Exploring Photos in Facebook

Mark D. Wood  
Eastman Kodak Company  
Rochester, NY, USA  
mark.d.wood@kodak.com

Minwoo Park  
Eastman Kodak Company  
Rochester, NY, USA  
minwoo.park@kodak.com

Abstract—Facebook has rapidly become for many the dominant means for sharing images, and the number of shared images accessible to any given Facebook user is easily in the tens of thousands. The sheer volume of pictures relegates most to obscurity, yet some of those pictures would be of great interest—if a person could only find them. This research explores ways to harness latent semantic information associated with pictures and interpersonal relationships to enable a person to browse for potentially interesting and germane images shared by people in their social network. The possibilities for semantic analysis are endless; this work illustrates two possible approaches while also highlighting future potential applications of semantic understanding.

Keywords—Facebook, social networks, image sharing, photo sharing, image understanding, image browsing, tag-cloud

I. INTRODUCTION

Facebook has in just a few years become a major vehicle for sharing personal photographs with friends and family, with Facebook reporting over 300 million photos uploaded per day, on average, during the first quarter of 2012 [1]. The Facebook environment supplies a rich platform for developing semantically aware applications, as the platform provides a means for recording and analyzing comments and likes associated with imagery, as well as enabling users to tag each other in images. A typical Facebook user might have hundreds of friends, and access to tens of thousands of images that have been shared to the user by those friends. The enormous volume of content makes it unlikely that any given person will see a shared photograph, especially if they did not notice it when it was first shared.

This research explores mechanisms for leveraging the information associated with photos—currently, the comments, tags and likes—as well as knowledge of user interactions to retrieve and highlight potentially relevant photos. In this work, the set of images available to be seen by a user consists of the user’s personal collection plus all the accessible images shared by that user’s social network on Facebook. (Note that individual Facebook privacy settings may prevent applications such as ours from having access to images on behalf of a user.) This set of images is “crawled” in advance of the user interacting with the system in order to retrieve and analyze the comments, tags, and, optionally, likes and interpersonal interactions associated with each image.

We describe two different paradigms for navigating the collection of images accessible to a person as part of their Facebook social network. In the first approach, the textual data associated with images is used to form a tag cloud. The user may then navigate the images by selecting a term in the tag cloud; the system retrieves and displays images corresponding to that term. In the second approach, the user selects a particular image, and the system automatically ranks all the other images according to their similarity to the selected image. The user may then pick one of the top ranked images as the new selection point, causing the ranking to be repeated. Both approaches are implemented as Facebook applications using the Facebook public API [2].

Clearly neither of these approaches harnesses all potential data, nor do they bring to bear all the latest in semantic understanding technology. However, they serve to illustrate some of what is possible in this environment; this paper further outlines potential future applications of new and emerging semantic technologies to further facilitate image exploration and related applications.

II. TAG-CLOUD BASED RETRIEVAL

In this approach, the application automatically constructs a tag cloud from the comments and captions associated with images accessible as part of one’s Facebook social network. The tag cloud serves as a text-based visual overview of the image collection. By clicking on a given tag, the user is presented with thumbnail photos associated with the clicked tag; the user may then select a particular image to see an enlarged view of the image. An interesting dimension added by this approach is that the system displays images that share a common theme across different photo albums in the social network. As a result, for example, images from similar or reoccurring events might be displayed together.

To construct the tag cloud, the application first crawls one’s social network, downloading comments, captions, and people tags. Stopwords are removed and then the frequency occurrence of each word in the comments and captions is computed. The set of words appearing in the title of an album is propagated to the individual images in the album, since every image in an album may not have been individually labeled. In addition, if an image was tagged as containing people, each person’s name is counted as a word
Finally, the tag cloud is created, with the frequency of a word determining the size of the font for that word. We use the algorithm [3] to fit a variable number of tags, each with a font size proportional to its frequency, into a prefixed layout while respecting alphabetic ordering of the tags.

Fig. 1 shows an example screen shot. The topmost part (Fig. 1a) represents the tag cloud specialized for the social activities of the current user and user’s social network. The middle part (Fig. 1b) shows the selected tag. The bottom part (Fig. 1c) shows images related to the selected tag. The application allows the user to horizontally scroll through the retrieved images if necessary. In this example, the user has selected the “ILLUSION” tag, and as can be seen in Fig. 1c, the application selected the images related to “ILLUSION”. The application could be easily extended to allow one to select multiple tags in the tag cloud to browse images.

