History Rhyme: Searching Historic Events by Multimedia Knowledge

Yifan Xiong[1], Jia Chen[2], Qin Jin[1]*, Chao Zhang[1]
[1]School of Information, Renmin University of China
{xiongyf, qjin, 2010202526}@ruc.edu.cn
[2]School of Computer Science, Carneige Mellon University
jiac@cs.cmu.edu

ABSTRACT

This demo shows a novel system "History Rhyme" which searches historic events by multimedia knowledge. Different from existing historic events related works, we focus on the retrieval of historic events by semantic related multimedia knowledge. Our system can not only search historic events based on keyword queries, but also retrieve similar historic events to a given event based on chosen facets. In both cases, the system returns top retrieval results and shows the image profiles of each historic event. To build such functions, we automatically mine knowledge from multimedia data and index each historic event. Our online demo is available at http://222.29.193.164/HistoryRhyme.

Keywords

Historic Event Search; Multimedia Knowledge

1. INTRODUCTION

A famous quote says "History does not repeat itself, but it does rhyme". It not only shows that history knowledge is valuable but also indicates that searching similar historic events is an important tool in exploring and studying history. Work [2] focuses on preserving multimedia history knowledge through image profiling. Work [1] links images to phases of historic events by semantic image profiling. In this demo, we propose a system to mine finer multimedia knowledge and use such knowledge to index historic events for retrieval. We name our system "History Rhyme". Different from work [3] which focuses on searching for historic events based on the location inferred from the query photo, our work focuses on retrieval of historic events by semantic related multimedia knowledge.

The multimedia knowledge includes three types: text knowledge, visual knowledge and cross-media knowledge. The text knowledge consists of exact detailed information related to the event, including person, location and time. Such knowledge could be used to measure similarity in event facets such

---

*Qin Jin is the corresponding author

as whether a certain person is involved in both events, or whether both events happened in the same year. The visual knowledge consists of rich concepts, including objects and scenes. Such knowledge could be used to measure the similarity in event facets such as whether planes are involved in both events, or whether both events happened in railway stations. The cross-media knowledge is represented by a set of image-phrase pairs with each pair describing which image in the historic event image profile corresponds to which phrase in the event text description. It is very useful when users are exploring a specific historic event. It links the knowledge in the image media to that in the text media.

We integrate the above three types of knowledge in one system for users to:

1. use keywords to search historic events. We provide a keyword query input interface since it is the standard input interface of the state-of-the-art search engines and users are already familiar with it. Furthermore, we provide keyword suggestion functions to ease the requirement of history background knowledge for users. Both text and visual knowledge is used in this function.

2. search similar historic events to a given event in the chosen facets/aspects. This function is particularly useful for browsing the historic event dataset. Users can use this function to discover the rhyme of history by fixed facets, or to compare events under certain facets. Both text and visual knowledge is used in this function.

3. explore the image profiles of a historic event Our system organizes images related a historic event based on the cross-media knowledge. To be specific, we group images according to image-phrase pairs which are mined from the historic events multimedia data. All images linked to the same phrase are grouped together, which is a concise interface for users to browse the image-phrase pairs under a historic event.

2. SYSTEM OVERVIEW

There are three key functions in our system: (1) keyword search function, (2) similar events search function and (3) image profile explore function. As shown in figure 1, all functions are based on events indexing and ranking using historic knowledge mined from multimedia data. The keyword search function and similar events search function use text and visual knowledge while image profile explore function uses cross-media knowledge.

Use keyword to search historic events: As shown in Figure 2, users can input a keyword query to search his-
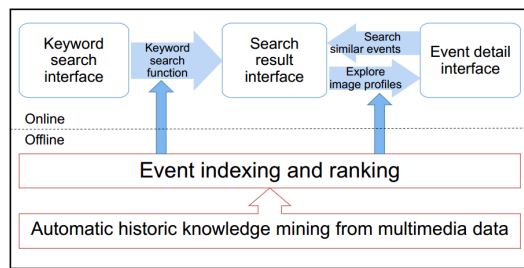
Figure 1: Framework of History Rhyme System

toric events in the keyword-search interface. There are three input fields in this interface, users can choose to input an object/scene class (area 1), or input a historical person's name (area 2), or select a history period (area 3). Inputs in one or more fields are allowed. Real-time suggestions are provided for a better experience if users choose to type the keywords. Additionally, to show the typical events in our dataset, we provide some query examples (area 4). Once the search-button is clicked, our system will return search results using both text and visual knowledge and users will be brought to the search-results interface as shown in Figure 3. The search-results interface shows the ranked results according to the user's query. The top input field shows keyword query (area 1), while the rest area shows the ranked results. Each result record corresponds to one historic event in our historic events corpus. Each event is shown by its text description (area 3) and key phrases (area 4), detailed image profile under this event will be shown after clicking on the header link (area 2).

