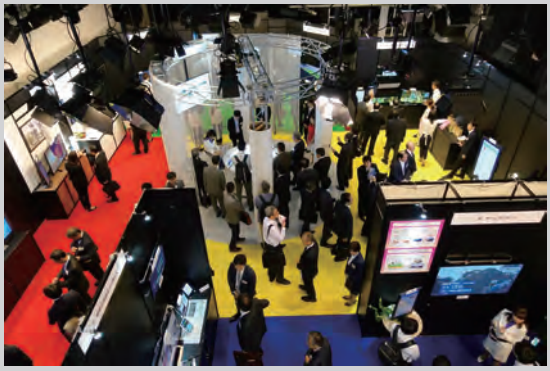


Figure 3-6. Exhibit of content-matching technology at NHK STRL Open House 2018

broadcasting and a car navigation service with the viewer's wish and exhibited it at the NHK STRL Open House 2018.

We conducted field experiments using actual TV programs in Tokyo and Hokkaido and identified expectations and challenges for providing services to link between TV experience and life activities through questionnaire- and interview-based evaluations.

We also studied a way of structuring program data and making it open. We prototyped a calendar application incorporating the linkage with program information and a function to interact with TV using Hybridcast Connect Library, and verified the feasibility of expanding access to program viewing and issues of content-matching technology.

We actively promoted examples of linkage between TV and smartphones, which are the outcomes of our research on content-matching technology, domestically and internationally through exhibitions such as Connected Media Tokyo 2018 and IBC 2018. We also proposed content-oriented IoT, a technology for controlling various IoT-enabled devices in synchronization with content, and researched content description methods[4].

Framework for producing content capable of service linkage

To construct a highly flexible and extensible content production environment by combining various media processes, we participated in the EBU Media Cloud and Microservice Architecture (MCMA) technical project that specifies common interfaces. We contributed to the development and publication of a software library that can be used for constructing an MCMA-compliant system on the cloud. We built a metadata tagging system that can link AI tools across multiple clouds by using the library and exhibited it at IBC 2018 as an MCMA activity group[5].

International promotion of Hybridcast

We are developing a system for producing equivalent applications that operate on both Hybridcast and HbbTV2, which are HTML5-based Integrated Broadcast- Broadband (IBB) systems used in Japan and Europe, respectively. In FY 2018, we tested feasibility of the system by creation of test applications by this system for various aspects. The application is designed to test multilingual support, dynamic graphics rendering and MPEG-DASH playback and the good performance and behavior of the applications is confirmed. In addition, we identified the conditions of the use of Hybridcast libraries to HbbTV2 in MPEG-DASH playback (Hybridcast (4K) video) through the detailed investigation of its applicability. We also presented our equivalent application production system at the HbbTV Symposium, a developer conference for HbbTV, and built cooperative relationships with research organizations in Europe. Additionally, we gave presentations on web technologies for Hybridcast Connect X and Hybridcast (4K) video at the World Wide Web Consortium (W3C).

TV-watching robot

We are researching a robot that watches TV with a viewer, serving as a partner to make TV viewing more enjoyable. We developed a TV-watching robot that has functions to understand the program being viewed, generate utterances autonomously and talk to the viewer and exhibited it at the NHK STRL Open House 2018.

In FY 2018, we developed a template-based utterance text generation method that generates an utterance text by extracting keywords from the program being viewed and combining them with a template sentence containing emotional expressions prepared from past subtitles[6][7]. We also added a function to generate a question that can be answered with a simple word such as "yes" or "no" to encourage the viewer's utterance. In addition to the template-based utterance text generation, we began studying a way to generate an utterance text automatically from program images. In FY 2018, we implemented a method for generating an utterance text from the images of the program being viewed by using an object detection technology using deep neural networks and a caption generation technology.

We also developed a cloud interface for using video and speech recognition services in external cloud environments to extract keywords from the program being viewed. This makes it possible to use different cloud infrastructures in a cross-sectoral manner and convert the outputs to a common format, enabling the selection of services to use according to the purpose of process. Moreover, the use of sequential-parallel processing achieved response within a certain delay time.

To realize dialogs triggered by utterances generated by the robot, we equipped our TV-watching robot under development with an external dialog system. We also organized issues of a speech-based conversation system, which involves the absence of clear response from the user and speech recognition errors, unlike a text-based dialog system.

For the robot to initiate an action to a person around during TV viewing, the recognition of a person and the timing of action are important. Images that can be captured with a camera used for recognizing a person vary greatly according to the angle of view, the installation position in the robot and the presence or absence of a rolling mechanism. We therefore conducted a comparative verification of person detection capability using images of multiple cameras including a fisheye camera, camera array and a camera mounted on the robot's head, and obtained knowledge about characteristics of each type.

Part of this research was conducted in cooperation with KDDI Research, Inc.

Viewing experiments using TV-watching robot

We conducted TV viewing experiments to identify the difference between conversations between persons and conversations involving a robot during TV viewing. We asked two people on good terms to watch any program they like

Figure 3-7. Scenes of viewing experiments using TV-watching robot

freely and observed their viewing behavior for cases with and without a TV-watching robot[8]. Using the observation data, we investigated the listening behavior of people in response to different types of utterances and the robot's motion of turning round. The results demonstrated that it was easier to get a response from people when the robot turned round and talked to them and that people often made utterances expressing their feelings while watching TV programs.

Questionnaire survey on communication robots

To develop a robot that watches TV in the viewer's company, we conducted a market survey by questionnaire to understand to what extent communication robots are recognized and accessed by general public. The survey results showed that 86% of respondents think of robots as something that support people, that there are still only a few opportunities to actually make contact with robots, and that people of higher age are more likely to expect a human-shaped robot[9].

[References]
(1) H. Ohmata and M. Ikeo: "User-Centric Companion Screen Architecture for Improving Accessibility of Broadcast Services," IBC2018 Conference (2018)
(2) H. Ohmata, M. Ikeo, H. Ogawa, C. Yamamura, T. Takiguchi and H. Fujisawa: "Companion Screen Architecture and System Model for Smooth Collaboration between TV Experience and Life Activities," IPSJ Journal, vol.60, no.1, 2019, p.223-239 (2018) (in Japanese)
(3) K. Hiramatsu, H. Ohmata, M. Ikeo, T. Takiguchi A. Fujii and H. Fujisawa: "A Proposal of Hybridcast Connect Library to enable Flexible Service Provision," Proc. of the ITE Annual Convention, 32D-3 (2018) (in Japanese)
(4) H. Ogawa, H. Ohmata, M. Ikeo, A. Fujii and H. Fujisawa: "System Architecture for Content-Oriented IoT Services," IEEE International Conference on Pervasive Computing and Communications (2019)
(5) https://tech.ebu.ch/docs/metadata/MCMA_IBC2018_Final.pdf
(6) Y. Kaneko, Y. Hoshi, Y. Murasaki and M. Uehara: "Utterance Generation based on Keywords of TV Programs," IPSJ SIG Technical Reports, Vol.2018-IFAT-131, no.2 (2018) (in Japanese)
(7) Y. Kaneko, Y. Hoshi, Y. Murasaki and M. Uehara: "A Study on Estimating Alternative Word of Unknown Keyword from Surrounding Words," Proc. of the ITE Annual Convention, 14B-1 (2018) (in Japanese)
(8) Y. Hoshi, Y. Kaneko, Y. Murasaki and M. Uehara: "Basic Survey on TV Watching with Robot by TV Watching Experiment," Proc. of the ITE Annual Convention, 14B-2 (2018) (in Japanese)
(9) Y. Murasaki, Y. Kaneko, Y. Hoshi and M. Uehara: "A Study on Impressions of Communication Robots," IPSJ SIG Technical Reports, Vol.2018-EIP-81 No.13 (2018) (in Japanese)

3.3 Security technologies

To make integrated broadcast-broadband services secure and reliable, we researched cryptography and security technologies.

Cryptography system for database access control

In a service taking advantage of the linkage between cars, mobile devices and TV, the user can receive personalized content in accordance with their preference and location by giving service providers their preference and location information. For such services, we previously developed a cryptography system that allows the user to receive recommended information while keeping their preference and location information closed to any entity. In FY 2018, we conducted a security analysis and implementation evaluation on this system and confirmed that it has high security and causes no major delay that affect the service[1]. To bring about a benefit also to data service providers, we proposed a mechanism for providing service providers whose data was used by users with user identifiers such as email address[2].

Cryptography algorithm for hybrid services

To provide more advanced convergence services of broadcasting and telecommunications, we researched cryptography technologies that can be used to provide secure and reliable services.

We previously developed a cryptography algorithm that can provide access control by specifying the attributes of providers who can access viewer information when the information is stored on the cloud. We compared the security and efficiency of this algorithm and other cryptography algorithms and demonstrated the effectiveness of our proposed system[3].

Post-quantum cryptography algorithm

We researched a cryptography algorithm that can maintain security even when quantum computers are put into practical use.

We developed a multi-party computation method for a hash function for generating a secret key and symmetric encryption for encrypting data. This method inputs multiple pieces of divided data (shares) generated from the data to be encrypted using a method called secret sharing into multiple computers and combines the results of individual calculations on each computer to generate one ciphertext as an output. With this method, the original data cannot be reconstructed from individual shares, meaning that the tolerance to quantum computers can be achieved. We extended our broadcasting service using this algorithm and developed a broadcasting system that can allow users to take out their secret key and that can protect privacy information even when quantum computers are put into practical use[4].

Watermarking algorithm for HEVC-compressed streams

For copyright protection of content, we researched a watermarking technology that can embed the receiver's identity into broadcast compressed content.

We studied the computational complexity and security of the embedding method for HEVC-compressed streams that we developed previously and confirmed its effectiveness[5].

[References]
(1) K. Kajita, K. Ogawa and G. Ohtake: "Security analysis and Implementation on the System using VABKS on IBB Service," ITE Annual Convention 2018, 14D-1 (2018) (in Japanese)
(2) K. Kajita, K. Ogawa and G. Ohtake: "For Realization of IBB Services – Privacy Preserving System considering Provider's merits," SCIS2019, 1C2-3 (2019) (in Japanese)
(3) N. Attrapadung, G. Hanaoka, K. Ogawa, G. Ohtake, H. Watanabe and S. Yamada: "Attribute-Based Encryption for Range Attributes," IEICE Trans. on Fundamentals of Electronics, Communication and Computer Science, Vol.E101-A, No.9 2018, pp.1440-1455 (2018)
(4) K. Ogawa and K. Nuida: "MPC: Improvement of Broadcasting Services," SCIS2019, 1C2-2 (2019) (in Japanese)
(5) K. Ogawa and G. Ohtake: "Efficient Watermarking scheme for Traitor Tracing Encryption Schemes," Journal of Electrical Engineering, Vol.6, No.5, pp. 245-255 (2018)

3.4 IP-based program production technologies

Studies have been conducted on a shift of the interface for broadcasters' program production systems from the conventional Serial Digital Interface (SDI) format to the Internet Protocol (IP) format. The IP format can reduce equipment costs and enable the efficient multiplexing transmission of various signals including video from multiple cameras and multichannel sound over a single cable. It is also expected to bring benefits such as a new program production workflow involving program production from remote areas (remote production).

To achieve an IP-based program production system, we are conducting R&D on IP transmission equipment for lightweight-compressed 8K signals, which enables efficient IP transmission of 8K program contributions in a limited bandwidth, an uncompressed 8K division transmission method for the shared use of program production equipment among 2K, 4K and 8K, IP flow real-time monitoring equipment capable of quality monitoring of IP signals, and the technical verification of IP live production using these systems.

IP transmission equipment for lightweight-compressed 8K signals

In FY 2018, we worked to improve the functions of the IP transmission equipment for lightweight-compressed 8K signals that we developed in FY 2017.

Using IP networks for transmitting program contributions has the advantage of keeping network costs low because other service providers' signals are also transmitted together over the same networks. Meanwhile, there is the challenge of how to ensure the stable transmission of program contributions even at the time of channel switchover due to network congestion or maintenance. As a measure to handle the degradation of network quality such as congestion, in FY 2018, we implemented a packet-loss correction coding scheme, which is resistant to burst-type packet loss, into our lightweight-compressed 8K IP transmission equipment. As a measure to deal with a channel switchover, we also implemented the redundancy of IP signals between the transmitter and receiver sides and a function that allows a variable setting for packet buffering time.

To verify the effectiveness of the lightweight-compressed 8K IP transmission equipment added with the two functions, we conducted laboratory experiments using a device that can simulate a channel error. The results demonstrated that stable video transmission can be achieved even when packet loss occurs.

Uncompressed 8K division transmission method

We developed a method for dividing uncompressed 8K video into multiple 2K video streams and transmitting them as multiple IP flows. This method enables the transmission of uncompressed 8K video using multiple wide-band lines such as widely available 10GbE/25GbE. It also allows the use of divided 2K video for preview and 8K switching process by synchronizing multiple 2K switchers. Moreover, it enables the shared use of program production equipment among 2K, 4K and 8K because it allows 2K program production equipment to be used for 8K program production.

We exhibited a switching system using this method at the NHK STRL Open House 2018 and Inter BEE 2018 and presented the research results at the SMPTE conference[1].

IP flow real-time monitoring equipment

While the IP network, unlike SDI, allows multiple video and sound materials to share a single network, it can cause jitter and packet loss in transmitted signals. This requires the monitoring of signal quality at the IP packet level and we therefore developed equipment that can monitor IP flows in real time (Figure 3-8). This equipment consists of hardware that extracts the header information from IP packets and software that analyzes the header information. As a feature, it can monitor 16 10GbE networks simultaneously by limiting the target of process to the header information. Also, the analysis software can visualize and present transmission latency and the communication status of controlling signals for each IP flow as well as basic indicators such as jitter and packet loss.

We used this equipment for the real-time monitoring of signals flowing over IP networks in the transmission experiments for IP remote production to be described later and demonstrated that it was helpful for identifying the cause of network congestion.

(a) Appearance (b) Examples of monitoring items

Figure 3-8. IP flow real-time monitoring equipment

Verification of IP live production technology

To verify the interoperability between IP live production devices and identify the issues of IP remote production, we conducted technical verifications of IP live production using IP-enabled program production equipment of various vendors in cooperation with the Outside Broadcast Engineering Division.

We demonstrated that broadcasting equipment of different vendors can be interconnected by using the SMPTE ST 2110 transmission standard for program production on IP networks. We also confirmed that an SDI system and an IP system can be interconnected for operation by using SMPTE ST 2059-1/2, which defines synchronized transmission and the phase of vertical video synchronization over IP networks. Meanwhile, we identified issues about monitoring and control technologies caused by IP implementation, such as network congestion and different versions of controlling signals in accordance with equipment. We shared information with the vendors to resolve these issues.

