Trajectory-based Visualization of Web Video Topics

Juan Cao
Institute of Computing Technology,
Chinese Academy of Science
caojuan@ict.ac.cn

Chong-Wah Ngo
Computer Science Department,
City University of Hong Kong
cwngo@cs.cityu.edu.hk

Yong-Dong Zhang
Institute of Computing Technology,
Chinese Academy of Science
zhyd@ict.ac.cn

Dong-Ming Zhang
Institute of Computing Technology,
Chinese Academy of Science
dmzhang@ict.ac.cn

Liang Ma
Institute of Computing Technology,
Chinese Academy of Science
maliang@ict.ac.cn

ABSTRACT

While there have been research efforts in organizing large-scale web videos into topics, efficient browsing of web video topics remains a challenging problem not yet addressed. The related issues include how to efficiently browse and track the evolution of topics and eventually locate the videos of interest. In this paper, we introduce a novel interface for visualizing video topics as evolution trajectories. The trajectory visualization is capable of highlighting milestone events and depicting the topical hotness over time. The interface also allows multi-level browsing from topics to events and to videos, resulting in search exploration could be more efficiently conducted to locate videos of interest. In addition, recommendation of topics accordingly to three-hots: content-hot, evolution-hot and potential-hot, can be easily supported by our system. A user study on three months' YouTube videos using our interface demonstrates the efficiency of our system in browsing web videos.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: [User Interfaces]

General Terms

Design

Keywords

Topic Trajectory, Web Video,Visualization

1. INTRODUCTION

A recent statistics by YouTube report [2] shows that 50% of users watch web videos through recommendations from friends, while no more than 22% of users indeed initiate search queries to explore videos of interest. This statistics implicates a strong demand on video recommendation. On the other hand, the unprecedented explosion in the volume of web videos has also made it difficult for web users to quickly access the videos of concern and for web administrators to conduct a systematic and thorough monitoring of web activities. Due to these reasons, most video-sharing websites provide a recommended list of videos ranked based on factors such as view count, download frequency and upload time. Generally speaking, the list gives a glance of the hottest videos, but the video-level recommendation is still limited in revealing the video relationship. Topic-level recommendation which tracks videos of interest and threads them according to topic evolution is more appealing. Such examples include CNN and Sina websites which manually categorize hot articles and videos into "hot topics". The manual effort nevertheless is time consuming particularly with the massive growth of web videos in the Internet.

To automate topic-level recommendation, two major issues are: collecting related videos into topics, and efficient visualization of topics. There have been numerous research efforts ranging from tracking news article [1] to news videos [8][6][10]. In text domain, Mei et al. [7] mined the meaningful evolution patterns from the topic distribution across time. The work however is for knowledge mining rather than for efficient user browsing. In video domain, Neo et al. presented as a topic as a ranked list of videos according to timeline [8]. Each video is further attached with its description and relevant articles. In [10], Wu et al. adopted a binary tree to model and thread the relationship of events in news topics. Ide el al. [5] developed a novel interface named Mediawalker to facilitate browsing of topic structure across the time dimensions.

These works [8][10][5], nevertheless, are built upon news videos with rich of speech transcripts and text captions, and follow the conventional TDT (Topic Detection and Tracking) [1] framework which focuses on the first story detection and topic tracking. Facing the web videos with sparse and noisy textual information, performing TDT by considering the local information only is greatly challenging. On the other hand, most works focus on the first issue, relatively few works address the visualization of video topics for user browsing. This paper presents a novel topic discovery and

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MM'10, October 25–29, 2010, Firenze, Italy.

Copyright 2010 ACM 978-1-60558-933-6/10/10 ...$10.00.

