HAL
open science

Temporal Web Image Retrieval

Gaël Dias, José G. Moreno, Adam Jatowt, Ricardo Campos

To cite this version:

Gaël Dias, José G. Moreno, Adam Jatowt, Ricardo Campos. Temporal Web Image Retrieval. 19th edition of the International Symposium on String Processing and Information Retrieval (SPIRE 2012)., Oct 2012, Cartagena de Indias, Colombia. 6 p. hal-01070166

HAL Id: hal-01070166

https://hal.science/hal-01070166v1

Submitted on 30 Sep 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Temporal Web Image Retrieval

Gaël Dias[1,4], José G. Moreno[1], Adam Jatowt[2], and Ricardo Campos[3,4]

[1]Université de Caen Basse-Normandie, UMR 6072 GREYC, F-14032 Caen, France
[2]Kyoto University and Japan Science and Technology Agency, Japan
[3]Polytechnic Institute of Tomar and LIAAD-INESC TEC, Portugal
[4]Center of Mathematics, University of Beira Interior, Portugal

Abstract. Temporal Web Image Retrieval can be defined as the process that retrieves sets of Web images with their temporal dimension from explicit or implicit temporal text queries. Supposing that (a) the temporal dimension is included in image indexing and (b) the query is explicitly expressed with a time tag (e.g. "Fukushima 2011"), the retrieval task can be straightforward as image retrieval has been studied for several years with success. However, text queries are usually implicit in time (e.g. "Second World War") and automatically capturing the time dimension included in Web images is a challenge that has not been studied so far to the best of our knowledge. In this paper, we will discuss different research issues about Temporal Web Image Retrieval and the current progresses of our research in temporal ephemeral clustering and temporal image filtering.

1 Introduction

In the image retrieval community, in contrast to the text retrieval one [1] [2], there has been little discussion about the concept of time-sensitive queries and temporal retrieval. In fact, many queries in image retrieval are atemporal as a user wishes to obtain images without any particular time range in mind (e.g. "Rabbits").

However, there is a range of queries which contain a temporal dimension. In this case, the user may like to obtain images, which show objects as they were in the past or they will be in the future. Within this context, there exist two different categories of queries. On the one hand, there are explicit temporal queries, which contain absolute dates or time periods (e.g. "Paris 1945", "Eiffel Tower 1889-1900", "World Cup 2014"). On the other hand, there are implicit temporal queries, whose time component is hidden in the semantics of the query (e.g. "Second World War", "Mayas").

So far, there are no efficient solutions that would chronologically order relevant images in large unconstrained collections such as the Web, where images have no trustworthy metadata attached and where many queries are not explicitly formulated in the time dimension. Thus, we propose to tackle this challenging task and show initial results through an architecture based on two different steps: temporal ephemeral clustering by query expansion and temporal image filtering by temporal image classification.

We divide the problem of returning images, which satisfy text queries with temporal dimensions to the following subtasks: (a) detecting and recognizing the temporal component of a user query, (b) estimating the timestamp of images and (c) finding relevant images. In particular, step (c) has been well studied and there exist many successful

methods, which are used by state-of-the-art Web image search engines [3]. However, to the best of our knowledge, there have been no solutions for steps (a) and (b) in the context of Web image retrieval. Note that step (a) is trivial in the case of explicit temporal components (e.g. "Berlin 1944") but difficult for queries with implicit temporal components (e.g. "Winter Olympics Sapporo", "California Gold Rush").

For step (a), we present a temporal ephemeral clustering strategy to temporally organize Web image search results for implicit temporal text queries. Ephemeral clustering, also known as post-retrieval clustering, aims to cluster Web search results on the fly as they are returned by a Web search engine (i.e. text or images). In the specific context of Temporal Web Image Retrieval (T-WIR), Web image results are retrieved by temporal query expansion and rearranged into temporal clusters as presented in Figure 1. This new paradigm can be useful for different activities (e.g. education, entertainment, environment) and different populations (e.g. kids, elderly). Indeed, users interested in the evolution of entities can obtain timeline-like overviews with representative images for significant years. For example, it is possible to see the changes in the appearance of persons (e.g., Madonna, Michael Jackson), places (e.g., Moscow, New York) or monuments (e.g. Twin Towers, Eiffel Tower).

