

---

# マルチメディアコンテンツ内容記述ツール -MovieTool- XML-SCHEMA BASED MULTIMEDIA CONTENT DESCRIPTION TOOL -MovieTool-

國枝 孝之\*      脇田 由喜\*      小山 毅\*  
Takayuki KUNIEDA      Yuki WAKITA      Takeshi KOYAMA

---

## 要 旨

MovieToolは、XMLベースのマルチメディアコンテンツ内容記述ツールで、ビデオ映像に対しマルチメディアコンテンツの内容記述インターフェースとして国際標準化されたMPEG-7準拠のコンテンツ記述ファイルを生成できるツールである。その特徴として、XML-Schemaを動的に読み込みマルチメディアコンテンツの記述スキームとマッピングを行う機能を有しており、XML-Schemaで定義されたメタデータ標準であれば、MPEG-7に限らずB-XMLなどにも対応でき、柔軟性と多様性を備えている。

## ABSTRACT

MovieTool is a Description tool for multimedia content with XML oriented representation. It can generate a content description file in the format defined MPEG-7, which is known as the multimedia content description interface. We adopted XML-Schema dynamic mapping mechanisms for the multimedia content description schemes. As a result, MovieTool can provide the flexibility and general versatility for various meta-data standards not only MPEG-7 but also B-XML etc.

---

\* 画像システム事業本部 ソフトウェア研究所  
Software Research Center, Imaging System Business Group

## 1. Introduction

The accumulation of multimedia data has created a strong demand for multimedia archiving systems. Content-based retrieval schemes are being standardized using several different approaches. Progress in storage technology and CPU power enable easy browsing of such multimedia content. However, only a few ways enable searching and selecting contents easily using user-friendly interfaces. Along with the progress in web- and XML-technology, several structure description languages have been proposed (e.g., SMIL, MPEG-7, BXML). MPEG-7 [1], a multimedia-content description interface, was starting to be standardized. We focus on contributing to the standardization process and started to develop the description tool.

## 2. MovieTool

We have started developing a multimedia-content description tool named MovieTool. MovieTool can read and write MPEG-7 formatted content-description files. First, this tool is used to either manually or automatically construct a structure. It can apply automatic segmentation methods using a content-based analysis. Then, a key frame is selected, content-based features are extracted, and a syntactic structure is generated. After that, attribute values can be input into each segment as well as into the overall content. MovieTool provides many attributes that are defined in MPEG-7 description schemes. It is one component of a multimedia content retrieval system we are developing. In this system, a content description value is registered to a database in an MPEG-7 file. In the easy web interface, attribute values are set as a condition, and then the required scenes or contents are displayed. They can be selected and played in specified order. To develop this system, we adopted XML-Schema dynamic mapping mechanisms for MovieTool.

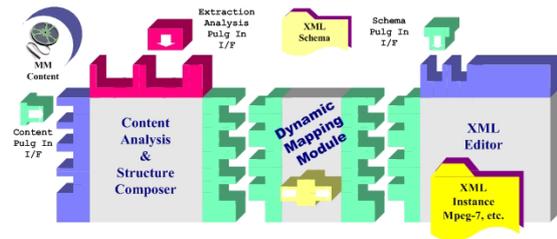


Fig.1 Module Structure of MovieTool

In MPEG-7 description, the description content is conceptually classified into two parts: syntactic, which is generated automatically or semi-automatically using existing extraction and analysis methods, and semantic, which demands human thought processes. To develop the description tool, the latter must be considered sufficiently. Taking this into account, we strictly separated the functionality of the modules. Fig.1 is a perspective representation of the module structure of MovieTool.

### 2-1 The Structure Composer

The Structure Composer module builds an audiovisual structure. Fig.2 shows the structure composer window. It can be used manually or automatically, and it can apply some automatic segmentation methods using content-based analysis. A key frame is selected, content-based features are extracted, and a syntactic structure is generated. All data are represented using the package-segment model (PS-Model) [2], [3]. The PS-Model is a data model that principally composed by package and segment nodes. A segment represents a cut, scene, or a logical duration in the content stream. A package plays a role in generalizing a segment set. The PS-Model structure resembles a tree structure. Each node can have an attribute node, which consists of some attributes of the parent node as well as user-defined supplemental information. This module also provides two plug-in interfaces: a content plug-in and an extraction-and-analysis plug-in. Content plug-in is used for adaptation of various formatted content such as MPEG-2, MPEG-4 and AVI. Extraction-and-analysis plug-in provides the interface where user can define new feature extraction modules. Therefore, the user can easily adopt various formatted media and new feature extraction. These functions enable a flexible application.