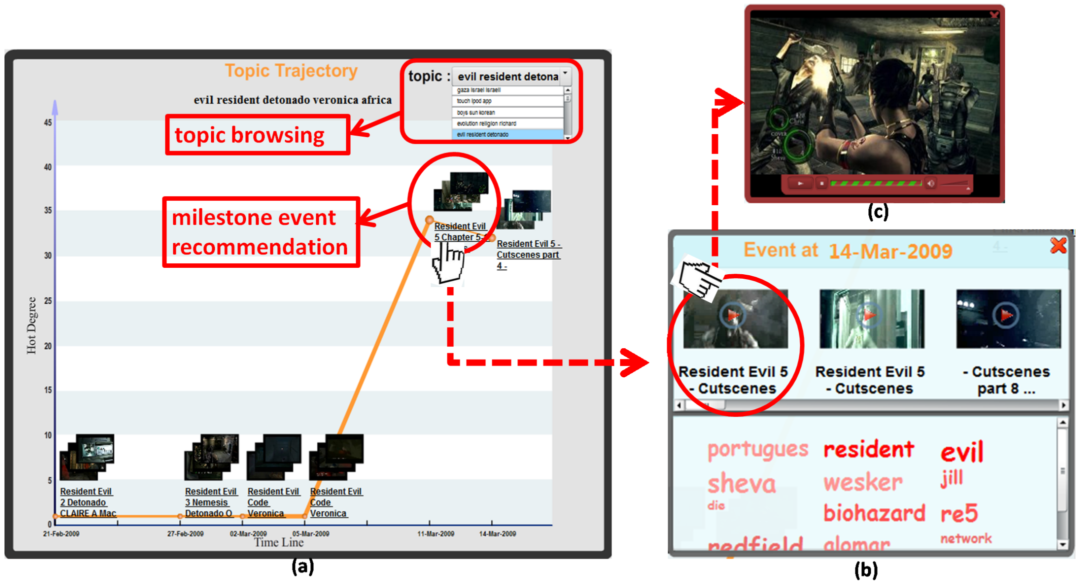

Figure 1: An example of browsing the web video topic "Resident Evil 5" using our interface. (a) Topic-level interface displays a trajectory showing the topic evolution in a 2D space of time and hot degree. The trajectory gives a glimpse of events happened at different time units. (b) Event-level interface provides a localized view of individual events with videos and tags. Videos are ranked according to popularity based on their view counts, while tags are displayed in different colors and sizes signifying their frequency and relevancy respectively. (c) Video-level interface plays the videos selected by users.

visualization system. Compare to the previous works, its main contributions are as follows:

- The system discovers the topics by extracting topic trajectories from the rough topic evolution graph. The global linking relationship can guarantee the robust topic detection than the local text-based first story detection strategy.

- The system designs an interface which visualizes video topic as a trajectory in a two dimensional space of hot-degree and time. It allows different levels of topical browsing. A global view gives a glance of topic in terms of evolution trend and the major events at each discrete time point. A localized view shows the details of a specific events by ranking or highlighting their videos, keyframes and tags.

2. TOPIC TRAJECTORY MINING

We first define the following two terminologies used in this paper. *Event* $\mathcal{E}$ is a group of related videos conveying a story and discovered at a time unit. *Topic* T is a group of topic-related events found over time. The system composes of the following three steps.

Event Mining: The web tag is temporal-effectiveness. Intuitive speaking, a tag only becomes salient and meaningful at several specific time points, and keeps flat during a long time. So we represent tag as a frequency trajectory on the time line, and detect bursty tags [4] which exhibit peak-like trajectory for every time unit. For the delay of news reports, we set three days as a time unit. The tag frequency is computed by the number of videos containing this tag at the corresponding time unit t_i. Then the bursty tags of every time unit are effectively clustered into meaningful events. Furthermore, the events are ranked based on the hot-degree measured by the view counts of the videos belonging to them.

Topical Linking: A hot topic will evolute from one event to another with time. So we measure the similarities between events in different time units based on tags and visual near-duplicates [9]. Events with more common tags and near-duplicate segments receive higher weights. Then a topic evolution graph $\mathbf{G} = (\mathsf{V}, \mathsf{E})$ is generated, where each node is an event and the edge between nodes signifies the similarity between two events.

