5 Technologies for advanced content production

As broadcasting media become more and more diversified, including Hybridcast and Super Hi-Vision broadcasting planned to start in 2018, we continued to research program production technologies that can provide attractive new content services by utilizing past raw video footage and content and to develop program contributions equipment capable of large-capacity transmission. We also began researching new technologies such one for incorporating tactile sensation into TV viewing and one for conveying sound effects other than speech to those with hearing difficulties.

We are researching content utilization technology that uses text information to make effective use of broadcast program contents. Hyponyms of an arbitrary word are obtained by using their semantic relations and are used for information retrieval and recommendation. We also developed a method for extracting the program topic from text information about the program and semantic relations between words.

In our research on video indexing technology, we conducted further research and development on a video asset management system, "Video Bank," for assisting with video search and manipulation, and developed a video retrieval method based on a learning procedure for recognizing a subject with a few training images. We also developed a method for estimating lighting conditions in real time on the basis of video analysis that can be used in video manipulation technologies.

In our research on bidirectional field pick-up units for high-speed wireless transmission of file-based video, we examined a congestion control algorithm that considers wireless transmission errors as well as automatic repeat request and high-speed adaptive modulation based on time division duplexing. The transmission characteristics were evaluated using an experimental prototype, and the features were verified in microwave-band field transmission experiments.

Field pick-up unit systems operating in the 700-MHz band, which are mainly used for live broadcasting of marathons, need to be phased out in favor of ones that operate in the 1.2-GHz and 2.3-GHz bands. We developed a transmitter and receiver using the space-time trellis coded multiple-input multiple-output (STTC-MIMO) scheme and used them for measuring and analyzing the propagation characteristics of the new frequency bands.

Similarly, the frequencies used by radio microphones need to be migrated from the 700-MHz band to the 1.2-GHz band or the "TV white space" UHF band, which was allocated for television broadcasting but can be used for other purposes depending on geographical and technical conditions. We developed a digital radio microphone using the orthogonal frequency division multiplexing scheme that enables sound transmission with high quality and low latency.

We downsized a millimeter-wave mobile camera capable of transmitting Hi-Vision video with high-quality and low latency to a practical size and used it for shooting various programs such as golf tournaments and the NHK Kouhaku year-end music show. We also worked on elemental technologies for future Super Hi-Vision wireless cameras including ones for large-capacity transmission using MIMO transmission technology and for higher output using single-carrier transmission technology.

5.1 TV contents indexing and recommendation technologies

TV contents utilization using text information

For effective utilization of the massive amount of broadcast TV program contents, we are researching ways to use natural language processing technology to obtain the relations among words and broadcast contents and to recommend diverse types of content. In our research on acquiring semantic relations between words, we developed a method for obtaining the hyponyms for an arbitrary word[1]. For hyponym relations automatically extracted from Wikipedia, which include errors, we calculated the reliability of the hyponymy on the basis of the degree of similarity between a concept and its multiple subordinate concepts as well as the degree of similarity between a concept and its multiple superordinate concepts. Doing this removed errors with high precision. This process enabled us to collect a large number of hyponyms for certain words such as "disease" and "medicine." We used information about the collected words co-occurring in text to obtain semantic relations between words such as a "medicine" effective for a "disease." We added them to the semantic relations dictionary we developed in 2013. The types of semantic relations in the dictionary have increased in number from 8 to 28, including a new relation, "disease - remedy." In addition, we used the co-occurrence frequency in text of two certain words included in these semantic relations and the degree of coupling in their link structure to numerically prioritize their semantic relations. This led to the development of information retrieval and recommendation applications.

In our research on relations acquisition for broadcast contents, we devised a method for identifying the topic of a program[2] by using text information about the program and the extended semantic relations. We demonstrated that the method is 6.8% more accurate than the conventional one. We incorporated the method into the system we developed in 2013 for constructing a contents map connecting TV programs with semantic links and created a TV program navigator to support viewer exploration of TV programs (Figure 1). The program navigator can guide a user from a single search word to a variety of programs. For example, it can guide a user from the search word "anemia" along the route shown in red arrows in Figure 1: "Anemia to Prevention to Clam" and to a program in which "Today's Menu" features clams.

