Metadata Management and Contents Archive System for Effective Use of Video Broadcast Assets

Susumu Yamamoto†, *Masanobu Manmoto, Kiyoshi Kurokawa, Hiroaki Udaka, Kazunari Yamayoshi, and Hiroyoshi Sasaki*

Abstract

The Contents Archive System developed by NTT Communications makes effective use of the automatic metadata generation, similarity-based image searching, and copyright management platform technologies developed by NTT Laboratories. It also provides flexible metadata management, which allows broadcasters to make full and efficient use of their video material assets.

1. Effective use of video assets by broadcasters

Diversification arising from the onset of digital broadcasting and the diversification of video distribution channels brought about by widespread use of broadband network access are steadily increasing the need for broadcasters to effectively use video material assets, such as broadcast programs, recordings, and other video contents. However, this is hindered by the following problems.

- It is difficult to search for and find a required video scene in a short time, because video content is archived on tape or other physical storage media and managed using metadata (information describing the video material) only for units of video titles.
- There are constraints on the quantity and locations of tape media and playback equipment for checking video.
- The labor cost for entering the metadata used for a video search is high.

There is also a wide range of different ways in which broadcasters make use of their video material assets, so various requirements regarding what kind of information is included in the metadata and the method used to search for it must be addressed in solving these problems.

† NTT Cyber Space Laboratories
Yokosuka-shi, 239-0847 Japan
E-mail: yamamoto.s@lab.ntt.co.jp

2. Contents Archive System

NTT Communications and NTT Cyber Solutions Laboratories have developed a contents archive system that solves the above problems to enable effective use of video material assets. An overview of the system is shown in **Fig. 1**. The system implements functions for a sequence of video archiving tasks, including (1) inputting and managing video content, such as complete video programs that have already been broadcast and recorded video materials, (2) assigning and managing metadata for the input video content, (3) searching and browsing the metadata, and (4) previewing and playing video materials obtained by searching. This system facilitates the reuse of archived video assets as materials for creating new programs within the broadcast station and as materials for sale to other broadcasters via a network using the search and preview functions. Its main functions are listed below.

- The required video material can be found quickly by using a search function that links the digitized video content and the metadata and a function for searching for scenes by similarity [1].
- The constraints on tape media and players are removed by video preview and play via an intranet, which makes the task of checking video more efficient.
- The cost of metadata input can be reduced by using voice recognition and other such automatic metadata generation functions [2], [3].