Fig. 1. Examples of Implicit Temporal Queries (a) Moscow (Cities); (b) Madonna (People); (c) Twin Towers (Monuments). Clusters are Automatically Time-Tagged.

Step (b) principally aims to filter out images obtained from step (a), which may not be representative of the required time period. Indeed, temporally incoherent images may be retrieved by state-of-the-art Web image search engines as indexing is usually made upon surronding text and as such prone to error. In fact, image filtering can be seen as a sub-task of a broader area, whose main aim is to correctly timestamp images and as such produce high quality image temporal indexation. Indeed, timely irrelevant images may be retrieved. To solve this problem, we propose a temporal image classification approach to automatically detect image timestamps based on visual features. It is important to note that in the case of digitized analog images, as opposed to born-digital images, there are two timestamps: the timestamp when the analog photo was taken in the real world and the timestamp when this photo was later digitized. In this paper, we focus on the former and our objective is to determine the actual time point at which the world view represented on the image was captured.

The main contributions of this position paper are (1) the definition of research challenges embodied by T-WIR, a new research topic and (2) the current progresses of our research in T-WIR through a proof of concept architecture tackling both temporal ephemeral clustering and temporal image filtering.

2 Temporal Ephemeral Clustering

The first idea is to retrieve and organize Web image search results based on a timeline. In the textual domain, promising results have been obtained for a historical point of view [4] or a futuristic vision [5]. However, temporal visual information retrieval has received very little attention. As far as we know, unless there exist image databases

populated with temporal metadata (e.g. Flickr) and a text query is explicitly expressed with its temporal expression (e.g. "New York from 2000 to 2012"), no solution has ever been proposed in a more realistic environment.

As a consequence, in order to recognize the temporal component of an implicit text query and organize Web image results on a timeline, we propose an ephemeral clustering strategy. In particular, we may find temporal cluster names based on temporal features mined from different Web resources such as[1]:

- Query suggestion engines (e.g. Google, Yahoo!, Wikipedia) following the regular expression "`query [1-2]+`", (e.g. "Olympic Games 2"),
- Web snippet implicit temporal tagging as proposed in [6],
- Web snippet knowledge-based temporal annotation using TAGME [7] for entity recognition and YAGO2 for year date retrieval [8].

Then, the query is expanded with its discovered temporal dimensions using one or a possible combination of the aforementionned strategies. For example, the query "Olympic Games" would be expanded like "Olympic Games 2012", "Olympic Games 2008", and so forth as year dates have been discovered. Finally, the expanded query is sent to an image search engine in order to retrieve temporally relevant images, as shown in algorithm 1. This bootstrapping approach allows to determine important dates related to images as well as increase the total pool of temporally relevant images as demonstrated in [9]. Indeed, at each expansion step, new relevant images may be retrieved. Positive results of this step are illustrated in Figure 1.

Although this methodology is straightforward, it has proved to lead to interesting results. However, discovering the time dimension of an implicit temporal query from text is a hard task and may be prone to error. As explained in [6], while extracting dates from query logs may lead to high recall, precision is surprisingly lower compared to their methodology based on Web snippet implicit temporal tagging. Moreover, the time dimension depends on the meaning of the query. For example, the query "Jaguar" may embody a temporal intent for the concept of "car", while the concept of "animal" may be atemporal. As a consequence, future work still needs to be developed for this specific task, in particular the third strategy aforementionned.

Algorithm 1 Temporal Ephemeral Clustering.

Input: $TextQuery$
Output: $TimeTaggedImgClusters$
1. $QuerySetDates$ = $ExtractTemporalIntents(Query)$
2. For each element $QueryDate_i$ in $QuerySetDates$
3. $TimeTaggedImgCluster_i$ = $ImgSearch(QueryDate_i)$
4. $TimeTaggedImgClusterName_i$ = $QueryDate_i$
5. return $TimeTaggedImgClusters$

[1] So far, the first and second solution have been implemented in our architecture.

3 Temporal Image Filtering

Supposing that the ephemeral clustering step is successful, the set of retrieved Web images may be temporally incoherent i.e. the images may not belong to their associated time cluster. For instance, a color photo may be retrieved for periods where color photography did not exist or was not yet widespread. As a consequence, in order to filter out incorrectly classified Web images, we propose an automatic methodology for temporal image filtering.

