Studies for Improving Conventional Broadcasting Services

OUTLINE

For the improvement of conventional broadcasting services, with a view to their future digitization, studies continued on technologies that support efficient and more expressive program production, and human-friendly services that make possible the provision of programs that are easy for all viewers to watch, listen to, and understand.

2.1 Efficient Program Production

Our research on efficient program production technology has promoted technologies to support the program production process from the planning to the editing stages. This has involved the construction of a database comprised of video materials and related information such as scripts or coverage memos that are generated at each stage of the program production process. In order to extract desired programs or data promptly from a massive broadcasting station database, we strove to develop a retrieval system that extracts and attaches indexing information to video data and other material, and then retrieves desired information based on the indexing. We also participated in and contributed to the global production of agent TV specifications, that allow for the customized selection and presentation of programs based on EPGs (Electronic Program Guides) and the viewer's personal preferences and interests.

Regarding new video expression technology, studies continued on an intelligent robotic camera system that can automatically shoot images by emulating cameramen's shooting techniques. We constructed a prototype system whereby two robot cameras shoot a talk show program by exchanging information between them. Research also progressed on an image-based virtual studio where actual video images can be compartmentalized and freely synthesized to create a studio set. We further improved the system to allow for easier manipulation and more natural synthetic video images. Enhanced functions related to picture quality and CG expression were also incorporated into TVML (TV program Making Language), a simple computer language used to describe TV program scripts, and the system was actually used for TV programs.

Efforts were also made to promote technological developments that correspond to the diversifying and advancing HDTV program production process which is expected to further expand with the implementation of digital satellite broadcasting. Aiming to achieve high-speed and high-density recording with an HDTV optical disk camera, research continued on optical disks, optical heads, and signal processing. Research was also conducted on an HDTV hard disk recording system for application to HDTV non-linear editing. Additionally, an HDTV OBL (Outside Broadcasting Link) was also developed for mobile relay broadcast use, such as for marathons.

Additional program production related research included the conversion of an “ice-zone microphone,” originally created for use at the Nagano Olympic skating competition, into an “insect microphone,” that is now capable of picking up micro-sounds from an insect. This insect microphone was used on scientific programs and received favorable comments. In order to contribute to the efficient use of stored programs, emphasis was placed on the development of a video archive retrieval system that can be simultaneously accessed by multiple users over station networks. Research also proceeded on computer technology to support the Japanese-English translation of broadcasting materials, especially for the efficient translation of newscasts.

2.2 Human-Friendly Broadcasting Services

Accelerating research on speech recognition systems that can automatically generate closed-captioning from a newscasters’ speech resulted in the achievement of a performance goal; a speech recognition accuracy of 95% was obtained when an anchor studio announcer read a manuscript. On the basis of this accomplishment, we constructed a practical system for experimental use in live news programs, and a broadcast news transcription system was put into service for the March 27th, 2000, edition of “News at 7.”

To make highly operable digital broadcasting available to all viewers, we also studied human-friendly interfaces such as remote control devices and operation screens. We also improved the performance of tactile communication devices which deliver digital broadcasting textual information to the visually-impaired or to individuals with both visual and hearing impairments. Regarding a speech rate converter capable of slowing down an announcers’ rapid speech to a more comfortable speed for elderly viewers, we developed PC software that can perform real time speech rate conversion.
**DTPP**

Desk Top Program Production (DTPP)

Research has continued on a prototype hierarchical TV program editing system based on TV program plan sheets to store video data together with various types of textual information such as TV program scripts that are generated at each stage of the program production process.

During fiscal 1999, enhancements were made to the video operating functions of a workstation version of this system and other functions were added including a network-based collaborative work environment. We started developing a data input camera with an efficient information input process after an application experiment conducted in 1998, using the TV program “Science Eye”, revealed the need for increased efficiency. The data input camera is capable of video management based on plan information of DTPP at the time of shooting, and automates the video material classification process, thus improving editorial efficiency. It is also equipped with a gyro and a GPS to automate the assignment of keywords such as camera operations and shot locations.

We developed a prototype program generation system for educational English conversation programs, which is based on a study of program information description methods using XML (eXtensible Markup Language) for secondary material use and material retrieval. Experiments on the production of conversational-style TVML programs and Internet web pages using basic English conversation information items in a template format (keywords of the conversation) were conducted.

With the aim of digest viewing or video retrieval in digital broadcasting, we carried out indexing production experiments on a sport (soccer) video, and produced a summary video of the game using the spectators’ cheering.

