Visualization of Video Content in Movie Generation Research

SHEN Jinhong Seiya MIYAZAKI Terumasa AOKI Hiroshi YASUDA
Faculty of Engineering, The University of Tokyo

1. Introduction

Nowadays' technologies enable us to access digital video streams (such as through the World Wide Web) and display videos (most of them are compressed) on personal computer. To manipulate video documents for video clips reuse, it is necessary to make indexing for video contents.

Varied indexing schemes have been put forward for different video retrieval goals. We are developing such a software tool DMP (Digital Movie Producer) that can interpret the textual screenplay into digital movie with effects of computer animation, real images and their simple composition [1, 2]. In this paper, we will describe how to apply current advanced video retrieval techniques and filmmaking theory to build DMPVR (DMP Video Retrieval), a subsystem of DMP, focusing on design multi-modal video indexing.

2. Ontology

Filmmaking techniques include four aspects:

- *mise-en-scène* (*what to shoot*) which involves setting, lighting, figures,
- *cinematograph* (*how to shoot it*) which involves camerawork – camera angle, camera movement and camera distance,
- *montage* (how to present the shots), e.g., fade in/out, parallel editing and
- *sound edition* (how to present the sounds), e.g., dialog, music, background sound from film theory.

In figure 1 those dark squares indicate the main contents should be extracted from video. They will be encoded in XML (extensible Markup Language) format that is called DMPML (DMP Markup Language) in our system.

3. Indexing Approach in DMPVR

(1) Definition of video indexing: some kind of information is extracted from video document and stored in its database in order to be queried later.

(2) Three kinds of approach: f*ully manual video content indexing* is a very time-consuming procedure so that automated indexing approach is required if possible. *Fully automatic semantic annotation* is still impossible with current VR technology. For the content that cannot be annotated automatically, *computer aided content indexing* may be chosen as a feasible way for complement. Their detailed explanations are showed in table 1.

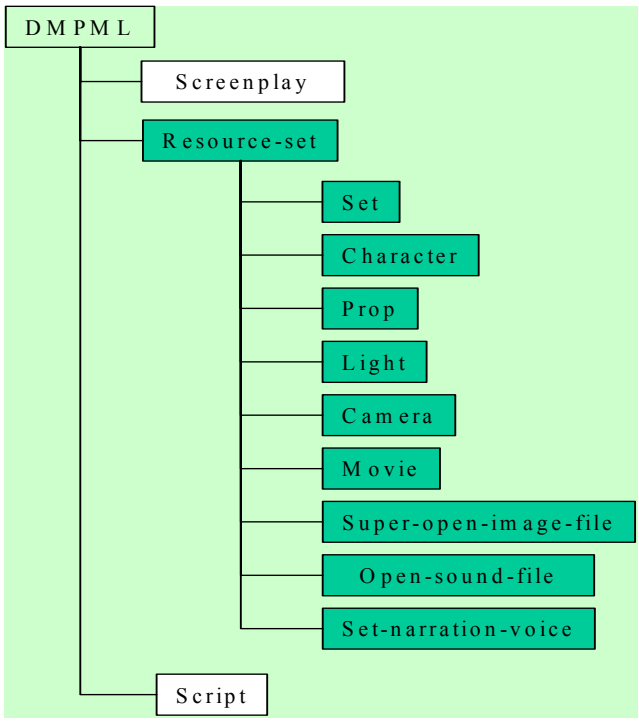

Figure 1. Contents Items Used in Filmmaking

Approaches	Tasks
Automatic annotation (figure 2)	1. Segment (vs. montage): Scene →Shot →Keyframe. 2. Semantic feature extraction (vs. mise-en-scène): –Set, character, and prop in specific domain; –Some camerawork like pan; –Sound: music, dialogue, etc. 3. Event extraction (vs. mise-en-scène & sound edition): e.g., sport type.
Computer aided annotation (figure 3)	User provides indices through interface of the software detector.

Table 1. Approaches of Video Content Indexing

(3) Our design: figure 2 and 3 demonstrate the architecture of our DMPVR video retrieval system.

Knowledge base KB contains domain knowledge like football game rules.

Inference engine The inference engine is used to reason with both the domain knowledge and data from video document and other information from the user.

Intelligent detector The virtual detector is responsible for the auditory-visual indexing aspect of video dependent on domain knowledge in KB. He will at least give domain terminals list if cannot give the required video contents by reasoning.

(4) System features: A suitable multi-category video modeling and multi-modal query mechanism with multi-modal video indexing were constructed based on MPEG-7 from the perspective of film director.

Multi-category video modeling By taking advantage of ontology as mentioned in the second section, it is possible to facilitate conceptual search.