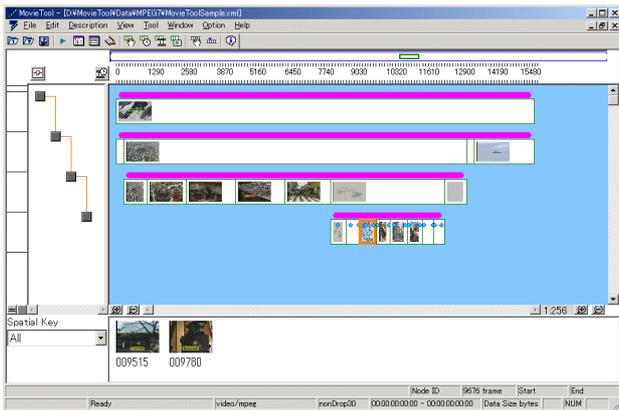


Fig.2 Structure Composer Window

## 2-2 The XML Editor

XML Editor module can input and edit body and attribute values in each segment as well as in the overall content. MovieTool provides all elements defined in MPEG-7 description schemes. This editor provides not only the basic functions for XML editing but also XML-Schema based validation. MovieTool allows the user to see and confirm description instances, as shown in Fig.3. On the right is the XML description, on the left is the schema which defines the XML-Schema syntax. The scenes displayed in the Structure Composer Window (Fig.2) and with the descriptions in the XML Editor Window (Fig.3) are correlated each other. For example, if the user selects a scene in the Structure Composer Window, the corresponding description is highlighted in the XML Editor Window. Conversely, selection of a description causes the corresponding scene to be highlighted.

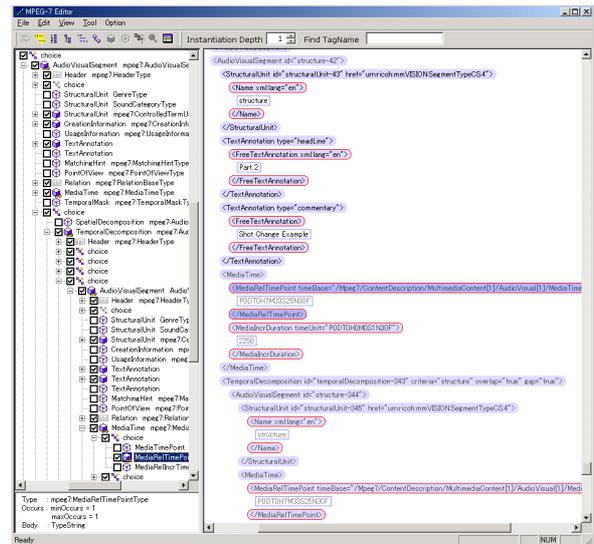


Fig.3 XML Editor Window

## 2-3 The Removable Connector and Dynamic Mapping

This module corresponds to meaning of XML-Schema. It provides bridging functions between the XML element and the PS-Model node. This module provides the dynamic mapping mechanism mentioned in the next paragraph for each XML-Schema definition. The removable-connector module plays a key role in the dynamic-mapping function. Here, we focus on the dynamic-mapping mechanism between the element defined by the XML-Schema and the node used in the structure-composer module. Fig.4 shows the mapping relations between the MPEG-7 and PS-Model spaces. The temporal decomposition element in MPEG-7 corresponds to the package node in the PS-model and the audio-visual segment corresponds to the segment node. Because these relations can be modularized as a removable connector, the tool can be easily adapted to various schemata.

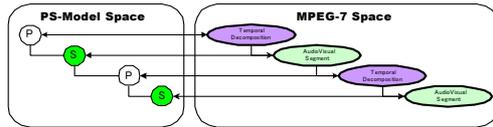


Fig.4 Mapping Relation between PS-Model and MPEG-7

### 3. Conclusion

The following window images are real-world tool interfaces. The left window shows the structure composer and the right one the XML editor. In reality, the removable connector is located between them, so the user can seamlessly change from applying automatic analysis to human annotation to describe the content. With the dynamic mapping mechanism, MovieTool provides flexibility and general versatility for use with various meta-data standards, not only in MPEG-7 but also in B-XML, and others. We plan to test these characteristics in the future.

#### References)

- 1) The MPEG Home Page: <http://www.cseit.it/mpeg/>
- 2) Kunieda T. and Wakita Y., "Package-Segment Model for Movie Retrieval System and Adaptable Applications," IEEE International Conference on Multimedia Computing Systems Proceedings," Volume.2, Florence, Italy, June 1999, pages 944-948.
- 3) Wakita Y., Kunieda T., Takahashi N., Hashimoto T., and Kuboki J., "Extended Package-Segment Model and Adaptable Applications," 6th International Workshop IDMS'99 Proceedings. Toulouse, France, October 1999, pages 163-176.