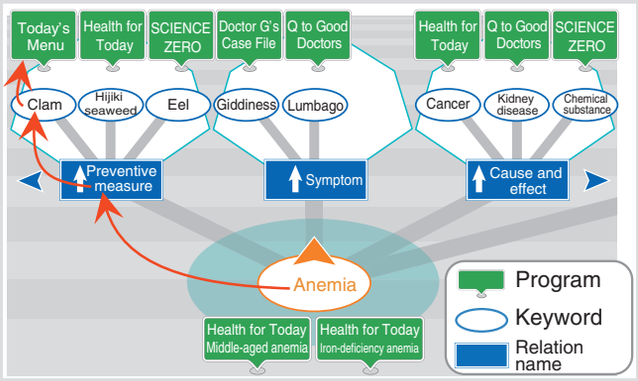

Figure 1. TV program navigator

Video indexing technology

To make use of raw video footage stored in video archives and to provide enriched video expressions, we are prototyping a video asset management system called "Video Bank." The system uses video analysis and sensing technologies to automatically add metadata useful for video production in terms of video search and manipulation.

For video retrieval, we developed a method for performing a learning procedure required for recognizing a subject on the basis of a few training images. It enables efficient, automatic assignment of a subject name. We built a video retrieval system that can add a searchable subject name at any time[3] by using this method and exhibited the system at the STRL Open House. This research was conducted in cooperation with IBM Japan.

For a more advanced video archive system, we further developed a visual-based image search technology for retrieving an image using the visual similarity of images. We had previously used image features such as color moment, hue histogram and fractal sequence to evaluate the similarity of images. To improve the search capability for images including objects such as buildings and cars, we introduced two new image features, a multiplex scale-edge direction histogram and a multiplex scale-edge connection angle histogram, that can emphasize a long edge and reflect the relationship with neighboring edges. This resulted in a precision rate of approximately 70% for the top ten items in the search results.

We also developed a system that can search from approximately 60,000 programs by linking these video retrieval technologies with the existing archive search system and conducted verification experiments.

We continued support for the experimental use of the earthquake disaster metadata system at the NHK Fukushima station for the purpose of organizing and managing a huge amount of reporting video on the Great East Japan Earthquake. Upgrading the visual-based image search functionality of the system enabled high-precision search of visually similar images even for disaster video, many of which include buildings, cars, and rubble. We also installed the same system at the NHK Morioka station to increase the variety of data, aiming at the development of a more versatile system.

We are also investigating ways to make video search more effective by extracting names, such as the name of a place or a person, from the image scene. We developed a technology for detecting a character area in the image and one for efficiently recognizing faces from a massive amount of video.

For video processing technology, we investigated ways to obtain lighting conditions (light position and intensity) and the camera parameters that are required for natural synthesis of a computer graphic (CG) image and a photographed image. We developed a method for estimating the lighting conditions in the studio in real time by using video analysis[4]. This method first calculates the light position from the video of two sensor cameras by using the principle of triangular surveying and images of the individual contribution of each light collected in advance. Simultaneous equations are then solved for the sensor camera images captured under the unknown lighting conditions to be estimated. In this process, the intensity of each light is estimated. We compared the image of an actual subject captured in the studio and the CG image of the subject rendered by using the studio lighting conditions obtained by this method. The results showed that the lighting conditions could be estimated sufficiently accurately (3% margin of intensity error) (Figure 2). This method enables production of natural synthetic video using a CG object drawn under the same lighting conditions as those for photographed images.

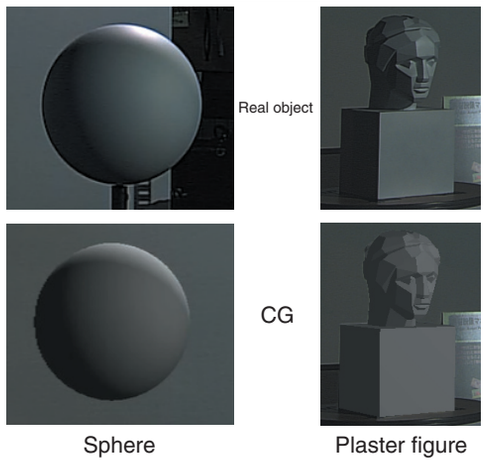

Figure 2. CG rendering using light estimation results

For outdoor lighting conditions, we developed a light sensor that uses multiple spectrophotometric sensors by upgrading a hybrid sensor technology, which uses multiple sensors to obtain information about a real space, and verified its basic principle. To collect camera parameters from this hybrid sensor, we devised a method for determining with high precision if the camera's motion is rotational or translational by using neural networks and applied the method to various types of cameras including a crane camera. We used this hybrid sensor on an experimental basis for producing actual programs to verify its effectiveness. Part of this research was conducted in cooperation with NHK Engineering System, Inc. and the Shimizu Corporation.