Search similar historic events to a given event in the chosen facets: Once users click on a historic event's header link, our system will show the event-details interface whose top part is shown in Figure 4. Users can read the detailed event description (area 1), choose certain facets in this event (area 2) and click on the search button (area 3) to retrieve similar events to discover the rhyme of history with respect to this event. Results will be shown in the search-results interface which is similar to Figure 3.

Explore the image profile of a historic event: To explore every facet of a certain event, users can look through the image profile part in the event-details interface as shown in Figure 5. In this interface, grouped images will be shown using cross-media knowledge. Some famous images are shown in a cascade layout (area 1) to represent this event. Image-phrase pairs are browsed by phrases (area 2) and images related to certain phrase are grouped together (area 3).

3. TECHNICAL DETAILS

The "History Rhyme" system is built based on our previous work [1], technical implementation and algorithms can be found at that work.

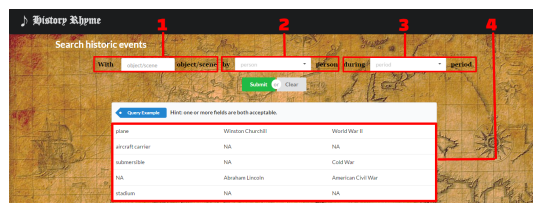**Automatic knowledge mining from multimedia data:**
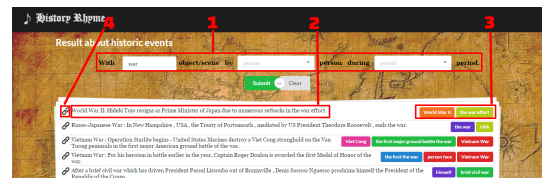

Figure 2: Keyword-search interface

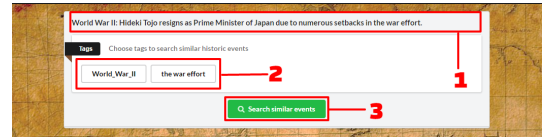**Figure 3:** Search-results interface

Figure 4: Text part in the event-details interface

Having crawled multimedia data of historic events from Internet, we extract person and period knowledge from text media using natural language processing method [4], object and scene knowledge from image media using pre-trained DCNN [6]. We design a score to automatically evaluate the similarity between a phrase and an image under certain event so that to link a phrase to an image for automatic knowledge mining. The evaluation details are available at our previous work [1].

Event indexing and ranking: Altogether 1,080 historic events and about 23,352 image-phrase pairs are indexed in our system using cross-media knowledge. We use word2vec [5] to calculate the similarity between the user query and key phrases and rank the top retrieved events.

4. CONCLUSION

The main contributions of the proposed system are summarized as follows:

1. It's the first application to retrieve historic events by semantic related multimedia knowledge.

2. Our system has automatically mined multimedia knowledge from 23,352 image-phrase pairs in 1,080 historic events.

3. Our system provides convenient interfaces for users to search historic events by keywords, explore image profiles of historic events and discover the rhyme of history by fixed facets.
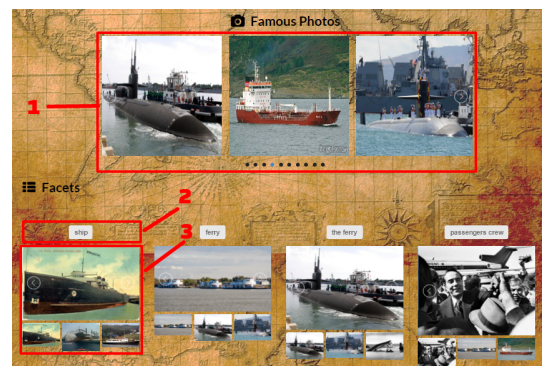
5. ACKNOWLEDGEMENTS

Figure 5: Image profile part in the event-details interface

6. REFERENCES

[1] J. Chen, Q. Jin, and Y. Xiong. Semantic image profiling for historic events: Linking images to phrases. In *Proceedings of the 24th Annual ACM Conference on Multimedia Conference, MM '16, Amsterdam, The Netherlands, October 15 - 19, 2016.*

[2] J. Chen, Q. Jin, Y. Yu, and A. G. Hauptmann. Image profiling for history events on the fly. In *Proceedings of the 23rd Annual ACM Conference on Multimedia Conference, MM '15, Brisbane, Australia, October 26 - 30, 2015*, pages 291–300, 2015.

[3] J. Chen, Q. Jin, W. Zhang, S. Bao, Z. Su, and Y. Yu. Tell me what happened here in history. In *ACM Multimedia Conference, MM '13, Barcelona, Spain, October 21-25, 2013*, pages 467–468, 2013.

[4] C. D. Manning, M. Surdeanu, J. Bauer, J. Finkel, S. J. Bethard, and D. McClosky. The Stanford CoreNLP natural language processing toolkit. In *Association for Computational Linguistics (ACL) System Demonstrations*, pages 55–60, 2014.

[5] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems 26*, pages 3111–3119. 2013.

[6] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014.