**TVML (TV program Making Language)**

TVML is a computer language that describes TV programs using text-based scripts. A script written in TVML is interpreted by computer software called the TVML Player, and a real-time program is automatically produced. During fiscal 1999, we completed a Windows version of the TVML Player and distributed it as freeware. This enabled the realization of a program production environment on an individual personal computer. As a professional version of the program with broadcasting quality output, we developed TVML Player PRO by making improvements in the picture quality, sound quality, CG expressions, and others. TVML Player PRO was actually used for three NHK broadcast programs (Figure 1). Other research projects related to TVML included the automated production of a news program and talk show, a camera simulator that allows for manual manipulation of a TVML camera by connecting TVML to a camera tripod, and automatic TVML production using a DTPP TV program plan sheet. TVML presentations and demonstrations were performed nationally and internationally to high acclaim.

**Broadcasting Agents**

Joint research on a European project, FACTS (FIPA Agent Communication Technologies and Services), advanced research and development on agent TV that utilizes agent technology (software which manages/retrieves/processes massive amounts of information for users) to realize program selection functions for viewers. The final agent TV version was completed during fiscal 1999.

This completed agent TV is a system which selects and presents programs for individual viewers employing user profile information, EPGs (Electronic Program Guides), and personal preferences and interests (Figure 2). The system consists of 20 different types of agents and it provides advanced functions such as automated user profile compilation, Natural Language Interfaces (NLI), and communication support between different Agent Platforms. Among these functions, NHK contributed to the development of a Natural Language Agent (NLA) for the provision of NLIs and a Provider Agent (PA) for the distribution of program information including EPGs.

This system is based on a FIPA (Foundation for Intelligent Physical Agents) international specification standard, and we have greatly contributed to the improvement of this specification by reporting on the actual operating characteristics of this new system.

**Program Indexing**

Concerning broadcasting station databases, we are conducting research on methods for program and data retrieval and browsing by utilizing indexing information attached to video and textual data.

As for the indexing support technique, a prototype system was employed that utilized a topic detection method invented during fiscal 1998 for TV-news manuscript databases. The prototype system was exhibited at the STRL open house in fiscal 1999, and was highly evaluated. Additionally, we applied this technique to an individual topic detection method that selects topics according to a user’s personal interests and needs.

Furthermore, by providing system suggestions based on knowledge we had acquired through our DTPP research, we contributed to the establishment of an international MPEG-7 standard that will realize the common utilization of data indexing stored in different databases by means of one search engine.

---

*Figure 1: Data input camera*

*Figure 2: FACTS Agent TV*
Image Expression Technology

Intelligent Robot Camera

We have continued the study of intelligent robotic camera systems that allow one cameraman to operate multiple cameras using the same shooting techniques as several experienced operators. In fiscal 1999, we clarified each camera’s individual role, the types of shooting cuts, and the technical director’s commands to the cameramen by analyzing the shooting techniques used with multiple cameras in a talk show program. Based on this data, we constructed a prototype two camera robotic operating system that can be used for a talk show involving two performers using a flip chart (Figure 1). Two sensor cameras (fixed) measure an object’s position and capture a complete studio image for the real-time estimation of each object’s 3-D coordinates. The robot cameras compose their shots using the following information along with the object’s 3-D coordinate data:

- Information concerning how and which object the other robot camera is shooting at any given moment.
- Sound information concerning which person is speaking and for how long.

We also test manufactured a man-machine interface that allows a cameraman to operate the system with voice commands and a pan pod.

Additionally, in order to apply the system to live sports programs, basic experiments were performed regarding sensing techniques for the detection of a small object moving about in a large area using a sensor camera’s pan, tilt, and zooming functions.

Image-based Virtual Studio

Aiming to construct a high quality and functional virtual studio, research has been progressing on an image-based virtual studio that can compose electronic studio props based on actual shot images. We integrated past research accomplishments such as high-resolution omnidirectional imaging (environmental video components) for the production of the surrounding scenery, and an element technology that breaks down video images into three-dimensional video components. In three-dimensional video components, we developed a technique to estimate object texture and specular-reflected coefficients by eliminating specific illumination information. As a result, it became possible to alter the illuminating conditions of three-dimensional video components as if they were in the environmental video components filmed under different lighting conditions.

In a prototype virtual studio, a virtual camera manipulator enabling the adjustment of camera position, direction (anywhere within 360°), and viewing angles, realizes a video production environment with a high degree of flexibility. Moreover, a graphical user interface (GUI) was constructed to make changes in the selection and arrangement of three-dimensional video components feasible, in addition to changes in illumination conditions. This system creates CG actors and actresses that replace actual performer images, and utilizes GUI to control the movement and speech of the characters. Through experiments on these tools, we confirmed they could be applied to HDTV program production, although camerawork and illumination conditions entail certain restrictions.

