3—4
Network-based Cooperative TV Program Production System

H.Sumiyoshi Y.Mochizuki S.Suzuki Y.Ito Y.Orihara, N.Yagi M.Nakamura S.Shimoda

NHK Science and Technical Research Laboratories

{sumiyosi,motiduki,shoichi,yitou,ori,yagi,shimoda} @strl.nhk.or.jp, masashi@nic.nhk.or.jp

Abstract - A new DTPP (Desk-Top Program Production) system has been developed that enables multiple program producers (directors) working at different locations to collaborate over a computer network and prepare a single program for broadcasting. In this system, information is shared among users by exchanging data edited on non-linear editing terminals in program post-production work over a network in real time. In short, the new DTPP system provides a collaborative work space for producing TV programs.

The system does not make use of a special server for collaborative work but rather multiple interconnected editing terminals having the same functions. In this configuration, data at a terminal which has just been edited by some operation is forwarded to all other connected terminals for updating. This form of information sharing, however, requires that some sort of data synchronizing method be established since multiple terminals are operating on the same data simultaneously. We therefore adopt a method whereby the system synchronizes the clocks on each terminal at the time of connection and sends an operation time stamp together with edited data. This enables most recently modified data to be identified and all information on all terminals to be updated appropriately.

This paper provides an overview of this new collaborative DTPP system and describes the techniques for exchanging edited data and synchronizing data.

1 Introduction

The authors have been researching a next-generation broadcast production system that can be applied to the multimedia era. We call this system DTPP (Desk-Top Program Production) since it has been developed by adding multimedia functions to DTP (Desk Top Publishing), the software used in publishing, and applied to the production of broadcast programs [1]. In the same way that books can be prepared on a computer using DTP, DTPP computerizes the program production process so that programs can be prepared on computer (Fig. 1).

In addition to video and sound materials typically used in program production, DTPP also integrates other program information such as scripts, graphics, captions, oral drafts, shooting information, edited data, etc. This makes it possible to be more visually creative compared to past production methods and to create multimedia programs through systematic joining of information.

Modern program production is collaborative work in which a number of people are given various roles and work is assigned accordingly with the objective of preparing a single program. For example, it is typical for video editing, captions preparation, and sound effects production to be performed separately and later combined into one program (Fig. 2). In DTPP that aims to computerize this program production process, construction of a distributed collaborative work environment by networking is indispensable. There is also a growing need for a collaborative work environment that transcends time and space for situations where it is difficult for all members of

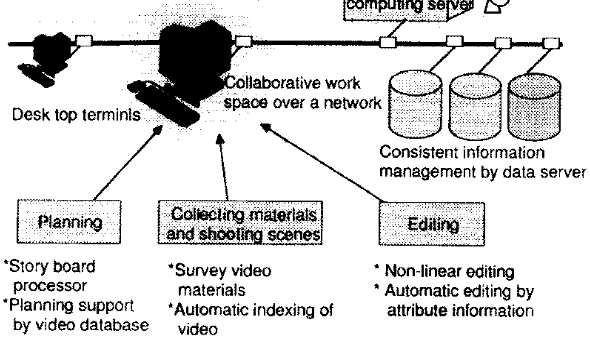


Fig.1 Program production support by DTPP

a project to sit down together at one place and at one time, such as in joint production with overseas companies. This kind of scenario is expected to occur more frequently in the future.

The authors have been performing studies on what kind of system would be most appropriate for facilitating collaborative program production work in a computer- and network-based environment. We have prepared a prototype system that supports post-production work, especially editing, where the division of work is quite clear, and have used it to perform experiments. This paper gives an overview of the prototype system and describes the data sharing technique and the synchronizing technique for program edition data in the collaborative work software that makes up the core of the system.

2 Present State of Collaborative Work Tools and Program Production

Although research into computer-supported collaborative work (CSCW) has actually been going on for some time, it can hardly be said that such systems are now in wide use. One reason for this is that no applications have been developed so far that can provide a more effective environment than gathering together at one place to perform collaborative work.

Systems that support collaborative work are now being developed [2] [3] , but these tend to center around people-to-people communication as in teleconferencing. As for research in CSCW for program production, there is a system by Gidney [4] that supports pre-production work in movie and television program production. It enables arrangements to be made for location shooting, casting, special effects, etc., using still pictures.

