Search for Images of Historical Objects Using Wikipedia

Yating ZHANG[1] Adam JATOWT[1] Katsumi TANAKA[1]

[1]Graduate School of Informatics, Kyoto University Yoshida Honmachi, Sakyo, Kyoto 606-8501, Japan

E-mail:{zhang, adam, tanaka}@ dl.kuis.kyoto-u.ac.jp

Abstract We investigate the temporal image retrieval problem by utilizing knowledge from Wikipedia to help retrieve images. Since current search engines do not explicitly consider the time related to the objects displayed in images, we believe that our proposed temporal image retrieval algorithm can more completely satisfy search intentions. A search is done by inputting an entity name (with/without location and time information), and chronologically sorted search results are output to capture the evolution of the queried entity and consider its co-occurrence with other historically important entities. We propose a novel way to map images to Wikipedia articles to utilize Wikipedia's structure to estimate the historical importance of displayed objects. A combination of criteria is proposed to measure the historical importance of images under entity evolution scenarios and co-occurrence with other historically important objects.

Keyword Temporal Image Retrieval, Historical Importance, Wikipedia

1. INTRODUCTION

Digital images continue to gain popularity as a form of delivering information and conveying ideas. Image search and retrieval play important roles. Within the image retrieval research community, two paradigms have been commonly followed: text-based and content-based. When applied to annotated image collections, the first, which is often called tag-based image retrieval (TBIR), is both efficient and effective. To alleviate the problem of missing and noisy tags, Wu et al. [7] proposed a tag completion strategy that automatically adds missing tags and corrects noisy ones. Other proposals [4, 5] solve this problem by applying neighbor voting algorithms to confer more credibility on visually representative tags. On the other hand, the content-based approach relies on visual descriptions extracted from images and presents users with the images that best match the example input image. However, in both approaches, the information employed to retrieve images is confined to the image itself, such as tags or the visual content of images and queries, preventing the usage of such external knowledge bases as Wikipedia. In addition, we were inspired by the inadequate handling of object temporality in state-of-the-art image retrieval as well as by the lack of effective/multi-view temporal image retrieval methods. Therefore, we address these problems by leveraging not only the context information of images but also Wikipedia[1] to reflect the evolution of objects over time and their interaction for retrieving historically important images.

Object evolution is a new facet discussed in temporal image retrieval problems. Intuitively, the evolution of objects displayed in images can be visualized by the changes in the appearance of entities over time. Among the few works [9, 10] that have addressed this issue, the notion of time associated with objects has never been explored explicitly. For example, for a query on Kinkaku-ji, users may not only want images of it each year, but also those taken during crucial moments of its long history. In fact, Kinkaku-ji burned down in 1950 and was rebuilt in 1955. We believe that its visual timeline representation could be automatically constructed with representative images taken from the three time periods separated by the above two dates. In particular, the expected results would contain: an image of the original Kinkaku-ji before 1950 (Fig. 1a), one taken when it was burned down (Fig. 1b), and one taken after it was rebuilt (Fig. 1c). In our proposed method, such crucial historical moments of an entity are extracted from Wikipedia to represent the significant evolution of queried objects. The advantage of utilizing Wikipedia reflects its authority and credibility as well as its capability to provide historical information related to many real-world entities.