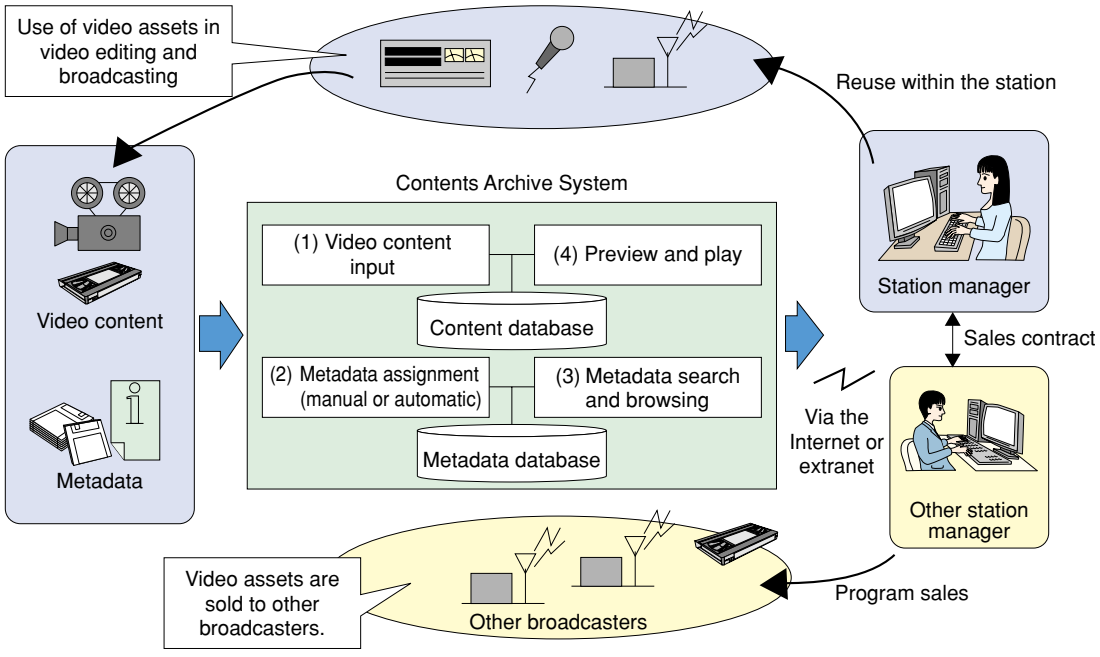

Fig. 1. Overview of Contents Archive System.

To meet various needs, we also considered extensibility of metadata items and search methods to allow customization of the system configuration to suit each application. The following sections explain metadata management and search in the Contents Archive System, the various video content management functions, and system extensibility.

3. Metadata management and search

3.1 Metadata structure

In this system, metadata is classified for management as described below in the units of the structural elements shown in **Fig. 2**, based on the task units of broadcasters.

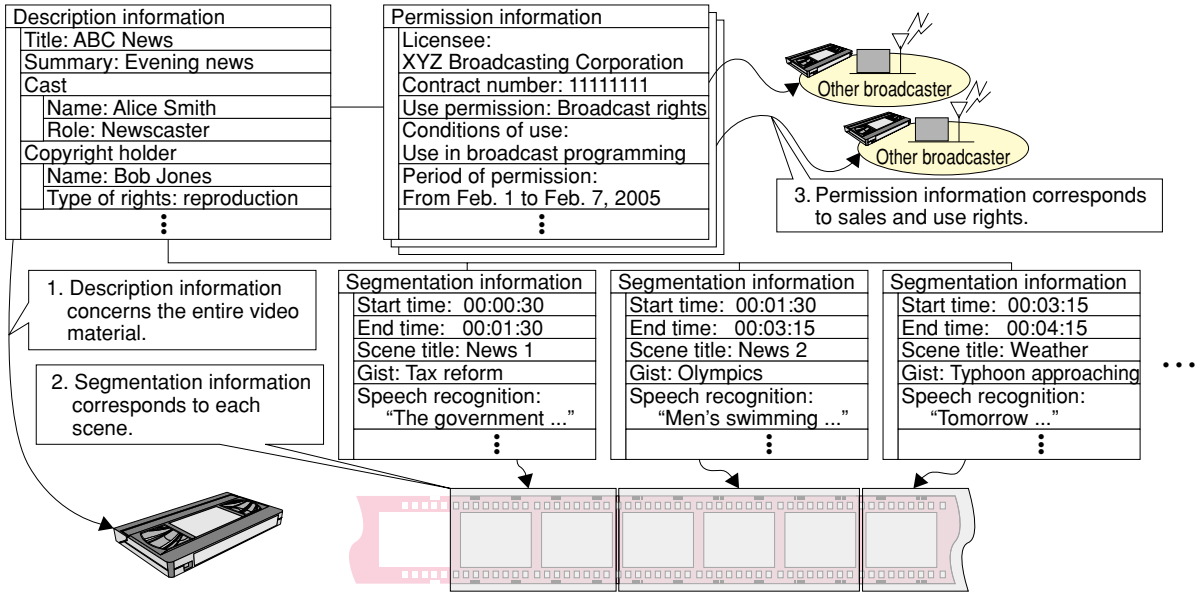

Fig. 2. Metadata structure.

