



Chapter VIII

Video Abstraction Techniques for a Digital Library

Hang-Bong Kang
Catholic University of Korea, Korea

The abstraction of a long video is often useful to a user in determining whether the video is worth viewing or not. In particular, video abstraction guarantees users of digital libraries with the fast, safe and reliable access of video data. Two approaches, such as summary sequences and highlights are possible in video abstraction. The summary sequences are good for documentaries because they give an overview of the contents of the entire video, whereas highlights are good for movie trailers because they contain only the most interesting video segments. The video abstraction can be generated by three steps: analyzing video, selecting video clips, and synthesizing the output. In the analyzing video step, salient features, structures, or patterns in visual information, audio information, and textual information are detected. In the selecting step, meaningful clips are selected from detected features in the previous step. In the output synthesis step, the selected video clips are composed into the final form of the abstract. In this chapter, we will discuss various video abstraction techniques for digital libraries. In addition, we will also discuss a context-based video abstraction method in which contextual information of the video shot is computed. This method is useful in generating highlights because the contextual information of the video shot reflects semantics in video data.

INTRODUCTION

Recently, the amount of video data stored in archives worldwide has been increasing rapidly. To effectively deal with the huge amount of video data for a digital library, efficient technologies of content-based video search and retrieval are crucial. In addition, for faster,

safer, more reliable access for users anywhere in the world, it is desirable to maintain the abstracted version of the video data on a digital library.

The abstraction of a long video is often useful to a viewer in determining whether the video is worth viewing or not. Video abstraction refers to the sequence of images in which the length of the sequence is shorter than the original but the essence of the content is preserved (Uchihashi, Foote, Girgenshon, and Boreczky, 1999). The video abstracts are useful in many application areas, such as multimedia archives for a digital library, home entertainment and movie marketing.

In video abstraction, there are two approaches, summary sequences and highlights. The summary sequences are good for documentaries because they give an overview of the contents of the entire video, whereas highlights are good for movie trailers because they contain only the most interesting video segments. In other words, summary sequences are particularly important for information content, such as documentaries and newscasts, but highlights are important for entertainment content, such as movies and TV dramas. However, making summary sequences or highlights from video data is not an easy task because the automatic extraction of semantics from raw video data is very difficult.

This chapter will focus on the various video abstraction techniques, such as summarization and highlight generation for a digital library. We also discuss context-based video abstraction. This chapter will be organized as follows. The next section describes the various techniques in constructing video abstractions. State-of-the-art research systems and their distinctive characteristics are described. The *Video Abstraction Process* section presents current issues in video abstraction. We will discuss how to effectively reflect semantics and affection. The section following that explains context-based video abstraction method. The *Future Trends* section explains promising future research directions in constructing video abstraction. And finally we provide some concluding remarks.

RELATED WORK

Significant efforts have been made on video abstraction in recent years. They were mostly about constructing summary sequences from video data. Various domains of video data, such as movies, TV dramas, newscasts, documentaries, presentation videos, sports video and even home videos, were tested in creating good quality of summary sequences. Information sources such as visual information, audio information, and textual information were also investigated.

Video skimming (Smith and Kanade, 1997; Christal, Smith, Taylor, and Winkler, 1998) was proposed for the abstraction of documentaries and newscasts in the CMU Informedia Digital Video Library Project. In this approach, the video and the transcript are aligned by the word spotting, and the language analysis is used to identify important words in the transcripts. For example, if a face is detected in the video data and a proper name appears in the transcript, it is assumed that a person is introduced and a 2-second video clip is selected into the summary sequences. Christal (1999) proposed a method of video digest for news video libraries. In this video digest, three types of digest, such as word relationships, timelines showing trends against time, and maps showing geographic correlations, were developed. Currently, Informedia II Digital Video Library Project is under research for the purpose of constructing video information summarization and demonstration testbed. In this project, a unified infrastructure for integration and demonstration of object detection, recognition and tracking, event understanding, query-by-example, and multi-modal fusion will be provided.

11 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/video-abstraction-techniques-digital-library/8618

Related Content

Virtual Sets: Concepts and Trends

Antonia Lucinelma Pessoa Albuquerque, Jonas Gomes and Luiz Velho (2001). *Design and Management of Multimedia Information Systems: Opportunities and Challenges* (pp. 247-267).

www.irma-international.org/chapter/virtual-sets-concepts-trends/8121

DMMs-Based Multiple Features Fusion for Human Action Recognition

Mohammad Farhad Bulbul, Yunsheng Jiang and Jinwen Ma (2015). *International Journal of Multimedia Data Engineering and Management* (pp. 23-39).

www.irma-international.org/article/dmms-based-multiple-features-fusion-for-human-action-recognition/135515

Constructing and Utilizing Video Ontology for Accurate and Fast Retrieval

Kimiaki Shirahama and Kuniaki Uehara (2011). *International Journal of Multimedia Data Engineering and Management* (pp. 59-75).

www.irma-international.org/article/constructing-utilizing-video-ontology-accurate/61312

A Distance-Window Approach for the Continuous Processing of Spatial Data Streams

Salman Ahmed Shaikh, Akiyoshi Matono and Kyoung-Sook Kim (2020). *International Journal of Multimedia Data Engineering and Management* (pp. 16-30).

www.irma-international.org/article/a-distance-window-approach-for-the-continuous-processing-of-spatial-data-streams/260962

Consumer Attitudes toward RFID Usage

Madlen Boslau (2009). *Encyclopedia of Multimedia Technology and Networking, Second Edition* (pp. 247-253).

www.irma-international.org/chapter/consumer-attitudes-toward-rfid-usage/17408