Given an arbitrary image, we wish to estimate the approximate date when it was taken. On the one hand, for born-digital images, the situation is relatively easy as there is only one timestamp, which in some cases can be retrieved from image metadata (e.g. EXIF). Note however, that for certain files, the timestamp is impossible to retrieve or can wrongly be assigned (e.g. camera erroneous time settings). On the other hand, for digitized copies of old analog images, discovering temporal timestamp is not trivial. Note also that if we would focus only on the born-digital images then, obviously, the length of the timeline that such approach covers would be very short (i.e. the last two decades during which digital photographing became popular).

In the prior work, different clues have been proposed to manually timestamp images mainly based on the physical properties of the media (e.g. paper type, size, coated, color scheme) [10]. In another approach, some references present hints based on visual characteristics of objects, places or people in the pictures. On the other hand, as far as we know, only one automatic method exists to temporally timestamp images, Kodak 2010 patent [11]. In particular, it focuses on dating photos based on distinguishing marks that may appear on the back of the photo (e.g. brand or written dates) or on the over-print of the image (e.g. logo). As a consequence, the scope of this methodology is limited and provides no real solution to our needs. Therefore, we propose to define a learning model with specific visual features capable of classifying the temporal dimensions of images. Note that in this research, we are not concerned about detecting the age of objects portrayed on photos, but on dating images based on their visual features.

To address initial experiments, we built a small dataset, where five city names were used as queries to extract images with their temporal metadata by querying Flickr[2] constrained by five temporal periods i.e. classes defined from historical progress of photography. The periods used were [1826, 1925), [1925, 1948), [1948, 1968), [1968, 1982), [1982, 2011] and correspond to intervals defined in chronological order by: first photo, camera democratization for black/white, camera democratization for color, first reflex camera, first digital camera and actual date. The number of images retrieved by each query-period was limited to 50. The final number of images in the dataset is 1170 and each period is considered a class.

After multiple experiments were performed based on different image features, three texture and color descriptors were selected: ScalableColor [12], FCTH [13] and CEDD [14]. The first descriptor is a color histogram extracted in the HSV color space and the others two descriptors are histogram representation of color distribution under different texture patterns. In total, four hundred low-level features were used to describe each image. Finally, we performed a 10-fold cross-validation using a Multiclass SVM classifier

[2] http://www.flickr.com (Last access on July 2012).

with a Linear Kernel and default parameters. The results show an average F-Measure of 0.509 ($\sigma = 0.042$) and average ROC Area of 0.757 ($\sigma = 0.046$). Note that the F-Measure of a random classifier is 0.2. Results are presented in Table 1.

	[1826, 1925)	[1925, 1948)	[1948, 1968)	[1968, 1982)	[1982, 2011]
# Images	237	205	228	246	254
Precision	0.500	0.528	0.456	0.545	0.525
Recall	0.624	0.512	0.408	0.520	0.488
F-Measure	0.555	0.520	0.431	0.532	0.506
ROC Area	0.796	0.798	0.688	0.768	0.735

Table 1. Precision, Recall, F-Measure and ROC Area values for Five Temporal Periods.

4 Conclusions and Perspectives

In this paper, we presented the foundations of a new research area called Temporal Web Image Retrieval. We developed a methodology as proof of concept based on (a) temporal ephemeral clustering and (b) temporal image filtering. The first results support our hypotheses and new challenges have emerged. Indeed, the perspectives of T-WIR are numerous. For the first step, many ideas can be proposed to recognize the temporality of implicit text queries based on Web snippets analysis [6], Web documents [15], query logs distributions [2] or Web archive studies [4]. For the second step, the introduction of query classification methodologies for Web image retrieval [16], content-based visual features [17] or the combination of textual and visual features in a multi-view paradigm [18] can be different issues to take into account to improve image timestamping.

Acknowledgments

We first offer our acknowledgments to Youssef Chahir, Antoine Doucet, Frédéric Jurie and Paul Martin for their productive comments and discussions. Second, this research was supported by MEXT Grant-in-Aid for Young Scientists B (#22700096) and by JST research promotion program Sakigake: “Analyzing Collective Memory and Developing Methods for Knowledge Extraction from Historical Documents”.

