ShoeBox: A Natural Way of Organizing Pictures According to User's Affinities

Bojan Blažica[1], Daniel Vladušič[1], and Dunja Mladenić[2]

[1] Xlab d.o.o., Pot za Brdom 100,
1000 Ljubljana, Slovenia
[2] Artificial Intelligence laboratory, Jožef Stefan Institute, Jamova 39,
1000 Ljubljana, Slovenia
{bojan.blazica,daniel.vladusic}@xlab.si
dunja.mladenic@ijs.si

Abstract. In this paper we tackle the problem of personalizing the experience of browsing through digital pictures. We address two questions: how to capture the user's personal affinity for a particular picture and how to visualize a large collection of pictures. We propose a novel approach towards organizing pictures called ShoeBox that aims for automatic capturing of the user's affinity for individual pictures.

Keywords: personal information management, digital photo collection, visualization, ranking, user-centered design.

1 Introduction

Back in the times of analogue photography some of us used to store photos in shoe boxes. When viewing them, we would gather around a table and turn the shoe box upside down, letting the photos scatter on the table. Nowadays, we know from personal experience, that a few weeks trip can result in a photo collection as large as a few thousand photos. Although each picture represents a certain memory like a place we visited or a person we met, not all of them bear the same importance to us.

Since the advent of digital photography many approaches have been taken in organizing large photo collections. Some of them use tagging [1], [2] or collaborative tagging [3], annotating using ontologies [4], automatic feature extraction [5], clustering according to the pictures' metadata like timestamp [6], [7], [8] and various combination of these.

Besides organization, visualization is also an important factor in personal information management. In contrast to the common visualization of pictures in a grid sorted by date or filename, Photofinder [9] sorts pictures according to a similarity measure and various demo multitouch application visualize pictures scattered on the desktop.

In our opinion the existing systems lack the ability to automatically capture the user's emotional aspect and taking it into account when sorting and visualizing pictures. Figure 1 presents a schematic representation of our solution to this problem. During the first time the collection of pictures is viewed the system builds a model of the users'

J.A. Jacko (Ed.): Human-Computer Interaction, Part III, HCII 2011, LNCS 6763, pp. 519–524, 2011.

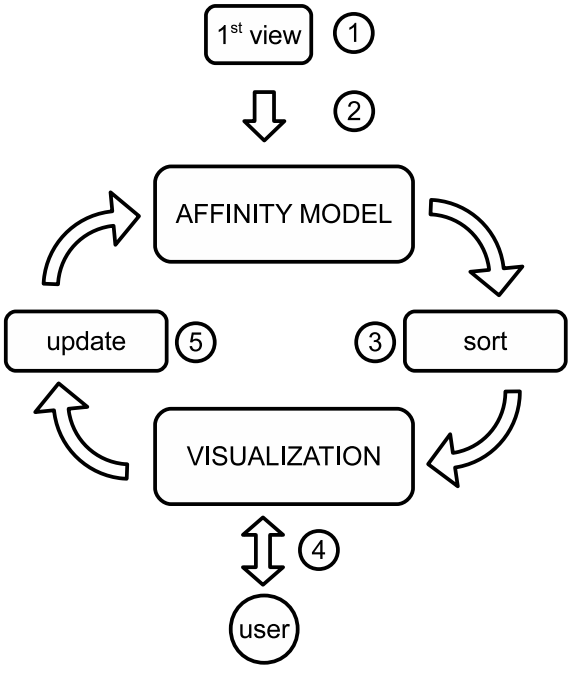

Fig. 1. A schematic overview of ShoeBox; the first time a collection is viewed (1) pictures are sorted according to the pictures' time-stamp. While the user is viewing the collection, the system builds the affinity model based on the time the user spends viewing each picture (2). At each next visualization of the collection, the affinity model is used to sort (3) pictures and visualize them accordingly. During interaction (4) the user's affinities are captured again and used to update the affinity model (5).

affinities. Each time the user views the collection, this model is used to organize it and at the same time the affinity model gets updated based on the user's interaction with the picture. In this paper we focus on using viewing time as a measure for affinity in the affinity model and on a natural way of visualization of large collections of photos.

