Chapter 11
Contextual In-Stream Video Advertising

Tao Mei
Microsoft Research Asia, China

Shipeng Li
Microsoft Research Asia, China

ABSTRACT

With Internet delivery of video content surging to an unprecedented level, online video advertising is becoming increasingly pervasive. In this chapter, we present a new advertising paradigm for online video, called contextual in-stream video advertising, which automatically associates the most relevant video ads with online videos and seamlessly inserts the ads at the most appropriate spatiotemporal positions within each individual video. Different from most current video-oriented sites that only display the ads at the predefined locations in a video, this advertising paradigm aims to embed more contextually relevant ads at less intrusive positions within the video stream nonlinearly. We introduce the following key techniques in this paradigm: video processing for ad location detection, text analysis for ad selection, and optimization for ad insertion. We also describe two recently developed systems as showcases, i.e., VideoSense and AdOn which support in-stream inline and overlay advertising, respectively.

INTRODUCTION

The proliferation of digital capture devices and the explosive growth of online social media (especially along with the so called Web 2.0 wave) have led to the countless private image and video collections on local computing devices, such as personal computers, cell phones, and personal digital assistants (PDAs), as well as the huge yet increasing public media collections on the Internet. Today’s online users face a daunting volume of video content. ComScore reports that in March 2006 alone consumers viewed 3.7 billion video streams and nearly 100 minutes of video content per viewer per month (ComScore). The most
popular video site—Youtube, drew 5 billion U.S. online video views in July 2008 (YouTube).

On the other hand, we have witnessed a fast and consistently growing online advertising market in recently years. Jupiter Research forecasted that online advertising spending will surge to $18.9 billion by 2010-up, which is about 59 percent from an estimated $11.9 billion in 2005 (Jupiter Research). To take the advantages of this increasing market share and effectively monetize video content, video advertising, which associates advertisements with an online video, has become a key online monetization strategy. By implementing a solid online video advertising strategy into an existing content delivery chain, content providers have the ability to deliver compelling content, reach a growing online audience, and generate additional revenue from online media. As reported by Online Publisher Association (Online Publishers), the majority (66%) of Internet users have ever seen video ads, while 44% have taken some action after viewing ads.

Many existing video-oriented sites, such as YouTube (YouTube), Google Video (Google Video), Yahoo! Video (Yahoo! Video), Metacafe (Metacafe), and Revver (Revver), have tried to provide effective video advertising services. However, it is likely that most of them match the ads with online videos only based on textual information and insert ads at the beginning or the end of a video 1. In other words, contextual relevance in these sites is only based on textual information, while less intrusive insertion points are fixed to the predefined locations, e.g., the beginning or the end of videos. For example, Revver selects one relevant ad (i.e., a static picture or a video clip) for each video clip, and shows it as the first or the last frame or segment of the corresponding video (Revver). Another example is Google’s AdSense for video advertising which overlays the ads at a fixed location in the videos (e.g., on the bottom fifth of videos 15 seconds in).

On the other hand, although there are a few systems for overlay advertising proposed recently in the research community (Chang et al., 2008) (Liao et al., 2008) (Liu et al., 2008), they are not practical for real-world application. For example, AdImage predominantly focuses on image-based ad matching while neglects the ad positions (the ads always appear on the right-bottom corner) (Liao et al., 2008). The virtual content insertion (Chang et al., 2008) (Liu et al., 2008) is not practical for user generated videos as these videos are typically with poor visual quality so that detecting smooth area in the frames and aligning the ad by geometric transformation is very challenging.

The following problems that significantly affect advertising effectiveness and impede user experience have not been investigated in existing video advertising:

- We believe ads should be inserted at appropriate locations within video streams rather than any predefined locations. The capability of discovering nonlinear ad locations within videos will lead to embedding not only a greater number of ads but also less intrusive ads within video content.
- We believe ads should be contextually relevant to online video content in terms of multimodal relevance rather than purely based on textual information. For example, when viewing an online music video, users may prefer a relevant ad with the similar editing style or audio tempo style to the video, which cannot be measured just by textual information. This capability will lead to delivering the ads with more relevance.

We present in this chapter a new advertising paradigm for online video, called contextual in-stream video advertising, which automatically associates the most relevant video ads with online videos and seamlessly inserts the ads at the most appropriate spatiotemporal positions within each individual video. Different from most current video-oriented sites that only display the ads at