References

1. Alonso, O., Gertz, M., Baeza-Yates, R.: Clustering and exploring search results using timeline constructions. In: Proceedings of the 18th ACM Conference on Information and knowledge Management (CIKM 2009). (2009) 97–106
2. Metzler, D., Jones, R., Peng, F., Zhang, R.: Improving search relevance for implicitly temporal queries. In: Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2009). (2009) 700–701
3. Lux, M., Chatzichristofis, S.: Lire: Lucene image retrieval: An extensible java cbir library. In: Proceedings of the 16th ACM International Conference on Multimedia (ACMMM 2008). (2008) 1085–1088
4. Au Yeung, C., Jatowt, A.: Studying how the past is remembered: Towards computational history through large scale text mining. In: Proceedings of the 20th ACM International Conference on Information and Knowledge Management (CIKM 2011). (2011) 1231–1240

5. Dias, G., Campos, R., Jorge, A.: Future retrieval: What does the future talk about? In: Proceedings of Workshop on Enriching Information Retrieval of the 34th ACM Annual SIGIR Conference (SIGIR 2011). (2011)
6. Campos, R., Dias, G., Jorge, A., Nunes, C.: Gte: A distributional second-order co-occurrence approach to improve the identification of top relevant dates in web snippets. In: Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM 2012). (2012)
7. Ferragina, P., Scaiella, U.: Tagme: On-the-fly annotation of short text fragments (by wikipedia entities). In: Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM 2010). (2010) 1625–1628
8. Hoffart, J., Suchanek, F., Berberich, K., Lewis-Kelham, E., de Melo, G., Weikum, G.: Yago2: Exploring and querying world knowledge in time, space, context, and many languages. In: Proceedings of the 20th International Conference Companion on World Wide Web (WWW 2011). (2011) 229–232
9. Moreno, J., Dias, G.: Using ephemeral clustering and query logs to organize web image search results on mobile devices. In: Proceedings of the International ACM Workshop on Interactive Multimedia on Mobile and Portable Devices of the ACM Multimedia Conference (ACMMM). (2011) 33–38
10. Coe, B.: Guide to Early Photographic Processes. Victoria & Albert Museum (1983)
11. Scalise, A., Blose, A.: Dating images from scanned watermarks. Technical report, Patent US 2010/0329575 (2010)
12. Cieplinski, L.: Mpeg-7 color descriptors and their applications. In: Computer Analysis of Images and Patterns. Volume 2124. (2001) 11–20
13. Chatzichristofis, S., Boutalis, Y.: Fcth: Fuzzy color and texture histogram - a low level feature for accurate image retrieval. In: Proceedings of the 2008 9th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2008). (2008) 191–196
14. Chatzichristofis, S., Boutalis, Y.: Cedd: Color and edge directivity descriptor: a compact descriptor for image indexing and retrieval. In: Proceedings of the 6th International Conference on Computer Vision Systems (ICVS 2008). (2008) 312–322
15. Strötgen, J., Alonso, O., Gertz, M.: Identification of top relevant temporal expressions in documents. In: Proceedings of the 2nd Temporal Web Analytics Workshop (TempWeb 2012) associated to the 2012 World Wide Web Conference (WWW 2012). (2012) 33–40
16. Krapac, J., Moray, A., Verbeek, J., Jurie, F.: Improving web-image search results using query-relative classifiers. In: Proceedings of IEEE Conference on Computer Vision & Pattern Recognition (CVPR 2010). (2010) 1094–1101
17. Tollari, S., Detyniecki, M., Marsala, C., Fakeri-Tabrizi, A., Amini, M.R., Gallinari, P.: Exploiting visual concepts to improve text-based image retrieval. In: Proceedings of the 31th European Conference on Information Retrieval (ECIR 2009). (2009) 701–705
18. Usunier, N., Amini, M.R., Goutte, C.: Multiview semi-supervised learning for ranking multilingual documents. In: Proceedings of the 2011 European Conference on Machine learning and Knowledge Discovery in Databases (ECML-PKDD 2011). (2011) 443–458