III. RETRIEVAL BY SIMILARITY

In this approach, the application automatically retrieves images similar to a given image. The user starts with a visual representation of their Facebook photo albums. By clicking on a given album, the user is then presented with a visual representation of the photos in that album; the user may then select a particular image to view an enlarged representation. An interesting dimension added by this approach is that the system ranks all rankable images in the social network according to their similarity to the currently selected image and displays the top-ranked images; the user may then select one of the similar images to view as the selected image, which in turn causes all the images to be reranked according to their similarity to the newly selected image. To provide the rankings, the application first crawls one’s social network, pulling not only comment and tag information but also likes and ownership information for each album and photo.

Fig. 3 provides an example screen shot. The leftmost strip of thumbnails (Fig. 1a) represents the current user’s album. The application could easily be extended to allow one to navigate to a particular image in a friend’s album. The next strip from the left (Fig. 1b) shows the images in the selected album. The rightmost strip (Fig. 1e) shows similar images to the currently selected image (Fig. 1c). In this example, the user has selected a series of images, with previously viewed images shown in the horizontal filmstrip (Fig. 1d).

A. Implemented Similarity Metrics

The current implementation leverages the following vectors of similarity:

- **Similarity between textual annotations and comments:** All text associated with an image, both comments and user-provided captions, is tokenized and stemmed using Porter stemming to build a document dictionary for the image. Common stop words are removed, including vernacular associated with Facebook such as “LOL”.
- **Similarity between the sets of people tagged in the images:** Facebook enables users to identify and tag the people in a given image, with persons identified using their Facebook identifier or by a text string.
- **Similarity between likes:** The rationale behind even including likes as a similarity metric is this: if a given user likes a particular image liked by a set of users, then it may be likely that the user will like other images tagged by that same set of users.
- **Similarity based upon interactions:** As part of its data analysis, the system keeps track for each users P and...
Given a selected image, similarity for annotations, tags and likes is computed using a vector of these four vectors. Sim-
ilarly for comments, tags or likes—in the user’s social network with the currently selected image using these four vectors, along with the number of returned results, through the interface illustrated in Fig. 4.

The demonstration ranks each rankable photo—those with comments, tags or likes—in the user’s social network with the currently selected image using these four vectors. Similarity for annotations, tags and likes is computed using a simple form of Tversky’s parameterized similarity ratio [4]. Given a selected image $P_A$ and the image being ranked $P_B$, let $C$ be the set of terms held in common, and $A$ and $B$ be the set of terms only present in the selected image, and $B$ be the set of terms only present in the image being ranked. The similarity ratio is simply:

$$\text{Sim}(P_A, P_B) = \frac{|C|}{|A| + |B| + |C|}$$ (1)

While this similarity metric works well for tags and likes, a couple of issues arise with comments. First, two images may share a common concept, but one image may have far more comments associated with it than the other. To mitigate the effect of differing comment stream lengths, log functions are applied to the $A$ and $B$ terms in the denominator. The second issue is that some words commonly appear more often than others, and therefore may have less semantic weight. Although the most egregious words are filtered out as stopwords, the problem still remains. To address this problem, the $C$ term is computed by considering the inverse document frequency of each word found in common between two images. Combining these two modifications, the following formulas are used to determine similarity between two photos’ comments:

$$C' = \sum_{t \in C} \log \frac{N_{\text{freq}_t}}{|A| + |B| + |C|}$$ (2)

$$\text{Sim}_C(P_A, P_B) = \frac{C'}{C' + \log(|A| + 1) + \log(|B| + 1)}$$ (3)

where $\text{freq}_t$ is the document frequency for $i$.

The demonstration allows a user to assign different weightings to the four currently supported similarity metrics. While not implemented, the computed similarity could easily be a function of both the current selected image as well as previously viewed images, as in the adaptive browser in [5].

IV. DISCUSSION

This research describes some initial steps towards harnessing semantic understanding to navigating the vast images present in one’s social network. Both paradigms uncover for users fun and surprising photos that the user may have never seen, and likely would have otherwise never noticed or would have found hard to find. These paradigms represent two very different approaches to retrieval. The approach based on tag clouds is a directed search. The system automatically infers the possible search vectors and displays them to the user in the form of a tag cloud; the user selects specific terms to be used as search terms, and the system retrieves and displays the images that match those terms. However, since the matching is based on comments, only an indirect connection may exist between a particular image and a given tag.

The similarity approach is more passive and less explicit; the system automatically shows the users more images similar to the selected image, with the user free to ignore those images, or to use one of them as a launching pad for exploring more images. The two approaches of course could be combined, providing both results that match a user-selected concept as well as showing results similar in one or more dimensions to the immediate matches.