1. Description information
   The system manages information concerning the content as a whole, such as the title and a brief description of the video content, the cast, and key words for searching. It also manages copyright information, including the name of the copyright holder and a description of the holder's rights.
2. Segmentation information
   The video content is divided into scenes, which are managed according to information such as the scene title, the music used for it, and the cast members that appear in the scene.
3. Permission information
   This indicates the usage rights concerning the video content. The system manages the licensee names, conditions of use, and other information about video content supplied to other broadcasters.

3.2 Task-conforming metadata life cycle

The various kinds of metadata are managed over a life cycle that conforms to the archive management work of broadcasters (**Fig. 3**).

1. Basic metadata for the registered content (mainly description information) is entered.
2. Segmentation information is generated automatically by the automatic metadata generation function.
3. Detailed metadata is entered in the generated segmentation data. A dictionary of people and music is provided to facilitate metadata entry.

In steps 1 to 3, browsing is restricted to the persons responsible for metadata entry.

4. When the metadata entry phase has been completed and the data has been opened to the creator, managers of other stations, or ordinary users, the publish function is executed to permit browsing. Individual metadata items can be made available or unavailable according to user classification.
5. The metadata can be searched by users. An example of the metadata access window is shown in **Fig. 4(a)**. In addition, the video content information for producing a program can be shared among users, such as arrangers and creators by mapping data called bookmarks to sets of useful scenes.
6. Permission information is managed for cases in which video materials are sold or rented to other broadcasters.
7. Metadata that is no longer needed is deleted using the metadata deletion function. Initially, the deletion is temporary: until it is deleted under the authority of the manager, deleted metadata can be recovered, which is useful if it was deleted by mistake or if the content is needed again.

3.3 Automatic generation of metadata and similar image search

The Contents Archive System uses the similar image search technology [1] and automatic metadata generation technology [2], [3] developed by the NTT Cyber Space and Cyber Solutions Laboratories to make the video archiving task even more efficient. Previously, the cost of metadata input was a major issue for a video archive construction. By applying automatic speech recognition to the speech of news announcers and other such persons, and automatic character recognition to the place names, person's names, and other telops* to be stored as metadata, the

* Telops: effects produced with a television opaque projector

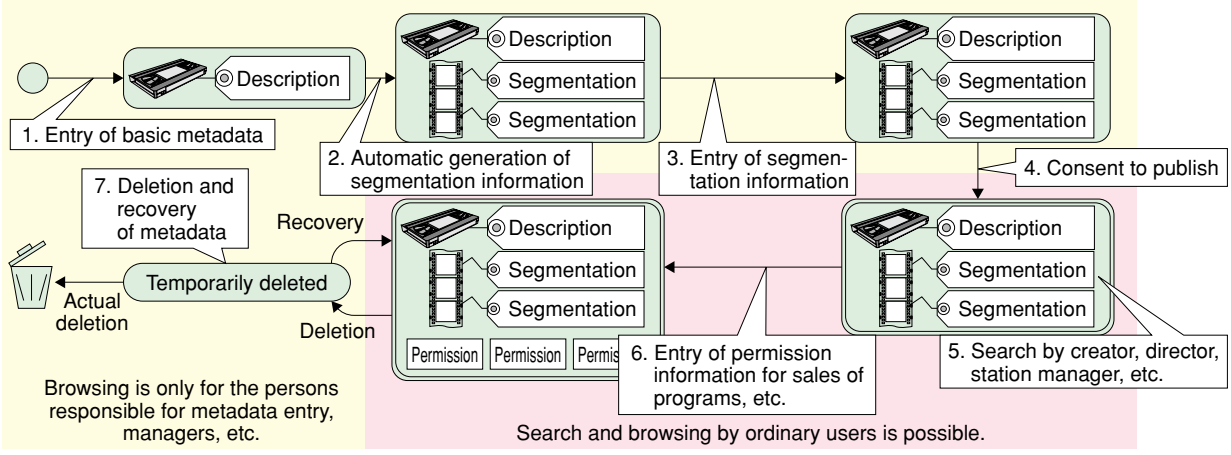

Fig. 3. Metadata life cycle.