Modern broadcast stations make use of computers in a variety of fields as part of the program production process. Typical examples are the preparation of scripts

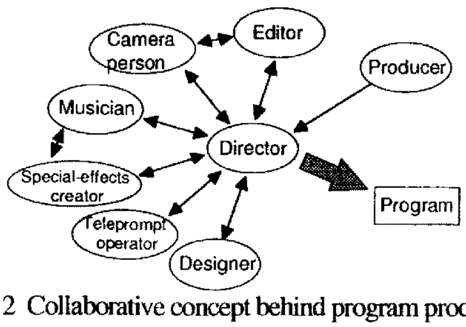


Fig. 2 Collaborative concept behind program production

and oral drafts using word processors and the use of electric teleprompting equipment and non-linear editing systems. The non-linear editing systems in particular first made their appearance around 1990 at about the same time that personal computers took a jump in performance, and NHK has been using off-line editing equipment on a test basis. At present, non-linear editing systems are not normally used for direct broadcasts due to poor picture quality, and instead broadcast-quality VTR video is re-edited and broadcast based on edited data from such systems. Although a system [5] has been developed that uses a high-speed network between non-linear editing systems, stores materials and results of editing on a server, and directly broadcasts from that server, it does not provide for collaborative work among multiple editing terminals.

Computer systems including non-linear editing systems now commonly used at broadcast stations are generally designed and introduced independently of each other, and not surprisingly, little progress has been made in standardizing data formats or interconnecting the systems by networking. As a result, information tends to be lost and its use inefficient. Moreover, because high-quality video is required for broadcasting, specialized video equipment is frequently used, but little consideration is being given to interfacing such equipment with computers.

3 Objectives and Features of System Construction

As one part of a DTPP system aiming to computerize program production work, we have constructed a system to allows users to perform post-production work, especially non-linear editing, in a collaborative manner from remote locations. The system enables multiple users to share not just video and sound data but all kinds of program edition data such as caption texts, superimpose timing data, in a collaborative work space supported by computers and a network, and makes possible broadcast-quality program editing.

As mentioned above, the system is not configured with a special server for collaborative work, and accordingly, terminal operation in a stand-alone configuration has been made essentially the same as that under a collaborative work environment. The system therefore combines the two main features of program production : individual (divided) work and group work, and supports program production in which individual work results contribute to a final program.

Conventional video editing and processing systems feature real-time editing functions of non-compressed video. In the DTPP concept of editing, however, materials are not directly edited, and instead, editable data (So called Script) independent of the actual materials becomes the object of editing, and to get the final program for broadcast, edit results based on the script and materials are generated. The prototype system expands on this idea and passes a script prepared with editing software to video processing equipment so that non-compressed video materials can be processed in real time.

4 System Configuration

The prototype system developed here consists of a user interface achieved by interconnecting functionally equivalent terminals without a special collaborative-work server, and real-time video processing equipment that processes broadcast-quality video (Fig. 3).

To be able to perform real-time non-linear editing of broadcast-quality video, one would need high-speed processing of multiple video data at a transmission rate of at least 270 Mbps for conventional TV. This rate, however, is unrealistic in terms of current computer and network performance. We have therefore adopted a system configuration that combines the two video systems below corresponding to non-compressed and compressed video to achieve real-time non-linear editing of broadcast-quality video.

1) Non-compressed real-time video processing system made up of a media server storing broadcast-quality video and sound and video processing equipment
2) User-interface system that enables video editing work to be checked using compressed video over operation terminals interconnected on a network.

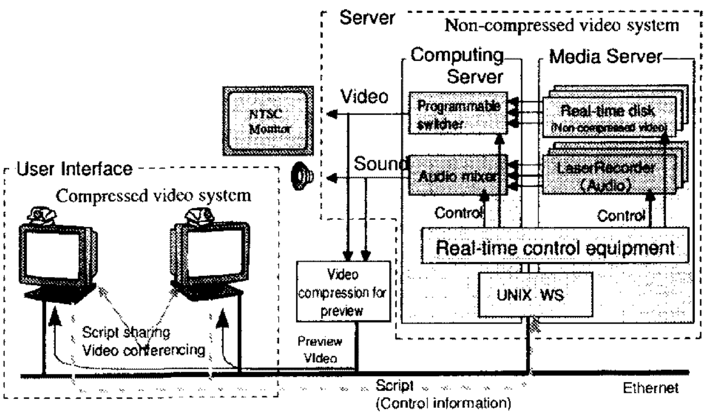


Fig. 3 System configuration

Non-compressed real-time video processing equipment is driven by Script prepared on the user-interface side while observing compressed video. Here, the connection between video and sound is defined in the Script using time data, which drastically reduces the amount of data to be transferred. This enables Script to be easily exchanged over current networks that have a narrow transmission bandwidth.