Opportunities for increased application of semantic understanding are manifold. The current algorithms do not carry out any pixel-based image understanding algorithms. Such algorithms, while typically computationally expensive, may be able to provide complementary semantic information. In particular, scene and material classifiers could be readily used to classify images according to scene type (e.g., beach) or material (e.g., blue sky), providing a new similarity metric and/or additional concepts for the tag cloud. Likewise, images could be classified into temporal events; knowledge of events and the photographer could be factored into the similarity computation in a user-defined manner. In some cases, the user may want images to be similar in some vectors but dissimilar in other vectors, e.g., more photos similar to my photo, but not captured by me.

The current work also associates some semantics with certain calculations where further research is needed to verify the validity of those assumptions. For example, in assessing the strength of user interactions, does a comment really have more weight than just a “like”? Do longer comments, on average, indicate a greater importance than shorter comments? Facebook users may self-report social relationships using means provided by Facebook, but we expect that actual interactions are a better measure of the strength of a relationship. Further research into how people interact in social networks is needed to understand how to interpret the raw interaction data available in Facebook.

More sophisticated natural language processing techniques can be applied to further measure the similarity be-
between comments. Note that comments are often incomplete sentence fragments, littered with misspellings and obscure references. Traditional techniques for part-of-speech tagging and concept expansion may need to be modified to be applicable. Simple synonym matching and reduction to canonical form would also be useful for tag cloud generation.

V. RELATED WORK

As the goal of this work is primarily to provide users with a fun and engaging way to browse and navigate content, a quantitative comparison with other work is difficult. However, we highlight some other relevant work in the browse and navigate space. First, we note that Facebook’s own timeline view provides Facebook users with a fun way to view a summary of their content. The timeline view only displays images from a user’s own collection, or where the user was specifically tagged. The timeline view is not intended to surface potentially interesting content across one’s broader social network, as done by this work. Various researchers have considered the general problems of search and navigate, including looking at both passive filtering mechanisms as well as more active retrieval mechanisms. The work of [6] looks at tag clouds as a retrieval technique, albeit not in the context of retrieving images from social media. They address the problem of improving the semantic value of tags; their work may be applicable to our tag-cloud based approach, especially since our tag cloud is formed from unstructured text, and the resulting terms may be very noisy. In [7], the authors present tag clusters as an alternative to tag clouds, where a tag cluster is visualized as a connected graph of concepts. The authors argue that tag clusters are preferable to tag clouds in terms of both utility and enjoyment as a general information retrieval mechanism. Although they did not specifically look at using tag clouds as a vehicle for navigating social media, one possible extension to our work would be to explore the use of a tag cluster for that purpose. In [5], the authors present a system for image querying, where the system refines its model of what is being searched for by implicit relevance feedback. The user initiates a search providing keywords and the system returns an initial set of results. The user then selects one of the results, and the system returns images similar to the selected results, using similarity-determining techniques comparable to the similarity-based navigation approach presented in this work. The system in [5] refines its search model by considering past selected images; this type of functionality could be added to the current system if desired. However, our similarity-based retrieval system is intended for wandering through social media rather than finding a specific image.

In [8], a system is presented for ranking images returned by a keyword-based query in terms of both visual content as well as the semantic information in associated textual annotations, where the system seeks to optimize a metric defined as the “average diverse precision.” Their approach results in greater diversity in the search results, which is especially important if the collection contains a large number of visually similar results. Our work does not specifically seek to introduce diversity in the results. In some cases, visually similar events from the same event would rank together in the similarity-based retrieval; in such cases, picking just a representative image might be preferable and this feature could be added in the future.

VI. CONCLUSIONS

The glut of imagery available on social networks such as Facebook can only be effectively tamed using semantic understanding techniques. Even relatively elementary techniques such as those employed by the algorithms described here expose interesting and germane content that would otherwise be hidden or forgotten. This work looks at two different approaches for exposing content to the user. The tag cloud approach provides a text-based summary giving the user different hooks into the broader collection of images; it enables directed search, albeit in a fun and somewhat abstract way. The similarity approach is intended to facilitate serendipitous browsing through a collection. Either approach shows promise as a means for helping users explore shared social media. User studies would be required to explore under what circumstances users prefer which approach. Combining the two approaches would enable both directed search and serendipitous browsing of social media.

ACKNOWLEDGMENTS

This work builds on earlier work by Graham Kiddle. Alex Loui provided very helpful feedback and comments on this work.

REFERENCES