

VIDEOFARM: Playing video and display associated information

Hsin C. Fu,

*Dept. of computer science
National Chiao-Tung University
Hsinchu, Taiwan, ROC
hcfu@cs.nctu.edu.tw*

Hsiao T. Pao

*Dept. of management science
National Chiao-Tung University
Hsinchu, Taiwan, ROC
htpao@cc.nctu.edu.tw*

Yeong Y. Xu

*Dept. of computer science
HungKuang University
Taichung, Taiwan, ROC
yyxuster@gmail.com*

Abstract—This paper addresses a new type of video sharing platform—VideoFarm. In addition to just playing user uploaded video clips, the VideoFarm also displays the associated information of currently playing video contents. The associated information includes the general description of the video clip, a dynamic electronic map which shows the location of the current video contents. The location associated information, such as local stores, restaurants, lodging, and near by tour attractions are provided and displayed in parallel with the playing video clip. Through the URLs of the associated information, a viewer may collect plenty of tour information, when (s)he is enjoying an interesting video clip. A prototype of VideoFarm has been implemented at <http://140.113.216.55/videomap> for public trial, evaluation, comments, and suggestions.

Keywords—video sharing, Web, multimedia, geography, content analysis

I. INTRODUCTION

In recent years, video sharing becomes one of popular services over Internet. For example, in January 2008 alone, nearly 79 million users had made over 3 billion views at *YouTube* Website [1]. There are many reasons that a Web site can attract a crowd of people. The main reasons are: it is easy to upload videos, it is easy to find desired movies, and the most important is to have interesting films. For example, an 8 minute video called "*Battle at Kruger*" [2] was posted on *YouTube* in 2007. Up to November of 2008, the video received more than 38 million views, and 41 thousand comments. This is an unprecedented record. However, after watching the short film, how many people know where on the earth the Kruger Park is?

Since most of videos on the web were made by amateurs, there is neither a map nor a description to depict where the movie contents were taken place like some commercial films such as Discovery Channel's "Pilot Guides" series [3]. Some people may say: "The video itself already possess both space and time scenarios, why additional maps are needed?" In fact, the space scenario includes local and macro views. In a video or a film, most of views are limited to local scenery; people can hardly deliberate macro-geographical aspects from local scenes.

In particular, some of foreign place names, which usually are transliterations, may often just a sequence of audio sound to most movie viewers. Therefore, video

contents involved in the unfamiliar names, dynamic changes in orientation and locations, often need supporting information to point audience's vision to a better sense of direction, to a solid sense of stereoscopic view on the Earth, and to bring some background stories for better understanding on the video contents. When a film is playing, if augmented information is presented in time, in particular the detailed location of a building, and its surrounding environment, and/or people's life style, even a not very eye-catching film may also be immense.

Generating associated messages with the contents of a movie is very difficult, because a number of qualifications and tremendous knowledge are required to provide proper information for films from all over the world:

- (1) One has to deeply understand the film's contents; otherwise it would not be possible to select impressive scenes or attractions for interesting messages.
- (2) One has to familiar with the background stories of a film, so as to associate scenes or attractions with the corresponding geographic locations and related information.

In the following sections, we will present a Web 2.0 based platform for video playing and augmented information displaying for the associated video contents. In Section 2, basic concepts of our approach are described. A walk through of the prototype *VideoFarm* platform is depicted in Section 3. Concluding remarks and future works are drawn in Section 4.

II. OUR APPROACH

Our approach to present online information with a video is to deploy a Web 2.0 based platform, called *VideoFarm*, which contains a video player and an online multi-information display board. We intend to congregate Internet users to contribute interesting movies or video clips, to contribute personal individual insights on selecting interesting and impressive content attractions, and to share their knowledge for the selected attractions by providing high quality associated information.

The *VideoFarm* platform provides the following features:

1. a user may upload interesting or impressive movies or video clips,

2. a user may suggest attractive scenes or frame images while watching a movie,
3. a user may annotate a scene or an image of its title, location, and brief description, and
4. a user may check associated information of the currently playing movie or a video clip.

The associated information includes an e-map of the scene location, local or near by business and geographic information, such as tour guides, transportation, food and lodging services, souvenir shops etc.

The *VideoFarm* contains three pages: **Home**, **Video** and **Upload**. In **Home** page, there are three categories of video list: (1) Promoted videos, (2) Fresh videos, and (3) Video list, as shown in Figure 1. At **Home** page, a user may click an image icon to view one of the recommended video clips, or may input keywords to search for desired movies or video clips in Video list.

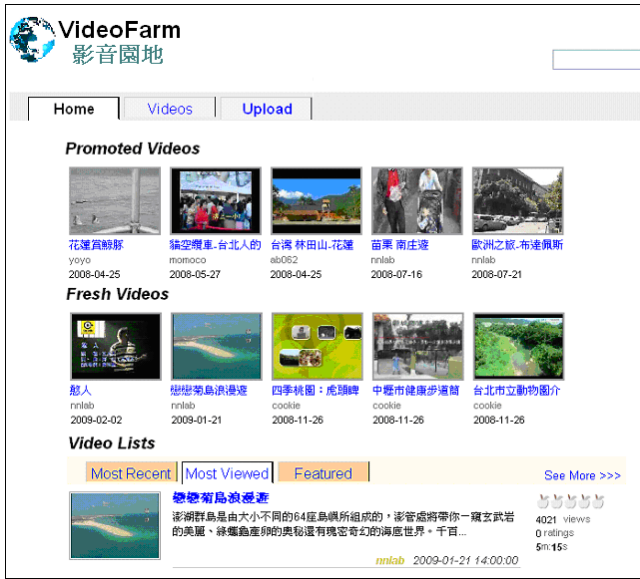


Figure 1: Home page of the *VideoFarm*.

