PhotoFall: Discovering Weblog Stories Through Photographs

Christopher Wienberg
Institute for Creative Technologies
12015 Waterfront Drive
Los Angeles, CA 90094
+1 310 574 5700
cwienberg@ict.usc.edu

Andrew S. Gordon
Institute for Creative Technologies
12015 Waterfront Drive
Los Angeles, CA 90094
+1 310 574 5700
gordon@ict.usc.edu

ABSTRACT
An effective means of retrieving relevant photographs from the web is to search for terms that would likely appear in the surrounding text in multimedia documents. In this paper, we investigate the complementary search strategy, where relevant multimedia documents are retrieved using the photographs they contain. We concentrate our efforts on the retrieval of large numbers of personal stories posted to Internet weblogs that are relevant to a particular search topic. Photographs are often included in posts of this sort, typically taken by the author during the course of the narrated events of the story. We describe a new story search tool, PhotoFall, which allows users to quickly find stories related to their topic of interest by judging the relevance of the photographs extracted from top search results. We evaluate the accuracy of relevance judgments made using this interface, and discuss the implications of the results for improving topic-based searches of multimedia content.

Categories and Subject Descriptors
H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval – *Relevance feedback, Selection process*

General Terms
Human Factors

Keywords
Weblogs, Storytelling, Photographs

1. INTRODUCTION
The growth of the web, social media, and user-generated content has created new opportunities to match the esoteric retrieval needs of one user with the esoteric content created by another. With millions of people posting to the web, there exists multimedia content for nearly every topic of interest. In the categories proposed by Jansen et al. [6], searches for these topics are *Informational*, rather than *Navigational* or *Transactional*, in that the aim is to locate relevant content, which may exist as text, images, or other forms of multimedia. When searching for user-generated content, however, there is a greater emphasis on the relevance of the content to the topic than on its popularity or the

authority of its author. Accordingly, content-based multimedia information retrieval technologies have become essential tools in these searches.

The multimedia nature of web content has led many to explore search technologies that exploit the relationships between different media types within web documents. Early work by Dunlop and van Rijsbergen [3] used textual context to enable retrieval of non-text media, e.g. images. Advances in image processing have enabled new content-based image retrieval approaches, along with methods for integrating content-based image retrieval with textual queries [10]. Zha et al. [11] further showed that photographs could be presented to users during the formulation of textual queries to reduce the use of ambiguous terms.

Campbell et al. [2] investigated whether the relationship between text and images in multimedia documents could be exploited in the opposite direction: using images to support the retrieval of textual content. In their experiment, photographs embedded in weblog posts were shown to subjects as a proxy for the text of the documents for the purpose of collecting relevance feedback. Although the accuracy of these relevance judgments were less than when subjects read the full text of the document, the speed at which photographs could be annotated greatly reduced the amount of time needed to converge on precise topic models.

In this paper, we investigate whether the finding of Campbell et al. [2] can be exploited to quickly curate large collections of relevant user-generated web content. We describe *PhotoFall*, a new interface for topic-based searches of the stories that people post to their weblogs. We leverage the speed in which users can provide relevance feedback on photographs to improve the precision of the stories that they spend their time reading.

2. PHOTOGRAPHS IN WEBLOG STORIES
In this research, we focus our efforts on one class of user-generated web content: the personal stories that people post to their Internet weblogs. Every day, millions of new posts are added to weblogs worldwide. The typical weblog takes the form of a personal journal, used to share information, opinions, and experiences with a small number of friends and family [7]. Among these posts are those that narrate the personal experiences of the blogger, i.e. weblog stories. These stories have the potential to meet the retrieval needs of a wider audience, enabling searchers to learn from the experiences of others. Needed are methods for identifying personal stories among other types of weblog posts, and an effective means of retrieving stories related to specific topics from large corpora.

Gordon and Swanson [5] estimated that only 4.8% of all non-spam weblog posts are personal stories, which they defined as non-fictional narrative discourse that describes a specific series of causally related events in the past, spanning a period of time of minutes, hours, or days, where the storyteller or a close associate is among the participants. Using supervised machine learning techniques, they constructed a text classifier to identify these personal stories in streams of weblog posts. Applying their classifier to the 25 million English-language weblog posts in the ICWSM 2009 Spinn3r Dataset [1], they identified nearly one million personal stories. Subsequently, Gordon et al. [4] used the same method to identify personal stories in Spinn3r.com's daily weblog feed, producing a corpus of tens of millions of stories posted since January of 2010.