Trajectory Extraction: In this step, the topic discovery has been transferred into path selection from the above topic evolution graph. Firstly, based on the context of graph, we can optimize the graph by adding the missing edges and remove the isolate weak links. Then, we decompose the huge topic graph into subgraphs with conventional depth first search algorithm, where the closely linked subgraph generally represents a topic or several related topics. Among the graph, every path is a candidate topic evolution trajectory with an order set of events which are linked chronologically, denoted as $\mathsf{T} =< \mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_{|\mathsf{T}|-1}, \mathcal{E}_{|\mathsf{T}|} >$ where events are placed in time order. Then a saliency measurement is proposed to extract meaningful topic trajectories from these subgraphs. The saliency of T is as

$$\text{Saliency}(\mathsf{T}) = \sum_{\mathcal{E}_t \in \mathsf{T}} \Lambda(\mathcal{E}_t) + \sum_{\{\mathcal{E}_{t-1}, \mathcal{E}_t\} \in \mathsf{T}} Sim(\mathcal{E}_{t-1}, \mathcal{E}_t) \quad (1)$$

The first term measures the social popularity of a topic based on the view count of events. The second term measures the topic compactness and evolution trend based on the event similarity. Basically, a topic with higher saliency score in-

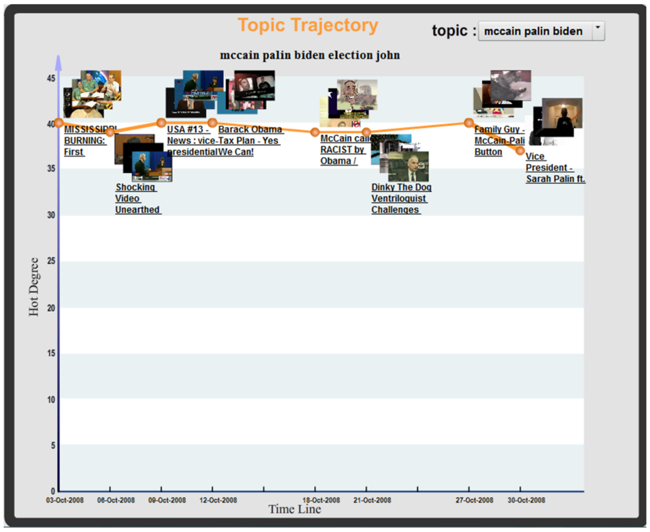

Figure 2: Trajectory visualization of the content-hot topic "US presidential election 2008".

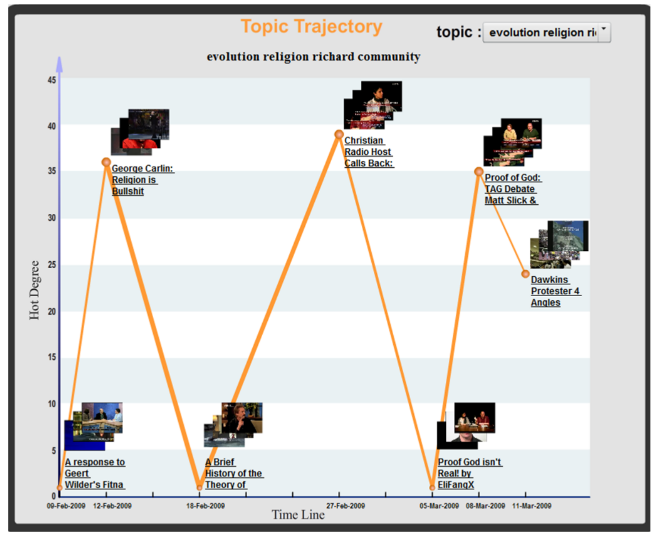

Figure 3: Trajectory visualization of the evolution-hot topic "Islamic belief".

dicates larger number of popular videos, most videos are tightly linked, and the topic evolves a longer period of time.

3. TOPIC VISUALIZATION

Each extracted trajectory can be vividly viewed in a 2D space, where the time axis indicates duration of a topic, and the hot-degree axis measures the popularity of events based on video view counts. With reference to Figure 1, the interface supports different levels of topical browsing.