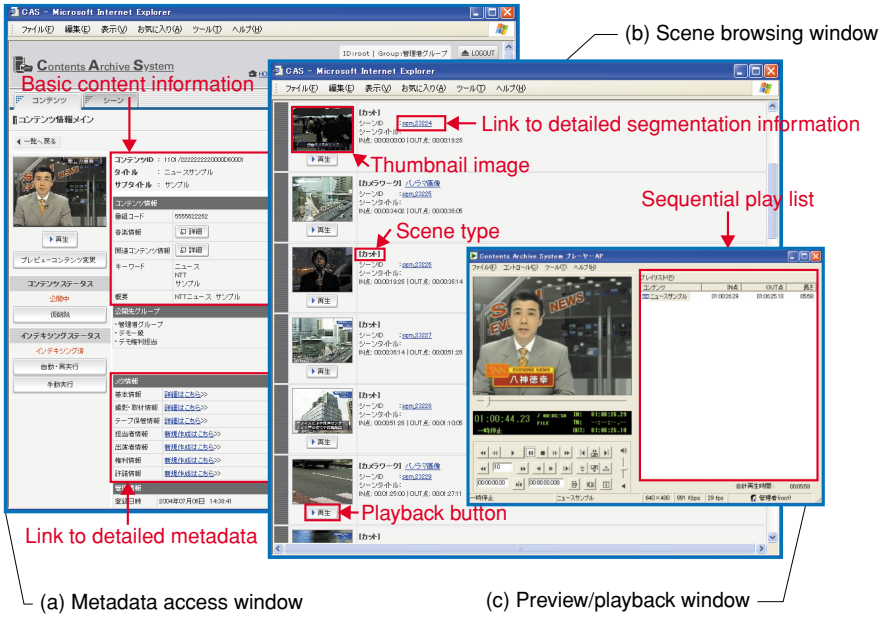

Fig. 4. Contents Archive System screen.

Contents Archive System can reduce the cost of metadata input. The recognition results of the automatic metadata generation function can also be used effectively as scene search conditions.

In addition, the video material is automatically segmented into scenes between cut points or telops displays, etc., and a thumbnail image is generated for each scene. The thumbnails can be searched visually to find the required parts within the video material. An example of a scene browsing window is shown in **Fig. 4(b)**. Six types of automatic scene detection are provided: cut point detection, camerawork detection, music section detection, speech recognition, telops recognition, and topic section detection.

The similar image search feature enables searching based on the similarity of the coloring or overall appearance of a thumbnail image. This feature is particularly effective when searching for landscape video and other such materials that suit the creator's concept better than video found by keyword search.

4. Video content management

4.1 Video content input

Video is acquired by connecting the system to a video tape recorder (VTR). Batch processing is also possible, in which the required sections on the tape are specified in advance by using the time code input and in/out point markers. This enables efficient content input and output. The acquired video is automatically encoded in WMV (Windows Media Video) format. Automatic metadata generation processing can be done at the same time as the video acquisition.

4.2 Preview and rough editing

The video content can be previewed on a personal computer terminal by following the directions on the metadata search result window. That allows efficient checking of the video without copying to video tape or needing a playback machine. Playback control using a jog dial to move through the video frames, etc., is supported, which facilitates the video checking operation. An example of the preview window is shown in **Fig. 4(c)**. In addition, consecutive playback of scenes selected from multiple stock video items according to a play list effectively produces a rough program edit that can be checked prior to the actual editing work, which is done using high-quality VTR equipment.

5. System architecture and extensibility

The system architecture of the Contents Archive System is shown in **Fig. 5**. The system comprises the following modules.

