

The effect of media richness factors on representativeness for video skim

Huey-Min Sun*, Chih-Wei Huang

Dept. of Information Management, Chang Jung Christian University, Tainan County, Taiwan 711, ROC

Received 9 November 2010; received in revised form 3 May 2011; accepted 7 July 2011

Communicated by P. Mulholland

Available online 14 July 2011

Abstract

The study examines the effect of four important aspects of film skimming, including segmentation process, proportion of total skimmed length (TSL), multiple cues available, and genre/domain of the film. We design three experiments to explore their effects on representativeness for video skim. The results of Experiment 1 show that the skimmed video combined with 10% of total skimmed length and 5 or 10 s of skimmed segment (SS) is more efficient for representativeness. The results of Experiment 2 show that the skimmed video with mostly ending part and multiple cues can significantly improve representativeness. The results of Experiment 3 reveal that the representativeness of skimmed video with different types of movie is significantly different.

In our experiments, the proportion of TSL is set to three levels, 5%, 10%, and 15%, while the size of SS is also set to three levels, 2.5, 5, and 10 s for the segmentation process. We observe that the skimmed video with the longer TSL and SS has the better representativeness of movie content, but the four combinations for 10% and 15% with 5 s and 10 s are insignificantly different. The finding is helpful for reducing the time cost of skimming video. Furthermore, we applied two important factors—personality focus of the medium and multiple cues, from media richness theory to our skimming method in order to raise the representativeness of video skim for different films. In the personality focus of the medium, we define a movie as having three parts—beginning, middle, and ending. In the multiple cues, the skimmed video with synchronized subtitle, audio, and video can assist our comprehension and reduce the uncertainty. We find that the skimmed video with mostly ending part and synchronized subtitle, audio, and video can raise the representativeness of movie content.

© 2011 Elsevier Ltd. All rights reserved.

Keywords: Media richness theory; Video skim; Subjective quality assessment

1. Introduction

With increase in computational power and memory storage capacity, the potential use of digital video is growing rapidly. Users will need a new technology in video abstraction to help them browse effectively and efficiently by skimming a key component. Video abstraction is a mechanism for generating a short summary of a video, which can either be a sequence of stationary images (keyframes) or moving images (video skims) (Truong and Venkatesh, 2007). A good video abstraction will enable users to get maximum information about video content in minimum time. Abstraction techniques are mainly designed to facilitate browsing for large digital video libraries, and they complement an automatic video retrieval approach

(i.e., searching). There are many semantic concepts that are very difficult or even impossible to extract automatically, as precise queries cannot be easily formed. Past works have focused on the improvement of techniques, but many have ignored some basic problems such as total skimmed length, skimmed segment, and their combination.

A video skim consists of a collection of video segments extracted from the original video. These segments are joined by either a cut or a gradual effect. In this study, we want to know whether the representativeness of a skimmed film is significantly different from the different proportions of total skimmed length. At the same time, we also want to know whether that has significant difference among the different sizes of skimmed segment. Aside from the basic problems, what are other factors that influence the representativeness of skimmed film? Two factors, personality focus of the medium and multiple cues, from

*Corresponding author.

E-mail address: prince@mail.cjcu.edu.tw (H.-M. Sun).

media richness theory (Daft and Lengel, 1986) are worth examining. In the personality focus of the medium, we know that different types of movie contents have different individual characteristics. For example, most films consist of a beginning, a middle, and an ending. Different parts of a film present different phases of a story. In multiple cues, a skimmed video is composed of subtitle, audio, and video. The more the cues the skimmed algorithm excerpts, the more the viewers comprehend.

In this study, we explored the representativeness of different levels (5%, 10%, and 15%) of total skimmed length, different levels (2.5, 5, and 10 s) of skimmed segment, and their combinations. Moreover, we applied two important factors—personality focus of the medium and multiple cues—from media richness theory to our skimming method in order to raise the representativeness of video skim for a film.