Electronic Video Correction on HDTV Drama Production

In the production of HDTV dramas, images are shown in high-definition, making incongruities and imperfections in studio sets and costumes more obvious, such as the borderline of wigs worn in Japanese historical dramas. Although efforts had been made to correct the problem by using special wigs designed for HDTV programming, this solution resulted in higher costs and longer dressing times. Therefore, we developed an electronic processing technology for after-shooting of HDTV images, specifically targeting the problem with actors’ wigs. By specifying the portion of the video that needs correcting, this technology automatically processes the massive amount of information contained in all of the picture frames.

In fiscal 1999, we realized high-speed processing by optimizing the developed algorithm. The processing time for one-second of HDTV material is less than 10 minutes. In cooperation with the Engineering Administration Department and the Broadcast Engineering Department of NHK, we further developed a practical verification system employing this technology and confirmed its effectiveness and simplicity by utilizing the system on a historical drama.
2.1.3 Archive Retrieval and Preview System

The NHK Archives are scheduled to open in 2003. They require program retrieval and previewing functions on a network level using a PC. We constructed a prototype video retrieval and preview system for the archive, employing a hierarchical recording system to evaluate the functionality.

For the basic prototype system manufactured in fiscal 1998, we proposed a hierarchical storage management system which combined a high-speed accessible hard disk with a large capacity, low cost tape streamer to store large volumes of highly compressed reference video programs. This basic prototype showed that the realization of a compact, low cost, high-speed, large capacity, and highly reliable system is possible.

In fiscal 1999, we applied this hierarchical storage management system to the reference video server for a retrieval and preview system. Advances were made in the development of contents management methods for multiple access as well as in the connection to a network (inserted figure). As a result, we demonstrated that multiple users could retrieve and preview programs from their own PCs on a network in this experimental system. This retrieval system is capable of referencing program-related information in addition to the instant previewing of key frames and video programs.

For program retrieval, the system uses a key word search, which is enhanced by the narrow down and synonym reference functions.

In order to improve the preview operability, a general-purpose program playback console can be connected to a personal terminal. It realizes easy variable forward and backward playback functions by its jog dial. The system can also display down-converted Hi-Vision (HDTV) for reference video programs in addition to standard definition television.
The optical disk camera's advantageous functions, such as quick field editing, make it more suitable as a camcorder for program coverage than conventional tape-type cameras.

An HDTV optical disk camera cannot be realized without attaining a high transfer rate for large amounts of high-density data. Therefore, we have promoted research on optical disks, optical heads, and signal-processing schemes.

Regarding high data bit rates, by making both the optical head, using red laserdiode (wavelength 635nm), and the signal processing system faster, we developed a phase-change optical disk evaluation system (inserted picture) that can measure a bit error rate (BER) of 100Mbps and higher, and conducted high-speed recording experiments.

Multiple beam parallel recording was effective in improving the data transfer rate, and we confirmed the possibility of near-infrared wavelength (780nm) use through the research of multi-beam optical sources. We also devised a control method with a feed-forward system that could follow the high-speed disk revolutions with high accuracy.

Regarding high-density characteristics, we started the trial manufacture of both optical heads and optical disks for blue laserdiode, and made recording experiments using them. This was possible due to the rapid progress being made on the development of blue laser diodes (wavelength 405nm). Moreover, we began research and characteristic evaluations on an optical element that is highly sensitive to the blue light wavelength. Additionally, we developed simulation software for the reproduction process, and other various software for reproduced signal processing. This made examinations possible in regard to the relationship between the bit error rate (BER) and necessary S/N ratios, and also in regard to the optimum signal-processing method.

We participated in the working group of the ARIB (Association of Radio Industries and Businesses), discussing requirements for an broadcasting-use optical disk. We also initiated studies on a suitable HDTV optical disk format for broadcasting.
To realize an HDTV nonlinear editing system, we started research on an HDTV video hard disk system in fiscal 1997. That same year, we proposed a hard disk evaluation tool for video applications, and we further proposed, in fiscal 1999, a technique to describe characteristics of hard disk drives with small number of parameters, using the tool. The parameters necessary for this technique are the known disk rotational speed, the physical location of the recorded data, and the seek time (the time required for the head to move to the desired track) obtained by means of the evaluation tool (inserted figure). Additionally, the creation of a virtual hard disk drive by parameter adjustments led to the establishment of guidelines for future performance improvements of conventional hard disk drives. Applying these accomplishments, we provided technical cooperation to manufacturers, thus contributing to the manufacture of a practical hard disk recording system capable of simultaneous playback and recording/playback, and synchronous simultaneous recording and playback for two HDTV systems.