- GUI (graphical user interface) module for controlling the user interface
- User application module that controls processing to match the broadcaster's workflow
- Content management application module for managing video content
- Metadata management application module for

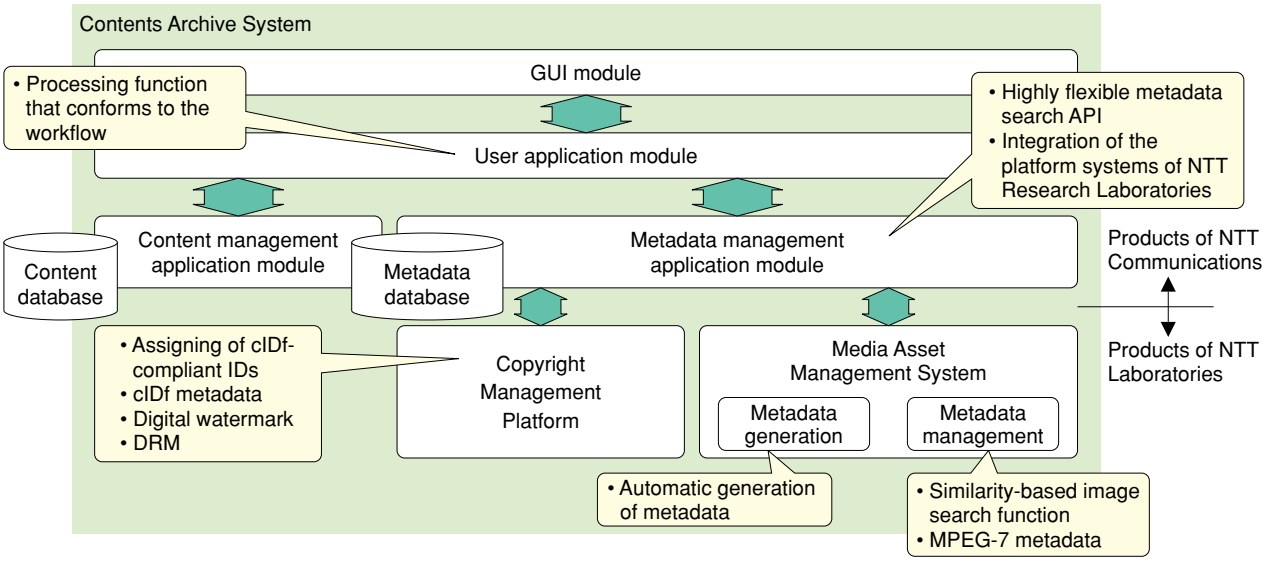

Fig. 5. System architecture.

managing metadata
- Platform systems developed by the NTT Laboratories (Copyright Management Platform [4], Media Asset Management System [5])

This system architecture allows the construction of an extensible and flexible system that can be reconfigured according to the needs of the broadcaster's customers.

5.1 Standard identifiers for video content distribution and sales

The copyright management platform can assign a content ID that conforms to the cIDf (content ID forum [6]) standard to serve as an identifier for publication. This provides extensibility for preventing unauthorized use of content when video content is distributed over the Internet, etc. in the future by allowing the incorporation of digital watermarking and digital rights management (DRM) functions of a copyright management platform based on this ID.

5.2 Mapping to metadata standards

Mappings are defined to map the metadata items designed to match the workflow of broadcasters to the MPEG-7 [7] and cIDf metadata standards. This allows easy extension of functions to allow cooperation with other systems that conform to various standards.

5.3 Flexibility for customization to meet the needs of customers

The system has a hierarchical architecture and is equipped with an internal application programming interface (API) that can perform searches on any metadata that is in XML (extensible markup language) format. This architecture can localize the scope of the effect of adding or changing functions or metadata items, thus reducing the development time for customization to meet customer needs. In addition, the system can be constructed with only the platform systems required to meet customer requirements.

6. Conclusion

Our system enables video material assets accumulated by broadcasters to be used in the following ways as well as for reuse within the studio and for sales of programs and stock materials between broadcasters: 1) video distribution by server-type broadcasting via the Internet or cell phones, 2) content digest creation by automatic metadata generation, 3) tapeless editing through digitization of stock video materials, and 4) customer support for checking video clips or video metadata of content that has been broadcast when viewers or sponsors have questions. In the future, we intend to introduce the archive system to broadcasters throughout the country and proceed with research and development aimed at the extensive use of video material assets.

