Finding Selfies of Users in Microblogged Photos

Dhiraj Joshi
FX Palo Alto Laboratory
3174 Porter Drive
Palo Alto, CA 94304
dhiraj@fxpal.com

Francine Chen
FX Palo Alto Laboratory
3174 Porter Drive
Palo Alto, CA 94304
chen@fxpal.com

Lynn Wilcox
FX Palo Alto Laboratory
3174 Porter Drive
Palo Alto, CA 94304
wilcox@fxpal.com

ABSTRACT

We examine the use of clustering to identify selfies in a social media user's photos. Faces are first detected within a user's photos followed by clustering using visual similarity. We define a cluster scoring scheme that uses a combination of within-cluster visual similarity and average face size in a cluster to rank potential selfie-clusters. Finally, we evaluate this ranking approach over a collection of Twitter users and discuss methods that can be used for improving performance in the future. An application of user selfies is estimating demographic information such as age, gender, and race in a more robust fashion.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Clustering, Retrieval

Keywords

Selfies, Photos, Social Media Analysis, Instagram, Twitter

1. INTRODUCTION

Microblog platforms such as Twitter have become the voice of millions of users on the Web today. Microblogs are somewhat different from traditional social networks such as Facebook in terms of shorter and more frequent posts by users and a more open ecosystem. While Twitter started largely as a text-based microblog service, it now supports images and video tweets. At the same time, many dedicated photo-centric social networks have come up lately (e.g. Instagram, Snapchat, Tumblr, and Path). Posting photos has become much easier with smartphones and is sometimes advantageous over typing. Moreover, a photo often allows a person to share their creativity, wonder, or emotion more simply than text. Not surprisingly, photo-centric social networks are drawing more users and are expected to see huge success in the near future [3]. One of the popular trends in social multimedia is the phenomenon of self-portraits or selfies [4]. A selfie is a picture taken of oneself while holding the camera at a close range. Besides the person taking photo, a selfie may also include other people. With the introduction of front facing cameras in smartphones, taking selfies has become especially easy and trendy. By identifying selfies, the demographics of a person, such as age, gender, and race, can potentially be estimated more robustly. Additionally, interests and social contexts of the person, such as whether they take many photos with a small set of the same people, presumably friends, can be inferred. Although selfies are very popular, we do not know of any automatic methods for identifying them. Currently, analysis of selfies has required either manual labeling of photos [1] or using photos tagged "selfie" [2]. Manual labeling is time-consuming, and relying on tagged photos ignores many selfies that are not tagged with a relevant keyword. In this paper we investigate the use of clustering for finding selfies in a user's collection of Twitter images. Our hypothesis rests on two observations (i) most people take groups of selfies that can be potentially discovered using visual clustering of faces, and (ii) faces in selfies are usually large in size compared to faces in non-selfie photos owing to the camera being at a closer range. Although methods for ranking document clusters have been studied, e.g., [6], we propose to rank face clusters based on features that capture these two observations.

SoMeRA'14, Jul 06 - July 11 2014, Gold Coast , Australia
ACM 978-1-4503-3022-0/14/07.
http://dx.doi.org/10.1145/2632188.2632209

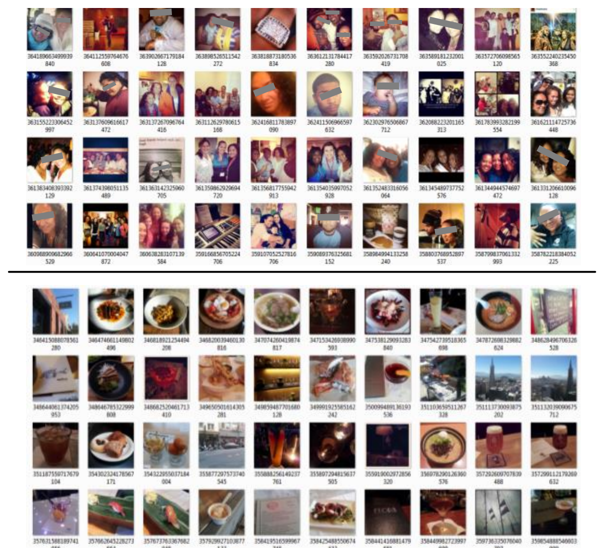

Figure 1. Photos of users with (top) most detected faces, (bottom) no detected faces.

2. DATA COLLECTION AND PROCESSING

Our data set contains the photos from the 190 Twitter users who posted the most Instagram photos in the San Francisco Bay area and surroundings from June to September, 2013. All the photos used in this experiment were tweeted from Instagram (i.e., posted on Instagram and tweeted as well). Face detection was performed using the OpenCV face detector. We used state-of-the-art visual

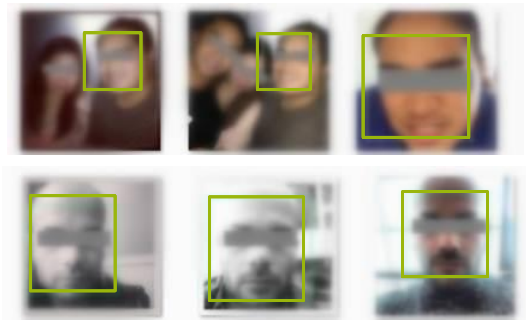

Figure 2. Example user-selfie clusters for two different users. Notice the visual variation among the selfies. Images have been blurred for anonymity.

features based on locality-constrained linear coding (LLC) as adapted in [5] to represent faces. Under the LLC framework, each face is represented by a 21504 dimension vector composed of the 1024 dimensional code for each of the 21 spatial pyramid grids. This was followed by computation of a similarity matrix consisting of similarity values between faces using the spatial pyramid matching framework. The computed similarity matrix and Affinity Propagation were then used to perform visual clustering of faces for each user.