[References]

(1) T. Koyama, J. Kawamoto, M. Kawaragi, T. Kurakake and K. Saito: "Implementation of 8K vision mixer that transports 8K image as multiple 2K SMPTE ST 2110-20 flows," SMPTE 2018 Annual Technical Conference & Exhibition, pp. 1-11 (2018)

4 Smart Production - Intelligent program production

4.1 Social big data analysis technology

Social media analysis technology

We are researching ways to collect the information useful for program production from a massive amount of text data handled by broadcasters ("text big data"). News producers collect newsworthy information from social media posts, which are a form of text big data. To support this effort, at NHK STRL, we are developing a system to automatically collect newsworthy posts in real time from Twitter by using the newsworthy posts previously obtained in news production sites as supervised data for machine learning. This system extracts newsworthy information from a huge amount of tweets by using neural networks (NNs). At present, however, this system identifies several thousands of tweets per day as newsworthy. We therefore developed a user interface to classify the extracted posts by news category automatically determined for the tweet and by the name of place (Figure 4-1)[1]. To improve the function to identify the place of an incident or accident, we studied a method for estimating the poster's location information using graph convolutional NNs with knowledge-based information [2].

Newsworthy tweets identified by the system can contain comments for or retweets of the information already reported by other news media. To determine such tweets, we developed a method for identifying reported information using convolutional NNs[3]. In social media, there are many accounts which make automatic posts, called "bots". Since posts by bots often contain copies of past posts, groundless rumors and so-called fake news, the information on whether the account of a certain post is a bot or not is important for program producers. To meet this demand, we developed a method for determining bot accounts on the basis of the regularity of the time of posting to social media[4].

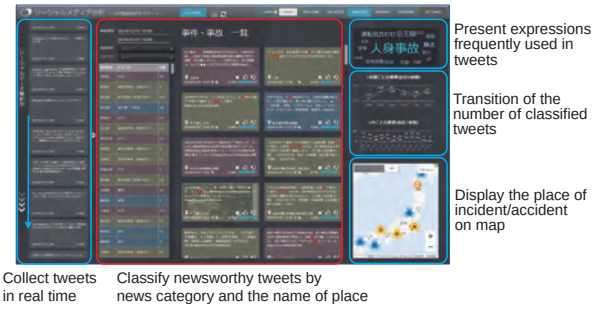


Figure 4-1. User interface for social media analysis system

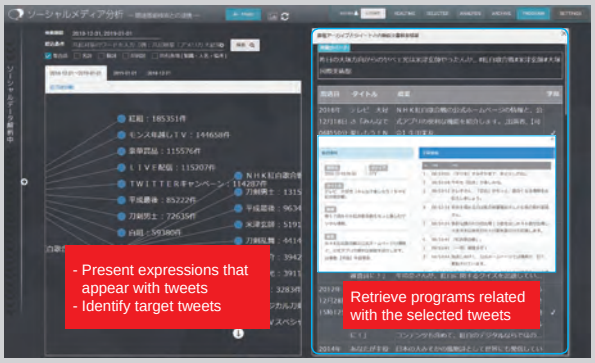


Figure 4-2. Retrieval of related programs on the basis of social media posts

Twitter contains tweets about a wide variety of incidents and accidents happening in the world and some of them may not be able to be handled by a fixed model. We therefore developed a system that allows its model to be updated using new training data prepared on the basis of the operation logs of system users[5]. Because some of the posts about incidents and accidents contain only images attached without text information indicating a fire or traffic accident, we researched a function to extract relevant tweets using not only text analysis but also image classification technologies to be introduced in 4.2 and implemented it into our system.

Broadcasters have accumulated information about programs that they produced in the past, which is useful for producing new programs. We developed a function to retrieve and present the information of relevant past programs on the basis of the collected social media posts by using a method for determining the relation between sentences using our semantic relations dictionary (Figure 4-2)[6].

We utilized the visualization function and analysis function of our system for live broadcast programs reflecting viewer opinions (Tengo-chan, Hajikko Revolution) and a year-end special program (Document 72 Hours). We also participated in the Incident Streams track at the TREC 2018 competition workshop held by the U.S. National Institute of Standards and Technology (NIST) and achieved excellent results (4th place in Information Type Categorization Task and 1st place in Information Priority Task)[7].

Opinion analysis technology

We are researching a technology to analyze opinions about programs comprehensively and constantly after broadcast. Opinions about programs contain not only opinions purely about program content but also requests for programming and business, covering a range of specific opinion types. We implemented an algorithm capable of classifying these opinions from the perspective of practical work of viewer opinion analysis into an analysis system that can be used by relevant departments.

We identified challenges for developing a technology for detecting the target of opinion, which is one of the important factors for determining that the opinion is concerning a program. Some of the conventional methods using social networking services (SNSs) use program hashtags, a notation peculiar to Twitter users, as clues. Program hashtags, however, are influenced by changes in how they are used in social media and the revisions of broadcasters' operation guidelines. It is therefore necessary to develop a method not excessively dependent on program hashtags in order to enable comprehensive and constant analysis. To meet this requirement, we investigated the advantages and disadvantages of methods using program hashtags and studied a method that uses a general language processing technology as clues.

[References]
(1) Y. Takei, K. Makino, T. Miyazaki, H. Sumiyoshi and J. Goto: "Social Media Analysis System for News Gathering Support," FIT 2018, E-015(2018) (in Japanese)
(2) T. Miyazaki, A. Rahimi, T. Cohn and T. Baldwin: "Twitter Geolocation using Knowledge-Based Methods," W-NUT2018, pp.7-16 (2018)
(3) K. Makino, Y. Takei, T. Miyazaki and J. Goto: "Classification of Tweets about Reported Events using Neural Networks," W-NUT2018, pp.153-163 (2018)

(4) H. Okamoto, T. Miyazaki and J. Goto: "Humans are Ambiguous and Capricious: Bot Detection Focusing on Statistical and Signal-Like Property of Tweeting Activities," IPSJ2019, 7K-02 (2019) (in Japanese)
(5) K. Makino, Y. Takei, H. Okamoto, T. Miyazaki and J. Goto: "A Study on Multitask Learning Methods for Extracting Important Tweets," ITE Winter Convention,12C-3 (2018) (in Japanese)
(6) T. Miyazaki, K. Makino, Y. Takei, H. Sumiyoshi and J. Goto: "TV Program Retrieval from Tweets for Supporting TV Program Creation," ITE Annual Convention, 13B-5 (2018) (in Japanese)
(7) T. Miyazaki, K. Makino, Y. Takei, H. Okamoto and J. Goto: "NHK STRL on the IS track of the TREC 2018," TREC2018 (2018)

4.2 Image analysis technology

Technology for automatic metadata assignment to video

Raw video footage stored in video archives and broadcast stations is a valuable resource for program producers. To make active use of such footage, we are researching automatic assignment of metadata to video.

Since program producers often retrieve video using personal names as keywords, a face recognition technology to identify persons appearing in footage is very important. A face detection technology for specifying face positions in footage is essential for face recognition and we worked to improve the accuracy by introducing a deep learning technology. Using cascaded convolutional neural networks (NNs), we developed a face detection method that can achieve both low computation cost and high detection accuracy, which improved detection omission by about 40% compared with conventional methods[1]. We developed a video editing support system with face recognition using this technology and used it for producing a documentary program, "Peeping Through 100 Cameras," which was aired in September. We also developed an automatic face blur system using face recognition in cooperation with relevant departments and exhibited it at the 48th NHK Program Technology Exhibition.

In news production sites, the trend to use tweets and SNS posts for news reporting activities is growing. To support this, we studied introducing a technology to improve the accuracy of extracting and classifying newsworthy tweets by analyzing posted images for our social media analysis system under development (see 4.1). We prepared unique training data considering Twitter's distinctive characteristic that a wide range of non-newsworthy images are posted and developed a method for classifying images into five categories related with accidents and disasters by using convolutional NNs[2]. We applied this technology to the social media analysis system and exhibited it at the NHK STRL Open House 2018.

We also worked toward the practical use of a system using technologies that we have developed. We continued with the experimental use of a video material management system equipped with an object recognition technology and a similar-image retrieval technology that we built on the intranet in FY 2017 at program production sites which handle CG synthesis and video effects, and modified the system on the basis of findings from the effort.

Automatic colorization technology for monochrome video

Monochrome video stored in broadcast stations is a valuable video resource. In response to the growing needs for colorizing and using it for programs, we are researching the automatic conversion of 4K-resolution-equivalent monochrome film video to color video.

For the automatic colorization system using three types of NNs (color estimation, color correction, color propagation over adjacent frames) that we developed in FY 2017, we evaluated its effect on the reduction in working hours and the suppression of color variations[3]. This system was used for producing an NHK special program, "Nomonhan: An Irresponsible Battle," which was aired in August, and a program, "That Day, That Time, That Program," which was aired in November, contributing to the reduction of working hours to about 1/60 that of the conventional process. We exhibited this system at the NHK STRL Open House 2018, IBC 2018 and Inter BEE 2018.

Additionally, we began training NNs using video obtained from production sites with the aim of using this technology for HDR (High Dynamic Range)-SDR (Standard Dynamic Range) conversion, which was proposed by a relevant department.

Video summarization technology

Distributing digest videos of programs via websites and social media is becoming more common in program production sites. To assist with the production of these videos, we are

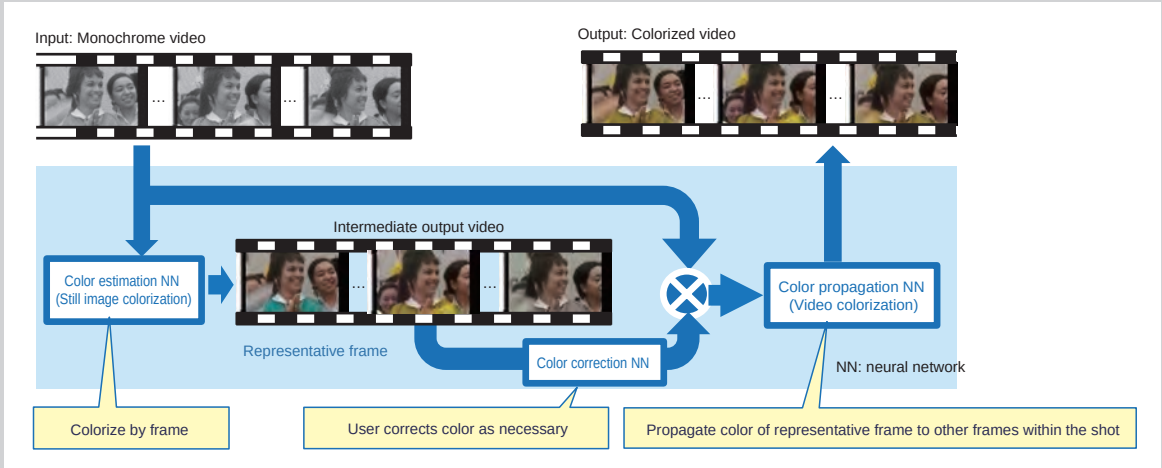


Figure 4-3. Automatic colorization technology for monochrome video

researching a technology to automatically summarize program video.

In FY 2017, we developed a demonstration system that can summarize a program on the basis of various information such as viewer responses via tweet analysis and image analysis results such as the faces of cast members, the size of open captions (telops) and the amount of camera work. In FY 2018, we increased the number of program genres targeted by the system and exhibited it at the NHK STRL Open House 2018 and Inter BEE 2018. Considering the convenience in production sites, we also developed an automatic video summarization tool that can operate on hardware equivalent to a laptop PC and asked relevant departments to verify its operation. Moreover, we introduced NN technology to improve the quality of summarization and constructed an automatic summarization model with a unique network structure using manually-edited summarized videos as training data.

As an application of the video summarization technology, we began developing a technology to automatically generate materials for program websites. As the initial study, we developed a technology to identify the location by comparing the frame images of program video with images displayed on map websites with the aim of achieving the automatic generation of route maps to be published on the websites of on-location TV programs[4].

[References]
(1) Y. Kawai, R. Endo, N. Fujimori, T. Mochizuki: "Prototype system for supporting TV program editing based on face recognition," Proc. of ITE Annual Convention, 22B-1 (2018) (in Japanese)
(2) N. Fujimori, T. Mochizuki, Y. Kawai, M. Sano: "Investigation of Dataset Construction Method and Decision Criterion for Classification of Images Posted on SNS," Proc. of ITE Annual Convention, 21B-2 (2018) (in Japanese)
(3) R. Endo, Y. Kawai, T. Mochizuki: "Monochrome Video Colorization System Taking Account of Color Consistency," Proc. of ITE Annual Convention, 22B-3 (2018) (in Japanese)
(4) A. Matsui, T.Mochizuki, N.Fujimori: "Multimodal Location Estimation of Videos based on Named Entity Extraction and Location Search using Image Matching," Proc. of ITE Annual Convention, 12B-1 (2018) (in Japanese)

4.3 Speech transcription technology

The transcription of speech in video footage is indispensable to produce programs. A system to produce transcription efficiently is needed for the more swift delivery of accurate programs to viewers. In FY 2018, we worked to increase the accuracy of speech recognition with the aim of realizing a transcription production system using speech recognition and began developing a transcription interface that can also handle the transcription of live content in addition to recorded video footage.

Speech recognition technology for transcription assistance

When the speech of a broadcast program is recognized for closed captioning, word sequences are usually estimated using conversational speech as input. Meanwhile, it is difficult to recognize conversational speech in video footage targeted for transcription assistance only with the sound information because the speech is not always recorded under a favorable sound recording condition and also contains many informal phrases. To address this problem, we researched a recognition technique that uses video information as well as sound information, focusing on the fact that footage used by broadcasters contains video information. We developed a new method that utilizes the information of the middle layer of a deep neural network (DNN) aimed at the object recognition of images and the information of video captions generated by DNNs for training a language model that expresses multiple word chains. This method improved the accuracy of the language model[1].

Since speech in video footage contains unclear sentence structures as well as many hesitations and repetitions, the recognition result lacks readability if displayed as is. It is therefore necessary to rewrite the speech recognition result in various ways, such as inserting appropriate punctuation marks and deleting unnecessary words, to improve the readability, but rule-based automatic formatting is not sufficient for improving the readability. To address this problem, we constructed a sentence format model that formats speech recognition results while making corrections using Encoder-Decorder networks based on DNNs. Using the closed captions of broadcast programs as the correct answer, we conducted experiments to see how this model improves the rate of disagreement with the closed captions[2]. We also conducted comparative experiments using training data with different rates of disagreement with the closed captions and the results demonstrated that there are criteria for selecting training data suited for sentence formatting.

Footage targeted for transcription also contains telephone speech, for which there is a strong demand for transcription assistance as with other materials. However, it was difficult to recognize telephone speech with the speech recognition system that we previously developed because there are restrictions on the frequency bandwidth for telephone speech materials. We therefore began an effort to convert the speech of our existing training data to telephone speech in a simulated manner and use it for training an acoustic model for telephone speech.

