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 sets based on actual shot images. This image-based virtual studio can freely arrange small object parts (three-dimensional video components) in a high-resolution omnidirectional image (environmental video components) for the production of the surrounding scenery. In fiscal 2000, we enhanced composition function with which a prototype HDTV image-based virtual studio that integrates actual performer video image into the studio image was constructed.

One of the new composition function made feasible is the generation of an electronic studio video image that incorporates a performer’s image, based on the standing position information on the performer in a video taken by a sensor camera. It is also now possible to realize video expressions in which a performer appears to be handling 3-D video components. A function was also added to express appropriate a 3-D video component’s shadow as cast on the environmental video component. Including such new functions, the HDTV image-based virtual studio’s functionality and effectiveness were confirmed.

This prototype virtual studio was exhibited at BS Digital Fairs held in various locations in Japan, including the one at the Broadcasting Center, and highly evaluated.

In order to increase the degree of freedom of this virtual studio, we further conducted research on the potential of environmental motion components and adjustment of illumination conditions. Regarding the environmental motion component, we constructed a special camera base to enable the capturing of motion images in a portion of the omnidirectional still image with no shooting condition adjustment. Regarding basic experiments on illumination adjustment, we invented and verified a technique to estimate the direction of the sun, the principal light source, based on an RGB depth information outdoor scene image measured at a fixed point. This principal light source information resulted in a natural video image with reduced illumination.

As another application of these image-based virtual studio technologies, we also constructed “Virtual Scope,” a device to allow any arbitrary portion that was specified by a sensor to be enlarged for viewing. The system won favorable comments when utilized in classes at a university, in addition to its exhibition at the “NHK Learning Fair 2000.”
TVML is a language that describes TV programs using text-based scripts. A script written in TVML is interpreted by computer software, call the TVML Player, and a real-time program is automatically produced.

In fiscal 2000, nearly broadcasting quality output was obtained from a PC using the Windows version of the “TVML Player PRO,” developed during fiscal 1999. This system has realized a production environment with extremely high cost-efficiency, even for professional use. Additionally, enhancements were made to the expressive power of its language by means of a TVML language specification update. We have nearly completed development of a new man-machine interface for TVML program production “TVML Author,” opening the way to full-scale TVML implementation (inserted figure). The addition of a new external control mode to the TVML Player has made the interactive use of TVML possible. Developments in relation to TVML research include a chat-based automatic talk show generation system, “Talk Mania,” which employed an automatic program generation technique, was successfully exhibited at the BS Digital Fair. We also promoted work on a prototype studio simulator, research on automatic media conversion, and presentations and demonstrations within Japan and abroad, which all received much recognition.
In object-based coding, objects are extracted from a video image, such as a person or the background scenery, and encoded individually. This will realize flexible individual object manipulations (enlargement, reduction, replacement, etc.) at the receiver, in addition to increasing the efficiency of the data compression by applying the optimum compression scheme for each object.

In fiscal 2000, we promoted studies of shape coding for arbitrarily shaped objects, as well as studies of scene description methods for the layout information so that multiple objects might be composed into a single scene. We also constructed a prototype model (a demonstration system using software) that verifies the performance and function of the entire coding system. We contributed to the establishment of the MPEG-4 studio profile (a video coding standard for studio use), through proposals to the MPEG based on the results of our research.

In regard to object shape coding, improvements were made on a method that utilized Wavelet descriptors to encode object contour information (developed in fiscal 1999), and we developed a method to reduce the bitrate by exploiting similarity to contour of the previous frame. We also recently developed HHC (Hierarchical Huffman Coding) for shape-coding, in which hardware can be easily constructed.

For scene descriptions, we developed an expression format that describes various necessary functions for the broadcast program editing process, and a prototype player that reconfigures the scenes described in such a format as a motion video image (inserted figure). This development realized a description method not only for the object layout, but also for complex scene changes, including DVE (Digital Video Effects).

In studies to attain efficient recordings of editing materials, etc., we initiated the development of a high-definition video image coding method based on JPEG2000.
In fiscal 2000, using the Generic Description Method for Contents Exchanges (developed in fiscal 1999) that describes video signal formats and equipment functions, we constructed a distributed video player system as a prototype, and then verified the Description Method by experiments. In this system, a user can use video signals without being aware of their formats. In this system, video contents (non-compressed video signal, MPEG-1/MPEG-2 compressed video signal) and equipments (displays, decoders, video players and video editing equipments) are distributed on a network, and a user selects the desired content on a computer terminal (in the inserted figure). The equipment necessary for playback is automatically connected (in the inserted figure), and the selected content is reproduced on the display (in the inserted figure).

To realize broadcasting stations using network technology, we studied a mechanism that will automatically combine equipment over the network, a management system that will enable multiple tasks to be accomplished simultaneously such as on-air transmission and editing on the same network and a software construction method to accomplish the broadcasting equipment’s functions on a computer. We classified the basic functions of a broker, a repository, and other management components. We then launched development of an experimental system.

We also studied the video transmission protocols on IP network. When data transfer from another device occurs during a real-time video signal transfer using TCP (Transmission Control Protocol), quality degradation at the receiver will occur such as video freezing. In order to remedy this, we fabricated a transfer scheme that adaptively alternates TCP and UDP (User Datagram Protocol) observing the amount of buffered data at the receiver, and verified the effect by computer simulation.
Research and development continued on the digital watermarking technology for the protection of intellectual property rights for digital contents.

The following three requirements are unique to video digital watermarking for broadcasting: 1) good picture quality after embedding watermarking, 2) embedded information can be detected even after repeated encoding and decoding processes, such as by MPEG-2, and 3) embedded information can be detected even after some kinds of signal processing, such as enlargement, reduction, or cropping for video and low-pass filtering, equalized for audio. The scheme that the STRL examined has impressive characteristics, such as 1) digital watermarking embedded in baseband signals can comply with various compression schemes, and 2) no original image is necessary for watermarking detection, allowing the elimination of original image retrieval from a massive video and audio database.

In regard to the digital watermarking of audio data, we participated in an international contest concerning audio digital watermarking, sponsored by the Japanese Society for the Rights of Authors and Composers (JASRAC). The results of acoustic tests on sound quality deterioration and various experiments to measure robustness of the system, have evaluated the scheme proposed by the STRL and determined that it meets the accreditation standard. Verification assessment of this audio digital watermarking scheme, using various sound sources, confirmed that no sound deterioration was detected after application of the present digital BS broadcasting coding and decoding process, yet digital watermarking could be detected accurately. We also did research to enhance reliability through the application of an error correction coding scheme.