1 http://download.wikipedia.org/enwiki/

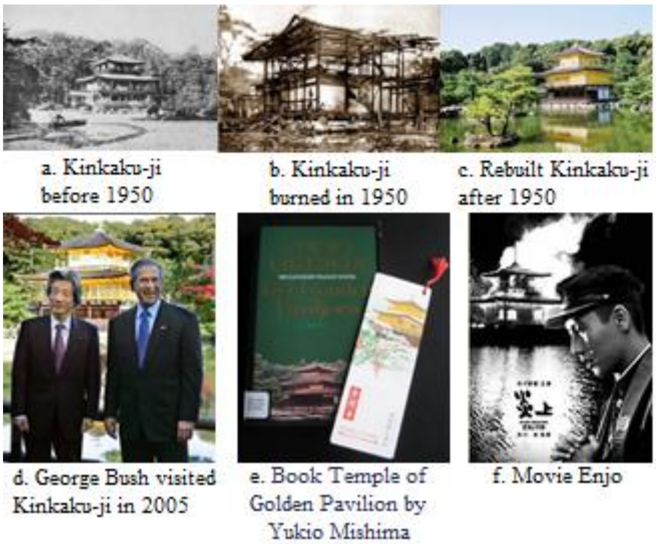

Fig. 1 Expected images of Kinkaku−ji

In addition to the images of an object showing significant changes over time, an image's historical importance is probably higher when it depicts the relations between a given object and other historically important entities. Visual co-occurrence can be considered one relation. For instance, Fig. 1d depicts that President Bush visited Kinkaku-ji with Prime Minister Koizumi. This image becomes important due to the appearance of President Bush, who is an important entity given Kinkaku-ji. A hidden co-occurrence relationship can exist between objects. For example, *The Temple of the Golden Pavilion*, (novel by Yukio Mishima) hows a novel called *The Temple of the Golden Pavilion (Kinkaku-ji)* by Japanese writer Yukio Mishima. Although the author does not visually co-occur with Kinkaku-ji in any image, this picture is deemed important since the book significantly addresses the temple's history. Accordingly, *Enjo* (1958 Japanese movie) is also chosen. *Enjo* is a movie based on Mishima's novel. Under this scenario, we developed a quantified model to study the historical importance of images by calculating the global importance of co-occurring objects based on Wikipedia's link structure (such as using PageRank) and analyzing the relative importance of other entities (both visually and non-visually) that co-occur with the queried object in a given image.

1.1 Problem Formulation

Our work is mainly composed of two stages: knowledge collection during the offline stage and image retrieval at the query time.

In the offline stage, our task gathers as much seminal information as possible about the entities and later utilizes it to retrieve images. We create a set of entities defined as E, where each entity e is associated with its Wikipedia article (we limit the entities to those with corresponding Wikipedia pages). For entity e, two facets of information are taken into consideration: *object evolution* and *entity co-occurrence*. We assume that the images depicting any significant snippets of the given object's evolution or essential interaction with other entities can be viewed as historical important images for a given object.

We formulate the problem of object evolution as detecting crucial moments throughout the entire history of a given entity. We define $\vec{T}$ as a vector of historically important dates related to entity e (we use "year" as a time unit for each t in $\vec{T}$); for each year t, we detect and extract $\overrightarrow{W^t}$, which represents a vector of representative *words* with *weights* to describe what happened in year t.

The other facet, entity co-occurrence, is approached by calculating the *global importance* of co-occurring entities based on Wikipedia's entire link structure and the *relative importance* of co-occurring objects that exist in local graphs "centered" on given entity e (details about defining local graphs are in Section 5). We utilize this co-occurrence information to measure the importance of interactions among entities for further exploring historically important images in which multiple entities visually co-occur or are non-visually related to each other. We call *co-occurring entity* coe, so its *global importance* is $I_g(coe|E)$, given all the entities under Wikipedia link structure and computing the importance of entity coe. For the *relative importance* denoted by $I_r(coe|e)$, the problem is formulated as follows. Given entity e, we compute the "importance" of entity coe with respect to e and obtain a vector of co-occurring objects $\overrightarrow{COE}$ with corresponding combined importance (weighted global importance and relative importance) of each entity. In this paper, the co-occurrence importance is independent from time. Intuitively, the interaction relationship among entities changes over time. However, to simplify the problem, we fix this relationship once the link structure of the Wikipedia pages is established.

At query time, when a user issues query q (for now we limit queries to entities in Wikipedia), we extend q with the top k most representative words in $\overrightarrow{W^t}$ for a given entity e and year t to build a new query. Then we extended queries to image search engine such as Flickr to get a collection of images that can be viewed as a candidate set of images. In terms of entity co-occurrence, query q is extended with the names of other entities with high importance given the queried object. The images returned from Flickr are then considered candidate images representing the interaction of other entities with the given object.