Transcription interface

We continued to develop an interface that allows the user to refer to speech recognition results efficiently and to modify the recognition errors as necessary with a minimum operation. In

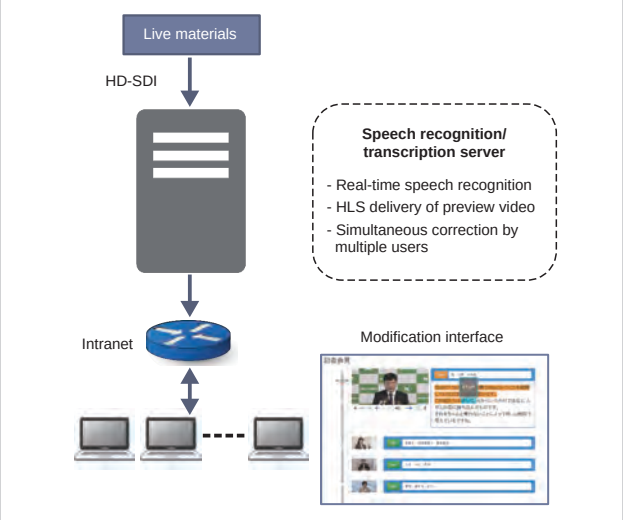


Figure 4-4. Real-time transcription system

FY 2017, we developed a modification interface for recorded content. In FY 2018, we developed a real-time transcription system for live content such as live broadcast programs and transmission contributions[3][4].

For the prompt transcription of live content, this system employs real-time speech recognition process and an HTTP Live Streaming (HLS) delivery technology for enabling the preview of arbitrary parts of live content (Figure 4-4). It also has a function to immediately deliver modified parts and corrected characters to all terminals because the transcription of live content is likely to involve cooperative work by multiple people.

To verify the effectiveness of this system, we introduced two sets of this system to news program production sites where the transcription of live content is frequently performed and began verification experiments.

[References]
(1) A. Hagiwara, H. Ito, M. Ichiki, T.Kobayakawa, T. Mishima and S. Sato: "Language model utilizing image features for automatic speech recognition," Autumn Meetings of the Acoustical Society of Japan, 2-Q-13, pp.1071-1074 (2018) (in Japanese)
(2) H. Ito, A. Hagiwara, T.Kobayakawa, T. Mishima, S. Sato and A. Kobayashi and S. Sato: "Transcription formation using encoder-decoder network," Autumn Meetings of the Acoustical Society of Japan, 2-Q-13, pp.1021-1022 (2018) (in Japanese)
(3) T. Mishima, M. Ichiki, A. Hagiwara, H. Ito, T.Kobayakawa and S. Sato: "A Real-Time Transcription System Using Speech Recognition," ITE Annual Convention, 21D-4 (2018) (in Japanese)
(4) A. Hagiwara, H. Ito, T.Kobayakawa, T. Mishima and S. Sato: "Development of Transcription System Using Speech Recognition for Video Footage," IPSJ SIG Technical Report, Vol. 2018-SLP-124, No. 5, pp.1-6, (2018) (in Japanese)

4.4 New image representation technique using real-space sensing

With a view to producing more interesting and user-friendly video content efficiently for live sports coverage and other live programs, we are researching new image representation technologies and advanced program production technologies using multiview images and sensor information.

New image representation by visualization technology

We researched new image representation using an object-tracking technology and an area extraction technology based on video analysis and sensor information.

Figure 4-5. Display of the golf putting trajectory at the 83rd Japan Open Golf Championship

Figure 4-6. Multi-motion display at the 60th NHK Trophy National Ski Jump Tournament

As new image representation using an object-tracking technology, we continued with our research on a "sword tracer" system that visualizes the movements of the tip of a sword in fencing in CG. The sword tracer system measures the positions of the two sword tips wrapped with reflective tape in real time by applying the object-tracking technology to video captured with an infrared camera[1]. In FY 2018, we made the tracking process more robust by distinguishing the two sword tips of both players using their attitude information obtained by analysis of visible camera images.

We researched a technology for visualizing the trajectories of a putt and a shot in golf events. In FY 2018, we developed a system that visualizes the putting trajectory in CG by detecting and tracking a golf ball in real time from images captured with a camera installed near the green. The system achieved tracking process that can robustly accommodate environmental changes by sequential learning the image features of a ball by frame. We used this system for a live program "The 83rd Japan Open Golf Championship," aired from October 11 to 14, 2018 (Figure 4-5). For shots, we achieved the superimposition of trajectory CGs on images from a handy camera, whose attitude changes, in addition to images from a fixed camera by using sensors for measuring the camera attitude angle.

We are researching a "multi-motion" system that extracts an athlete region from camera images and shows the athlete's motion with stroboscopic effect. In FY 2018, we increased the speed of generating images by optimizing the program and improved the operation interface assuming its use for commentaries in programs. This system was used in commentary scenes[2] of the live broadcasting of "The 60th NHK Trophy National Ski Jump Tournament," aired on November 4, 2018 (Figure 4-6).

We continued with our development of "Sports 4D Motion," a new image representation technology that combines multiview images and 3D CGs. In FY 2018, we developed an online processing system for Sports 4D Motion[3]. This system can generate Sports 4D Motion images within several tens of seconds after the completion of capture by high-speed image processing, such as projective transformation of 4K multiview images, and 3D CG synthesis. We exhibited this system at the NHK STRL Open House 2018 and also confirmed its effectiveness by conducting performance evaluation through demonstration experiments.

Advanced live program production technology

We began research on key technologies for "meta-studio" to achieve new video representation and improved efficiency in live program production by utilizing various object information obtained from images and sensors.

In our research on automatic program production technologies, in FY 2018, we began developing an AI robotic camera that performs automatic shooting in accordance with the situation. As a scene analysis technology for soccer events, we prototyped software that enables the conversion of the subjects into objects and the simultaneous extraction of image features such as the positions and velocities of players and a ball, players' face directions and the colors and types of their shirts. We improved the accuracy of face direction estimation and increased the speed for online processing by introducing a classifier using deep convolutional neural networks to a method that we previously developed. We also developed a rule-based camera work generation algorithm on the basis of shooting know-how that we learned from experienced camera persons for live program production. In addition, we verified the operation of automatic camera work along with a robotic camera simulator using cutout from 8K video.

We made progress in our research on a studio robot that enables natural joint performance between live performers and CG characters in a studio. In FY 2018, we developed a real-time unwanted object removal technology to hide unnecessary parts when the robot is too large to be completely hidden behind the CG superimposed on the robot position in the video. Removing an unwanted object requires the background information such as the structure and surface pattern behind the unwanted object and the position and attitude information of the camera. This technology measures the background information in advance as a CG object and stores it in the CG rendering device. During capturing, it generates a CG of the object under the same position and attitude conditions as the camera and thus generates background images in which the unwanted object does not exist. It then replaces the area of the unwanted object in photographed images with the CG background images using the robot position information. This allows only the unwanted object to be removed in real time.

We developed a general-purpose camera attitude sensor that can be used for various capture equipment such as handy cameras and crane cameras. In FY 2018, we prototyped a system that measures the camera attitude angle with a micro-electro-mechanical-systems (MEMS) inertial sensor and a fiber optic gyroscope (FOG), the horizontal position with a laser sensor, and the height with an RGB-D sensor, and conducted verification experiments. We demonstrated that these sensors equipped on a camera can stably measure real-time data required for CG synthesis.

[References]
(1) M. Takahashi, S. Yokozawa, H. Mitsumine, T. Itsuki, M. Naoe, and S. Funaki: "Sword Tracer: Visualization of Sword Trajectories in Fencing," Proc. of the ACM SIGGRAPH 2018 Talks (2018)
(2) A. Arai, H. Morioka, H. Mitsumine, D. Kato: "Development of Multi-Motion System Considering Live Production," Proc. of ITE Winter Annual Convention, 13D-4 (2018) (in Japanese)
(3) K. Ikeya, J. Arai, T. Mishina, M. Yamaguchi: "Capturing method for integral three-dimensional imaging using multiviewpoint robotic cameras," J. Electron. Imaging 27(2), 023022 (2018)

4.5 Promotion of the use of AI technologies

Promotion of the use of AI technologies

In FY 2016, we aggregated our research outcomes of AI technologies for over 30 years and established Smart Production Lab in NHK STRL as a base for promoting a working-style reform by efficient program production and accelerating the practical application of our research results for user-friendly broadcasting. The laboratory has been actively releasing information through domestic and international events.

In FY 2018, we reinforced the development structure of the laboratory to swiftly respond to requests from NHK broadcast stations across the nation including regional stations for using AI-related technologies and supported the development of various technologies, such as an automatic transcription system using speech recognition technology, trial live closed-captioning for regional stations and an AI announcer using speech synthesis technology, for practical use.

To promote the extensive practical use of AI technologies, we launched a smart production project participated by diverse program production-related departments. We also held an event hosted by this project in February with the aim of matching program producers' needs for AI utilization with technologies developed by NHK STRL and achieved beneficial results such as the identification of new opportunities for practical application (Figure 4-7).

Figure 4-7. Event hosted by the smart production project

5 Smart Production - Universal service

5.1 Automatic closed-captioning technology

Closed-captioning service not only serves the needs of people with hearing disabilities and the elderly by conveying speech in TV programs in text but also offers a useful function to general viewers watching programs in a noisy environment or an environment where audio cannot be played such as public places. Also, there is a demand from viewers for more programs with closed captions, including live broadcast programs and programs produced by regional broadcasting stations. Speech recognition technology is used for real-time closed captioning, but many regional broadcasting stations are faced with issues such as a shortage of staff to correct speech recognition errors and a considerable amount of time required to prepare necessary equipment and system. To address these issues and expand closed-captioning service, we began a trial service that distributes speech recognition results on the internet without correcting recognition errors with the aim of evaluating the extent to which uncorrected speech recognition results can help viewers understand programs.

Speech recognition technology for internet delivery

A service to deliver uncorrected recognition results on the internet requires a higher recognition accuracy. We therefore employed a speech recognition technology that we developed for the transcription of video footage, which requires a higher level of recognition than for program speech[1]. This technology can generate highly precise recognition results sequentially for speech inputs streamed over network without waiting for the end of a sentence. We verified the recognition accuracy of this speech recognition technology when used for news and information programs produced by regional broadcasting stations and demonstrated that a recognition accuracy of 90% to 97% can be achieved for news items and live coverage and information items.

Meanwhile, the recognition accuracy declined significantly for interviews with local residents, which contain inarticulate speech, an accent peculiar to the region or different phrases from those used in standard Japanese. The recognition results of these parts make it difficult to understand the program due to incorrectly recognized words. Also, many interviews in news and information programs are open captioned and do not need additional closed captions. We therefore elaborated a way of presenting closed captions by stopping speech recognition and displaying "..." in closed captions for parts where the recognition accuracy is expected to decline.

The Chinese characters of personal names used in programs produced by regional broadcasting stations cannot be correctly identified in some cases because they are not always the same as those of names used in programs produced in Tokyo, which are used as training data for speech recognition. Therefore, we decided to display personal names in katakana, which is the Japanese phonetic characters.

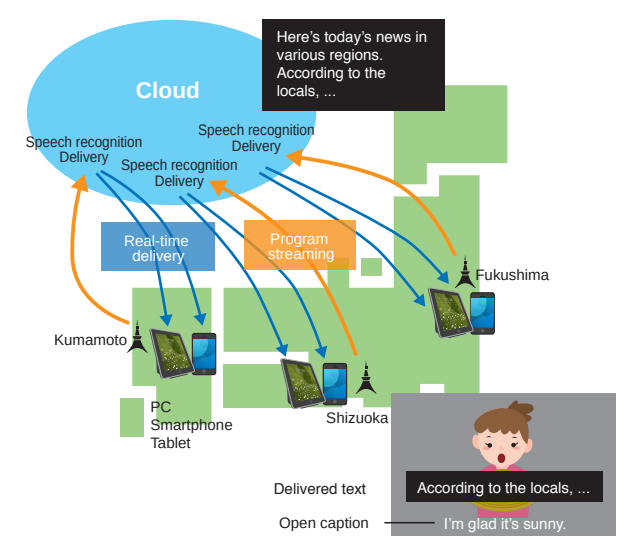


Figure 5-1. System overview

NHK broadcasts regional programs of each regional broadcasting station in a broadcast time frame starting 6:10 PM on weekdays. Providing this trial service across the nation would require as many sets of speech recognition equipment as there are regional broadcasting stations, incurring a large-scale capital investment and operation costs for the installation and maintenance. We therefore built a system efficiently by aggregating speech recognition and delivery equipment on the cloud (Figure 5-1). We began this trial service in three NHK broadcast stations of Shizuoka, Kumamoto and Fukushima in February 2019.

[References]
(1) A. Hagiwara, H. Ito, T.Kobayakawa, T. Mishima and S. Sato: "Development of Transcription System Using Speech Recognition for Video Footage," IPSJ SIG Technical Report, Vol. 2018-SLP-124, No. 5, pp.1-6, (2018) (in Japanese)

5.2 Audio description technology

We are researching "audio description" technologies, which produce voice explanations for live broadcast programs so that people with visual impairment can enjoy live sports programs better. We studied "automated audio description," which supplements human commentaries with auxiliary voice explanations for visually impaired people, "robot commentary," which provides commentaries for internet services in place of human announcer, and a speech synthesis technology, which is the base of these technologies.

Automated audio description

As with manually produced audio descriptions, an automated audio description should not overlap with program speech to the extent possible. We studied a technique to estimate a desirable timing for inserting a commentary from the variation trend of acoustic features of speech. We developed a prototype system that operates in real time and exhibited it at the NHK STRL Open House 2018. The results of comparison between

the timings estimated by this technique and the timings which are visually confirmed to be ready for an insertion demonstrated a certain level of effectiveness[1][2]. We also conducted evaluation experiments participated by visually impaired people on favorable hearing conditions under a situation where speech overlaps cannot be avoided. The results clarified the conditions which give little influence even if an overlap occurs and the conditions which must be prevented, demonstrating guidelines for system design.

Additionally, we investigated the acoustic features of easy-to-hear commentary speech and studied the feasibility of expanding service to non-sports programs in cooperation with visually impaired people and program producers.

Robot commentary

Robot commentary is a technology to automatically generate commentary manuscripts describing the situation in games from real-time competition data provided in international sports competitions and to read out the manuscripts with synthesized speech[3]. We provided a robot commentary service for some sports via internet delivery at a sports event held in 2018. With the aim of enriching the content of speech and providing commentaries for more sports in an event to be held in 2020, we improved the speech algorithm and studied a way of collecting necessary data.