4.1 Real-time Editing and Processing Equipment

As shown in Fig. 3, real-time editing and processing equipment consists of a media server that stores non-compressed video and sound materials and a computing server that can process the non-compressed video and sound.

In the prototype system, the media server consists of a real-time hard disk that can record up to 26 minutes of moving pictures and a laser disc for recording sound. Both of these devices are controlled by time codes.

The computing server combines a programmable switcher, an audio mixer, and other server devices such as a character generator all configured for control by computer. These devices can be manipulated in a uniform way by the script prepared on the user interface.

4.2 User Interface

In the prototype system, a video and sound interface is provided via personal computers interconnected on a network. The computers are equipped with non-linear editing software having communication functions developed at NHK for collaborative work. This software enables the same editing operations to be performed whether the computer is being used in a stand-alone configuration or connected to a collaborative work environment.

The user interface enables video to be exchanged between users for communication purposes. In this implementation, video communication functions already possessed by the personal computers are expanded instead of using special hardware for transmitting video over the network.

5 Achieving Collaborative Work

Figure 4 (a) shows a common method for sharing information among distributed terminals in which a data server is provided for managing data at one location. Our system, on the other hand, does without a special server for collaborative work and instead interconnects multiple editing terminals on an equal basis with no master-slave relationship, as shown in Fig. 4 (b) . In this configuration, information is shared by sending data edited at one terminal to the other connected terminals. This method has the following advantages :

1) A simple configuration can be achieved since no special server is required ; and

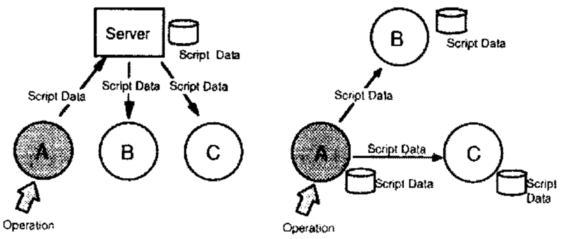
2) Operations can be performed even in a non-networked stand-alone configuration.

However, since each terminal holds shared data for editing, the information must be made consistent (synchronous) . The authors have therefore adopted a method whereby information that has been modified by some operation at a terminal is immediately sent to all other terminals for display on each editing screen.

5.1 Shared Information

The system developed here centers around linear editing work and has the editing functions listed in Table 1.

To be able to use a non-linear editing environment with such functions in common among multiple editors, information on the temporal arrangement of video and sound, switching effects from one scene to another, etc.,



(a) Server format (b) Parallel distributed format (this system)

Fig. 4 Configurations for sharing data

```
A00004,008,00:00:00:00,00:00:00:00,00:00:06:24,00:00:00:25,00:00:20:02,184,HD:008,00:02:14:14,00:00:27:21
B00008,BGM5,00:00:00:00,00:00:02:20,00:00:11:04,00:00:02:00,00:00:00:00,240,HD:BGM5,00:18:13:09,00:00:13:13
V00005,008,00:00:00:00,00:00:20:02,00:00:07:19,1,E00001,HD:008,00:02:14:14,00:00:27:21
T00002,XXT00002,00:00:00:19,00:00:00:00,00:00:03:16,"This is Comment."
S00007,XXS00007,00:00:01:18,00:00:00:00,00:00:04:20,"This is Subtitle.",1
A00005,010,00:00:06:16,00:00:01:25,00:00:12:08,00:00:00:27,00:00:00:00,240,HD:010,00:02:46:00,00:00:06:19
V00006,010,00:00:06:16,00:00:00:00,00:00:06:19,0,NULL,HD:010,00:02:46:00,00:00:06:19
E00001,dummy,00:00:06:16,00:00:01:03,1
```

Fig. 5 Sample Script

as needed for program editing must be shared for joint operation.

Our editing software features the use of text information (Script) to internally represent the data to be operated on. Fig. 5 shows a sample Script. Because all information needed for program editing like the temporal arrangement of video and sound is included here, sharing this edit information enables the results of operations made at any one connected terminal to be reflected in real time on the screens of all users engaged in the collaborative work.

The specific mechanism for information sharing will be explained later in this paper.

5.2 Communication Functions

1) Face-to-face Communication

A function that provides communication between users involved in collaborative work is indispensable. Our system is equipped with a teleconferencing function that can display on the screen a video image of other users collaborating on the program production. This provides an exceedingly effective means of communication since graphics, video and other materials possessed by a user can be displayed for viewing by all other users in addition to users' facial images.