The final step is to select historically important images for the queried object from the candidate image set obtained from Flickr. The knowledge gathered in the offline stage is leveraged to filter the candidates by comparing the contextual information of the image with the extracted snippets annotating the object evolution and entity co-occurrence.

We clarify our **query model** as follows:

- *Input*: entity name e (optionally with location l and time t).
- *Output*: list of top k historically important images of entity e such that each image illustrates the evolution of entity e or annotates the interaction with other entities, and overall the results visually represent the history of e.

The remainder of the paper is structured as follows. In Section 2 we briefly discuss related work followed by clarifying our definitions in Section 3. Sections 4 and 5 elucidate how to utilize Wikipedia to collect important temporal information and co-occurring entities. In Section 4 we propose a novel algorithm for detecting the evolution of objects. Section 5 introduces composite historical importance considering both the global and relative importance of co-occurring objects with respect to a given entity. Section 6 discusses a two-step image retrieval stage by gathering candidate images from Flickr and selecting a final list of images with high historical importance. Conclusions are presented in Section 7.

2. RELATED WORK

Detecting the object evolution displayed in images can be generalized as a *temporal image retrieval and re-ranking* problem, which is one of our main objectives. In contrast to temporal text retrieval, temporal image retrieval is a fairly new research topic. [9] first proposed the idea of automatically capturing the time dimension included in web images from explicit or implicit temporal text queries based on query logs. However, even though extracting dates from query logs may lead to high recall, the precision is generally low. In addition, although the authors mentioned the notion of the evolution of entities, they failed to implement any method for finding historically important images. In contrast, we solve this problem by utilizing knowledge from Wikipedia for detecting not only the historically important time periods of entities but also utilizing the surrounding knowledge to support the significance. Another work by Kim and Xing [10] investigates a time-sensitive image retrieval problem by customizing search results depending on the time when a user issues a query. For example, if a query is issued for Kinkaku-ji in mid-January, the system will return images showing it in winter. Although such an approach is intuitive, we focus on the temporal information related to displayed objects rather than on the timestamps of images to portray the objects at different seasons. Our objective provides users with important images of a queried entity considering its entire timeline and its interaction with other entities over history.

Measuring the *historical importance of entities* is another important component of our system, and we use it as a facet to compute the historical importance of images. A similar approach was proposed by Takahashi [3] to evaluate the significance of historical entities using Wikipedia's link structure. However, this approach only focuses on the global measures of entity importance in which each entity is ranked relative to all other entities mapped in a graph. We propose a composite importance measurement that additionally explores the relative importance of entities in a local graph with respect to a given object. Evaluation of relative importance is often essential. For example, regarding Kinkaku-ji, the book *The Temple of the Golden Pavilion* is of low importance compared to other novels. However, since it makes a seminal contribution to Kinkaku-ji's history, it is highly important to the temple itself, even if it is not as crucial in general as many other books. White and Smyth focus on defining and computing the importance of nodes in a graph relative to one or more root nodes [11]. Their approach, which defines weighted paths and explicitly computes the relative importance, is also employed in our work. It is assumed that the longer a path is from the root node to the other node, the less importance is conferred along that path. For the computation of global importance, we applied Topic-sensitive PageRank [1]. We modified their work by introducing Wikipedia as a category and a knowledge basis and extended the topics to capture both time and location dimensions to evaluate the global importance of entities with respect to a specific topic, time, and location. Considering our final objective, applying the historical importance of entities to quantify the importance of images is a challenge that has not been studied so far to the best of our knowledge.

The idea of utilizing Wikipedia resources to retrieve images was also introduced [8] to enhance the performance of search engines to retrieve the most relevant facets of given object query logs and associated tags. Our work is different from theirs in two aspects. First and foremost, its scope is different: we solve a temporal image retrieval problem, while they focused on retrieving diverse facets with high relevance to an object (without considering temporal notion). Second, our work enriches the contextual information of images by mapping each image to several Wikipedia pages that are highly related to the displayed objects in images. This projection allows us to extend our work in the future to automatically create historical summaries of entities for better understanding the story behind images.