To further promote utilization of raw video footage, we researched video texturing technology for interpolating and extrapolating part of a video in the spatio-temporal domain. We developed a basic technology for interpolating between frames to change the scale of the video to the temporal axis dimension. This technology enables high-quality interpolation between frames by analyzing the temporal transition of a high-density pixel area in the video.

[References]
(1) I. Yamada, T. Miyazaki, M. Miyazaki, K. Miura, A. Matsui and H. Tanaka: "Class-Instance Acquisition using Automatically Extracted Hyponymy Relations," IEICE Technical Report, Vol. 114, No. 444, NLC2014-44, pp. 1–6 (2015)(In Japanese)
(2) K. Miura, I. Yamada, T. Miyazaki, M. Miyazaki, A. Matsui, N. Kato, H. Sumiyoshi and H. Tanaka: "Link Generation for TV Programs based on Semantic Relations between Words," IEICE Technical Report, Vol. 114, No. 366, NLC2014-42, pp. 105–110 (2014)(In Japanese)
(3) Y. Kawai, T. Mochizuki, H. Sumiyosh and S. Fujiwara: "Video retrieval system based on object recognition technique," ITE Annual Conference, 2-5 (2014) (in Japanese)
(4) H. Morioka, H. Okubo and H. Mitsumine: "Real-time Estimation Method of Lighting Condition with Color Sensor Cameras for Image Compositing," ITE Annual Conference, 11-3 (2014) (in Japanese)

5.2 Bidirectional field pick-up unit (FPU) transmission technology

We are researching bidirectional field pick-up units (FPUs) for use in high-speed wireless transmission of file-based video footage. We researched a high-speed adaptive modulation scheme and automatic repeat request using time division duplexing (TDD), upgraded our dual-polarized parabolic antenna, examined a congestion control algorithm that takes into account wireless transmission errors , and evaluated file transmission capabilities through field experiments.

We evaluated the high-speed adaptive modulation scheme implemented in an experimental prototype operating in a radio propagation environment and demonstrated that it effectively prevents transmission errors by switching modulation methods automatically even when propagation conditions vary rapidly[1]. We also refined the frequency conversion circuit and carrier filter coefficient of the transmission/reception control unit to ensure sufficient transmission quality even when a high-order modulation method such as 256 QAM or 1024 QAM is used. As a result, the modulation error ratio was improved by more than 4 dB.

For automatic repeat requests, we studied the Hybrid Automatic Repeat reQuest (HARQ) protocol, which adaptively combines error corrections and repeat requests to adjust the code rate[2]. Evaluations demonstrated that using HARQ improves the bit error rate when the channel characteristics are degraded and prevents transmission bit rate degradation (Figure 1).

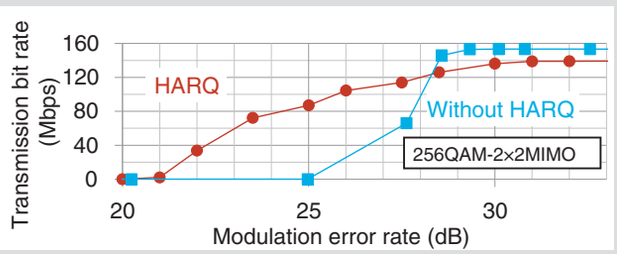

Figure 1. Comparison of transmission bit rate characteristics using actual equipment

We also upgraded the dual-polarized parabolic antenna that we first prototyped in FY 2012 (Figure 2) to increase the gain and widen the bandwidth. As a result, the frequency characteristics were improved, and the gain difference in the 6/7-GHz bands was reduced to within 1.2 dB[3]. We also devised a structure that facilitates setting up the antenna for practical use.