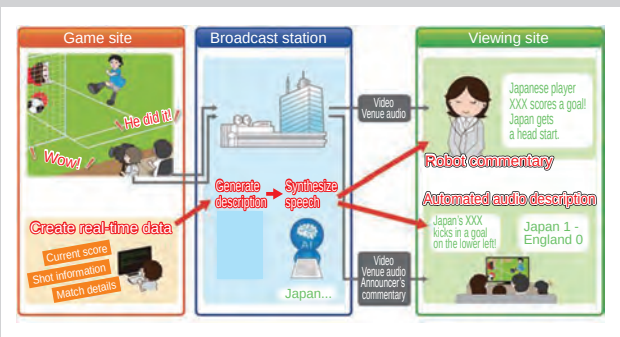


Figure 5-2. Mechanism of automated audio description and robot commentary

Speech synthesis technology

As a practical application of the speech synthesis technology using deep neural networks (DNNs) that we developed in FY 2017, we developed the voice of CG reporter "Yomiko" for a program "News Check 11." We prepared necessary data and modified the operation program to enable the handling of diverse speech expressions according to production effect needs and improved utterance skills such as intonations, poses and conversational tones through actual use in programs. This technology is also utilized for news reading service on the internet. Concurrently, we began developing a new speech synthesis technology that could bring further quality improvement of synthesized speech and higher efficiency of the preparation of training data for a speech model, which currently takes significant costs and time[4][5].

We developed a DNN speech synthesis technology to realize speech with an announcer-equivalent quality for a trial service in which the radio weather information of regional broadcasting stations will be partly provided by speech synthesis. We achieved high-quality speech by limiting the content of speech to weather forecast programs. We began test broadcasts in prefectural-area radio broadcasting by the Kofu station in March 2019.

Additionally, we helped promote the use of Yomiko for internet content by using a speech rate conversion technology that we previously developed. We added a production effect to read tongue twisters fluently and a function to shorten the speech of news commentary video to fit in a specified time by varying the speech speed. We also continued to support the operation of Chinese learning applications, "Seicho Kakunin-kun (Tone Checker)" and "Sorijita Kakunin-kun (Retroflexion Checker)," in an Educational TV program "Learn Chinese on TV" by applying a speech processing technology that we previously developed.

[References]
(1) M. Ichiki, T. Shimizu, A. Imai, T. Takagi, M. Iwabuchi, K. Kurihara, T. Miyazaki, T. Kumano, H. Kaneko, S. Sato, N. Seiyama, Y. Yamanouchi and H. Sumiyoshi: "A Study of Overlap between Live Television Commentary and Automated Audio Descriptions," Miesenberger, K. et al. eds., Springer, 2018, pp.220-224 (2018)
(2) M. Ichiki, T. Kumano, A. Imai and T. Takagi: "Timing determination method to insert an automated audio description in live television broadcast," IEICE Technical Report vol.118, no.270, WIT2018-29, p.45-50 (2018) (in Japanese)
(3) K. Kurihara, A. Imai, N. Seiyama, T. Shimizu, S. Sato, I. Yamada, T. Kumano, R. Tako, T. Miyazaki, M.Ichiki, T. Takagi and H. Sumiyoshi: "Automatic Generation of Audio Descriptions for Sports Programs,"SMPTE Motion Imaging Journal (2019)
(4) K. Kurihara, N. Seiyama, T. Kumano and A. Imai: "Study of Japanese End-to-End speech synthesis method that inputting kana and prosodic symbols," Proc. of ASJ fall general conference 1-4-1 (2018) (in Japanese)
(5) K. Kurihara, N. Seiyama, T. Kumano and A. Imai: "Study on switching method of speaking styles in Japanese end-to-end speech synthesis," Proc. of ITE winter conference 12C-1 (2018) (in Japanese)

5.3 Machine translation technology

To provide information to foreigners promptly and efficiently, we are conducting research on machine translation for texts of news reports, broadcast programs and newspaper articles and on news scripts with reading assistance information to help understand Japanese news.

Machine translation of news and program content

Broadcasters translate Japanese into foreign language to provide information for non-native speakers. For the speedy and efficient production of foreign language content, we are researching machine translation[1]. The mainstream of recent machine translation technologies is a method that collects a huge amount of parallel translation data and trains a translation model using neural networks. In FY 2018, we created 500,000 pairs of Japanese and English news sentences by translating Japanese news manuscripts manually. Using this data, we prototyped a Japanese-English machine translation system. In addition, we developed an interface that allows easy manual correction of the machine translation results by mapping Japanese and English words and assigning the degree of translation reliability for English words. This research was conducted in cooperation with the National Institute of Information and Communications Technology (NICT).

Furthermore, we are researching English-Spanish machine translation with the aim of supporting the production of Spanish closed captions for video-on-demand (VOD) service programs on NHK WORLD. In FY 2018, we newly prepared

40,000 pairs of closed caption sentences from VOD programs and 10,000 pairs of sentences from scripts of NHK WORLD RADIO JAPAN as parallel data of English and Spanish.

Machine translation of newspaper articles

With the aim of facilitating communication between non-Japanese and Japanese in business scenes, we are researching machine translation technologies for conversations and small talk in meetings and social occasions and newspaper articles in cooperation with five external institutions. NHK is in charge of a newspaper article translation technology. In FY 2018, we conducted Japanese-English machine translation experiments using parallel data generated by automatically mapping Japanese sentences and English sentences of Jiji Press. The experiments achieved better results both in objective evaluation and subjective evaluation than existing translation systems. We also identified research subjects that we should work on in the future, such as support for translation of new words, under-translation, mistranslation, over-translation, and translation with context. This research was supported by NICT as part of a project titled "R&D of Deep Learning Technology for Advanced Multilingual Speech Translation."

News scripts with reading assistance information

In FY 2018, we produced 64 articles of news scripts with reading assistance information for experimental purpose and investigated the effect of such news scripts for non-native speakers in Japan. The results demonstrated that reading assistance information helps improve their understanding of news. This research was conducted in cooperation with Utsunomiya University.

[References]
(1) I. Goto and H. Tanaka: "Detecting Untranslated Content for Neural Machine Translation," Journal of Natural Language Processing, Vol. 25, No. 5, pp.577-598 (2018) (in Japanese)
(2) I. Goto, H. Mino, T. Kumano and H. Tanaka: "Proposal of News with Reading Assistance," ITE Journal, Vol.72 No.8, p. J113-J115 (2018) (in Japanese)

5.4 Information presentation technology

To convey information to all viewers including those with vision or hearing impairments in an easy-to-understand manner, we made progress in our research on a technology to generate sign language computer graphics (CG) from sports information and a technology to convey the status of sports events and motion information to tactile sensation. We also began research on information presentation using the sense of smell.

Sign language CG for presenting sports information

To enrich broadcasting services for viewers who mainly use sign language, we are researching a technology for automatically generating sign language animations using CG for information about sports events.

We developed a system that automatically generates sign language CG and Japanese closed captions by combining competition data delivered during a game and templates prepared in advance and displays them on the web browser in synchronization with the game status. We conducted experiments on the real-time automatic generation and presentation for ice hockey and curling matches and exhibited the results at the NHK STRL Open House 2018. Evaluation experiments using the generated content demonstrated guidelines for the information that should be presented and a screen layout in actual service (Figure 5-3)[1].

Figure 5-3. Example of the screen layout of sports sign language CG service

In our research on machine translation from Japanese sentences to sign language CG for sports news, we prototyped an automatic translation system using a syntax transfer method, which changes the order of words by converting Japanese syntactic structures to sign language ones, to support sentences with complicated syntactic structures. The conversion of syntactic structures uses data applied with machine learning from the results of Japanese syntax analysis and the results of sign language syntax analysis that we developed in FY 2017. We also developed a sign language CG production assistance system that allows the user to modify translation errors manually by changing the order of sign language words or replacing an incorrect word with a correct one. The results of evaluation experiments using sign language CG animations generated by this system demonstrated the need for a function to appropriately present the words or phrases to be replaced for efficient modification[2]. We also began developing a function to add appropriate facial expressions in accordance with the impression of the context and words and studying evaluation experiments on the production assistance system assuming actual operation.

With the aim of expanding the parallel corpus necessary for improving the accuracy of machine translation, we began R&D on a technology to convert sign language video into text using image recognition. We verified previous studies on sign language recognition using deep learning and conducted evaluation experiments using training data for Japanese Sign Language (JSL). We confirmed the effectiveness of deep learning and obtained knowledge about training data necessary for improving the recognition rate of JSL.

We have been gathering user feedback about the understandability of our weather report sign language CG for the Kanto Region via an evaluation website that we released in NHK Online in February 2017. To expand the coverage area of scheduled forecasts from the prefectural capitals of the seven prefectures in the Kanto Region to those of the nation's 47 prefectures, we prototyped a weather report sign language CG generation system and verified its operation. Part of this study was conducted in cooperation with Kogakuin University.

Haptic presentation technology for touchable TV

We are researching a technology for conveying the information of movements in video to people's skin. Our main purpose is to convey information such as the moving directions of a ball and players and the timings when the ball hits the floor, wall or racket in fast-moving sports content that is difficult to convey with speech information. For the tactile presentation of motions and timings, we decided to use three types of stimuli, vibration, sliding and acceleration. In FY 2017, we identified fundamental conditions for the perception and discrimination of these stimuli and demonstrated the feasibility of conveying motions and timings. In FY 2018, we conducted experiments by artificially adding tactile stimuli to the video and speech of virtual sports content to verify whether it is possible to understand the flow of the game and which team has scored a point. The experiments, participated by visually impaired people, demonstrated that it is possible to understand the game status even without visual information. An experiment in which the ball moving back and forth in tennis was presented by sliding, the linear movement of a pressure stimulus to the skin, identified perception and recognition characteristics such as the difficulty in conveying the changes of motion speed while demonstrating the feasibility of conveying the direction of ball motion.

We developed a wristband haptic device that gives stimuli to the wrist using four vibrators through such improvement of the cube vibration device developed in FY 2017 as unnecessary to hold it by hand (Figure 5-4)[3]. Using this device, we expressed player actions in volleyball (serve, receive, set, attack) and the ball hitting in or outside of the line by vibration. The results of evaluation experiments participated by people with visual impairment and people with hearing impairment demonstrated that a certain level of information can be conveyed.

For an experiment on simulating and conveying impact by using vibration with large acceleration, we developed a ball-type haptic device that presents vibrations on the palm (Figure 5-5) and confirmed the possibility of understanding the game status without visual information by varying and combining parameters such as the size of amplitude to be presented and the volume and type of sound information[4]. We plan to continue our research to achieve more enriched content representation using AR/VR technology and real-time presentation from sports event information. Part of this research was conducted in cooperation with the University of Tokyo and Niigata University.

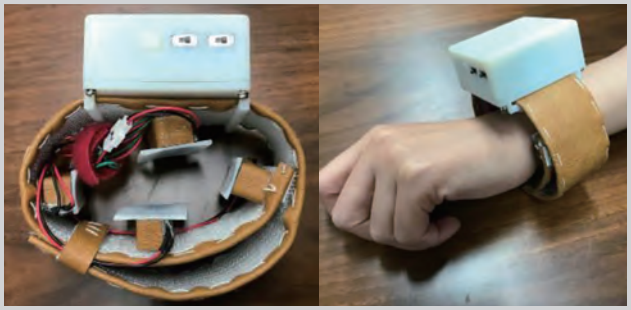


Figure 5-4. Appearance of wristband haptic device

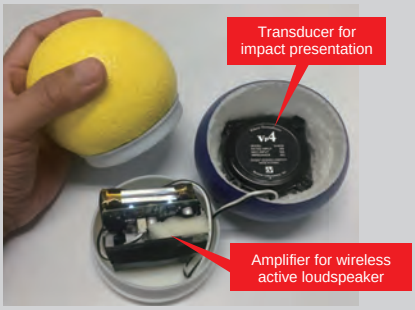


Figure 5-5. Internal structure of ball-type haptic device

In our research on a technology for effectively presenting 2D information that is difficult to describe in words, such as diagrams and graphs, to people with visual impairment, we continued with our development of a finger-leading presentation system with a tactile display. This system combines a tactile display that conveys information using the unevenness and vibrations of pin arrays that move up and down and a method for conveying important points by leading fingers with a kinetic robot arm[5]. In FY 2018, we conducted development to enable the system to be used for many purposes, including helping visually impaired people learn characters and assisting deaf-blind people with their communication. A common way to convey characters to visually impaired people who have not mastered braille or deaf-blind people is for a caregiver to trace characters with his/her finger on the deaf-blind person's palm. We developed a method for guiding the user's finger according to the stroke order of characters and demonstrated through evaluation experiments that this method achieves a higher character recognition rate than that of character-tracing on the palm. This showed that this system is effective not only as a learning tool but for helping deaf-blind people with their communication. Part of this research was conducted in cooperation with Tsukuba University of Technology.

Olfactory information presentation method

To provide richer viewing experiences, we began research on an olfactory information presentation method. We investigated previous studies and the latest trend of olfactory information presentation technologies and studied broadcast content for which the addition of olfactory information is effective. Also, we began studying a method for effective olfactory information presentation with simultaneous presentation of video.

[References]

(1) T. Uchida, H. Sumiyoshi, T. Miyazaki, M. Azuma, S. Umeda, N. Kato, Y. Yamanouchi and N. Hiruma: "Evaluation of the Sign Language Support System for Viewing Sports Programs," International ACM SIGACCESS Conference on Computers and Accessibility, pp.361-363 (2018)

(2) S. Umeda, N. Kato, M. Azuma, T. Uchida, N. Hiruma, H. Sumiyoshi and H. Kaneko: "Sign Language CG Production Support System to Modify Signing Expression with Machine Translation," Human Communication Group Symposium, A-7-5 (2018)

(3) M. Azuma, T. Handa, T. Shimizu and S. Kondo: "Hapto-band: Wristband Haptic Device Conveying Information," Lecture Notes in Electrical Engineering, Vol.535 (3rd International Conference on Asia Haptics), Kajimoto, H. et al eds., Springer, D2A17 (2018)

(4) T. Handa, M. Azuma, T. Shimizu and S. Kondo: "A Ball-type Haptic Interface to Enjoy Sports Games," Lecture Notes in Electrical Engineering, Vol.535 (3rd International Conference on Asia Haptics), Kajimoto, H. et al eds., Springer, D1A16 (2018)

(5) T. Sakai, T. Handa, T. Shimizu: "Development of Prop-Tactile Display to Convey 2-D Information and Evaluation of Effect -Affection of Presentation Conditions on Powered Mechanical Leading Method for Spatial Position Cognition-, IEICE Trans. Information and Systems, Vol.J101-D, No.3, pp.628-639 (2018)

6 Devices and materials for next-generation broadcasting

6.1 Imaging technologies

Three-dimensional integrated imaging devices

We are researching imaging devices with a 3D structure in our quest to develop a next-generation image sensor having more pixels and a higher frame rate. These devices are fabricated by bonding a photodetector and a signal processing circuit which are formed on different substrates. They have a signal processing circuit for each pixel directly beneath the photodetector (Figure 6-1). Since this structure can maintain a high frame rate even when the number of pixels increases by enabling signals from all pixels to be read out simultaneously, it is expected to be helpful for capturing future 3D images.