References

[1] K. Kushima, H. Akama, S. Kon'ya, and M. Yamamuro, "ExSight:

Highly Accurate Object Based Image Retrieval System Enhanced by Redundant Object Extraction," Proc. of the First International Conference on Web-Age Information Management (WAIM'2000), LNCS, Vol. 1846, pp. 331-343, 2000.
[2] H. Kuwano, Y. Matsuo, and K. Kawazoe, "Reducing the Cost of Metadata Generation by Using Video/Audio Indexing and Natural Language Processing Techniques," NTT Technical Review, Vol. 2, No. 8, pp. 68-74, 2004.
[3] Y. Hayashi, S. Matsunaga, and Y. Matsuo, "Speech and Language Processing for Generating Content Description Metadata for Broadcast News," NTT Technical Review, Vol. 1, No. 3, pp. 62-65, 2003.
[4] H. Ohmura, K. Kurokawa, T. Yamada, and Y. Matsuura, "Digital content copyright management and protection platform: copyright distribution platform," NTT Technical Journal, Vol. 14, No. 10, pp. 16-19, 2002 (in Japanese).
[5] M. Tsunakawa, F. Konishi, and T. Nakanishi, "Media Asset Management (MAM) System for Efficient Content Management Using Metadata," NTT Technical Review, Vol. 2, No. 9, pp. 62-67, 2004.
[6] http://www.cidf.org/
[7] http://www.chiariglione.org/mpeg/

Susumu Yamamoto
Researcher, First Promotion Project, NTT Cyber Solutions Laboratories, (currently working in the Visual Media Communications Project, NTT Cyber Space Laboratories).
He received the B.E. degree in electronic engineering and M.E. degree in computer science from Keio University, Yokohama, Kanagawa in 1997 and 1999, respectively. He joined NTT in 1999 and is engaged in R&D of information security and digital content copyright management systems. He is a member of the Information Processing Society of Japan (IPSJ).

Masanobu Manmoto
Researcher, First Promotion Project, NTT Cyber Solutions Laboratories, (currently working in the Visual Media Communications Project, NTT Cyber Space Laboratories).
He received the B.S. and M.S. degrees in planetary material sciences from Yamaguchi University, Ube, Yamaguchi in 1997 and 1999, respectively. He joined NTT in 1999 and is engaged in R&D of enterprise system for broadcasters and digital content copyright management system. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan.

Kiyoshi Kurokawa
Senior Research Engineer, First Promotion Project, NTT Cyber Solutions Laboratories, (currently working in the Visual Media Communications Project, NTT Cyber Space Laboratories).
He received the B.E. degree in computer science and M.E. degree in electronic engineering from Kyushu Institute of Technology, Fukuoka in 1988 and 1990, respectively. He joined NTT in 1990 and researched database design, information resource management, and data visualization. He is currently engaged in R&D of digital content copyright management systems. He is a member of IPSJ.

Hiroaki Udaka
Chief, Innovative IP Architecture Center, NTT Communications Corporation.
He received the B.E. degree in resource chemistry from Ehime University, Matsuyama, Ehime in 1991. He joined NTT in 1991 and engaged in development of the facsimile communication network service (F-net) and enterprise systems for video content.

Kazunari Yamayoshi
Innovative IP Architecture Center, NTT Communications Corporation.
He received the B.E. degree in mechanical engineering from Kobe University, Kobe in 1991. He joined NTT in 1991 and engaged in the development of frame relay services and OCN services.

Hiroyoshi Sasaki
Innovative IP Architecture Center, NTT Communications Corporation.
He received the B.E. degree in information engineering from Iwate University, Morioka, Iwate in 1993. He joined NTT in 1993 and engaged in the development of a videotex network service (CAPTAIN) and electronic commerce services.