Topic-level Browsing includes a scroll box which summarizes the list of detected topics ranked with their trajectory saliency in a video collection. Clicking a topic will show the corresponding trajectory in a 2D space of time and hot degree, as shown in Figure 1(a). Events are distributed along the trajectory and tagged with keyframes and a short description of texts, allowing users to rapidly trace the event sequence. The trajectory-based visualization gives a glance of the whole topic evaluation, while signifies the importance of events at different time points. For example, in Figure 1(a), the topic "Resident Evil 5" initially keeps a low profile, and reaches a peak on 11-March-2009 for an official announcement that this game would be released on 13-March-2009. Users can easily locate the events of interest with the displayed trajectory, or conveniently track backward and forward to see surrounding events.

Event-level Browsing supports the efficient means of visualizing tags and videos as shown in Figure 1(b). The tags are displayed in different colors and sizes representing the frequency and relevancy of the tags to the events. Tag relevancy is determined based on the number of videos with this tag in the corresponding event. By catching the attention with large and deep red tags, the display allows users to efficiently grasp the representative or key tags about the main content of an event. In addition to tags, videos are ranked according to the social popularity. Generally speaking, the most view video of an event is likely, though not absolutely, to be more representative. Each video is represented with a keyframe and its title. By clicking the keyframe, the video will be displayed as shown in Figure 1(c).

4. TOPIC RECOMMENDATION

Topic trajectory vividly display the birth, growth, decay and death of a topic along the timeline. By employing simple classification scheme based on primitive features such as number of peaks, degree of hotness, event compactness (similarity) and topic duration, the trajectories can be broadly categorized into three-hots as following.

Content-hot includes topic trajectories which keeps a high-level of hot degree for a certain period of time. This type of topics is concerned by most web users, and the content is often surrounding about a center theme with different peripheral events.

Evolution-hot typically includes topics with strong evolution trend, where the trajectories exhibit up-and-down trend over different periods of time. The contents of these topics are usually related to some sensational and sensitive news or discussion in the Internet.

Potential-hot includes topics initially concerned by a small group of web users, but increasingly capture the public attention and eventually end up with an erupt trend. This kind of topics is typically very focused and narrowed in the scope of discussion. Web monitors are especially interested in predicting the tendency of these topics.

In addition to automatic classification of three-hot topics, the developed interface indeed also offers efficient means of monitoring and recommending topics. Figure 2 shows a content-hot topic "US presidential election 2008". This topic becomes hot because the uploaded videos received many view counts over times. This kind of video topics captures short-term hot issues and could be recommended to users who are querying "*What's hot now?*". Figure 3 shows a evolution-hot topic about "Islamic belief". The topic did not keep hot throughout the whole life span. Instead, it has strong evolution trend and was repeatedly concerned by the public. This kind of topics is generally about sensitive political issues or super-stars, which periodically trigger public concerns. These topics are often welcomed by TV broadcasters who care about "*What's going on?*". The trajectory shown in Figure 1 is indeed an example of potential-hot topic about "Resident Evil 5". The topic was concerned by a small group of users initially, and abruptly broke out after the official announcement of the game. This kind of topics can be recommended to web monitors who concern "*What's going to be hot next?*".

5. USER STUDY

We conduct a subjective evaluation of the developed interface on a real world video dataset MCG-WEBV [3]. The dataset consists of 80,031 web videos, which include the most viewed videos uploaded to YouTube during 13 December 2008 to 13 March 2009 and their related videos. Based on our approach, a total of 122 topic trajectories are mined from the dataset. Table 1 shows the top-10 topics ranked based on trajectory saliency. Ten assessors were invited to evaluate the interface based on the following criteria.

1. Is the interface user-friendly and attractive?
2. Is the displayed information rich?
3. Can you locate the videos of interest efficiently using the interface?
4. Is the classification of topics into three-hots useful for exploring topics of interest?