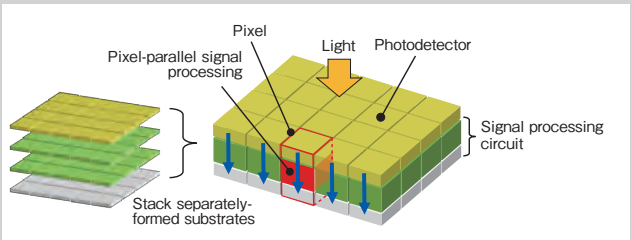


Figure 6-1. 3D integrated imaging device

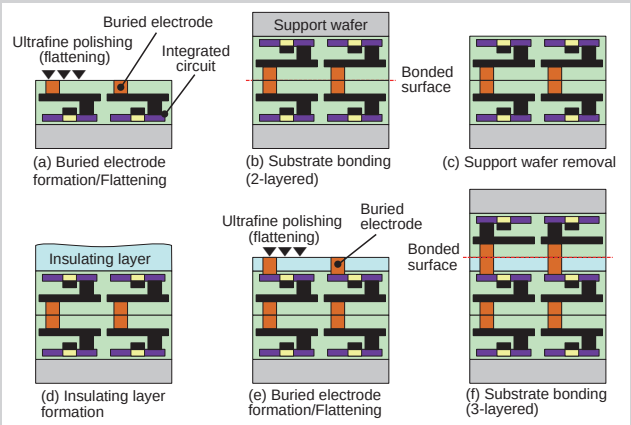


Figure 6-2. Multilayering process

In FY 2018, we developed a multilayering process to realize a structure of three or more layers[1] with the goal of achieving a highly integrated device (Figure 6-2). In addition to a technology for forming a two-layered structure (Figure 6-2 (a) to (c)), we newly developed a technology for forming an insulating layer and a buried electrode on the back of an integrated circuit and a technology for flattening the backside (Figure 6-2 (d) to (e)). This led to successful prototyping of a device with a three-layered structure. We demonstrated that a structure with more layers could be achieved by repeating this process and that transforming a conventional signal processing circuit with a planar formation into a three-dimensional one can reduce the pixel size and increase the degree of integration of a signal processing circuit.

This research was conducted in cooperation with the University of Tokyo.

RGB-stack-type image sensors

We are conducting research on RGB-stack-type image sensors with the goal of realizing a single-chip color camera that is small, lightweight and highly mobile. These sensors consist of alternating layers of organic photoconductive films (organic films) sensitive to each of red (R), green (G) and blue (B) and thin-film-transistor (TFT) arrays for reading the signals generated from each organic film. In FY 2018, we worked to miniaturize TFT arrays to achieve a high-definition device.

The conventional TFT array used indium–gallium–zinc oxide (IGZO) for its semiconductor and had an etch-stop structure with a layer for protecting the semiconductor during the processing of the source and drain (S/D) electrodes (Figure 6-3 (a)). This structure, however, needed to be simplified to further miniaturize the TFT array. We therefore employed a back-channel-etch structure (Figure 6-3 (b)), which can eliminate the barrier layer and shorten the TFT channel length compared with that for the etch-stop structure. Since the S/D electrodes need to be processed directly on the semiconductor for this structure, we used indium–tin–zinc oxide (ITZO) semiconductor material, which has tolerance against electrode formation process. By introducing the above fabrication method and material, we developed a prototype TFT with a channel length of 2 µm, shortened from conventional 6 µm, and demonstrated that it achieved an ON-OFF ratio of 10^8 or more, which is sufficient for a signal readout TFT for image sensors. Using this TFT, we successfully miniaturized the pixel pitch of the signal readout TFT array from 50 µm to 20 µm (Figure 6-4)[2], showing the feasibility of a high-definition RGB-stack-type image sensor.

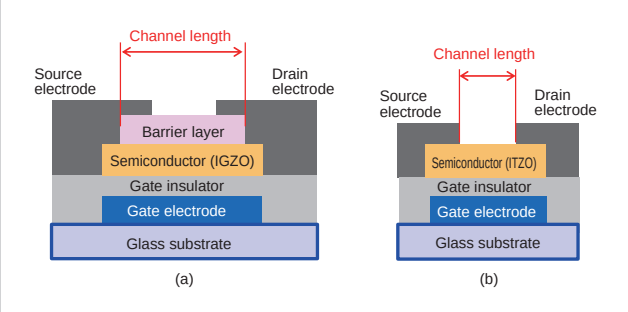


Figure 6-3. (a) Etch-stop structure TFT and (b) back-channel-etch structure TFT

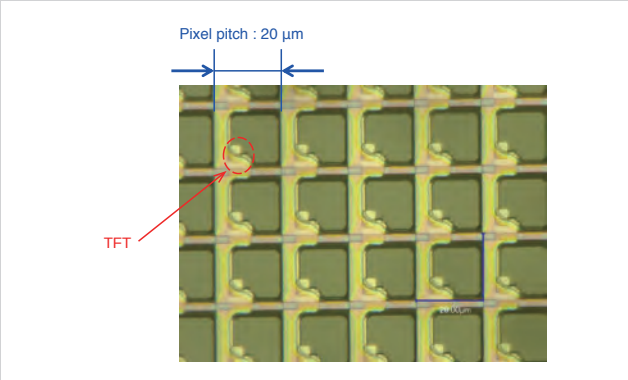
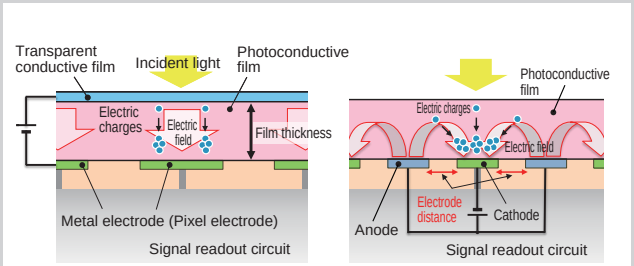


Figure 6-4. Microscope image of prototype TFT array

Planer-type photoconductive film structure for solid-state image sensor with ultra-high sensitivity

In 2.2, we described our research on an 8K solid-state image sensor overlaid with a multiplier film for 8K cameras with high sensitivity. This device applies a voltage in the film thickness direction to form electric fields necessary for multiplying electric charges using avalanche multiplication phenomenon in the film (Figure 6-5 (a)) and thus aims to achieve an electric charge multiplication factor of about 10 times. While the electric field in the film needs to be further strengthened to increase the multiplication factor for future devices with ultra-high sensitivity, there is a limit to the voltage that can be applied to a film overlaid on a solid-state image sensor, making it necessary to reduce the thickness of the multiplier film in the case of the conventional structure. However, the thinner the film is, the less light can be absorbed, deteriorating the use efficiency of light. To solve this problem, we began researching a new device structure that can strengthen the inner electric field without depending on the film thickness (Figure 6-5 (b)). This structure can obtain an electric field which is laterally strong in the film by forming an electrode that applies a voltage (anode) in the same plane as a pixel electrode (cathode). This should make it possible to achieve a high electric charge multiplication factor without deteriorating light absorption characteristics while maintaining the thickness of a photoconductive film.



(a) Conventional electrode structure (b) New electrode structure (planer-type)

Figure 6-5. Structure of image sensor overlaid with multiplier film

In FY 2018, we estimated the electric field intensity in the film necessary for achieving an electric charge multiplication factor of about 100 times on the basis of results of preliminary experiments and investigated an electrode structure for obtaining a desirable electric field intensity through simulations using electric field analysis. Figure 6-6 shows an effective electrode structure that we obtained from the analysis. Since the size of a single cathode-anode pair is expected to be very small, we used multiple cathode-anode pairs to constitute one pixel and arranged many sets of a cathode surrounded with an anode with a minute distance of submicron (0.1 μm or less) in between. We found that a strong electric field (10^8 V/m) for achieving a multiplication factor of about 100 times can be formed near the cathodes uniformly by optimizing the electrode size using this electrode structure.

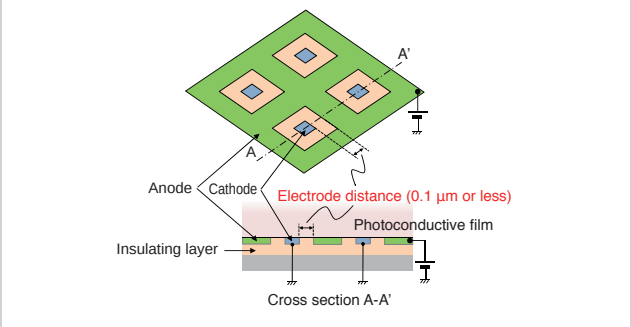


Figure 6-6. Electrode structure used for analysis

[References]
(1) M. Goto, Y. Honda, T. Watabe, K. Hagiwara, M. Nanba, Y. Iguchi, T. Saraya, M. Kobayashi, E. Higurashi, H. Toshiyoshi, T. Hiramoto: "Design of Three-Layered Pixel-Parallel Image Sensors," The 79th JSAP Autumn Meeting, 19p-432-7 (2018) (in Japanese)
(2) Y. Hori, T. Takagi, T. Sakai, M. Nakata, H. Sato, H. Ohtake and S. Aihara: "Miniaturization of TFT readout circuit for organic image sensor," ITE Annual Convention, 13C-1 (2018) (in Japanese)

6.2 Recording technologies

Multi-level holographic memory

Archiving storage system for 8K video signal, that has a very high data-transfer-rate and large capacity, are required. We have been researching on holographic memory using multilevel recording to meet these requirements. In FY 2018, we studied a crosstalk reduction method that is effective for amplitude multilevel recording using four-level modulation and a demodulation technology for reproduced data based on machine learning.

Two-dimensional data page in which symbol pixels of four type luminance are arranged are recorded and reproduced by laser in amplitude four-level holographic memory. Amplitude four-level recording data in holographic memory are recorded and reproduced by laser irradiation. If the luminance value of symbol pixels while data reproduction is different from the value at the time of recording due to noise, an error occurs with reproduced information. The major cause of noise is leaked light between symbol pixels (crosstalk). To reduce this crosstalk, we made symbol pixels smaller and inserted black symbol pixels between neighboring symbol pixels (to reduce the apparent fill factor) (Figure 6-7). The results of recording and reproduction experiments using data pages created by the proposed method demonstrated that the method can reduce the error rate of reproduction signals to a completely correctable level[(1)].

For a demodulation technology for reproduced data, we previously developed a demodulation method using convolutional neural networks and demonstrated its effectiveness for two-level recording. In FY 2018, we conducted numerical simulations to study the application of this method to amplitude four-level recording. The results demonstrated that demodulating 3×3 symbol pixels constituting a modulation block by using convolutional neural networks can reduce demodulation errors to 1/4 those of the conventional hard decision method, which makes 0-1 judgment with a single threshold value. We also increased the size of a modulation block to 5×5 symbol pixels to increase redundancy considering the influence of noise from the surroundings. We confirmed that this can further reduce demodulation errors to about 1/10 those of the hard decision method[(2)].

Additionally, we conducted fundamental experiments on amplitude multilevel recording with multiple exposure by using a holographic memory prototype drive that we developed in FY

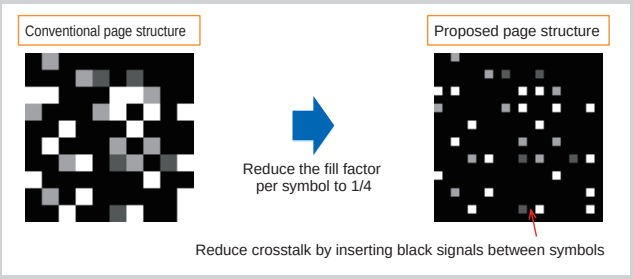


Figure 6-7. Page structure for reducing inter-symbol pixel crosstalk

2015.

Part of this research was conducted in cooperation with Hitachi-LG Data Storage, Inc.

Magnetic nanowire memory utilizing current-driven magnetic nano-domains

With the goal of realizing a high-speed magnetic recording device with no moving parts and a high reliability, we are conducting R&D on a magnetic recording device that utilizes the high-speed-motion characteristics of nanosized magnetic domains formed in magnetic nanowires. In FY 2018, we developed a process technology for the integration of a recording head on a magnetic nanowire medium to form magnetic domains accurately. We also studied a new recording condition that can reduce currents necessary for magnetic domain formation.

The integration of a recording head on a magnetic nanowire medium requires a five-layer structure which consists of a marker layer for position alignment ([1]), magnetic nanowire layer ([2]), insulator interlayer ([3]), recording head layer ([4]) and a top electrode layer ([5]) on the substrate. We employed a method that forms patterns in each layer without misalignment and stacks the layers by depositing thin films and lifting unnecessary parts/constituents off repeatedly for making prototype device. We used a laser lithography system, which can uniformly expose a large area, for [1], [3] and [5], which have relatively large structures, while we used an electron beam lithography system for [2] and [4], which have very minute patterns. Since each layer is formed by using both systems alternately, we used a unique alignment marker to suppress the misalignment of the magnetic nanowire and recording device to 40 nm or less.

Using the Landau–Lifshitz–Gilbert (LLG) equation, which describes magnetization dynamics and damping in general magnetic materials, we analyzed the process of magnetic domain formation when currents applied to the recording head. The results showed that magnetic domain is not stable due to large polarity fluctuations at the both edge of generated magnetic domains if magnetic domains are recorded by a magnetic field generated by applying a current to a single pole wire as a recording head. To address this problem, we devised a recording method that utilizes sudden magnetic field changes

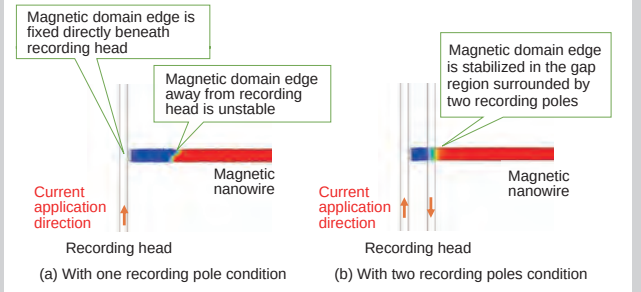


Figure 6-8. Difference in magnetic domains formed using recording head with one or two recording poles

generated by applying currents to two recording poles arranged in parallel in opposite directions (Figure 6-8)[3]. We found that this method could achieve a high-speed and stable magnetic domain recording and halve the current density per recording head.

Creation of spin-orbit-torque magnetic memory using topological insulator

We conducted research on the application of topological insulators to magnetic nanowire memory in cooperation with Tokyo Institute of Technology and the University of Tokyo as a commissioned project from Japan Science and Technology Agency for a strategic basic research programs titled "Creation of Core Technology based on the Topological Materials Science for Innovative Devices." A topological insulator is a new material that can produce strong spin torque (rotating power for local magnetic moments) using currents with aligned spins flowing on the surface, though it is an insulator in its bulk. It is expected that connecting bismuth-antimonide (BiSb), which is a topological insulator having a special crystal orientation, with a magnetic nanowire can largely reduce electric power necessary for the magnetic domain driving in magnetic nanowires to about 1/100. In FY 2018, we investigated a method for epitaxial growth between this bismuth-antimonide and magnetic nanowire materials in good crystallite condition.