For the congestion control algorithm, we developed a transmission control protocol (TCP) for transmitting files at high speed even on wireless networks. The conventional TCP, which determines the congestion status of a network on the basis of increasing transmission latency, makes an erroneous determination if latency increases due to unrelated congestion such as retransmission delay to recover bit errors and waiting time for wireless transmission in a time division duplex system. The result is a reduction in the transmission rate. We thus developed a method for estimating the amount of latency increase unrelated to congestion. Experiments using actual networks demonstrated that the method prevents erroneous determination of congestion status and thus can reduce the time needed for data transmission[4].

Figure 2. Dual-polarized parabolic antenna with splash plate (left: front, right: back)

We conducted field experiments on an experimental bidirectional FPU prototype using a microwave band (6–7 GHz) to evaluate its file transmission capability and adaptive modulation capability. It performed bidirectional communication over 21 km (overland) using TDD. With the transmission parameters set to 2×2 multiple-input multiple-output transmission and 256 QAM modulation and using a (180,164) Reed-Solomon code for error correction, it provided a throughput of approximately 145 Mbps, twice that of the conventional method for video file transmission using the file transfer protocol. The adaptive modulation scheme enabled transmission with stable throughput by using an appropriate modulation method automatically selected to match the conditions of the propagation path[5].

[References]
(1) K. Mitsuyama, T. Koyama, F. Uzawa, N. Iai and K. Aoki: "Performance Evaluation of High Speed Adaptive Modulation Scheme for Bidirectional Digital FPU System," ITE Technical Report, Vol. 38, No. 30, pp. 29–32, (2014) (in Japanese)
(2) F. Uzawa, K. Mitsuyama, K. Aoki and T. Hiraguri: "A Study on Applying HARQ Scheme for Bidirectional FPU," ITE Journal, Vol. 68, No. 5, pp. J210–J212 (2014) (in Japanese)
(3) T. Kumagai, K. Mitsuyama, N. Kogo and N. Iai: "A Dual-polarized Parabolic Antenna with Splash-plate for Bidirectional FPU," IEICE General Conference, B-11-5 (2015) (in Japanese)
(4) T. Koyama and K. Aoki: "Slow Start Algorithm for Mobile Broadband Networks Including Delay Unrelated to Network Congestion," ICNC2015, CNC-1-5, (2015)
(5) T. Koyama, F. Uzawa, T. Kumagai, K. Mitsuyama, N. Iai, K. Aoki and T. Okayama: "Field Experiment of File Transfer by Bidirectional Digital FPU System," IEICE General Conference, B-11-4 (2015) (in Japanese)

5.3 700-MHz-band frequency migration

A government policy has required field pick-up units (FPUs) and professional wireless microphones (specified radio microphones) operating in the 700-MHz band to change to a new frequency band. We investigated ways to achieve a smooth frequency migration.

FPU transmission technology

We previously developed the space-time trellis coded multiple-input multiple-output (STTC-MIMO) scheme for FPUs intended for use in mobile transmission using a new frequency band (1.2 GHz or 2.3 GHz). We conducted mobile transmission experiments on this scheme around our laboratory using the 2.3-GHz band and a transmitter and receiver we prototyped for supporting large-scale relay broadcast programs such as road races. We compared the transmission characteristics of different antenna configurations in terms of polarization type such as vertical, oblique, and circular and the number of reception antennas and investigated an antenna configuration suitable for the STTC-MIMO scheme[1]. We also upgraded the prototype transmitter and receiver to implement a new parameter that has been included in the ARIB STD-B57 standard for new-frequency FPUs. We also continued our support from FY 2013 of FPU manufacturers developing MIMO FPUs for the new frequency bands.

Specified radio microphone transmission technology

Although digital radio microphones resistant to interference were developed for a smooth frequency migration, there was quality degradation due to compression of the sound signals and transmission latency due to the digital signal processing. We developed prototype devices for practical use of digital radio microphones using the orthogonal frequency division multiplexing (OFDM) scheme that would enable audio transmission with high quality and low latency. We succeeded in fabricating prototypes of a handheld transmitter, a bodypack transmitter, and a receiver capable of simultaneous reception of two channels in sizes as compact as existing devices (Figure 1). The devices, which are compliant with ARIB STD-T112 ver. 1.2 (established in March 2014), have enabled transmission of high-quality pulse code modulation (PCM) audio with latency below 1 ms.