3. PROBLEM DEFINITION

3.1 Object Evolution

Object evolution can be viewed as the conceptual or physical changes in an object's attributes, such as name changes, object splittings, object mergers, or temporal variations in shape and color. Intuitively, the object evolution displayed in images can be visualized by the changes in the appearance of entities by time or identified as the variations in the contextual information of images.

3.2 Historical Importance of Entities

We define the historical importance of entities as a combination of global importance and relative importance.

The web can be defined as a very large graph where nodes represent web pages and directed edges represent hyperlinks among pages. We construct graphs from Wikipedia and consider each Wikipedia article an entity (node in the graph). In this context, the *global importance* of an entity is the "importance" of the corresponding article given the entire link structure of Wikipedia. Unlike global measures, the *relative importance* represents the interaction relations among entities in the graph.

3.3 Historical Importance of Images

The historically important images of a queried entity are identified as (1) the images depicting the significant evolution of an entity in time; (2) the images in which queried objects co-occur with other entities with high historical importance with respect to the queried object.

4. DETECTING OBJECT EVOLUTION

This section describes our method that extracts temporal information related to a given entity and detects the changes in an entity's attributes over time.

For entity e, we generate a vector of words $\overrightarrow{W^t}$ at recorded time point t (a year) to describe what happened with the entity in a particular year. Among those words, only those are needed which distinguish the event experienced by the entity at time t from other events. The "representativeness" of these words can be computed by the following equations:

$$P(w_j|e,t) = \frac{P(w_j,e,t)}{P(e,t)} = \frac{N(w_j,e,t)}{N(e,t)} \quad (1)$$

$$P(w_j|e) = \frac{P(w_j,e)}{P(e)} = \frac{N(w_j,e)}{N(e)} \quad (2)$$

$$\sigma(w_j|e,t) = \frac{P(w_j|e,t)}{P(w_j|e)}. \quad (3)$$

Equation (1) denotes the probability of the occurrence of w_j given the occurrence of entity e and time t. $N(w_j,e,t)$ is the number of sliding windows in which w_j,e,t are recorded together. Here, a sliding window is defined as a fixed number of sentences. Eq. (2) denotes the probability of the occurrence of w_j given only the occurrence of entity e. In this sense, "representativeness" σ is defined in Eq. (3) as the division of the above two probabilities. Intuitively, the word is more representative if it is more likely to appear in a particular time and less likely to appear over the entire timeline of the entity.

We select top c words with high "representativeness" to refer to the events that occurred in time t. Representative words $\overrightarrow{W^t}$ and corresponding "representativeness" score $\overrightarrow{R^t}$ are denoted as follows:

$$\overrightarrow{W^t} = [w_1, w_2, w_3 \dots\dots w_c]$$
$$\overrightarrow{R^t} = [r_1, r_2, r_3 \dots\dots r_c].$$

In this sense, pairs of words and representative scores $\{\overrightarrow{W^t}, \overrightarrow{R^t}\}$ imply the evolution of an object over time.

5. COMPUTING HISTORICAL IMPORTANCE OF CO-OCCURRING ENTITIES

The historical importance of co-occurring entities is approached by computing the *global importance* of entities based on the entire link structure of Wikipedia pages and the *relative importance* of the co-occurred entities with respect to a given entity. Thereafter, we obtained a vector of co-occurred objects $\overrightarrow{COE}$ with corresponding combined importance (weighted global and relative importance) for a given entity.

5.1 Computing Global Importance of Entity

Based on the link structure of Wikipedia, we compute the historical importance of all the entities in E within a specific topic, time, and location by developing a model of topic-time-location sensitive PageRank.

5.1.1 Outline of Historical Importance Measurement

In our approach to measure the historical importance of entities, we precompute the importance scores offline as in topic-sensitive PageRank [1]. However, instead of taking only the topic into consideration, we introduce two other dimensions for each entity: temporal and spatial. For each dimension, we compute offline a vector of the scores of the importance of an entity with respect to this dimension's categories. To be more precise, for each entity, we compute a vector of scores with respect to various topics, a vector of scores during different time periods, and a vector of scores over distinctive regions. At the query time, the historical importance scores are generated based on the topics, time periods, and regions identified from the query to compose a mixture of scores for those entities that are closely related to the query. Here, we assume that the link structure of the Wikipedia pages implicitly confers the importance of the entities. Under this hypothesis, we regard each Wikipedia article as an entity, and the page with the highest topic-time-space sensitive PageRank score has the highest historical importance within a specific topic, during a unique time period, and in a particular region.