[References]

(1) T. Muroi, Y. Katano, N. Kinoshita and N. Ishii: "Superimposed Spatial Guard Interval on Data Page for Reducing Inter-Symbol Interference in Amplitude Multi-Level Recording Holographic Memory," Tech. Dig. ISOM '18, Mo-C-03, 2018, pp.13-14 (2018)

(2) Y. Katano, T. Muroi, N. Kinoshita and N. Ishii: "Demodulation of Multi-Level Data using Convolutional Neural Network in Holographic Data Storage," 2018 International Conference on Digital Image Computing: Techniques and Applications (DICTA), IEEE Xplore, pp.728-729 (2018)

(3) M. Kawana et al.: "Micromagnetics simulation of magnetic domain formation in magnetic nanowire in various recording element shapes," 42nd Annual conference on magnetics in Japan, 11pPS-30, p.58 (2018) (in Japanese)

6.3 Display technologies

Flexible OLED displays with longer lifetime and higher color purity

Organic light-emitting diode (OLED) devices use active materials such as alkali metals for their electron injection layer. Since these materials are sensitive to moisture and oxygen, the devices deteriorate over time when used on a film substrate in air. This poses the greatest challenge in putting a flexible OLED display to practical use. To address this issue, we are researching and developing an OLED that does not use alkali metals and can better withstand oxygen and moisture, called an inverted OLED. In FY 2018, we worked to identify the principle of materials and formation process for an inverted OLED with higher performance.

To achieve an inverted OLED with a lower voltage, higher efficiency and a longer lifetime, the selection of electron

injection materials is significantly important. The results of consideration of various materials demonstrated that a special organic material containing boron is suitable for improving the electron injection performance and lifetime because of factors such as electron orbit energy. We also found that an inverted OLED with a lower voltage and longer lifetime can be achieved when using solution deposition for the formation process of an electron injection layer, rather than vacuum deposition. We examined the reason for this from the results of elementary analysis and found that a small amount of zinc oxide in the lower layer dissolved during the solution deposition process of the electron injection layer, changing the energy levels of impurities formed in the electron injection layer and thus contributing largely to the improvement of device performance[1]. On the basis of these results, we demonstrated that an inverted OLED with a solution-processed electron injection layer causes no luminescent degradation even when using a film with low gas-barrier properties, achieves a lifetime in excess of 3,000 hours for red and 1,500 hours for green even when it is continuously on, and shows much higher atmospheric stability than usual OLEDs, which cause significant luminance degradation[2].

To reproduce a wide color range of SHV, it is necessary to develop a power-saving and long-lasting OLED with a high color purity. We previously demonstrated that an OLED with a high color purity can be achieved by using a platinum complex having a rigid network molecular structure as the luminescent emitter. A platinum complex is capable of high-color-purity luminescence when used as a single material, but the color purity declines when it is mixed with a host material, which transmits electric energy in the light-emitting layer. To address this problem, in FY 2018, we systematically investigated what kind of host materials are suited for a display with high color purity. The results showed that color purity is deteriorated by unnecessary luminescence generated on the long wavelength side when a part called triazine contained in a host material and a platinum complex, which is a luminescent emitter, approach and interact with each other[3]. On the basis of this knowledge, we plan to design a new host material and utilize it for the development of a wide-color-gamut OLED display.

Technologies for increasing image quality and lowering driving power consumption of large OLED displays

We progressed with our R&D on thin-film transistors (TFTs), which are driving elements of pixels, to increase the image quality and lower the power consumption of sheet-type displays. In FY 2018, we developed a high-mobility TFT that uses zinc oxynitride (ZnON) as the semiconductor material. We found that doping ZnON with a minute amount of impurities improves the electrical characteristics of ZnON-TFT significantly. More specifically, we demonstrated that the switching characteristics of TFTs are improved by adding an element having a large binding energy with nitrogen (such as Ta) and that the mobility is improved by adding an element that contributes to high conductivity (such as In). A ZnON-TFT doped with In achieved a high mobility of 59 cm^2/Vs at the maximum.

As a video signal processing technology for increasing image quality, we studied a luminance control method for HDR video. We previously devised a method for controlling the driving power while maintaining the tone representation of light and dark regions when displaying HDR video with a high average luminance. However, since this method controls only light regions, image degradation occurs in intermediate gradations, which has posed a challenge. To address this problem, we devised a method for preventing image degradation by controlling luminance using an optimum signal processing for HDR video and demonstrated its effectiveness with evaluation images[4].

Solution-processed devices for large flexible displays

With the goal of realizing a large flexible display that is thin, lightweight and rollable, we are conducting R&D on oxide TFTs that can be fabricated by solution process without using a large vacuum chamber and on light-emitting diodes using semiconductor nanocrystals (quantum dots (QDs)) called QD-LEDs.

Solution-processed oxide TFTs had the problem of low mobility due to many impurities and defects contained in the semiconductor film formed by the solution method. We therefore improved the film quality of semiconductor by using hydrogen injection and oxidation (HIO) process as well as doping the oxide semiconductor (In-Zn-O) with fluorine. This enabled a TFT fabricated by a low-temperature (300°C) process that can be applied to film substrates to achieve a high mobility (12.8 cm^2/Vs) comparable to that of a TFT fabricated by the conventional vacuum process[5].

QD-LEDs can control the wavelength and full width at half maximum of the emission spectrum by grain size control of QDs. In FY 2018, we prototyped a QD-LED that uses an indium phosphide material (ZnInGaP/ZnS) as low toxic QDs. We introduced a hole transport layer that can inject holes into the light-emitting layer efficiently while blocking electrons to a QD-LED using ZnInGaP/ZnS QDs that emit green light. This achieved an external quantum efficiency of 3.4% (Figure 6-9)[6]. The research on the QD-LED using ZnInGaP/ZnS was conducted in cooperation with ULVAC, Inc.

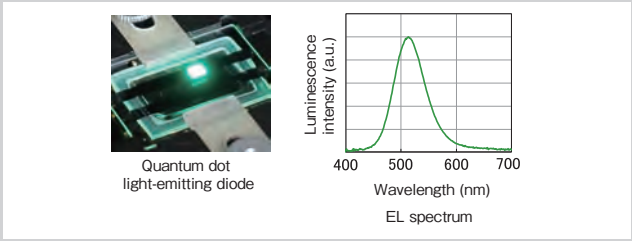


Figure 6-9. ZnInGaP/ZnS QD-LED and EL spectrum

[References]
(1) H. Fukagawa, T. Sasaki, T. Tsuzuki, Y. Nakajima, T. Takei, G. Motomura, M. Hasegawa, K.Morii, and T. Shimizu: "Long-Lived Flexible Displays Employing Efficient and Stable Inverted Organic Light-Emitting Diodes," Advanced Materials, DOI: 10.1002/adma.201706768 (2018)
(2) T. Sasaki, H. Fukagawa, K. Kuwada, M. Hasegawa, K. Morii, T. Shimizu: "Demonstration of Long-Term Stable Emission from Inverted OLED with Imperfect Encapsulation," SID Digest, pp. 811-814 (2018)
(3) Y. Iwasaki, H. Fukagawa, T. Shimizu: "Host Materials Suitable for Organic Light-Emitting Diodes with High Color Purity," ITE Winter Annual Convention, 21D-2(2018).
(4) T. Okada, T. Usui and Y. Fujisaki: "Picture-level Control for Managing HDR Images and Power Consumption for Large OLED Displays," IDW'18 VHF1-2, pp.937-940 (2018)
(5) M. Miyakawa, M. Nakata, H. Tsuji and Y. Fujisaki: "High-performance solution-processed thin-film transistors using fluorine-doped aqueous metal oxides," The 25th International Workshop on Active-Matrix Flatpanel Displays and Devices (AM-FPD'18), 3-2 (2018)
(6) K. Ogura, G. Motomura, T. Tsuzuki, Y. Fujisaki, J. Nagakubo, M. Hirakawa and T. Nishihashi: "Efficient Green Light-Emitting Diodes Using ZnInGaP/ZnS Nanocrystals as Cadmium-Free Quantum Dots," Materials Research Society (MRS) Fall Meeting, EP04.10.17 (2018)

7 Research-related work

7.1 Joint activities with other organizations

Participation in standardization organizations

NHK STRL is participating in standardization activities at international and domestic standardization organizations and projects, mainly related to broadcasting. In particular, we are contributing to the development of technical standards that incorporate our research results.

We have made a number of contributions to the ITU Radiocommunication Sector (ITU-R). As part of Study Group 4 (SG 4) for satellite services, we proposed contributions on a phased array antenna for 21-GHz-band satellite broadcasting and on evaluation of wide-band transmission characteristics, and our proposals were incorporated into ITU-R Reports. At Study Group 5 (SG 5) for terrestrial services, we proposed frequency sharing conditions for high altitude platform stations (HAPS), which are an item on the agenda for World Radiocommunication Conference 2019 (WRC-19), and 6-GHz-band FPUs used in Japan, and our proposal was incorporated into a draft Conference Preparatory Meeting (CPM) report and an ITU-R Report. As part of Study Group 6 (SG 6) for broadcasting services, we submitted a number of contributions on various subjects including advanced technologies for terrestrial TV broadcasting, an IP interface for program production, an integrated platform for broadcasting and telecommunications networks, an integrated broadcast-broadband system utilizing second screens, a transmission method for audio metadata, an audio coding scheme for 4K/8K program contributions, a VR/360° video system, operational practice for high-dynamic-range TV program production, and the application of Artificial Intelligence (AI) for broadcasting, and our proposals were incorporated into ITU-R Recommendations, Reports and research subjects.

At the Moving Picture Experts Group (MPEG), which is a working group of a joint committee of the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC), we proposed elemental technologies for a next-generation video coding scheme, for which standardization efforts began, and our proposal was adopted into a Working Draft. We also participated in standardization activities for a new 3D video coding and submitted the result of a comparison of the data compression amount between the case of applying the existing coding scheme to elemental images and the case of applying it to multi-view images.

At the Society of Motion Picture and Television Engineers (SMPTE), we contributed to the update of draft standards for a method for multiplexing serialized audio metadata used for object-based audio with an AES3 digital audio signal interface.

As part of the Asia-Pacific Broadcasting Union (ABU), we presented our R&D efforts for the advanced satellite broadcasting for 4K/8K and 8K program production during the Soccer World Cup and the Olympic Games at the technical committee meeting held in Ashkhabad, Turkmenistan. We also contributed as a project leader to the sharing of the R&D status of technologies such as next-generation terrestrial broadcasting, hybrid broadcasting, metadata for program production and user-friendly broadcasting. NHK's paper on program production utilizing AI was awarded the ABU Technical Review Best Article Award.

In addition to the above activities, we engaged in a number of standardization activities, including the European Broadcasting Union (EBU), the Advanced Television Systems Committee (ATSC), the Audio Engineering Society (AES), the 3rd Generation Partnership Project (3GPP), which discusses standards for next-generation mobile communications, the Advanced Media Workflow Association Networked Media Incubator (AMWA NMI), which standardizes connection management methods of IP program production systems, the World Wide Web Consortium (W3C), which develops Web standards, the Association of Radio Industries and Businesses (ARIB), the Japan Electronics and Information Technology Association (JEITA) and the Telecommunication Technology Committee (TTC) of Japan.

Leadership activities at major standardization organizations

International Telecommunication Union (ITU)

Committee name	Leadership role
International Telecommunication Union Radiocommunication Sector (ITU-R)	
Study Group 6 (SG 6, Broadcasting services)	Chair
Radio-wave propagation committee	Expert member
Satellite and scientific services committee	Expert member
Broadcast services committee	Expert member
Terrestrial wireless communications committee	Expert member

Asia-Pacific Broadcasting Union (ABU)

Committee name	Leadership role
Technical committee	Chair (to September 2018) Vice-Chair (from October 2018)

Information and Communications Council of the Ministry of Internal Affairs and Communications

Committee name	Leadership role
Information and communications technology subcommittee	
ITU section	
Spectrum management and planning committee	Expert member

Association of Radio Industries and Businesses (ARIB)

Committee name	Leadership role
Technical committee	
Broadcasting international standardization working group	Chair
Digital broadcast systems development section	Committee chair
Multiplexing working group	Manager
Download methods TG	Leader
Data MMT transmission JTG	Leader
Video coding working group	Manager
Data coding working group	Manager
Copyright protection working group	Manager
Digital receivers working group	Manager

Digital satellite broadcasting working group	Manager
Advanced satellite broadcasting demonstration experiments TG	Leader
Mobile multimedia broadcasting systems working group	Manager
Digital terrestrial broadcasting channel coding working group	Manager
Studio facilities development section	
Studio sound working group	Manager
Next-generation sound services study WG	Leader
Sound quality evaluation methods working group	Manager
Broadcast contribution file format study working group	
Data content exchange methods JTG	Leader
Intra-device interface working group	Manager
Contribution transmission development section	
Terrestrial wireless contribution transmission working group	Manager
Millimeter-wave contribution transmission TG	Leader
New frequency FPU study TG	Leader
Microwave UHDTV-FPU study TG	Leader
Promotion strategy committee	
Digital broadcasting promotion sub-committee	
Digital broadcasting experts group (DIBEG)	
International technical assistance task force	Manager
Next-generation broadcast study task force assisting Japan-Brazil joint work section, etc.	Manager
Standards Assembly	Acting committee chair

Telecommunication Technology Committee (TTC)

Committee name	Leadership role
Multimedia application working group	
IPTV-SWG	Leader

Collaboration with overseas research facilities

We participated in sub-groups (Renderer, Augmented Reality (AR), Artificial Intelligence (AI)/Machine Learning (ML)) of the Broadcast Technology Futures (BTF) group under the technical committee of the European Broadcasting Union (EBU) and investigated efforts of BTF members.

We participated in the activities of EBU's Media Cloud and Microservice Architecture (MCMA) project for the standardization of a future content production infrastructure and discussed a change from the SOAP specification established by the Framework for Interoperable Media Services (FIMS) to the REST specification, which assumes linkage among multiple clouds. We also conducted demonstrations of released MCMA libraries and sample applications at the International Broadcasting Convention (IBC) 2018.

Collaborative research and cooperating institutes

In FY2018, we conducted a total of 21 collaborative research projects and 31 cooperative research projects on topics ranging from system development to materials and basic research.

We collaborated with graduate schools in seven universities (Chiba University, the University of Electro-Communications, Tokyo Institute of Technology, Tokyo Denki University, Tokyo University of Science, Tohoku University and Waseda University) on education and research through activities such as sending part-time lecturers and accepting trainees.