Multi-modal query mechanism Visual content may be conveyed in both narrative (language) and image.

Multi-modal video indexing Systems that combine visual features, sound, text as well as structured descriptions can get powerful retrieval. We will use textual information (such as closed captions) whenever available for video indexing.

4. Conclusion

In DMP, if there are suitable video clips in video database or video web library, the required clips will be extracted from the database/library, otherwise, 3D animation will be created based on cinematic knowledge, so that at present it is feasible to automatically make motion picture with visual effects like computer animation and real images, and their simple composition. Along with the improvement of computational power and human knowledge representation, we will combine new automated video indexing algorithms and intelligent reasoning into our DMPVR.

Reference

[1] SHEN Jinhong, Seiya MIYAZAKI, Terumasa AOKI, Hiroshi YASUDA, “Filmmaking Production System with Rule-based Reasoning”, Image and Vision Computing New Zealand (IVCNZ 2003), Palmerston North, New Zealand, Nov. 26-28, 2003

[2] SHEN Jinhong, Seiya MIYAZAKI, Terumasa AOKI, Hiroshi YASUDA, “Hierarchical Video Modeling for Indexing and Retrieval Based on MPEG-7”, CS2003-146, pp 35-18, 2003 AVM Conference of IEICE, Sapporo, Dec., 2003.

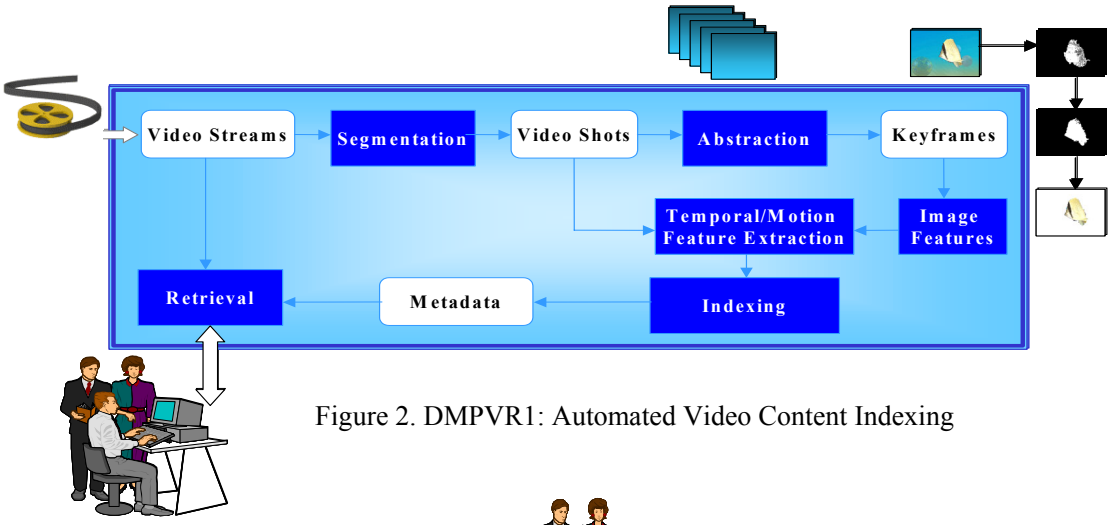

Figure 2. DMPVR1: Automated Video Content Indexing

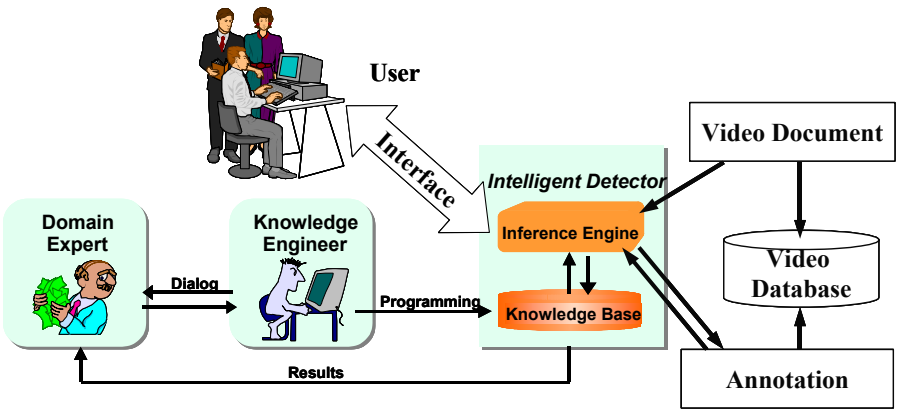

Figure 3. DMPVR2: Computer Aided Video Content Indexing