5.1.2 Identification of Topic Category of Entity

Since we utilize Wikipedia pages as webpage resources to compute the historical importance scores of entities, we employed a list of Wikipedia's major topic classifications [2] to map each Wikipedia page: each entity to one of these categories to further compute the PageRank vectors of the topics. This list contains 26 categories and is used throughout Wikipedia to organize the presentation of links to articles [2]. Therefore, these 26 categories are guaranteed to cover all Wikipedia articles.

For each topic category c_j, we also compute a bag of words representation W_j, consisting of all the nouns displayed in Wikipedia articles belonging to this category. Let W_{jk} denote the number of occurrences of word k in Wikipedia articles under category c_j.

5.1.3 Computation of PageRank Vector of Topics

We adopt the topic-sensitive PageRank and compute rank vectors over the topics:

$$\overrightarrow{Rank_j} = (1-\beta)M \times \overrightarrow{Rank_j} + \vec{s}_j \frac{\beta}{|s_j|}.$$

Let S_j be the set of pages in Wikipedia-topic category c_j. We compute PageRank vector $\overrightarrow{Rank_j}$ for topic c_j. $1-\beta$ is a damping factor added during the rank propagation to guarantee convergence to a unique rank vector. $\vec{s}_j$ is considered a *teleport* vector in which the user is designed to jump to a random page chosen non-uniformly where

$$s_{jk} = \begin{cases} 1 & k \in S_j, \\ 0 & k \notin S_j. \end{cases}$$

$|s_j|$ is the number of 1s in $\vec{s}_j$.

5.1.4 Computation of PageRank Vector of Time Periods

As for the dimension of time, we consider decade the category to divide time periods [0-2013] into 202 clusters, $t_0 \dots t_{201}$. We also apply the methodology discussed in Section 5.3 to compute the PageRank vector over different time periods.

5.1.5 Computation of PageRank Vector of Regions

We chose "country" as a unit of regions to denote spatial information. The historical importance of an entity in most cases varies across countries. For example, Yoshida Shōin, a distinguished intellectual in Japan, wielded great influence during the Meiji Restoration. However, he had no influence on Western countries. In this scenario, we employ the list of administrative divisions by country [6] based on the ISO 3166 standard[2] published by the International Organization for Standardization in which there are 198 countries. We use $r_0 \dots r_{197}$ to represent 198 non-overlapped regions. When considering the definition of countries, we must consider the changes in borders over time as well as a country's existence, changes in its names, and so on. However, for simplicity, we assume that countries existed more or less within their current borders and use fixed regions $r_0 \dots r_{197}$ throughout time. Obviously, since this assumption is quite strong, in the future, we plan to extend this approach using distinctive lists for different periods to capture certain aspects of country evolution over time.

Finally, we compute PageRank vectors within different regions based on the methodology discussed in Section 5.3.

5.1.6 Computation of Historical Importance Score Based on Query

The final historical importance score of an entity is computed at the query time. Our system provides users with a query box with three input areas: entity, time, and region. The time and region query parts are optional. Given query q, we compute the class probabilities for each

---

[2] ISO 3166 is the International Standard for country codes and codes for their subdivisions. http://www.iso.org/iso/country_codes

category under three dimensions: topic, time, and region.

a. For topics, $P(c_j|q)$ can be easily computed from the topic bag of words W_j, and part of $P(c_j)$ is treated as uniform:

$$P(c_j|q) = \frac{P(c_j) \cdot p(q|c_j)}{p(q)} \propto P(c_j) \cdot \prod_i p(q_i|c_j).$$

b. To calculate $P(t_j|q)$, for the time period, we map the time in the query box to one of the time period categories. The input of time is currently limited to a year format, such as 1889, a period format, such as 1889-1950, or a century format, such as 20th. Then we map the time information to one of the time classes.

c. For regions, we limit the queries to city and country names. Just as the case of time periods, we let queries fall into one spatial category, $P(r_q|q) = 1$, and the probabilities for other classes are 0.