Visiting researchers and trainees and dispatch of STRL staff overseas

As part of a program for hosting young researchers from ABU (Asia-Pacific Broadcasting Union) member institutes, we hosted one researcher from Iran to honor our commitment to information exchange with other countries and the mutual development of broadcasting technologies. We also hosted three visiting researchers (one from the Brazilian TV broadcaster TV Globo and two from domestic broadcasters). We also took on one post-doctoral research project (Table 7-1).

We provided guidance to a total of nine trainees from four universities (Waseda University, Tokyo University of Science, the University of Electro-Communications and the Tokyo University of Agriculture and Technology) in their work towards their Bachelor's and Master's degrees.

Five STRL researchers were dispatched to research institutions in the United States, Belgium and Australia (Table 7-2).

Table 7-1. Visiting researchers

Type	Term	Research topic
ABU visiting researcher	2018/12/6 to 2019/3/29	Content production assistance system utilizing AI
Visiting researcher	From 2019/1/15	R&D on a next-generation terrestrial broadcasting system
Visiting researcher	2018/5/14 to 2019/3/31	Integrated broadcast-broadband technology utilizing area information
Visiting researcher	From 2018/9/3	Services utilizing Hybridcast Connect
Post-doctoral student	2017/9/1 to 2019/3/31	Development of high-speed and high-precision phase detection technology for hologram reproduction (2017/9/1 to 2018/8/31) R&D and application of high-speed and high-precision phase detection technology for hologram reproduction (2018/9/1 to 2019/3/31)

Table 7-2. Dispatch of NHK STRL researchers overseas

Location	Term	Research topic
IMEC International, Belgium	2017/9/25 to 2018/9/20	High-pixel-density and high-functionality imaging sensor using cutting-edge semiconductor manufacturing technology
The University of Melbourne, Australia	2017/10/24 to 2018/4/19	Cutting-edge research on natural language processing
MIT Media Lab, USA	2018/1/21 to 2018/7/31	Cooperative interactive content production by multiple people using 8K displays
MIT Media Lab, USA	From 2018/9/25	Interactive content production of 3D audio using sensor networks
Stanford University, USA	From 2018/10/21	Investigative research on flexible wearable electronics that harmonize with broadcasting

Commissioned research

We are participating in research and development projects with national and other public facilities in order to make our research on broadcast technology more efficient and effective. In FY 2018, we took on three projects commissioned from the government and public institutions (the Ministry of Internal Affairs and Communications, NICT* and JST**).

- Research and Development for Advanced Digital Terrestrial Television Broadcasting System
- Research and Development on Deep Learning Technology for Advanced Multi-language Speech Translation
- Creation of Spin-Trajectory-Torque Magnetic Memory Using Topological Surface State

* : National Institute of Information and Communications Technology
** : Japan Science and Technology Agency

Committee members, research advisers, guest researchers

We held two meetings of the broadcast technology research committee and received input on research activities from academic and professional committee members. We held 16 sessions to obtain advice and opinions from research advisers. We also invited researchers from other organizations to promote five research topics with us.

Broadcast Technology Research Committee Members

March 2019
** Committee chair, * Committee vice-chair

Name	Affiliation
Kiyoharu Aizawa**	Professor, University of Tokyo
Toshiaki Kawai	Senior Managing Representative Director, Tokyo Broadcasting System Television Inc.
Tadahisa Kawaguchi	Executive Director, TV Asahi Corporation
Katsuhiko Kawazoe	Senior Vice President and Head of Research and Development Planning, Nippon Telegraph and Telephone Corporation
Yasuhiro Koike	Professor, Keio University
Tetsunori Kobayashi	Professor, Waseda University
Yoichi Suzuki	Professor, Tohoku University
Junichi Takada	Professor, Tokyo Institute of Technology
Atsushi Takahara	Professor, Kyushu University
Fumihiko Tomita*	Vice President, National Institute of Information and Communications Technology (NICT)
Yasuyuki Nakajima	President/CEO, KDDI Research, Inc.
Ichiro Matsuda	Professor, Tokyo University of Science
Yukinobu Miki	Senior Vice-President, National Institute of Advanced Industrial Science and Technology (AIST)
Masayuki Murata	Professor, Osaka University
Satoru Yanagishima	Director, Ministry of Internal Affairs and Communications

Research Advisers

March 2019

Name	Affiliation
Makoto Ando	Director, National Institute of Technology
Makoto Itami	Professor, Tokyo University of Science
Susumu Itoh	Professor, Tokyo University of Science
Tohru Ifukube	Emeritus Professor, University of Tokyo
Hideki Imai	Emeritus Professor, University of Tokyo
Tatsuo Uchida	Emeritus Professor, Tohoku University
Juro Ohga	Emeritus Professor, Shibaura Institute of Technology
Tomoaki Ohtsuki	Professor, Keio University
Jiro Katto	Professor, Waseda University
Yoshimasa Kawata	Professor, Shizuoka University
Satoshi Shioiri	Director, Research Institute of Electrical Communication, Tohoku University
Takao Someya	Professor, University of Tokyo
Fumio Takahata	Professor, Waseda University
Katsumi Tokumaru	Emeritus Professor, University of Tsukuba
Mitsutoshi Hatori	Emeritus Professor, University of Tokyo
Takayuki Hamamoto	Professor, Tokyo University of Science
Hiroshi Harashima	Emeritus Professor, University of Tokyo
Takehiko Bando	Emeritus Professor, Niigata University
Takefumi Hiraguri	Professor, Nippon Institute of Technology
Timothy John Baldwin	Professor, University of Melbourne
Masato Miyoshi	Professor, Kanazawa University

Guest Researchers

March 2019

Name	Affiliation
Masayuki Ikebe	Professor, Hokkaido University
Mamoru Iwabuchi	Professor, Waseda University
Tokio Nakada	Visiting Professor, Tokyo University of Science
Toshiaki Fujii	Professor, Nagoya University
Tetsuya Watanabe	Associate Professor, Niigata University

7.2 Publication of research results

STRL Open House

The NHK STRL Open House 2018 was held over six days from May 22 under the theme of "Connecting to you smartly and realistically." It featured a presentation of NHK STRL 3-Year R&D Plan (FY 2018-2020), which was announced in April, 24 exhibits on our latest research results and five interactive exhibits. The event was attended by 20,500 visitors. At the entrance hall, we presented future diverse viewing styles including 3D TV and our research concept to realize them, "Diverse Vision."

We also exhibited our latest research outcomes such as program production assistance technologies using AI, broadcasting service technologies utilizing the internet, full-featured 8K technologies and universal service technologies.

Two keynote speeches on NHK STRL 3-Year R&D Plan and the realization of IoA (Internet of Abilities) and lectures on our three R&D themes of "Reality Imaging," "Connected Media" and "Smart Production" were delivered in the auditorium.

Schedule
- May 22 (Tuesday) Opening ceremony
- May 23 (Wednesday) Open to invitees
- May 24 - June 27 (Thursday to Sunday) Open to the public

エントランスの様子


将来の多様な視聴スタイル

Keynote speeches

Title	Speaker
NHK STRL 3-years R&D Plan (FY2018-2020)	Toru Kuroda, Director, Science & Technology Research Laboratories, NHK
Challenges in Achieving IoA (Internet of Abilities) and the Future of Broadcasting	Junichi Rekimoto, Professor, The University of Tokyo/Deputy Director, Sony Computer Science Laboratories, Inc.

Lectures

Title	Speaker
Reality Imaging: Toward the Ultimate Tele-Experience	Shinichi Sakaida, Head of Advanced Television Systems Research Division
Connected Media: Changes in Broadcasting with Connectivity	Toshio Nakagawa, Head of Internet Service Systems Research Division
AI-Driven Smart Production	Masakazu Iwaki, Head of Human Interface Research Division

Research exhibits

E1	NHK STRL 3-Year R&D Plan (FY 2018-2020)	12	OB Van for Full-Featured 8K Program Production
E2	Diverse Viewing Styles	13	8K 4x-Speed Slow-Motion System
1	A Real-Time Transcription System for Program Production	14	8K/120 Hz Video Codec
2	Automatic Video Summarization System	15	Technology for Realizing High-Sensitivity 8K Solid-State Image Sensor
3	Automatic Colorization Technology for Monochrome Video Using Artificial Intelligence	16	Optical Phased Array Using Electro-Optic Polymers
4	Text Big Data Analysis	17	Next Generation Audio Services using Object-based Sound System
5	Automated Sports Commentaries and Audio Descriptions	18	Sports 4D Motion
6	Sign Language Synthesis Systems for Sports Information	19	Mobile Transmission Technology for 8K Program Contribution
7	Media-Unifying Platform	20	8K Transmission Technology for IP-Based Program Production
8	"Unified Experience" Bridging TV, Internet, and Daily Life	21	Satellite Broadcasting in the 21 GHz Band
9	Home Companion Robot to Enjoy Watching TV Together	22	Advanced Digital Terrestrial TV Broadcasting System
10	Sheet-Type 8K OLED Display	23	3D Video System with Resolution of 300,000 Pixels
11	Highly Air-Stable Inverted Organic Light-emitting Diode	24	Sword Tracer

■ **Interactive exhibits**

1.	8K Living Room	4.	Free-Viewpoint Service Using MMT
2.	Home Reproduction System of 22.2 Multichannel Sound	5.	VR Using 8K Display
3.	Let's Win a Game Using Images by Hologram		

Overseas exhibitions

The world's largest broadcast equipment exhibition, the National Association of Broadcasters (NAB) Show 2018, was held in April. We screened 8K content in an 8K theater, demonstrated three different styles of 8K viewing at home, and exhibited our latest research outcomes mainly on 8K-related technologies. The show attracted about 93,000 registrants from around the world.

The International Broadcasting Convention (IBC) 2018, the largest broadcast equipment exhibition in Europe, was held in September. We exhibited an 8K living room theater consisting of an 88-inch 8K OLED display and a 22.2 ch sound system, where we screened content. We also exhibited our latest research results mainly on 8K-related technologies. The convention drew about 56,000 visitors from around the world.

■ **Two overseas exhibitions**

Event name	Dates	Exhibits
NAB Show 2018 (Las Vegas, USA)	4/9 to 4/12	8K theater, Demonstration of three 8K home viewing styles, 4× slow-motion system using 8K/240-Hz single-sensor camera, 8K IP lightweight compression transmission equipment, High-resolution VR using 8K display, Multiview demonstration using MMT delivery technology
IBC 2018 (Amsterdam, Netherlands)	9/14 to 9/18	8K living room theater with 88-inch 8K OLED display and 22.2 ch sound system, 4× slow-motion system using 8K/240-Hz three-chip camera, 8K/240-Hz production system, Hybridcast Connect, Automatic colorization technology for monochrome video using AI

Exhibitions in Japan

Throughout the year, NHK broadcasting stations all over Japan hosted events and exhibitions of broadcast technologies resulting from our R&D. In particular, at events to publicize the advanced satellite broadcasting for 4K/8K, which was launched in December 2018, we exhibited an "8K living room theater," which embodied a future home viewing environment with ultra-high resolution 8K video and 22.2 multichannel sound, and presented immersive video content to many visitors.

■ **30 exhibitions in Japan**

Event name (Only major events)	Dates	Exhibits
Hiroshima Flower Festival (NHK Hiroshima Station)	5/3 to 5/5	Augmented TV, etc.
Okhotsk Mirai Festival (NHK Kitami Station)	7/14 to 7/15	Domo's Slapstick Race, etc.
4K/8K Super Hi-Vision Park	8/1 to 8/4	8K living room theater, etc.
ITE Annual Conference	8/29 to 8/31	8K content screening
N Spo! 2018	9/15 to 9/16	Sword tracer
Saga Saiko Festival 2018	10/20 to 10/21	Domo's Slapstick Race, etc.
ITE Special Exhibition	10/27 to 12/15	8K content screening, etc.
Seto City Digital Festival 2018	11/4	IP 3D, etc.
Inter BEE 2018	11/14 to 11/16	Colorization of monochrome video, etc.
4K/8K Super Hi-Vision Park	11/30 to 12/4	8K living room theater
NHK Science Stadium 2018	12/1 to 12/2	IP 3D, etc.

Academic conferences, etc.

We presented our research results at many conferences in Japan, such as the ITE and IEICE conferences, and had papers published in international publications such as Advanced Materials, IEEE Transactions, IEEE Access and Scientific Reports.

Academic journals in Japan	47 papers
Overseas journals	27 papers
Academic and research conferences in Japan	220 papers
Overseas/International conferences, etc.	195 papers
Contributions to general periodicals	48 articles
Lectures at other organizations	58 events
Total	595

Press releases

We issued nine press releases on our research results and other topics.

Dates	Press release content
2018/4/5	Announcement of the STRL Open House 2018
4/5	Development of automatic colorization technology for monochrome video
5/22	Development of 3D imaging system “Aktina Vision”
5/22	Development of an 8K Super Hi-Vision mobile relay FPU
5/22	Development of a 120-Hz full-featured 8K video encoder compliant with broadcasting standards
5/22	Development of an 8K high-speed camera and a slow-motion player
5/22	Development of “Hybridcast Connect Library” for easy linkage among TV, smartphones and IoT-enabled devices
2019/1/18	Announcement of the 73rd STRL Open House
3/29	The World’s First 8K Broadcasting System and Future Media Technologies Featured at NAB Show 2019

Visits, tours, and event news coverage

To promote R&D on 8K Super Hi-Vision, integral 3D television, internet utilization technology and smart production, we held tours for people working in a variety of fields including civil service, manufacturing, broadcasting, movies and academic research. We welcomed visitors from around the world, including officials of international broadcasting conference organizations, broadcasters from various countries and JICA trainees.

Inspections, tours	55 (18 from overseas) 974 visitors (145 from overseas)
News media	15 events

Bulletins

We published bulletins describing our research activities and achievements and special issues on topics such as such as Super Hi-Vision imaging and sound technologies, broadcast-broadband services and sports image representation technologies.

The Broadcast Technology journal, which is directed at overseas readers, featured in-depth articles about our latest research and trends such as the development of Super Hi-Vision FPUs, high-definition VR using 8K displays, and video quality improvement of OLED.

Domestic Publications

STRL Dayori (Japanese, monthly)	No.157 to No.168
NHK STRL R&D (Japanese, bimonthly)	No.169 to No.174
Annual Report (Japanese, annually)	FY2017 Edition

Publications for overseas readers

Broadcast Technology (English, quarterly)	No.72 to No.75
Annual Report (English, annually)	FY2017 Edition

NHK STRL R&D

STRL Dayori

Broadcast Technology

Website

NHK STRL website describes our laboratories and their research and posts reports and announcements on events such as the Open House and the organization’s journals. For the website for the Open House 2018, in particular, we implemented user-friendly page designs for smartphones and tablets as well as PCs and included URLs to relevant journals in the page of each exhibition item so that users can access detailed information easily.