Aiming to improve the reliability of video hard disks, we have also developed an ultrasonic vibration spindle that vibrates the disk during start and stop functions to reduce the friction between the head and disk. By ultrasonic head levitation experiments, we analyzed the optimum vibrating mode. As a result, the friction coefficient was reduced more than 75%, thus confirming the system’s effectiveness in reducing head friction. We also discovered the possibility of miniaturizing the spindle by increasing sound levels with a reflector plate. Furthermore, we developed a sound sensor to estimate system hard disk problems caused by head crashes, and disk damage detection experiments were carried out.

![Hard disk readout function image](image-url)
Digital satellite broadcasting which is to start in December, 2000, will provide digital HDTV broadcasting. To prepare for this advance, we are pursuing research to achieve high quality images for material converted by NTSC-HDTV up-converters (devices used to convert NTSC signals to digital HDTV signals) in order to utilize existing and archive material in an HDTV format.

This system, differently from conventional linear interpolation, first learns the corresponding relationship between the waveform edge of HDTV images and that of converted NTSC data signals. Applying this information, the system estimates the HDTV image contours that are to be produced based on the NTSC signal edge data, thus realizing a high-resolution image. The effectiveness of the system on still pictures had already been confirmed.

In fiscal 1999, we examined the problems related to still image processing and extracted future research agendas, with the aim of applying the results to video images. We proposed an estimation technique utilizing high-resolution elements that correspond to motion information and a technique involving the addition of high-resolution elements to the converted image, and confirmed the basic performance of the system (inserted figure).
Selective Sound Receiving Microphone System

Research is ongoing on a new sound pickup system that can isolate a desired sound source even when the sound source and other noise sources are located in the same direction. Simulations have shown that approximately 40dB of noise can be suppressed when the sound sources are located within 30° of the microphone’s main axis and the desired sound and noise have no correlation. Digital filter processing was adopted to correct the characteristic differences between the two microphones used for the system, one of the factors that decreased the system’s separation performance, and this improved the characteristic differences, maintaining them under 1% of the waveform signal amplitude. Additionally, we manufactured a dedicated signal processor for obtaining the separation coefficient, and field tested the device with its possible commercialization in view.

Insect Microphone

With the aim of picking up small sounds, such as those generated by insects, we constructed a microphone consisting of a dynamic electroacoustic transducer attached to a contact pin. Although its operating principle is similar to that of a record needle, we fabricated a contact pin supporting mechanism with the high compliance and realized a high-sensitivity and wide-band operation. Its sensitivity is -33dBV/0.1G, and it successfully picked up the sound of a snail’s heartbeat, the sound of a snail eating a carrot, and the sound of an ant’s footsteps, etc. This new microphone was introduced in TV programs (inserted picture).

IC Microphone

The development of an ultra-small and high-performance condenser microphone using silicon chip processing technology is now underway. In fiscal 1999, we fabricated a 5 μm thick silicon diaphragm using a photoresist pattern design based on parameters obtained from simulations done in 1998. We also began fabrication of the microphone structure.
Machine Aided Translation System: From Japanese to English

We are investigating a machine aided translation system from Japanese to English (Translators’ Workbench) which supports such translation work. This system includes various functions such as (a) a translation example browser, (b) a term retrieval function, and (c) a bilingual Web retrieval function. In the future and as a supporting function, this system will be integrated with Japanese-English machine translation, which is also being promoted. In fiscal 1999, we realized the above-mentioned (a) and (b) functions, as well as part of the (c) function, and applied them to the construction of a prototype bilingual editor which helps a translator to work on a project.

Especially, a “translation example browser” which includes the (a) and (b) functions was unified in the English article writing system that the International Planning & Broadcasting Department and the Multimedia Development Department of NHK have been developing (inserted figure).

Machine Translation from Japanese to English

Regarding machine translation from Japanese to English, research continued on a rule-based translation system. This system performs translation in three distinct phases: the analysis of Japanese; the transfer of Japanese structures into English ones; and the generation of English script. Regarding Japanese-English transfer, we constructed a pre-transfer function that links a Japanese bunsetsu dependency analysis to the Japanese-English transfer process. To achieve high-speed transfer, advances have been made on the processing procedure, including the limiting of applicable rule selections. We produced a bilingual dictionary consisting of 300 function words for Japanese-English transfer. The vocabulary of content words has been increased from 2,000 words in 1998 to 5,000 words presently. Regarding the generation of English script, we constructed a prototype of the generation engine which determines the English word order. As a result, the English translation of Japanese news can be accomplished, although the system is still somewhat limited in its expressions.
HDTV OBL (Outside broadcasting link)
OFDM-OBLs (Orthogonal Frequency Division Multiplexing -Outside Broadcasting Link) for standard television broadcasts that use the 800 MHz-band and HDTV service links in the 6-7 GHz-band have been developed with the aim of utilizing them for live relay broadcasts of races, such as marathons.