Campbell et al. [2] noted the prevalence of photographs in weblog stories, estimating that 19.5% of these posts contained at least one. Using a simple heuristic based on the aspect ratio of linked images, they were able to accurately distinguish photographs from all other image types (precision=1.0, recall=0.94). Sampling from the set of photographs in stories, they investigated the degree to which the narrative text of the post was related with the visual content of the photograph. By their estimate, 82% of the photographs that accompany the text of weblog stories were taken during the events of the narrative, typically by the bloggers themselves. It is this strong connection between text and photographs in weblog storytelling that we capitalize upon in the design of our PhotoFall interface.

3. PHOTOFALL

We developed the PhotoFall interface to be an efficient and visually appealing way to find topic-relevant stories in streams of weblog posts. The name is meant to evoke the concept of a waterfall, but where the materials flowing down the stream are photographs rather than water. This metaphor continues into the layout of items in the interface.

In PhotoFall, users are presented with a webpage with hundreds of photographs extracted from weblog stories that the system expects are related to their search topic (Figure 1). Photographs are organized in reverse chronological order by the date they were posted, with markers on the page to indicate the month and year of the photographs that follow. When a user clicks on a photograph, a new web browser window opens to display the full weblog post from which the photograph was extracted. In the simple use case, users browse the photographs to find ones that are particularly interesting or obviously relevant to their search interest, and click them to read the associated stories.

Generating a topic-specific PhotoFall page requires a number of steps. First, each PhotoFall page requires the specification of a textual search query for the topic of interest. These queries are submitted to a text retrieval engine that indexes the full text of a large corpus of weblog stories. In our prototype, we employed the Terrier IR Platform [8] to search 20 million weblog stories from the collection of Gordon et al. [4]. Second, we analyze the HTML of the top 1000 search results to identify posts that include photographs along with the narrative text. We download each image file linked in the text of the post, and use the aspect-ratio heuristic of Campbell et al. [2] to distinguish photographs from other image content. Where multiple photographs appear in a post, we select the middle-most photograph. Third, we generate a new HTML page containing all of the photographs extracted from search results, resized and bordered in uniform manner, and ordered such that the most recently posted photographs appear first.

Figure 1. The PhotoFall interface, displaying photographs extracted from weblog stories about the topic *Surfing*.

In the simple use case, where PhotoFall provides a visual retrieval interface to search results, it provides users with an appealing way to browse and access the content of hundreds of items from the top 1000 search results. However, the real utility of the PhotoFall interface comes when used in conjunction with textual search tools that incorporate relevance feedback from users. In this advanced use case, PhotoFall becomes an interface for quickly providing relevance feedback to improve an existing query, or to quickly curate large numbers of positive examples once a query has been refined.

To support this advanced use case, PhotoFall allows users to quickly annotate the relevance of individual photographs to the search topic. When a user hovers their cursor over a photograph, they are presented with the option of marking it as relevant (the "YES" button) or not relevant (the "X" button), as seen in Figure 2. Photographs that are annotated as relevant are displayed with a solid white border, as seen in Figure 1, while those marked as not relevant are darkened so as to be nearly invisible. Using these controls, hundreds of items on a PhotoFall page can be annotated in a matter of minutes.

Annotations made through PhotoFall can be used in conjunction with textual search tools in different ways. Where the goal is to improve the topic model of an existing query, the text of a weblog post containing an annotated photograph can be directly used for relevance feedback using traditional means [9]. Where the goal is to curate large collections of positive examples of stories all related to a specific topic, users can reduce the amount of time spent reading non-relevant search results by considering only

Figure 2. Detail of the PhotoFall interface, showing the annotation controls that are revealed to users when they move the cursor over a photograph.

those that have been annotated as relevant using the PhotoFall interface. A third approach is to combine elements of both of these strategies, where the posts with relevant photographs are read to verify their relevance to the topic, and then used to refine the existing query.

