Threading Stories and Generating Topic Structures in News Videos across Different Sources

Xiao Wu
Department of Computer Science
City University of Hong Kong
83 Tat Chee Avenue, Hong Kong
wuxiao@cs.cityu.edu.hk

ABSTRACT
News videos delivered from different sources constitute a huge volume of daily information. These videos, overall, form a huge collection of news stories that are intertwined with various novel and old topic themes. To date, it remains a challenging task on how to automatically extract a concise view of news stories according to topic themes. This doctoral thesis studies the issues in story dependency threading and topical auto-documentary in news stories. Initially, a co-clustering algorithm is proposed to perform the news story clustering by exploiting the duality between stories and multi-modal concepts. Then, the novelty and redundancy detection is performed to capture the relationship among stories of a topic. To facilitate the fast navigation of news topic, a novel topic structure is then proposed to chains the dependencies of stories. A main thread is extracted to highlight the important aspects of a theme. A news video editing optimization algorithm can be directly applied to automatically select suitable video and speech contents from the original video source to create an edited video documentary.

Categories and Subject Descriptors
H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval – Information filtering, Search process.

General Terms
Algorithms, Design, Experimentation, Performance

Keywords
Novelty/redundancy detection, topic structure, topic threading

1. INTRODUCTION
News videos broadcast from different channels each day constitute an important but huge source for topic tracking and documentation. Most people are interested in keeping abreast of the main development and the highlights that happened in the news. Topic threading and auto-documentary appears as one prominent solution to this problem. The high-level semantic analysis and retrieval are meaningful for users. For example, when searching the topic “London Bombing”, it is more interesting if the system can provide a concise overview or a fresh development of the topic, rather than showing a list of items and letting users find out the dependencies among them. Automatic tracking of the evolution of news stories over the course of a topic can facilitate indexing and retrieval, and facilitate summarizing the topic into a documentary. The process of identifying dependencies among news stories is called as topic threading. The outcome is the semantic organization of news stories that allows users to rapidly interpret and analyze a topic.

This doctoral works consist of studies in story co-clustering and dependency threading, with particular focus in the fusion of multi-modal features. The investigated issues can ultimately lead to the construction of topic structure and the automatic generation of documentary. A news story consists of a sequence of shots in the visual track, a series of words extracted from speech transcript in the audio track, and a set of low-level (e.g., motion, audio and caption information) and high-level (e.g., face, anchor-person, dialog) concepts. Topic themes across different sources and time are identified by considering various features under a co-clustering algorithm. News stories clustered under a specific topic are then undergone a series of content validation such as novelty and redundancy detection. The outcome is a set of stories threaded according to their degree of novelty. To construct topic structure, story dependency pairs which indicate the dependency relationship among stories are further extracted. A topic structure binary tree is constructed by carefully chaining the dependency pairs. The topic structure can be further compressed by removing peripheral and redundant stories. We name the resulting “compressed tree” ---- main thread which basically summarizes the highlight of a topic. Based on main thread, optimization-based editing algorithm can be applied to automatically select suitable video shots and speech contents from the main thread tree to create an edited video documentary.

2. RESEARCH ISSUES
2.1 Story Co-clustering
Clustering continuously reported or time-evolving stories in news videos is a critical step for topic tracking and summarization. Instead of managing multimedia information with one-way clustering, we propose a two-way clustering algorithm, namely co-clustering. We exploit both textual and visual concepts for more effective story-level clustering.

Currently, we model the relationship of stories with textual and visual concepts under the representation of bipartite graph [3]. The textual and visual concepts are extracted respectively from speech transcripts and keyframes. Co-clustering algorithm is employed to exploit the duality of stories and textual-visual concepts based on spectral graph partitioning. The problem of co-
clustering is to partition the bipartite graph into sub-graphs by considering the co-occurrence between stories and textual-visual concepts. We have conducted experiments on TRECVID-2004 corpus [2]. Empirical results show that the co-clustering of news stories with textual-visual concepts get good performance, compared with using only visual or texture concept.