Next, we compute the query-sensitive historical importance score of each Wikipedia page. Let rank_{jd} be the rank of Wikipedia page d given by the rank vector. Weights β_1, β_2 and β_3 denote the weight of each dimension. In most cases, they are supposed to be equal, and when the query contains no time or space information, the corresponding weight is set to 0:

$$\begin{aligned} Score_{topic \cdot time \cdot space} &= \beta_1 \cdot \sum_j p(c_j|q) \cdot rank_{cjd} + \beta_2 \\ &\cdot \sum_j p(t_j|q) \cdot rank_{tjd} \ + \beta_3 \\ &\cdot \sum_j p(r_j|q) \cdot rank_{rjd}. \end{aligned}$$

5.2 Computing Relative Importance of Entity Using Weighted Path

Wikipedia is mapped to a directed graph $G = (V, E)$ in which nodes V are defined as Wikipedia articles (also entities) and edges E are the hyperlinks among articles. We use $\mathcal{P}(e, coe)$ to denote a particular set of paths between a given entity and its co-occurring entities.

According to [11], the following are the two main properties that should be satisfied using the approach of weighted paths: 1) two nodes are related by the paths that connect them and 2) the longer a path is, the less importance is conferred along that path. To achieve these we define $I_r(coe|e)$ as follows:

$$I_r(coe|e) = \sum_{i=1}^{|\mathcal{P}(e,coe)|} \lambda^{-|\mathcal{P}_i|}, \qquad (4)$$

where λ is a scalar coefficient, $1 \le \lambda \le \infty$, that determines how much importance is conferred from e to coe. In this formulation, the amount of importance decays exponentially with the increase in the path length.

Before exploring the set of paths from e to coe, we define our paths as the shortest path, which are the sets of K-short Node-Disjoint Paths that have neither edges nor nodes in common. We use a set of paths instead of one shortest path between two vertices because multiple paths add more importance to coe relative to e than only considering the shortest path. A breadth first search algorithm is employed to find a good set of paths $\mathcal{P}(e, coe)$.

The final historical importance is given by linearly combining global importance $I_g(coe|E)$ and relative importance $I_r(coe|e)$:

$$I(coe|e) = (1-\theta) \cdot I_g(coe|E) + \theta \cdot I_r(coe|e). \qquad (5)$$

Here, θ represents the "closeness" between e and coe. In Eq. (5), when θ is larger than 0.5, then such coe closer to e will be preferred. For example, in the Kinkaku-ji case, according to Fig. 2, *The Temple of The Golden Pavilion* has higher relative importance than George W. Bush, relative to Kinkaku-ji, since there is no link between them in the Wikipedia pages. In this sense, if θ is set to 1, Fig. 1d will not be returned.

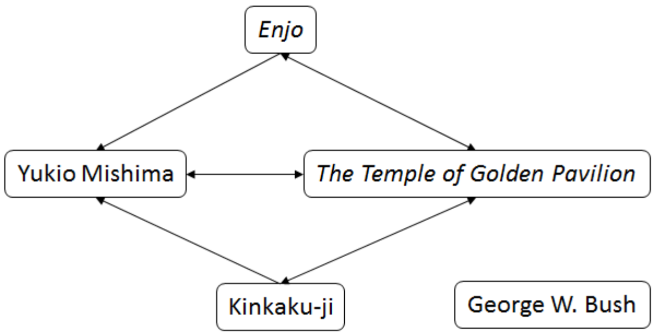

Fig. 2 Interaction among Kinkaku-ji and other entities

6. IMAGE RETRIEVAL

In this paper, we proposed a two-step image retrieval approach by first utilizing Flickr API to obtain as many images as possible related to the queried entity and retrieved a final list of images with high historical importance.

6.1 Creating Image Collection

At the query time, when a user issues query q (i.e., an entity with a corresponding Wikipedia page), we extend q with words in $\overrightarrow{W^t}$ and with years in $\vec{T}$. Then we call Flickr API with such a query to get a set of candidate images S. Each image is represented by a vector of tags.