Figure 1. Transmitter and receiver of developed OFDM digital wireless microphone

OFDM digital radio microphones use the 1.2-GHz band, which is one of the new frequency bands. There was concern that a signal using the this band would be affected by interference from radar waves, so we developed a method to eliminate such interference and thereby enable more stable signal reception[2].

[References]
(1) T. Nakagawa, F. Uzawa and N. Iai: "Transmission characteristics and antenna configuration of STTC-MIMO scheme," ITE Technical Report Vol. 38, No. 49, pp. 7–12 (2014) (in Japanese)
(2) F. Ito, T. Kuki and H. Hamazumi: "A Study on Pulse Interference Suppression with Symbol Elimination for OFDM Digital Wireless Microphone Systems," ITE Winter Annual Conference, 1-2 (2014) (in Japanese)

5.4 Wireless cameras

Toward the development of a practical millimeter-wave mobile camera capable of wireless transmission of high-quality Hi-Vision video with low latency, we investigated ways to make such equipment more compact. We also researched elemental technologies for future 8K Super Hi-Vision wireless cameras including large-capacity transmission by using a multiple-input multiple-output (MIMO) scheme with four transmitters and higher output power by using a single-carrier frequency domain equalization (SC-FDE) scheme.

Millimeter-wave mobile camera

We improved the performance of the demodulator by enhancing frame synchronization and downsized the return link demodulator for a more practical millimeter-wave mobile camera. We also continued to support program production using the millimeter-wave mobile camera. Our millimeter-wave mobile camera was used for shooting a wide range of programs such as gymnastics competitions, golf tournaments, the NHK Kouhaku year-end music show, and the New Year parade of fire brigades, where agile shooting and powerful imagery using wireless cameras are essential.

Multiple-input multiple-output (MIMO) with four transmitters

Extending the two-transmitter QPSK-MIMO system we previously developed to a four-transmitter system doubled the transmission capacity, but it also exponentially increased the computation volume required for demodulation by 16 times. To address this problem, we developed in FY 2013 an algorithm for MIMO signal detection with a reduced computation volume so that it can be built into hardware. In FY 2014, we prototyped an OFDM modulator and demodulator for the four-transmitter MIMO scheme using this algorithm and evaluated the computation volume of the hardware. We also prototyped a 10-GHz-band radio frequency transmitter to evaluate the transmission characteristics.

Single-carrier frequency domain equalization (SC-FDE)

We conducted laboratory experiments with radio-frequency (RF) signal connections to evaluate the transmission characteristics of the SC-FDE scheme when using a millimeter-wave-band power amplifier, which is prone to distortion at high output power. The results showed that the required received carrier-to-noise ratio for quasi-error-free transmission does not degrade when an amplifier using the SC-FDE scheme is operating at a 6-dB-higher output power compared to the conventional OFDM scheme. We conducted wireless transmission experiments on the SC-FDE scheme in the 42-GHz band for the first time and demonstrated that this scheme is effective for mobile transmission in multi-path environments[1]. We also evaluated the mobile transmission characteristics of the scheme by using amplifier output as a parameter. We obtained better amplifier output than with the OFDM scheme both in laboratory and wireless transmission experiments, meaning that the SC-FDE scheme can increase the transmission power even with a millimeter-wave-band power amplifier without a pre-distortion function.

[References]
(1) Y. Matsusaki, T. Nakagawa and H. Hamazumi: "Indoor Transmission Experiments of SC-FDE in 42-GHz Band," IEICE General Conference, B-5-45 (2015) (in Japanese)

5.5 Innovative projects and proposals for the Tokyo 2020 Olympics

NHK STRL has embarked on research with a great deal of novelty and originality for the purpose of finding new areas of study. We made progress in six such research projects, including a TV viewing system that provides the viewer with tactile sensations in addition to conventional audiovisual information and a user-friendly, enriched broadcasting service that conveys environmental sound including sound effects other than speech to the elderly and those with hearing difficulties.