2) Overlay Whiteboard

An overlay type of whiteboard is provided so that users in a collaborative work environment can convey their opinions and feelings to each other. This is not simply a shared whiteboard but rather a semi-transparent "sheet" that appears to be laid over the screen enabling users to write as desired on the non-linear editing workspace (Fig. 6.) A user can also use the mouse cursor to indicate places

Table 1 Editing Functions for Program Production

Type	Function
Video	· Clip placing · Time setting (start time, end time, length) · Switching effect between two video scenes
Sound	· Clip placing · Time setting (start time, end time, length) · Level adjustment · Fade in/fade out · Mixing of two sounds
Captions	· Placing and moving · Time setting (start time, end time, length) · Character input
Comments	· Placing and moving · Time setting (start time, end time, length) · Character input

on the screen to other users when pointing is more effective than words.

5.3 Information Sharing by Exchanging Messages

As explained above, this system adopts a configuration whereby multiple editing terminals are interconnected on an equal basis, as opposed to a master-slave relationship, as a method for sharing information in a distributed environment. In this method, the system sends data that has been operated on (edited) at one terminal to all other connected terminals.

The user interface software that we have developed has been given a function for exchanging control information between terminals in the form of messages as a mechanism for sharing information. Such messages may consist of connection procedures and Script. While connected, any non-linear editing operation (move video clip, change period, etc.) performed on a terminal's user interface results in the sending of a control message.

A message consists of an event ID and operation information. For example, for some operation performed on a video track, the event ID "UPDT" (update) indicating video track update and the video-track Script would be sent out as a message to all connected terminals. A terminal receiving this message would replace its data with the updated data and refresh the display accordingly. See Fig. 7.

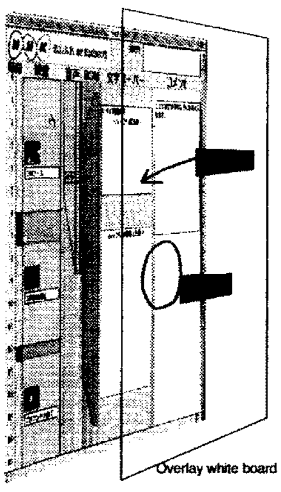


Fig. 6 Overlay whiteboard

5.4 Data Synchronization

When unifying information between multiple terminals by the above message-exchanging technique, the problem may arise as to what data to treat as the most current when multiple terminals exchange messages after performing different operations on the same object either simultaneously or within a very short time of each other. In particular, since processing by the user-interface software is event driven, messages received during an operation cannot be processed until processing for that operation completes. Various methods for solving these problems can be considered, but here we simply adopt the rule that the last performed operation is reflected at all terminals.

The clock on each terminal is set to 0 when establishing connections between multiple work terminals, and thereafter control messages that are exchanged between the terminals are stamped with the time of the corresponding operation before being sent out. If a message for unifying data is then received at a terminal, the time stamp of that message will be compared with the time stamp of the last data modification at that terminal. If the time stamp of the received message is older, then the data on that terminal is updated accordingly, and if the time stamp of the received message is earlier, nothing is done.

Time stamps are presently prepared in units of 1/60 of a second. In addition, to prevent as much as possible such control conflicts, video and sound transmission is performed on a track by track basis to ease congestion on the network.

6 Implementation and Experimants

We employed Apple Macintosh computers for the personal computers providing the user interface, and

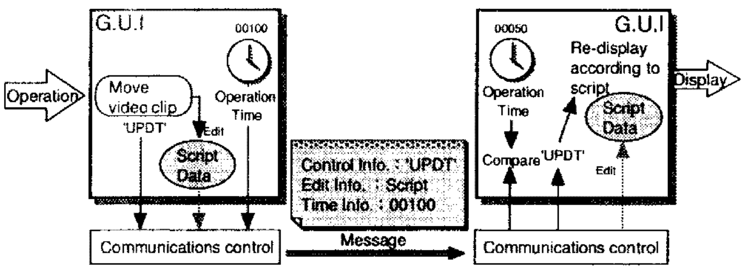


Fig. 7 Data sharing by messages

developed the interface software in-house on SuperCard (TM). We also used Apple QuickTime Conferencing for the communications interface to exchange video, sound and control messages between terminals, and developed an original command group for performing control from the graphical user interface (GUI). See Figs. 8 and 9.