2 Affinity Model

In defining a measure of the user's affinity for a picture we took a similar approach to the one used in web analytics: time spent viewing a picture. The first time a user views a collection, the pictures are sorted according to the time they were taken on. The user then assigns focus to a picture he wants to view by selecting it. We interpret this as a declaration of affinity for that particular picture and the time spent viewing it (viewing time) is considered a measure of strength of the affinity for that picture. When the picture loses focus, the time elapsed is stored in exchangeable image file format (EXIF). After the first view of the collection the system is already able to sort pictures according to viewing times. The next time the picture is viewed its viewing time will be added. The maximum amount of viewing time that can be added per

viewing is set to 30 seconds in case the picture remained in focus while the user went away from the computer.

$$PAS = w_{vt}f(vt) + w_i f(i). \quad (1)$$

The introduction of the Picture Affinity Score PAS, described by equation 1, allows for a broader definition of affinity. The affinity score takes into account viewing time and also interaction with a picture (e.g. zooming, rotating) and is calculated as a weighted sum of functions of viewing time (vt) and interaction (i). Besides the affinity score, the algorithm also stores a unique collection ID and the time of the first view into EXIF format. This data is needed when visualizing the collection and updating the affinity model.

Algorithm 1 shows the implementation of decay. This feature foresees the use of ShoeBox for viewing multiple collections simultaneously and is intended to prevent the algorithm to be biased towards pictures from older collections.

To prevent relevant pictures from being completely forgotten, after sorting, an amount of pictures with short viewing times is placed among those with high viewing times. This is done by randomly picking 10% of the 20% less relevant pictures and randomly placing them among the 20% most relevant pictures. We argue that this not only helps the system not to misplace relevant pictures but also aids the user's recall of the memories represented by all the pictures, even the ones he doesn't like.

Algorithm 1. Visualization in ShoeBox with the application of decay based on current time (*CT*) and decay speed (*DS*)

```
if picture doesn't have EXIF tag 1st_view_time then
  generate collection_ID
  sort according time-stamp
  for picture in collection do
    picture.1st_view_time • CT
    picture.collection_ID • collection_ID
  end for
else
  for picture in collection do
  apply decay:
  AS=picture.AS exp(-DS (CT - picture.1st_view_time))
 picture.AS • AS
 end for
 sort according AS
 apply recall aid
end if
calculate center points
apply random rotation
```

3 Visualization

As mentioned before, we consider scattering pictures on a table to be a natural way of browsing through a large collection of photos. We propose something similar for viewing digital pictures; instead of viewing pictures in a grid, we scatter them on the desktop and allow them to partially occlude one another. When the pictures are ordered in a list, they are positioned in a spiral and stacked in layers so that the first

picture in the list is in the center of the screen and in the top layer and the last somewhere at the border in the bottom layer. A picture in a higher layer can partially occlude a picture in a lower layer. All pictures are randomly oriented, as shown on figure 2. To be able to identify which picture the user is looking at, we blur all of them, except one. That picture has focus. As future work, this could be replaced with automatic gaze detection.

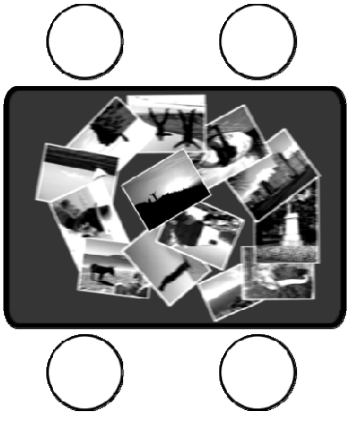

Fig. 2. An example of the visualization of a collection of pictures on a multitouch tabletop display with four users around it. The spiral organization and the random orientation of the pictures puts all the users in an equal position for browsing through the collection.

Although not strictly required, a multitouch tabletop display is used to enhance resemblance to viewing of analogue photos. When more users are sitting around a table, the top left corner of that table for a user is the bottom right corner for another. Placing the pictures in a grid puts users in an unequal situation. This is the reason we place pictures in a spiral with a random orientation.

4 Preliminary Results and Future Work

The above described model needs experiments that will help us evaluate our ideas. First of all we must prove that there is a link between viewing time and interaction with the user's affinity for a certain picture. Additionally, by experimenting with users we will also be able to refine the model and its variables like the weights used in calculating PAS and decay speed.

Our first experiments focused mainly on two things: on validating the main idea behind ShoeBox (viewing time as a measure for the users' affinity for a picture) and on gaining feedback from users and these preliminary experiments themselves on how to design further experiments that will help us to properly quantify the variables of the model. During the experiments we also observed how users interacted with the pictures in order to get some clues on how interaction could be interpreted as a measure for affinity.

Six users were involved in the experiment, three females and three males aged between twenty-five and thirty. After getting accustomed to multitouch interaction they started the experiment in which they were presented 50 photos obtained from Flickr with the keywords “sport car", "flower", "mountain", "winter", "party", "portrait", "war", "landscape", "dog", "see"; 5 pictures for each keyword.

