Framework for Supporting Metadata Services

Mitsuaki Tsunakara†*, Ryoji Kataoka, and Masashi Morimoto*

Abstract

Metadata-sharing businesses have been attracting considerable attention recently. These include high-speed search services based on metadata that describes various object attributes and service-linking using metadata as a means of conveying information. This article introduces basic metadata creation, sharing, and application technologies for such services.

1. Basic technologies for processing metadata

Services that utilize metadata such as multimedia-content delivery services and broadband television services using electronic program guides (EPGs) are beginning to spread. The following methods for using metadata can be considered.

- Describe and utilize the attributes of targeted contents, users, and services to facilitate efficient and high-speed navigation through large volumes of content and among users.
- Let services share user and content attributes, access logs, and other types of metadata and form inter-service links to provide enhanced value-added services.

Basic technologies for constructing services that use metadata in this way can be generally classified as follows (Fig. 1).

- Metadata creation technology
  Enables multiple entities (content providers, service providers, users, *etc.*) to create metadata efficiently for various kinds of contents.
- Metadata sharing technology
  Manages metadata and manipulates (converts, integrates, extracts, *etc.*) it in accordance with service systems.
- Metadata application technology
  Enables service users to use metadata to facilitate efficient searches for content and lets them perform genre searches, content recommendations, similarity searches based on feature quantities, and other types of advanced searches.

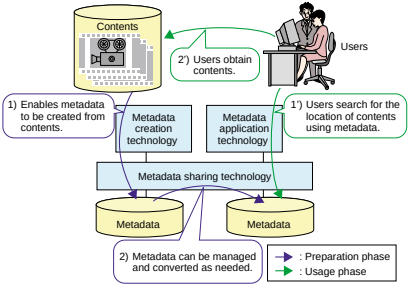

Fig. 1. Basic metadata technologies.

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

2. Metadata creation technology

Needless to say, metadata associated with content must be prepared before services utilizing it can be made. It must be possible to create metadata that can represent the content in an appropriate and diversified manner, to create it easily and efficiently, and to create it in accordance with how it will be used. To meet these needs, NTT Cyber Space Laboratories have been researching and developing basic "content-metadata creation technology," [1] as shown in Fig. 2.

2.1 Creation of diversified metadata

Content titles and summaries are typical examples of content-related metadata. However, when you attempt to produce metadata that appropriately represents the makeup of a video program consisting of multiple scenes (video segments), using only these types of metadata, which describe the overall video, is simply not sufficient. For this reason, the "metadata editing program" and the "contents archive builder" provide an indexing function that automatically extracts representative frames and scene makeup that very accurately represent segment information based on changes in images, sound, and music in a video [2]. This makes it possible to produce metadata that can represent content in a more suitable and diversified manner. Various standards of notation exist for metadata of this type such as the metadata framework standard for server-type broadcasting in the TV-Anytime Forum [3].

2.2 Efficient metadata creation

In the production of scene-related metadata as described above, a huge amount of labor would be required if it were all done manually. In addition to an automatic extraction function using a plug-in type of indexing engine, the metadata editing program provides other functions to help make metadata production easier and more efficient, such as an import function for using existing metadata and a production interface for viewing and assigning metadata [2], [4]. Automatic synthesis and layout of panoramic images from videos are superior examples for efficient metadata viewing. In addition, the contents archive builder provides Web-based functions for efficient metadata editing over the network: a metadata viewing/editing function, a metadata-creation support function using super-imposed text recognition technology, and a distributed-parallel indexing function.

2.3 Metadata creation according to use

We can imagine situations in which metadata outside of standards might be needed in accordance with its use as intended by a service provider. The "properties editor" provides a function for extending meta-

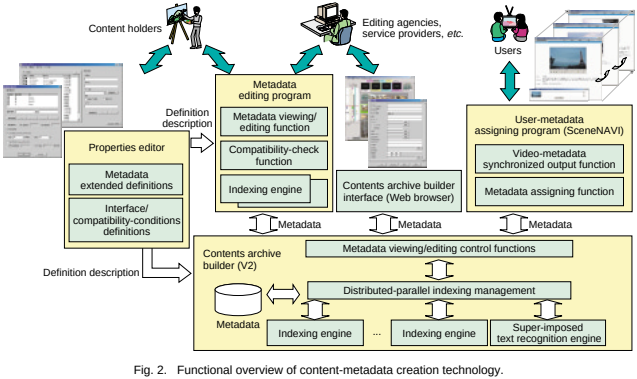

Fig. 2. Functional overview of content-metadata creation technology.

data definitions as needed and outputs definition descriptions in Extensible Markup Language (XML) schema format. Moreover, the "metadata editing program" provides a function for customizing the production interface according to such extended definitions and a function for performing compatibility checks on assigned metadata. These functions enable appropriate and efficient metadata production while maintaining extendibility.