Relatively simple collaborative program production experiments were carried out on the prototype system between our laboratories in Tokyo and ATR (Advanced Telecommunications Research Institute International) in Kyoto with using both a local Ethernet network and a remote ISDN (128 kbps) connection. (Fig. 10).

7 Conclusion

The experiments showed that distributed collaborative work involving non-linear editing in program post production could be performed with the prototype system. They demonstrated on the whole the feasibility of a collaborative work environment for future program production such as joint production by remotely located people.

The following describes problems encountered when using this system and measures proposed for solving them.

1) Insufficient network transmission capacity and CPU power

Even the 10-Mbps Ethernet network used in our experimental system configuration provided insufficient performance. For example, screen size to view another user's face had to be small, and the frame rate for displaying the results of processing was too low making it difficult to check edited video. And because network and CPU resources were often occupied with transmitting video, phenomena like congestion in exchanging control messages could be observed. To solve these problems, a high-speed network is imperative, but a practical network with sufficient capacity to transmit video data does not yet exist. Research is proceeding into bandwidth-control type of networks that dynamically allocate transmission capacity according to data amount, priority, etc.

2) Media Server Data Management

In the current system, data on the media server is not computerized, and as a result, information that must be managed together such as video clip time data has been distributed on the user interface side. Computer power must be significantly raised in order to store broadcast-quality video data as computerized data. To solve this problem, we are studying the construction of a media server that accommodates both compressed and non-compressed video simultaneously, and through computer processing, enables video to be treated as a single entity from the user interface side.

In addition to working on solutions to the problems encountered in the above experiments, the authors will be carrying out research with the aim of applying a collaborative work environment to the entire program production process. For example, we plan to analyze the

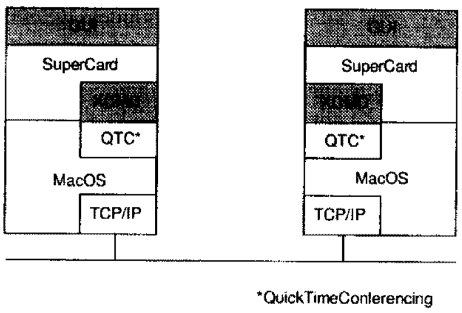


Fig. 8 User-interface software configuration

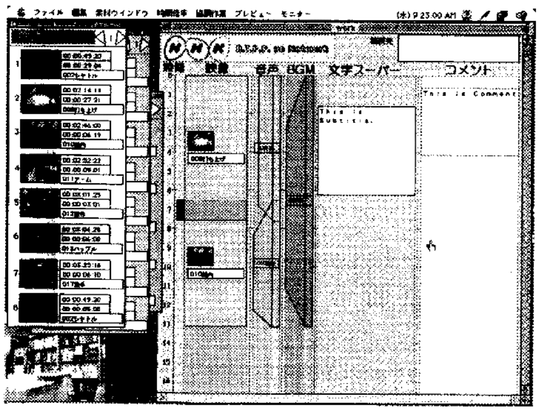


Fig. 9 Configuration of operation screen

Fig. 10 Experimental format

characteristics of program information and decide on a framework for all program information so that it can be handled as computerized data. We also plan to analyze the needs of users in detail to construct a more powerful and user-friendly DTPP system.

Reference

[1]K.Enami,"Desk Top Program Production (DTPP) - A Scenario of Sutudio Digitalization -", 18th International Television Symposium (June 1993), Montreux

[2]K.Watanabe et al, "Multimedia Desktop Conferencing System : MERMAID", Transactions of Information Processing Society of Japan,Vol 32,No.9

[3]H.Ishi, "Design of Realtime Groupware", Journal of Information Processing Society of Japan, Aug. 1993,Vol.34 , No. 8, pp1017-1027

[4]E.Gidney, A.Chandler, G.McFarlane,"CSCW for Film and TV Preproduction.",IEEE Multimedia,Fall 1994,pp16-26

[5]H.Owari, K Asari," Nonlinear Editing and Playback System", ITE Technical Report, Vol.19 No.58, PP 7-10

[6] Y.B.Kim, M.Shibata,"Content-Based Video Indexing and Retrieval -A Natural Languege Approach -",IEICE TRANS.INF.& SYST.,Vol.E79-D,No6 June 1996

[7] H.Sumiyoshi et al.,"Network-based Cooperative TV Program Production System", IEEE TRANSACTIONS ON BROADCASTING, Vol.42, No.3, Sep 1996