6.2 Computing Historical Importance of Images

The final list of images is classified in two groups: one that implies the evolution of the queried object, and another that depicts the interaction with other entities. For object evolution, we compute the *TF-IDF* score for each tag of the image, based on the frequency of tag d in image I_i compared to the frequency of tag d in the entire image set S. Then for image I_i, we obtain a vector of tags with their *TF-IDF* scores that consist of pairs of tags and scores denoted as $\{\vec{D}, \vec{V}\}$. Remembering the pair of values we computed in Section 4, $\{\overrightarrow{W^t}, \overrightarrow{R^t}\}$ for queried entities, we compare these two values by calculating their cosine similarity. If the similarity exceeds threshold ε, then image I_i is identified as a historically important image that is useful to depict the evolution of a queried object.

With respect to co-occurrence with other entities, we detect the appearance of co-occurring entity names in image tags and rank the image by $\boldsymbol{I}(\boldsymbol{coe}|\boldsymbol{e})$ discussed in Section 5. Candidates with high $\boldsymbol{I}(\boldsymbol{coe}|\boldsymbol{e})$ scores are then returned.

7. CONCLUSION

We investigated a variant of temporal image retrieval, in which object evolution and interaction among entities are explicitly considered and the historical knowledge derived from Wikipedia is utilized to help retrieve images. This paper's contribution is providing a novel way to enable users to visualize the evolution of objects as well as the relationships among entities. In this context, we developed a two-stage temporal image retrieval model by gathering the knowledge of entities by time and chronologically retrieving and re-ranking historically important images.

ACKNOWLEDGEMENTS

This work was supported in part by Grants-in-Aid for Scientific Research (No. 24240013) from MEXT of Japan.

REFERENCE

[1] T. H. Haveliwala. Topic-sensitive PageRank: A context-sensitive ranking algorithm for web search. *IEEE Transactions on Knowledge and Data*

[2] Wikipedia: Category: Main topic classifications http://en.wikipedia.org/wiki/Category:Main_topic_classifications

[3] Y. Takahashi, H. Ohshima, M. Yamamoto, H. Iwasaki, S. Oyama, K, Tanaka. Evaluating Significance of Historical Entities Based on Tempo-Spatial Impacts *Analysis Using Wikipedia Link Structure. In HT2011: Proceedings of the 22nd ACM Conference on Hypertext and Hypermedia*, pages 83-92, June 6-9, 2011.

[4] X. Li, C. G. M. Snoek, and M. Worring. Neighbor Voting: Learning Tag Relevance by Neighbor Voting for Social Image Retrieval. In *Proc. of MIR*, pages 180-187, 2008.

[5] A. Sun and S. S. Bhowmick. Quantifying Tag Representativeness of Visual Content of Social Images. In *Proceedings of the international conference on Multimedia*, Pages 471-480, 2010

[6] Wikipedia. Table of administrative divisions by country http://en.wikipedia.org/wiki/Table_of_adminstrative_divisions_by_country

[7] L. Wu, R. Jin and J, A.K. Tag Completion for Image Retrieval. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 35(3), 2011.

[8] R. Zwol and B. Sigurbjornsson. Faceted Exploration of Image Search Results. In *Proceedings of the 19th international conference on World wide web*, Pages 961-970, 2010.

[9] G. Dias, J. G. Moreno, A. Jatowt and R. Campos. Temporal Web Image Retrieval. 19th International Symposium on String Processing and Information Retrieval 7608(2012), pages 199-204, 2012.

[10] G. Kim and E. P Xing. Time-Sensitive Web Image Ranking and Retrieval via Dynamic Multi-Task Regression. In *Proceedings of the 6th ACM International Conference on Web Search and Data Mining*, pages163-172, February 4–8, 2013.

[11] S. White and P. Smyth. Algorithms for Estimating Relative Importance in Networks. In *Proceedings of the 9th ACM SIGKDD international conference on Knowledge discovery and data mining*, Pages 266-275, August 24-27, 2003.