The assessors rate each question with a three-degree score: strongly agree (1.0), agree (0.5) and disagree (0). Table 2 summarizes the result of user evaluation. Overall, the result is very encouraging. All assessors agreed that the interface is novel and the presented information is more concise if compared to traditional way of browsing videos in sequential order. First, the topic trajectory can precisely tell the evolution and important events at different time units. Second, important events can be located effortlessly from the keyframes and short descriptions attached to peaks of trajectories. The three-hot classification scheme is also found to be a useful feature. Some assessors indicated that they indeed preferred browsing evolving topics such as evolution-hot and potential-hot, and expressed their opinion about uncertain issues, rather than the videos posted by officials in content-hot topics. Under the platform of Web 2.0, most videos under evolution-hot and potential-hot are indeed created and uploaded by Internet users rather than by officials. The classification scheme facilitates them in locating the desired topics to browse.

On the other hand, some assessors commented that understanding the events with sparse text information only is insufficient. Possible external information should be introduced to rich the event description. The other critical comment is about the use of color to emphasize tag importance, which probably requires further perceptual study. We will look into both aspects for future extension of our system.

6. CONCLUSION

In this demo paper, we have presented our interface for visualizating video topics extracted from a large collection of web videos. The trajectory-based visualization offers several advantages which are also verified in our user studies on a 3-month web video dataset. First, the trajectory display vividly captures the time-dependent information of a topic and is capable of highlighting the milestone events. Second, classification of three-hots provides a meaningful way of topic recommendation for Web 2.0 platform. Third, by different granularity of browsing at topic and event levels, search exploration could be more efficiently conducted.

Table 1: Top-10 hot topics discovered by our system.

Topic ID	Class	Description
1	content-hot	US presidential election 2008.
2	content-hot	A Conflict between Israel.
3	evoluton-hot	Tutorial of Apple's Iphone.
4	evoluton-hot	Drama of flower boys.
5	evoluton-hot	Discussion of Islamic belief.
6	potential-hot	Resident Evil 5.
7	content-hot	Angeles Lakers.
8	evoluton-hot	Films of Alex. Jones.
9	evoluton-hot	Discussion of Game Xbox.
10	content-hot	American Idol Season 8.

Table 2: Subjective studies for trajectory-based visualization.

	Strongly Agree	Agree	Disagree	Average Score
User-friendly?	2	7	1	0.55
Informative?	6	3	1	0.75
Efficient?	8	2	0	0.9
Classification is useful?	7	3	0	0.85

7. ACKNOWLEDGEMENTS

This work was supported by the National Nature Science Foundation of China (60902090, 60802028), National Basic Research Program of China (973 Program, 2007CB311100).

8. REFERENCES

[1] http://www.nist.gov/speech/tests/tdt/.

[2] *YouTube Report 2009.* http://youtubereport2009.com/, 2009.

[3] J. Cao, Y. D. Zhang, and et al. Mcg-webv: A benchmark dataset for web video analysis. In *Technical Report,ICT-MCG-09-001*, 2009.

[4] Q. He, K. Chang, and E. P. Lim. Using burstiness to improve clustering of topics in news streams. In *DM*, 2007.

[5] I. Ide, T. Kinoshita, T. Takahashi, S. Satoh, and H. Murase. mediawalker: a video archive explorer based on time-series semantic structure. In *ACM MM*, 2007.

[6] L. Liu, L. Sun, Y. Rui, Y. Shi, and S. Yang. Web video topic discovery and tracking via bipartite graph reinforcement model. In *WWW*, 2008.

[7] Q. Mei and C. Zhai. Discovering evolutionary theme patterns from text: an exploration of temporal text mining. In *KDD*, 2005.

[8] S. Y. Neo, T. Ran, and et. al. The use of topic evaluation to help users browse and find answer in news video corpus. In *ACM MM*, 2007.

[9] H. K. Tan, C. W. Ngo, and et al. Scalable detection of partial near-duplicate videos by visual-temporal consistency. In *ACM MM*, 2009.

[10] X. Wu, C. W. Ngo, and Q. Li. Threading and autodocumenting news videos. *IEEE Signal Processing Magazine*, 23(2):59–68, 2006.