To understand if viewing time is a meaningful measure for affinity for a picture we allowed the users to view the pictures for two minutes. After that, the most and least viewed ten were selected and arranged in pairs; the most viewed picture with the least viewed and so on. The user was then asked which picture he or she likes the most for each pair. We expected pictures with higher viewing times to be chosen.

The users offered a lot of insight on how to plan next experiments, such as to allow them more time for browsing or give them more space to move the pictures around. A user suggested that pictures should implement inertia so that he could swipe away the ones he doesn’t like. Another thing we noticed was that some users spontaneously started stacking pictures in two piles; one with pictures they like and one with pictures they don’t like. This reinforced our assumption that it is possible to extract the user’s affinity for a certain picture not only from viewing time, but also from how he or she interacts with the picture. Despite first getting accustomed with multitouch interaction, some users still experienced problems while browsing pictures.

Besides providing us with guidelines for future work, the experiment also showed that viewing time is indeed related with the users’ affinity for a picture as five out of six users were predominantly choosing the picture with the higher viewing time. Table 1 shows the results for all six users. Overall, in 60 pairs the picture with the higher viewing time was chosen 37 times (61%). Taking into account that the only user that choose more pictures with lower viewing times is the one that experienced the most problems with multitouch interaction, this preliminary results confirm our assumption about viewing time being a measure for the user’s affinity for a picture.

Future work will focus on defining the experiments and conducting them on a larger scale so that the experiments will have real statistical value. Furthermore in the next experiments we will ask the users to bring their own photos to see whether there is a difference in the relation between viewing time and the affinity for personal or general photos. After that we will be able to fully define the affinity model and the focus of our research will shift from the definition of the model to the exploitation of the model in personal information management and in image retrieval tasks.

Table 2. Number of pictures with higher and lower viewing times chosen for each user

User	No. of pictures chosen (higher viewing time)	No. of pictures chosen (lower viewing time)
1	7	3
2	3	7
3	8	2
4	6	4
5	7	3
6	6	4

5 Conclusions

In contrast to currently available software that organizes photo collections according to folder structures, metadata or semantic technologies we provide a method that takes into account the users' emotional aspect by organizing pictures according to his or her affinities, without requesting the user to explicitly express his or her affinities. This is done by measuring the time spent viewing a picture. Also, the organization obtained this way is dynamic, meaning that it changes through time resembling changes in the user's affinities. Furthermore we present an alternative, natural way of visualizing photo collections. As it is desirable of photo managing software to be able to sort photos according to preference, date etc. we see our work not replacing, but complementing current organizing methods.

Acknowledgments. This research was funded in part by the European Union, European Social Fund, Operational Programme for Human Resources, Development for the Period 2007-2013 and by the Slovenian Research Agency.

We would also like to thank our friends who volunteered for experimenting with ShoeBox; for their time spent testing and for valuable feedback on the implementation of ShoeBox itself.

References

1. Picasa, `http://picasa.google.com`
2. Flickr, `http://www.flickr.com`
3. Facebook, `http://www.facebook.com`
4. Latif, K., Mustofa, K., Tjoa, A.: An Approach for a Personal Information Management System for Photos of a Lifetime by Exploiting Semantics. In: Bressan, S., Küng, J., Wagner, R. (eds.) DEXA 2006. LNCS, vol. 4080, pp. 467–477. Springer, Heidelberg (2006)
5. Liu, D., Chen, T.: Unsupervised image categorization and object localization using topic models and correspondences between images. In: International Conference on Computer Vision (ICCV), pp. 1–7. IEEE Computer Society, Los Alamitos (2007)
6. Platt, J.C., Czerwinski, M., Field, B.A.: PhotoTOC: automatic clustering for browsing personal photographs. In: Proceedings of the Fourth IEEE Pacific Rim Conference on Multimedia, pp. 6–10. IEEE, Los Alamitos (2003)
7. Loui, A.C., Savakis, A.E.: Automatic image event segmentation and quality screening for albuming applications. In: ICME 2000, vol. 2, pp. 1125–1128. IEEE Press, Los Alamitos (2000)
8. Graham, A., Molina, H.G., Paepcke, A., Winograd, T.: Time as essence for photo browsing through personal digital libraries. In: Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries, pp. 326–335. ACM, New York (2002)
9. Kang, H., Shneiderman, B.: Visualization methods for personal photo collections: Browsing and searching in the photofinder. In: IEEE International Conference on Multimedia and Expo (III), pp. 1539–1542. IEEE, Los Alamitos (2000)