The "user-metadata assigning program" provides a function for obtaining user-annotated metadata about a scene by synchronizing the playback of the scene and the display of annotated metadata. It enables easier annotation according to a user's specific interest about the scene. This novel function makes it possible to assign metadata from viewpoints other than those of the content holders and service providers.

3. Metadata sharing technology

Much metadata has been standardized according to the application field and intended usage, and there are even metadata specifications unique to certain services. The role of metadata sharing technology is to enable efficient and flexible processing of such metadata. To this end, NTT Cyber Space Laboratories are researching and developing an "XML Sharing Platform System" (Fig. 3). The basic technologies of this system, which can handle all types of metadata by XML, cover metadata management and conversion.

3.1 Metadata management technology

We have developed an ultrafast XML search engine called LiteObject, characterized by an XPath interface, a high-speed multi-dimensional index [5], a memory-resident database, and an object relational data model. LiteObject's XPath search processing is up to ten times as fast as commercially available DBMS products.

We are also in the process of developing a search/update language based on XQuery, which is now being standardized as an XML query language by the World Wide Web Consortium (W3C) as well as an XML data management function targeting Java Database Connectivity (JDBC) information sources and the LiteObject engine. These products will facilitate metadata management by enabling metadata to be tuned according to system applications and environments while maintaining service portability through an already standardized uniform interface.

3.2 Metadata conversion technology

When two or more systems share data, it is necessary to control data conversion, transfer, transaction,

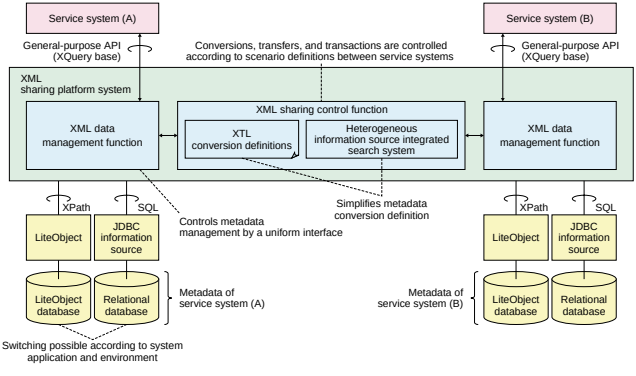

Fig. 3. Functional configuration of XML sharing platform system.

and other processes. We are developing an XML sharing control function to control these processes through scenario specification. For example, you can specify a scenario for converting metadata in TV-Anytime format to MPEG-7, transferring the converted metadata to a system targeted for sharing, and storing it directly in a database on that system. Another scenario might call for transferring metadata directly to another system without conversion and passing it to a program running on that system. In short, scenarios can be used to specify various types of control according to the metadata in question.

Our system supports XML Transfer Language (XTL) [6] as a metadata-conversion definition language. Although XSLT (Extensible Stylesheet Language Transformations) has come to be widely used for XML conversion, its complicated grammar makes XSLT difficult to learn, and while other conversion techniques exist, their functions are insufficient, among other problems. It is for these reasons that XTL was developed. XTL makes it easy to define metadata conversions using about 1/10 as much code as for XSLT and describe conversion functions at a level equivalent to or better than those of XQuery.

We are also developing an XML conversion engine using our in-house-developed heterogeneous information source integrated search system. This engine enables easy and flexible conversion (e.g., item-name conversion, structure conversion, data-expression format conversion, and multiple XML binding) by specifying only the format after conversion. Storing conversion definitions needed for development by many services in the XML sharing control function should help shorten the development time for linking service systems.

4. Metadata application technology

From the user's standpoint, the purpose of examining metadata assigned to content is not so much to check what certain content is all about but rather to find desirable content. However, examining metadata to check the makeup of different items of content could require an awful lot of effort if the number of content entries were large. Accordingly, to discover desirable content quickly using metadata, we need technology for filtering metadata according to user needs. NTT Cyber Space Laboratories are preparing basic technologies to meet this need and are also researching and developing a system called "MetaConcierge" [1] to simplify the construction of services using content metadata (Fig. 4).

4.1 Metadata retrieval

To enable a user to find desired content with ease, an effective technique is to store the content metadata in a database and to enable searching by keyword, directory, or the like. In general, however, the number of search results increases as the volume of contents becomes larger, so such a simple technique alone cannot easily provide the desired content. The MetaConcierge system helps to solve this problem by combining keyword and directory searches with a function that presents contents according to user preferences in a prioritized or selective manner (recommendation search). With this system, a service provider can specify contents that can be accessed or the range of directories to be presented for user groups segmented by sex and/or age. Recommended contents can also be chosen based on the access history of each user group.