The paper is organized as follows. Section 2 presents the related work. Section 3 states our research approach. Experimental design is described in Section 4. Results of the experiment are shown in Section 5. The discussion and conclusion are stated in Section 6.

2. Literature

Our study is based on media richness theory to explore the effect of media richness factors on representativeness for video skim by the research method of subjective quality assessment. We reviewed the video skim field, the factors of media richness theory, and the methods of subjective quality assessment.

2.1. Video skim

Video skim generation is a relatively new area of research and often requires high-level content analysis. The simplest method for generating video skims is uniform sampling, which extracts fixed-duration excerpts of the original video at fixed intervals. There are four major aspects of a skim generation technique: (1) segmentation process, (2) proportion of total skimmed length, (3) multiple cues available, and (4) genre/domain of the film to be summarized. Descriptions of these aspects are as follows.

2.1.1. Segmentation process

Skimmed segment excerpts the whole video sequence into separate units. Some of which will be included in the skim. However, the segmentation of the video sequence is divided into different types of lengths, such as shot (Li et al., 2006; Le et al., 2008), scene (Ngo et al., 2005; Li et al., 2006; Chasanis et al., 2009), interesting event (Ariki et al., 2003), and the segmentation of complete speech (Taskiran et al., 2001). These authors used different separating schemes to form a basic unit of segmentation. The length of a skimmed segment ranges between 2 and 10 s (Christel et al., 1998). In the skimmed segment, one of the most basic methods excerpts a source video, keeping

the first 10 s at fixed intervals of 100 s. In addition, Christel et al. (1998) adopted 2.5 and 5 s of size as tested skimmed segment to experiment.

2.1.2. Proportion of total skimmed length

Typically, the total length of a video skim can be specified as *a priori* or left unknown until the end of the extraction process. Some studies (He et al., 1999; Ma et al., 2002) were among video skimming works that require the length of skim to be specified by the user. Most of them adopted around 10% of total skimmed length (Christel et al., 2008; Benini et al., 2007; Chen et al., 2009). Christel et al. (2008) concentrated on playable video summaries, experimenting with summaries that have durations of one-twenty-fifth (4%) or smaller compared to the target video. Benini et al. (2007) generated some dynamic summaries with their related soundtracks. For each video two associated skims were produced: one with 10% of the original video length and the other with 25%. In Chen et al. (2009), the results demonstrated that preserving 10–20% of the total shots in the test videos yields satisfactory performance. In our study, we chose three levels, i.e., 5%, 10%, and 15%, to evaluate the representativeness of skimmed video.

2.1.3. Multiple cues

Video skim can be combined by the integration of subtitle, audio, and video. The more cues the skimming algorithm excerpts, the more the viewers comprehend. In addition, they can be assembled by modal synchronization and modal asynchronization. In the former, the subtitle, audio stream, and visual stream are synchronized by the time line of the original video sequence (Kim et al., 2003). In the latter, the subtitle, audio stream, and visual stream are not synchronized. The synchronized type of skim is essential to video programs such as movies, dramas, and talk shows. It is easy to comprehend the content since what we hear from the audio track directly corresponds to what we see on the screen.

2.1.4. Genre/domain of the film to be summarized

Movie content abstraction can produce a static storyboard, a summary sequence, or a highlight (Li and Kuo, 2003). A summary sequence provides users a small taste of the entire video, such as movies, while a highlight contains only the content that may appear interesting to viewers such as sports games. In our study, we focus on summary-oriented skim and explore movie content skim.

2.2. Media richness theory and comprehension

Media richness theory is a model that can be used to describe a communication medium by its ability to reproduce the information sent over it (Daft and Lengel, 1986). The model includes the functions, the medium's capacity for immediate feedback, multiple cues available, language variety, and personality focus of the medium. The theory proposes two important constructs, uncertainty and equivocality, to be

Download English Version:

<https://daneshyari.com/en/article/401217>

Download Persian Version:

<https://daneshyari.com/article/401217>

[Daneshyari.com](https://daneshyari.com)