Knowledge Representation for Video Assisted by Domain-Specific Ontology

Dan Song\textsuperscript{1}, Miyoung Cho\textsuperscript{1}, Chang Choi\textsuperscript{1}, Juhyun Shin\textsuperscript{1}, Jongan Park\textsuperscript{2}, and Pankoo Kim\textsuperscript{3}\footnote{Corresponding author.}

\textsuperscript{1} Dept. of Computer Science
Chosun University, 375 Seosuk-dong Dong-Ku Gwangju 501-759 Korea
songdan@stmail.chosun.ac.kr, irune@chosun.ac.kr, eduranceaura@gmail.com, jhshinkr@chosun.ac.kr

\textsuperscript{2} Dept. of Information and Communication Engineering,
Chosun University, Korea
japark@chosun.ac.kr

\textsuperscript{3} Dept. of CSE, Chosun University, Korea
pkkim@chosun.ac.kr

Abstract. Video analysis typically has been pursued in two different directions. Either previous approaches have focused on low-level descriptors, such as dominant color, or they have focused on the video content, such as person or object. In this paper, we present a video analysis environment not only to bridge these two directions but also can extract and manage semantic metadata from multimedia content autonomously for addressing the interaction between browsing and search capabilities. Concretely speaking, we implemented a tool that links MPEG-7 visual descriptors to high-level, domain-specific concepts. Our approach is ontology-driven, in the sense that we provide ontology based domain-specific extensions of the standards for describing the knowledge of video content. In this work, we consider one shot (episode) in the billiard game of video as the specific domain and we will be through the practical works to explain the process of representation of video knowledge. In the experiment part, we prove our approach effectiveness by comparing with the video content retrieval based on only key-word.

1 Introduction

Although new multimedia standards, such as MPEG-4 and MPEG-7 [1], provide the needed functionalities in order to manipulate and transmit objects and metadata, their extraction, and that most importantly at a semantic level, is out of the scope of these standards and is left to the content developer. Extraction of low-level features and object recognition are important phases in developing multimedia database management systems.

There has been a research focus to develop techniques to annotate the content of images on the Web using Web ontology languages such as RDF and OWL. Past efforts have largely focused on mapping low-level image features to ontological
concepts [2] and have involved the development of tools that are closely tied to domain specific ontologies for annotation purposes [3,4]. Additionally, the lack of precise models and formats for object and system representation and the high complexity of multimedia processing algorithms make the development of fully automatic semantic multimedia analysis and management systems a challenging task. This is due to the difficulty that often mentioned as the semantic gap. The use of knowledge domain is probably the only way by which higher level semantics can be incorporated into techniques that capture the semantic concepts. So, in this paper, a comprehensive method for video content analysis based on the specific knowledge domain was proposed using on the tools of Protégé which is the classical ontology editor and PhotoStuff that is the most promising annotation software that allows users to makeup of an image/video key-frame with respect to concepts in an ontology.

We organize the remainder of the paper as follows: Section 2 is about the overview for video analysis. Section 3 introduces the infrastructure of domain knowledge. As the major part, section 4 shows us how to present video content through one specific domain ontology. It contains two sub-sections: ontology building and mapping from the low-level features to high-level semantics for video knowledge representation. And, Analysis results for video content retrieval are showed in Section 5. After these comprehensive explanations, we will conclude in section 6.

2 Overview for Video Analysis

Video is a structured medium in which actions and events in time and space convey stories, so, a video program (raw video data) must be viewed as a document, not a non-structured sequence of frames.

From Figure 1, we can see the second layer: video conceptual feature which was represented by video shots that are the basic units used for accessing video and a sequence of frames recorded contiguously and re-presenting a continuous action in

Fig. 1. Video Modeling and Representation