Regarding the latter HDTV system, we have researched a synchronous demodulation scheme using pilot carriers which is expected to have superior characteristics in comparison to the differential demodulation scheme. Optimizing the configuration of pilot carriers resulted in preferable mobile transmission characteristics in the microwave band with a high Doppler shift.

On the other hand, in the case of road race relay broadcasts, including non line-of-sight propagation, utilization of the 800 MHz-band is more advantageous, requiring fewer receiving stations for transmission. Thus, we investigated a transmission scheme which secures a necessary transmission capacity for digital HDTV material through the application of two channels and multilevel QAM formats, jointly promoting the development of the transmitter with the NHK Engineering Administration Department.

SNG
In order to realize a portable SNG RF terminal, research on a low-profile active antenna with which active devices such as power amplifiers can be integrated is being promoted. In fiscal 1999, we calculated the influence on the phase and amplitude of the radiated waves due to mutual coupling between antenna elements and confirmed that the beam pointing error will be at an allowable level for our portable SNG RF terminal.

We studied the structure of transmitting active antennas using EM simulations and found that the electromagnetically coupled slot antenna is a promising antenna element for a portable SNG RF terminal.

We also studied a feeding method which can reduce active devices by appropriate coupling between antenna elements.

### OFDM-OBL under development

<table>
<thead>
<tr>
<th>Frequency band</th>
<th>6-7 GHz-band</th>
<th>800 MHz-band</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major modulation schemes</td>
<td>16QAM-OFDM</td>
<td>16QAM-OFDM</td>
</tr>
<tr>
<td>64QAM-OFDM</td>
<td>64QAM-OFDM</td>
<td></td>
</tr>
<tr>
<td>Occupied bandwidth</td>
<td>15.5 MHz</td>
<td>8.5 MHz x 2</td>
</tr>
<tr>
<td>Maximum transmission capacity</td>
<td>75 Mbps</td>
<td>87 Mbps</td>
</tr>
<tr>
<td>Bit rate</td>
<td>25 Mbps</td>
<td>29 Mbps</td>
</tr>
<tr>
<td>(16QAM, r=1/2)</td>
<td>(16QAM, r=1/2)</td>
<td></td>
</tr>
<tr>
<td>57 Mbps</td>
<td>57 Mbps</td>
<td></td>
</tr>
<tr>
<td>(64QAM, r=3/4)</td>
<td>(64QAM, r=3/4)</td>
<td></td>
</tr>
<tr>
<td>Application</td>
<td>Fixed to mobile relay within the range of sight</td>
<td>Mobile relay including non line-of-sight propagation</td>
</tr>
</tbody>
</table>

r: Coding rate
In TV program shooting, a color glass filter is inserted into the camera optics to compensate color temperature when the light source is changed. However, it is difficult to insert a color filter during shooting, and the amount of compensation can not be changed for the glass filter. Therefore, in cooperation with NHK Nagoya Broadcasting Station, we developed a liquid crystal (LC) color-temperature compensating filter that can continuously adjust color temperature at high speed. The filter consists of two glass substrates coated with a transparent electrode and nematic LCs mixed with dichroic dyes imbedded between the substrates. The dyes' light absorption spectra change depending on the direction of the molecular alignment. In fiscal 1998, we realized an absorption spectrum characteristic that converts solar light into incandescent by the use of two types of dyes with absorption peaks at 398nm and 556nm. We designed a filter using two LC cells so that the quantity of color temperature compensation is independent of the polarization of the incident light.

In fiscal 1999, we optimized the dye concentration ratio. In addition, we changed the LC orientation from homogeneous to homeotropic alignment. As a result, the filter can continuously control the color temperature of transmitted light between 3700K and 5600K with an applied voltage of up to 10V. Response time is approximately 80ms. A camera equipped with the filter (inserted picture) was used in the live program ‘Lunchtime Nippon’ in November and December of 1999, and accomplished the shooting of natural video images under varied lighting conditions with different color temperatures. For automatic operation, we made a prototype system that drives the filter according to the spectral change in lighting, detected by a small sensor installed on the camera. The basic operation of the system was verified, however; the correction error increases if the light source is out of sight of the sensor. Further improvement in the system is under way.