This third approach affords a new workflow to curate collections of relevant weblog stories over long periods of time (weeks, months, or years). By continuing to collect new weblog stories over time, the photographs displayed in each PhotoFall page would be continually updated with new items. Likewise, the query used to generate each PhotoFall page would be continually refined as additional high-quality relevance feedback is obtained from a careful reading of stories with annotated photographs. The broad appeal of both photographs and weblog stories suggests that the use of the PhotoFall interface in this workflow could be delegated to larger populations of web users, particularly to communities of interest that align with the specific search topics.

4. EVALUATION

In all use cases of PhotoFall, its utility is dependent on the users' ability to assess the relevance of a weblog story based solely on the content of a photograph included in the post. Accordingly, we evaluated the PhotoFall interface by determining the accuracy of annotations assigned to photographs across seven different search topics:

1. Airline travel: Stories of people flying on commercial airlines
2. Yoga class: Stories about practicing yoga
3. Protest rallies: Stories about political protest rallies and marches.
4. Voting: Stories about casting votes in elections
5. Lutherie: Stories about the fabrication and repair of stringed musical instruments
6. Drive-In Movies: Stories about watching movies in a drive-in theater
7. Surfing: Stories about catching waves on a surfboard, as seen in Figures 1 and 2.

For each of these seven topics, we created a search query for use in the PhotoFall interface by authoring an initial description of the topic, written as a short first-person fictional narrative. We then refined this query using traditional relevance feedback techniques [9], annotating a small number of the top search results returned using the Terrier IR Platform [8] to search 20 million weblog stories from the collection of Gordon et al. [4]. We then generated seven PhotoFall pages using these refined queries.

We, the authors of this paper, then used the PhotoFall interface to annotate the relevance of hundreds of photographs related to each topic, spending only a few minutes in each case. Our annotation rule was to only assign an annotation to a photograph if it was obviously relevant or not relevant to the given topic, i.e. ambiguous photographs were skipped. Following this rule, roughly half of all photographs across all seven topics were annotated. Subsequently, we reviewed each of the associated weblog posts to make a judgment as to their relevance based on the full text and additional media as it appeared on the web. In this step, stories were again annotated as relevant or not relevant, and skipped if the post was no longer available on the web or contained objectionable content. In all, 1511 photographs were extracted from the top 1000 search results across these seven topics (22% of posts). 774 photographs were annotated as relevant or not relevant (51%). Of these, we judged the relevance of 687 stories (89%) based on a reading of the full post as it appears on the web.

The results of this evaluation are presented in Table 1. Averaging

Table 1. Accuracy of PhotoFall annotations

Topic	Photos on PhotoFall page	Photographs Annotated	Stories Judged	Accuracy	Accuracy of Positive Annotations	Accuracy of Negative Annotations
Airline Travel	165	75	70	92.3%	66/66 (100%)	0/4 (0%)
Yoga Class	169	148	137	83.2%	52/54 (96.3%)	62/83 (74.7%)
Protest Rallies	144	90	53	90.6%	44/44 (100%)	4/9 (44.4%)
Voting	98	66	53	62.3%	32/37 (86.5%)	1/16 (6.3%)
Lutherie	412	292	281	95.7%	38/49 (77.6%)	231/232 (99.6%)
Drive-In Movies	262	25	22	100%	5/5 (100%)	17/17 (100%)
Surfing	261	78	71	93.0%	66/71 (93.0%)	0/0 (-)
Total	1511	774	687	-	303/326	315/361
Macro average	-	-	-	88.2%	93.3%	54.2%
Micro average	-	-	-	90.0%	92.9%	87.3%

across all topics (macro average), 88.2% of all annotations made using PhotoFall agreed with annotation made after reading the corresponding weblog post. However, we observed considerable variability across topics, as well as between the accuracy of relevant (positive) annotations and not relevant (negative) annotations. In most topics, large numbers of photographs were annotated as relevant, and these annotations were highly accurate (93% macro and micro average). However, the number and accuracy of negative annotations (photographs judged not relevant) varied dramatically across topics.