Example of the exhibition item page for NHK STRL Open House 2018

7.3 Applications of research results

Cooperation with program producers

Equipment resulting from our R&D has been used in many programs. Our system that measures and displays the putting trajectory in golf and our system that generates multi-motion images of the athlete's motion in ski jumping were used in sports programs. Our 8K 2×-speed slow-motion system was used in the production of a figure skating program for the BS8K channel launched in December.

In addition, our system for colorizing past monochrome video using AI technology was utilized for the program production of NHK Special. We collaborated in the production of 46 programs in FY 2018.

Patents

We participated in the establishment of a new patent pool* for UHDTV digital broadcasting over CATV, for which licensing was started in July 2018. We also participate in patent pools for 2K/satellite UHDTV digital broadcasting and high-efficiency video coding standards. These pools especially promote the use of patents held by NHK to help with the promotion of broadcasting services. We are protecting the rights to our broadcasting and communications-related R&D as part of our intellectual property management efforts. We are also actively promoting contracts on transfers of patented NHK technologies by enhancing our Technology Catalogue, which summarizes NHK's transferrable technologies, and at events such as the STRL Open House 2018, CEATEC JAPAN 2018, Technical Show Yokohama 2019 and other events we held in cooperation with local governments and other organizations.

* Patent pool: A mechanism that bundles licenses of multiple patents required by standards under reasonable conditions

Patents and utility model applications submitted

	Type	New	Total at end of FY
Domestic	Patents	300	1,100
	Utility models	0	0
	Designs	2	2
Overseas	Patents	12	84
Total		314	1,186

Patents and utility models granted

	Type	New	Total at end of FY
Domestic	Patents	264	1,997
	Utility models	0	0
	Designs	0	0
Overseas	Patents	22	123
Total		286	2,120

Patents and utility models in use **(NHK Total)**

Type	New	Total at end of FY
Contracts	25	300
Licenses	49	493
Patents	27	248
Expertise	22	245

Technical cooperation **(NHK Total)**

Type	Total (including projects continued from previous year)
Technical cooperation projects	24 (2 from previous year)
Commissioned research projects	3 (1 from previous year)

Prizes and degrees

In FY 2018, NHK STRL researchers received 24 prizes, including the Maejima Award and the Meritorious Award on Radio. Two researchers obtained a doctoral degree in FY 2018, and at the end of FY 2018, 84 STRL members held doctoral degrees.

Award Winner	Award Name	Awarded by	In recognition of	Date
Go Ohtake	Maejima Award	Tsushinbunka Association	Development of a privacy protection system for safe and secure utilization of personal data	2018/4/10
Shoji Tanaka, Yoichi Suzuki, Akinori Hashimoto (Engineering Dept.)	The Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology, Prize for Science and Technology	Ministry of Education, Culture, Sports, Science and Technology	Development of a transmission scheme for ultra-high-definition television satellite broadcasting	2018/4/17
Atsuro Ichigaya, Shunsuke Iwamura, Shimpei Nemoto	ITU-AJ Award	The ITU Association of Japan	Contribution to the finalization of HDR extension and the supplementary document for ITU-T Recommendations H.265 and 264 in the joint collaborative team between ITU-T SG16/WP3/Q.6 and MPEG, Contribution to the revision of BT.1872 for the HD/UHD digital ENG	2018/5/17
Keiichi Kubota	Niwa-Takayanagi Award, Contribution Award	The Institute of Image Information and Television Engineers (ITE)	Contribution to the R&D, practical application and international promotion of Hi-Vision and Super Hi-Vision	2018/5/30
Hiroyuki Hamazumi	Niwa-Takayanagi Award, Achievement Award	The Institute of Image Information and Television Engineers (ITE)	Contribution to the R&D and practical application of OFDM digital radio microphone	2018/5/30
Chiaki Mori, Toshiyuki Nishiguchi (NHK Engineering System, Inc.), Kazuho Ono	Technology Promotion Award, Advanced Development Award (R&D Division)	The Institute of Image Information and Television Engineers (ITE)	3D reverberator	2018/5/30
Yutaro Katano	Suzuki Memorial Award	The Institute of Image Information and Television Engineers (ITE)	SN ratio improvement of holographic memory using roll-off filter, A study on efficient training methods for convolutional neural networks for holographic memory	2018/8/30
Yoshikuni Hirano, Yasushi Motoyama, Kenji Machida, Hiroshi Kikuchi, Katsu Tanaka (NHK Engineering System, Inc.)	Image Information Media Future Award, Frontier Award	The Institute of Image Information and Television Engineers (ITE)	Pioneering research on optical phased array using EO polymer	2018/5/30
Takuya Handa	Image Information Media Future Award, Frontier Award	The Institute of Image Information and Television Engineers (ITE)	Development of a haptic device for presenting the shape and hardness of objects in the air	2018/5/30
4K/8K Test Broadcast Transmission Equipment Development Group	Image Information Media Future Award, Next-Generation TV Technology Award	The Institute of Image Information and Television Engineers (ITE)	Development of a 4K/8K test broadcast system	2018/5/30
Hirokazu Kamoda, Kenji Murase, Yoshifumi Matsusaki, Fumito Ito, Tomofumi Koyama, Jun Tsumochi, Naohiko Iai, Koichiro Imamura, Hiroyuki Hamazumi, Takayuki Nakagawa (Engineering Dept.), Shinichi Suzuki (Engineering Dept.), Takashi Kumagai (Engineering Dept.), Naoto Kogo (NHK Nagoya station), Kazuhiko Shibuya (NHK Engineering System, Inc.)	Meritorious Award on Radio, Ministerial Commendation	Association of Radio Industries and Businesses (ARIB)	Development of an FPU for 4K/8K broadcast program contributions	2018/6/27
Tsubasa Sasaki	IEEJ Excellent Presentation Award	Institute of Electrical Engineers of Japan (IEEJ)	Presentation: “Inverted organic light-emitting diodes fabricated by low temperature process using inorganic-organic hybrid material”	2018/9/6
Kazuto Ogawa	IEICE Engineering Sciences Society Contribution Award (Committee management)	The Institute of Electronics, Information, and Communication Engineers (IEICE)	Contribution to the management of the IWSEC 2017 international conference	2018/9/12
Hideki Mitsumine	Compilation Achievement Award	The Institute of Image Information and Television Engineers (ITE)	Contribution to the compilation of submitted papers	2018/9/20
Simon Clippingdale	FIT Encouragement Award	The Institute of Electronics, Information, and Communication Engineers (IEICE), Information Processing Society of Japan	Presentation at FIT 2018: “Detection of weak expressions using Kinect during video viewing”	2018/9/21
Fumiya Yamagishi	Poster Award for Best Young Researchers	The Institute of Electronics, Information, and Communication Engineers (IEICE) Communications Society	A study on the boost ratio of pilot signals considering the distortion of a millimeter-wave-band power amplifier using the SC-FDE scheme	2018/9/28
Yuko Yamanouchi, Yoshihiko Kawai, Takahiro Mochizuki, Jun Goto, Shoei Sato, Atsushi Imai	ABU Technical Review Best Article Prize	ABU	Research on smart production	2018/10/1
Shuichi Aoki	International Standardization Encouragement Award (Commendation by Industrial Science and Technology Policy and Environment Bureau Chief)	Ministry of Economy, Trade and Industry	Contribution to international standardization activities	2018/10/2

Award Winner	Award Name	Awarded by	In recognition of	Date
Takehiro Sugimoto	International Standard Development Award	Information Processing Society of Japan/Information Technology Standards Commissions of Japan	Contribution to the issuance of ISO/IEC 14496-3:2009/Amd. 6/2017	2018/10/9
Masafumi Nagasaka, Susumu Nakazawa (Broadcasting Satellite System Corporation), Masaaki Kojima, Shoji Tanaka (Broadcasting Satellite System Corporation)	APMC 2018 Prize	The Institute of Electronics, Information, and Communication Engineers (IEICE)	Paper and presentation: “Prototype of 12/21GHz-band Dual-circularly Polarized Receiving Antenna for Satellite Broadcasting”	2018/11/9
Tetsuya Hayashida	The Telecommunications Industry Achievement Award	The Telecommunications Association (TTA)	-	2018/11/22
Yuki Honda, Masahide Goto, Toshihisa Watabe, Masakazu Nanba, Yoshinori Iguchi, Takuya Saraya (Univ. of Tokyo), Masaharu Kobayashi (Univ. of Tokyo), Eiji Higurashi (Univ. of Tokyo), Hiroshi Toshiyoshi (Univ. of Tokyo), Toshiro Hiramoto (Univ. of Tokyo)	9th Integrated MEMS Technology Research Workshop Best Poster Award	Japan Society of Applied Physics (JSAP), Study Group of the Integrated MEMS	Three-Layer Stacked Au/SiO_2 Hybrid Bonding with 6-µm-pitch Au Electrodes for 3D Structured Image Sensors	2019/3/10
Yuki Honda, Masahide Goto, Toshihisa Watabe, Masakazu Nanba, Yoshinori Iguchi, Takuya Saraya (Univ. of Tokyo), Masaharu Kobayashi (Univ. of Tokyo), Eiji Higurashi (Univ. of Tokyo), Hiroshi Toshiyoshi (Univ. of Tokyo), Toshiro Hiramoto (Univ. of Tokyo)	35th Sensor Symposium on Sensors, Micromachines and Applied Systems/10th Integrated MEMS Symposium Best Paper Award	Japan Society of Applied Physics (JSAP), Study Group of the Integrated MEMS	Two-layered 320×240 pixel parallel CMOS image sensor using direct connection of SOI wafer	2019/3/10
Takehiro Sugimoto, Yasushige Nakayama (Engineering Dept.), Tomoyasu Komori	Telecommunications Advancement Foundation Award, Telecom System Technology Award	The Telecommunications Advancement Foundation	22.2ch Audio Encoding/Decoding Hardware System Based on MPEG-4 AAC	2019/3/20

NHK Science & Technology Research Laboratories Outline

The NHK Science & Technology Research Laboratories (NHK STRL) is the sole research facility in Japan specializing in broadcasting technology, and as part of the public broadcaster, its role is to lead Japan in developing new broadcasting technology and contributing to a rich broadcasting culture.

History and near future of broadcasting development and STRL

2018 : 4K/8K satellite broadcasting begins

2016 : 8K Super Hi-Vision satellite test broadcasting

2013 : Hybridcast broadcasting begins

2011 : Analog television broadcasting ends

2006 : One-Seg service begins

2003 : Digital terrestrial broadcasting begins

2000 : BS Digital broadcasting begins

1995 : 8K Super Hi-Vision research begins

1991 : Analog Hi-Vision broadcasting begins

1989 : BS Analog broadcasting begins

1982 : Digital broadcasting research begins

U.S.-made television purchased for the home of the first subscriber

1966 : Satellite broadcasting research begins

1964 : Hi-Vision research begins

1953 : Television broadcasting begins

1930: NHK Technical Research Laboratories established

1925: Radio broadcasting begins

STRL Open House

The STRL Open House is held every year in May to introduce our R&D to the public.

STRL by the numbers

Established in June 1930		
June 1930 - January 1965	Technical Research Laboratories	
January 1965 - July 1984	Technical Research Laboratories, Broadcast Science Research Laboratories	
July 1984 - Present	Science & Technology Research Laboratories	
Employees	256	(including 226 researchers)
Degree-holding personnel	84	
Patents held : Domestic	1,997	
International	123	

(at end of FY 2018)

Current research building

Completed March 2002
High-rise building:
14 floors above ground, two below ground
Mid-rise building:
6 floors above ground, two below ground
Total floor space: Approx. 46,000 m²
Total research area: Approx. 16,000 m²
Total land area: Approx. 33,000 m²

NHK STRL Organization

Director of STRL Kohji Mitani

Deputy Director of STRL Toru Imai

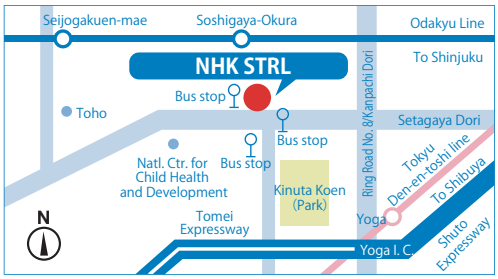
Deputy Director of STRL Masakazu Iwaki*
Executive Research Engineer Takashi Kato
Executive Research Engineer Tetsuomi Ikeda

*Head of Secretariat for AI Promotion

		Head
Planning & Coordination Division	Research planning/management, public relations, international/domestic liaison, etc.	Keiji Ishii
Patents Division	Patent applications and administration, technology transfers, etc.	Kenji Nakashima
Internet Service Systems Research Division	Integrated broadcast-broadband technology (Hybridcast, etc.), IT security, broadband video service technology, etc.	Kiyohiko Ishikawa
Advanced Transmission Systems Research Division	Satellite/terrestrial/cable broadcast transmission technology, multiplexing technology, 8K contribution/IP transmission technology, etc.	Masayuki Takada
Advanced Television Systems Research Division	8K program production equipment, video coding technology, highly realistic sound systems, etc.	Shinichi Sakaida
Smart Production Research Division	Video content analysis, speech recognition/synthesis, machine translation, social media analysis, automatic sign language CG system, automatic audio description generation system, etc.	Yuko Yamanouchi
Spatial Imaging Research Division	Spatial 3D video system, 3D display device, AR/VR technology, novel video representation technology, cognitive science, etc.	Tomoyuki Mishina
Advanced Functional Devices Research Division	High-sensitive imaging and functional imaging devices, high-capacity and high data-transfer-rate recording technology, sheet-type display technology, etc.	Hiroshi Shimamoto
General Affairs Division	Personnel, labor coordination, accounting, building management, etc.	Ryoji Takahashi
Secretariat for AI Promotion	Short term introduction of AI technology to program production	

(at end of FY2018)

Access to NHK STRL



Directions

■Odakyu line, from Seijogakuen-mae station, south exit:

[Odakyu Bus/Tokyu Bus]

・Shibu 24（渋24） toward Shibuya Station

[Tokyu Bus]

・To 12（等12） toward Todoroki-soshajo

・Yo 06（用06） toward Yoga Station（weekdays only）

・Toritsu 01（都立01） toward Toritsu Daigaku Station, north exit

■Tokyu Den-en-toshi line, from Yoga station:

[Tokyu Bus]

・To 12（等12） toward Seijo-gakuen-mae station

・Yo 06（用06） toward Seijo-gakuen-mae station（weekdays only）

In all cases, get off the bus at the "NHK STRL"（NHK技術研究所）bus stop

Edited and Published by:

Nippon Hoso Kyokai（NHK）Science & Technology Research Laboratories（STRL）

1-10-11 Kinuta, Setagaya-ku, Tokyo

Tel:+81-3-3465-1111

http://www.nhk.or.jp/strl/index-e.html

Annual Report

NHK
Science & Technology
Research Laboratories

2018

September 2019