To identify selfies based on the observations described earlier, face clusters were ranked using a combination of average visual similarity among faces and average size of faces in cluster.

$$\alpha \frac{\sum_{i\in Cluster}\ \sum_{j\in Cluster,\, j<i} Sim(i,j)}{NumPairs} + \beta \frac{\sum_{i\in Cluster} Size(i)}{NumFaces}$$

For normalization α and β are the inverses of standard deviations of the corresponding measures (similarity or size) computed across all clusters. Clusters with less than 3 photos were eliminated. We also eliminated clusters containing exact duplicates of a face (wherein standard deviation of within cluster similarity matrix is zero). This was done to eliminate certain random faces (such as celebrities) that people post and often repost. A total of 781 face clusters were obtained.

3. RESULTS AND DISCUSSION

Users in our collection varied widely in terms of number of detected faces. In Fig. 1, we show photos of users from different ends of the spectrum (with the most detected faces and with no detected faces). Absence or presence (and number) of detected faces in a user's collection can both be important determinants of users' personalities. For example, in Fig. 1 (top) we can see that the user with the most detected faces seems to be a group/party lover whereas in Fig. 1 (bottom), we notice that a user with no detected faces (in this case) happens to take many photos of food.

In order to evaluate our cluster ranking method, we manually labelled the faces in 80 top-scoring clusters (about 10% of all the clusters). These clusters belonged to a total of 27 different users. Fig. 2 shows two examples of highly ranked selfie clusters. Note that the top user exhibits more visual diversity in his selfies than the bottom user. In labeling whether each facial image in a user's photos is that of the user, we hypothesize that the most dominant person in a user's selfies is the user. We also define our goal to be identification of enough relatively large images of a user, rather than identification of all images of a user. Based on these criteria, labeling is performed by tagging as positive examples all images

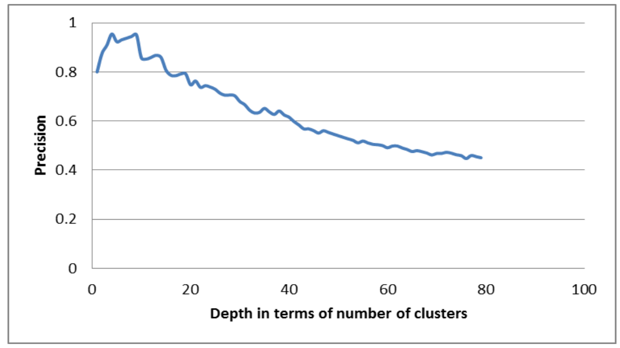

Figure 3. Precision values at different ranks (depth).

of the user's face in each cluster that contains at least one selfie. All other images, including those in clusters without a selfie, are tagged as negative examples.

We ordered the clusters by score and then in Fig. 3 we plot precision (for the selfie-finding task) at different cluster ranks. At each rank the precision is computed using the following classification: all clusters with a score greater than or equal to the current cluster are classified as selfies. The generally decreasing trend of the plot indicates that our scoring measure provides a meaningful ranking. We also observed that some clusters of kid photos taken at close range got reasonably high scores. These are clearly not selfies of the user and are counted as false positives. However, we believe that in the future, age recognition algorithms may be used to eliminate such cases and improve performance.

4. SUMMARY AND FUTURE WORK

Our results indicate that ranking clusters of facial images based on average similarity and average facial image size is a promising direction for identifying images of a social media user's face. In the future we would like to explore extracting demographic information, such as age and gender, using the identified selfie clusters. Age, gender or other user demographic attributes could be computed for all faces in a cluster and aggregated across them. We hypothesize that by using multiple larger images of a user's face, our demographic estimates will be more robust. We would also like to use other criteria such as presence of arms and size of torso in photos to more robustly identify user selfies.

5. REFERENCES

[1] Selfiecity. http://selfiecity.net.

[2] The Selfiest Cities in the World: TIME's Definitive Ranking" http://time.com/selfies-cities-world-rankings/.

[3] http://www.forbes.com/sites/jaysondemers/2013/09/24/the-top-7-social-media-marketing-trends-that-will-dominate-2014/

[4] K. Losse, The Return of the Selfie. http://www.newyorker.com/online/blogs/elements/2013/06/the-return-of-the-selfie.html

[5] H. Chen, A. Gallagher, and B. Girod, "What's in a Name: First Names as Facial Attributes", *In Proc. CVPR*, 2013.

[6] F. Raiber, and O. Kurland, "Ranking document clusters using Markov random fields", *In Proc. of the 36th Int. ACM SIGIR*, 2013.