2.2 Story Threading
There is a great deal of redundancy in news stories. Users of news channels do not want to view every piece of videos over and over again. They are primarily interested in learning how the topic is evolved or what the highlight of this topic is. Novelty/redundancy detection is the second step for topic threading and auto-documentary which can identify the dependencies among stories and reducing the amount of redundant material presented to the user.

Previous attempts of novelty/redundancy detection [1, 4] are based entirely on textual cues. Here, we combine textual-visual information under Cosine Distance for story similarity measurement. The cosine of the angle between a news story vector and each previously delivered news story vector determines the redundancy score for that news story. Our experimental results show that Cosine Distance measure with textual and visual concepts is a robust redundancy measure compared with other popular measures such as Simple New Word Count, and Set Difference.

2.3 Topic Structure Generation
We propose and experiment a threading method based on binary tree to construct the topic structure. The topic structure for a news topic helps people get a quick and clear overview of the topic dependencies and also allows them to navigate through the topic faster.

The topic structure is represented by a binary tree. We are given a set of news stories on a given topic, and these news stories are ordered by their time of publication. Each news story is compared with all previously delivered news stories one by one to find the most similar news story with it by novelty/redundancy detection. After that, story dependency pairs which indicate the story development and dependency are formed based on their publication orders. Based on the story dependency pair set of a topic, a topic structure binary tree can be constructed.

The topic structure offers several unique features to facilitate browsing and auto-documentary. Through this topic structure, users can easily find the dependencies among these news stories which are useful for content summarization.

3. APPLICATION: DOCUMENTARY

3.1 Highlight Extraction
An effective way to attract a viewer is to present a documentary that is as compact as possible, yet preserves the most critical features required to tell a topic. We propose a method beneficial for auto-documentary by extracting the main thread of a topic structure. In a topic, there usually exist many peripheral and redundant news stories. To generate a documentary, these stories should be automatically detected and removed from the main thread topic structure. The redundant news stories have minor development, and carry duplicate information to users. Typical examples are news reviews and news stories with same contents but from different sources. For peripheral stories, they contain fresh information, but without further evolution. They usually depict isolated but not so important themes which can be pruned when other interesting and evolving themes exist. In order to give users a concise view of the main thread of a topic, the peripheral and redundant nodes should be removed to obtain a clear main thread and provide it for next step auto-documentary.

3.2 Automatic Editing
News summary can be generated by assembling the news stories in the main thread for auto-documentary. However, this is just the beginning step towards auto-documentary. Although inter-story redundancy has been removed, intra-story redundancy such as repeated shots and spoken sentences still exist.

A news video editing optimization algorithm is then proposed to reduce the intra-story redundancy of main thread, while still preserving the coherence of audio and visual modalities. Specifically, the whole procedure is a selection-elimination process, where the upshot of this process guarantees the coherence and conciseness of final edited video. We formulate this process as a knapsack problem. The redundancy detection of shots and sentences are performed at visual and audio tracks respectively. A news video editing optimization algorithm is applied to automatically select suitable video shots and speech contents from the original video source to create an edited video documentary. After removing the selected redundant shots and sentences, the remaining parts at visual and audio tracks are assembled to form a concise documentary.

4. FUTURE WORK
For co-clustering, though encouraging, further analysis such as how to select the best number of clusters and how to improve precision-recall by other concepts (e.g., audio, motion) is still required. Outlier detection in news video can improve the performance of clustering, which is worth exploring further. The First Story Detection (FSD) of a topic is still a hard problem. Typical topic structure types need to be explored. The synchronization of speech and visual tracks needs to be further addressed, especially when the speech is come from the interviewers. Furthermore, effective and efficient near-duplicate image detection is another critical issue to consider. Once these techniques are solved, the news video auto-documentary system will be applied to the mobile and wireless area.

Acknowledgement
I would like to thank my research supervisor Dr. Chong-Wah Ngo for his guidance on my Ph. D works described in this paper.

5. REFERENCES