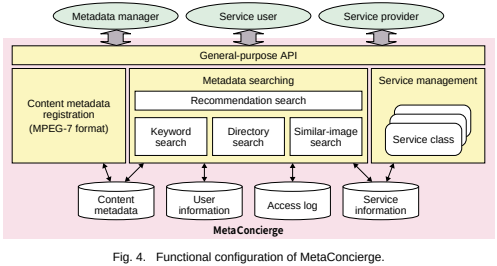

Fig. 4. Functional configuration of MetaConcierge.

On the other hand, it is not necessarily easy to describe content that one wants to see by an appropriate keyword. To enable content searches to be performed using ambiguous conditions such as "a scene with this kind of feeling," MetaConcierge enables search conditions to be input in an intuitive manner using images based on similar-image search technology, which enables the similarity between images to be determined based on feature quantities such as colors and shapes (low-level metadata) extracted from image data. Although high-speed processing is generally difficult for such technology, MetaConcierge achieves high speeds using an image search engine [7] developed by NTT Cyber Space Laboratories. A new multi-dimensional indexing technology called C-tree can search for 20 similar images in an image database containing about 100,000 objects within only 0.82 seconds. If low-level metadata extracted from representative images in a video is registered in this engine, a user can specify an image and then retrieve scenes having a similar feeling, something that could not be done in the past with keywords.

4.2 Service management

The best content metadata search functions to combine, or service menus to configure, differ from one service to the next. Accordingly, a system that provides metadata services requires functions for configuring service menus in a flexible manner according to service requests. MetaConcierge addresses this need by providing functions that enable various search functions to be managed in groups of service classes and that allow the service provider to edit these classes as necessary to configure service menus with ease. It also provides a function for packaging multiple items of content so that content can be retrieved in units of packages.

5. Conclusion

Research and development proceeds on basic technologies to support metadata services, whose need is now being felt in a wide variety of business situations. For here on, we will focus our efforts on promoting metadata sharing businesses and on achieving high-speed, multi-functional, and automatic metadata production.

References

[1] Y. Umemoto, A. Nambu, M. Abe, S. Seo, T. Hounishi, N. Kobayashi, M. Sakai, and T. Sadakata, "Overview of the Content Distribution Platform," NTT REVIEW, Vol. 14, No. 4, pp. 4-22, 2002.

[2] H. Kuwano, Y. Taniguchi, H. Arai, M. Mori, S. Kurakake, and H. Kojima, "Telop-on-demand: Video structuring and retrieval based on text recognition," IEEE Int. Conf. on Multimedia and Expo 2000, New York, U.S.A., pp. 759-762, July 2000.

[3] http://www.tv-anytime.org/

[4] Y. Taniguchi, A. Akutsu, and Y. Tonomura: PanoramaExcerpts, "Video Cataloging by Automatic Synthesis and Layout of Panoramic Images," Vol. E83-D, No. 12, pp. 2039-2046, 2000.

[5] Y. Sakurai, M. Yoshikawa, S. Uemura and H. Kojima, "The A-tree: An Index Structure for High-Dimensional Spaces Using Relative Approximation," Proc. of the 26th Intl. Conf. on Very Large Data Bases (VLDB), Cairo, Egypt, pp. 516-526, Sep. 2000.

[6] M. Onizuka, "XTL: An XML transformation language," Proceedings of International Conference on Extreme Markup Languages, Montreal, Canada, pp. 167-180, Aug. 2001.

[7] K. Kushima, H. Akama, S. Konya, 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), Shang Hai, China, pp. 331-343, June 2000.

Mitsuaki Tsunakawa
Senior Research Engineer, Information-base Project, NTT Cyber Space Laboratories.
He received the B.S. degree in mathematics from Tsukuba University, Japan, in 1990. He joined NTT Communications and Information Processing Laboratories in 1990. He has been researching DBMS technology and the integration of heterogeneous information sources. He is currently engaged in research and development of metadata exchange technologies. He is a member of the IEICE (Institute of Electronics, Information and Communication Engineers).

Ryoji Kataoka
Senior Research Engineer, Supervisor, Information-base Project, NTT Cyber Space Laboratories.
He received the B.E. and M.E. degrees in electronic engineering from Chiba University, Japan, in 1985 and 1987, respectively. He then joined NTT Communications and Information Processing Laboratories, where he engaged in research on transaction processing and multimedia databases. He is currently engaged in research and development on multimedia asset management technologies. He is a member of the IEICE and IPSJ (the Information Processing Society of Japan).

Masashi Morimoto
Senior Research Engineer, Supervisor, Information-base Project, NTT Cyber Space Laboratories.
He received his B.E. and M.E. degrees in information engineering from Kyoto University, Kyoto, Japan in 1986 and 1988, respectively. He joined NTT in 1988, working on image analysis, image recognition, video analysis, and human-computer interaction via networks. From 1996 to 1997 he was a visiting researcher at Stanford University working on image retrieval algorithms. He is a member of IPSJ, IEICE, and ITE (the Institute of Image Information and Television Engineers).