There are several factors that may account for variability in the number and accuracy of negative annotations. Topics that are well represented in weblog stories will have a higher precision in the top 1000 search results used to generate a PhotoFall page. This will lead to fewer negative annotations overall, as in the case with surfing, and a greater chance of false negatives, as in the case with airline travel. Stories for other topics, such as voting in elections, appear prone to the use of photographs that are not obviously related to the events of the narrative, exacerbating the problem of false negatives. Conversely, numerous and highly accurate negative annotations are possible in other topics. Photographs of lutherie, showing musical instruments under construction or repair, were easily distinguishable from other arts and crafts photographs, which appear in PhotoFall due to similar vocabulary. The most accurate use of PhotoFall was seen in the topic of drive-in movies, a rare topic with photographs that easily distinguish relevant posts from others with similar terms, such as going to drive-in restaurants and traditional indoor movie theaters.

5. CONCLUSIONS

We draw two primary conclusions from our evaluation of the PhotoFall interface. First, PhotoFall is an extremely effective interface for quickly identifying relevant weblog stories across a range of different topics. With an accuracy of 93% for positive annotations, PhotoFall serves a variety of use cases. In the simple use case, where users aim to navigate to relevant stories, their judgments of photograph relevance will be a good guide. Where the aim is to improve topic models through relevance feedback, this level of accuracy affords the use of PhotoFall annotations directly, without the need to review the full text of the weblog post. Likewise, where the aim is to curate large collections of relevant stories over time, annotations in PhotoFall can be effectively used to quickly discover new positive instances.

Second, the mixed levels of accuracy seen in annotations of photographs not relevant to the search topic suggest that a more nuanced use of these negative annotations is required. For some topics, these annotations can be used directly in all use cases. For other topics, the frequency of false negatives would be detrimental, e.g. when used as relevance feedback in modeling the search topic. We believe that many of these barriers could be overcome with a better understanding of the relationship between topics and photography. Some ability to predict the relevance of photographs given an arbitrary topic would be extremely useful, and is a promising direction for future research.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] Burton, K., Java, A., and Soboroff, I. (2009) The ICWSM 2009 Spinn3r Dataset. Proceedings of the Third Annual Conference on Weblogs and Social Media (ICWSM 2009), May, San Jose, CA.

[2] Campbell, A., Wienberg, C., and Gordon, A. (2012) Collecting relevance feedback on titles and photographs in weblog posts. International Conference on Intelligent User Interfaces (IUI 2012), Lisbon, Portugal, February 2012.

[3] Dunlop, M. and van Rijsbergen, C. (1993) Hypermedia and free text retrieval. Information Processing and Management 29(3):287-298.

[4] Gordon, A., Bejan, C., and Sagae, K. (2011) Commonsense Causal Reasoning Using Millions of Personal Stories. Twenty-Fifth Conference on Artificial Intelligence (AAAI-11), August 7–11, 2011, San Francisco, CA.

[5] Gordon, A. and Swanson, R. (2009) Identifying Personal Stories in Millions of Weblog Entries. Proceedings of the Third Annual Conference on Weblogs and Social Media (ICWSM 2009), Data Challenge Workshop, San Jose, CA.

[6] Jansen, B., Booth, D., and Spink, A. (2008). Determining the informational, navigational, and transactional intent of Web queries. Information Processing and Management 44:1251-1266.

[7] Munson, S. and Resnick, P. (2011) The Prevalence of Political Discourse in Non-Political Blogs. Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media, July 17-21, 2011, Barcelona, Spain.

[8] Ounis, I., Lioma, C., Macdonald, C., and Plachouras, V. (2007) Research Directions in Terrier: a Search Engine for Advanced Retrieval on the Web. Upgrade 7(1):49-56.

[9] Rocchio, J. (1971) Relevance Feedback in Information Retrieval. In G. Salton (ed.), The SMART Retrieval System: Experiments in Automatic Document Processing. Prentice Hall, 313-323.

[10] Villa, R., Halvey, M., Joho, H., Hanna, D., and Jose, J. (2010) Can an intermediary collection help users search image databases without annotations? 10th Annual Joint Conference on Digital Libraries (JCDL-10).

[11] Zha, Z., Yang, L., Mei, T., Wang, M., Wang, Z., Chua, T., and Hua, X. (2010) Visual Query Suggestion: Towards Capturing User Intent in Internet Image Search. ACM Transactions on Multimedia Computing, Communications and Applications, Vol. 6, No. 3, Article 13.