At **Video** page, two display areas are shown: (a) a video player, and (2) an on-line multi-information display board. During the playing of a movie, the video contents associated multimedia information is presented in the display board as shown in Figure 2. The contents associated information include: a dynamic e-map, key-frames of attractions, general description of attractions, URL of geographic information, related business and social networks.

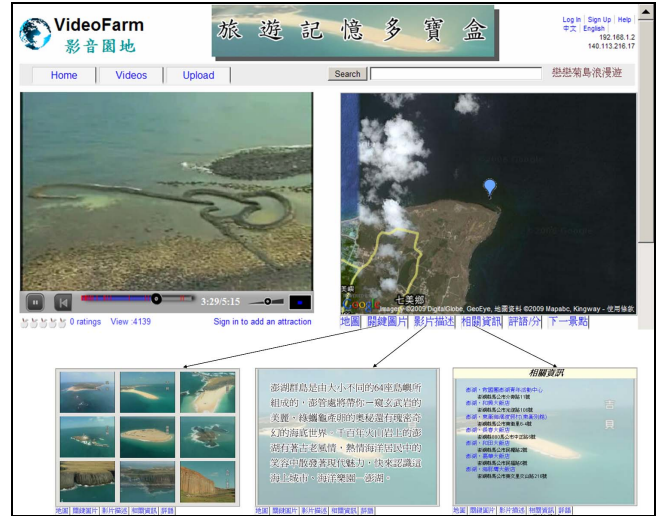


Figure 2: Video page of the *VideoFarm*. On the right side, a multi-information board displays an e-map, key frames, descriptions, and local business information of the currently playing video contents.

At **Upload** page, a registered user may upload a movie or a video clip to share with other Web browsers. The procedure of uploading a video clip is similar to most of video sharing web sites. When an upload request is asked by a user, a dialogue box will be displayed, as shown in Figure 3. In addition to upload a video clip, the user may input a brief *description* of the uploaded video.



Figure 3: Dialogue box for video upload.

III. WALK THROUGH OF VIDEOFARM

A prototype of VideoFram has been implemented at <http://140.113.216.55/videmap> to show the proposed four features described in Section 2. When a user clicks on a preferred video icon in either Home page or in Video page, the video begins playing and the information board shows the e-map of the associated video contents first. Then, it may switch to a URL list of geographic information, such as local or nearby scene spots, tour guides, lodging and food services, souvenir shops, or even associated commercial ads, as shown in Figure 2.

While watching a video, a user sees a scene, such as a house, a car, a celebrity, a fashion show or anything that is impressive or motivating background stories, commercial ads, or something worthy to be shared with friends, or other people, (s)he may click the “Add an Attraction” button to add an attraction picture and associated information to the *Attraction list*, and to signal the information displaying mechanism to include new attraction related information contents. Also, an attraction dialogue box as shown in Figure 4, is opened to receive input from the user. The input includes: 1) the *title* of the attraction, 2) the *location* of the attraction, and 3) brief *introduction* or *description* of the attraction.

Among the request information, only the title of the attraction is mandatory to initial an attraction suggestion, and the rest of information can be filled in by other users who know more or better of the suggested attraction. When the user finished and closed the attraction suggestion dialogue box, the video continues playing, and the user may enjoy the movie and/or make a new suggestion of another attraction. In the meantime, the provided text information of the attraction is analysed to extract keywords for the searching of various associated multimedia contents.

It is common to see that a person’s favored attractions may be other’s aversion. Thus, presenting attractions to a Web video browser needs to be carefully screening to avoid counter effect. In other words, it is needed to match attractions with browsers’ preference. Tagging mechanism is deployed in both attraction description and user registration.

On the attraction side, keywords that are extracted as tags from the attraction descriptions by using IR techniques [4], and recently developed content tagging techniques [5, 6] are also used to build tags from the frame images of selected attractions. In order to tagging user’s preference, a user is asked to fill in the preference-related tag words during the registration. In addition, when a user suggests attractions, the tags associated with these attractions are collected and attached to the user’s tag list.



Figure 4: Dialogue box for attraction description.

IV. CONCLUDING REMARKS

The walk through presentation in Section 3, shows a great success and future potential of the augmented information board in extending a video sharing web site into a multimodal information provider and various e-commerce agents. The prototype system also validates the combination of recently advanced WWW search techniques, and newly developed content tagging schemes to achieve a satisfactory level of personalized multimedia information retrieval platform. Currently, we are working on the English language version of the *VideoFarm*, and in the near future, we plan to include channels and community pages to enhance the social networking capability of the *VideoFarm* platform.

ACKNOWLEDGEMENT

This research was supported in part by the National Science Council under grant NSC96-2221-E-009-167. and in part by the Microsoft under grant FY08-RES-THEME-092.

REFERENCES

- [1] Yi-Wyn Yen, “YouTube Looks For the Money Clip,” <http://techland.blogs.fortune.com/2008/03/25/youtube-looks-for-the-money-clip/>, March 25, 2008.

- [2] http://tw.youtube.com/watch?v=LU8DDY_z68kM
- [3] http://www.pilotguides.com/tv_shows/globe_trekker/index.php
- [4] Amit Singhal, "Modern Information Retrieval: A Brief Overview," *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering* **24** (4): 35–43.
- [5] Xin-Jing Wang, Lei Zhang, Xirong Li, Wei-Ying Ma, "Annotating Images by Mining Image Search Results," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 11, pp. 1919-1932, November, 2008.
- [6] Jia Li, James Z. Wang, "Real-Time Computerized Annotation of Pictures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 6, pp. 985-1002, June, 2008.