FreeEye - Intuitive Summarisation of Photo Collections

Kan Ren
CCSR, University of Surrey
Guildford GU2 7XH, UK
+44(0)1483684704
k.ren@surrey.ac.uk

Risto Sarvas
Helsinki Institute for
Information Technology HIIT
PO Box 9800, 02015 TKK, Finland
risto.sarvas@hiit.fi

Janko Calic
CCSR, University of Surrey
Guildford GU2 7XH, UK
+44(0)1483684739
j.calic@surrey.ac.uk

ABSTRACT

This paper presents user evaluation of the FreeEye tool for intuitive browsing and summarization of large-scale photo collections. The tool was tested in three different personal photo selection scenarios: a short-time event, a vacation and a yearbook. The experiments were conducted with five participants, evaluating their satisfaction with the summarization result and the overall process. The results demonstrate good usability of the FreeEye tool and improvement when compared to the standard methods of the participants for selection from large personal photo collections.

Categories and Subject Descriptors

H.3.7 [**INFORMATION STORAGE AND RETRIEVAL**]: Digital Libraries – *User issues*; H.5.2 [**INFORMATION INTERFACES AND PRESENTATION**]: User Interfaces – *Graphical user interfaces, Screen design, User-centered design*;

General Terms: Algorithms, Design, Experimentation, Human Factors, Performance

Keywords: Multimedia systems, image browsing, HCI, image clustering

1. INTRODUCTION

The work presented in this paper is a user-centric evaluation of the FreeEye photo-browsing interface presented in [1] for CeWe Challenge, and the demonstration can be found at [2]. The FreeEye tool makes a shift towards user-centered design of interactive photo browsing interfaces by augmenting user's interaction rather than extracting the related semantics. This shift enables efficient and intuitive navigation though large personal photo collections [3], thus facilitating familiarisation with the content and effortless selection of a thematic subset.

2. BROWSING INTERFACE

In order to interactively browse large photo collections, the browsing interface follows the idea of ranked image representation, where more relevant images should be more apparent and thus displayed bigger. User selects an image from the dataset, which is positioned centrally, and the remaining data is hierarchically represented on the screen. By doing this, the user practically moves the center of perspective from which the collection is explored.

The image browsing system comprises two main modules: image clustering engine and the interface generation, as depicted in Figure 1. The main driver for the image importance is defined by the level of similarity between the user-selected central image and other images from the database by using perceptual features. In this paper a three-dimensional HSV colour histogram has been chosen as the similarity metric, but the choice of the metric is completely independent of the presented interface. To achieve system scalability and algorithm complexity nearly linear to the number of photos, a specific graph based clustering algorithm is utilized [3]. The interactive interface is generated as follows. The centre image is displayed at 100% of its size. If the user clicks on an image, the image will move the centre of the next screen, and the display will adapt itself to represent perceptual neighborhood of the central image. The second level of image sizes contains 12 images from the same cluster that are most similar to the central image. These images are displayed at 50% their original size. The third level contains 36 images displayed at 25% size, separated into two parts: four edges and four corners. The 32 images located at the four edges are the centers of clusters closest to the central image. To support knowledge discovery and help users locating other areas of interest, four random pictures are located at four corners of the screen. Every time the user clicks, the system re-arranges all images as described above.

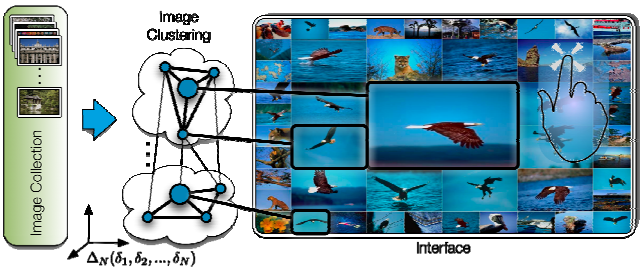

Figure 1. Core modules of the FreeEye browsing interface

3. USER EVALUATION

To evaluate the designed photo selection tool we conducted five user trials. The recruited participants were 3 women and 2 men aged 24-32, and all but one had a computer science background. For each trial the participant brought a set of their own digital photos. The number of photos brought by each participant ranged from 1385 to 1664. For each participant there were three separate tasks. The first task was to select photographs from a short-time event (1-2 days) to be sent by email to someone. The second task was to select photographs from a long time event (more than two days) to be uploaded to a web page or shown to someone. The third task was to select photographs for a book representing events and happenings in the past 6-12 months. For each task the participants were asked to think about specific people they would show the photographs to. The selected photographs were not actually sent or shown to anyone outside the trials. After each task

MM'09, Octber 19–24, 2009, Beijing, China.
ACM 978-1-60558-608-3/09/10.

the participants were asked a set of questions about the tool, the event, and photographs. The participants were also asked to give a score from 1-5 on how well the tools represented the events, how well the tool helped them to find photographs, and how the tool compared to their regular ways of selecting photographs. The answers to these questions are summarized in Table 2. For each task the number of clicks and the time spent was measured ,as well as the number of photos selected (see Table 1).

Table 1. Quantitative resulats of the user study

	selected photos	time spent	clicks	sec/click	sec/photo
Task 1	10.4	1:52	16.8	6.65	10.7
Task 2	15.6	5:36	49.2	6.82	21.5
Task 3	23.4	6:16	56.6	6.65	16.1

3.1 Evaluation Results

The short events the participants searched photos for were a birthday party, roller skating, and holiday trips. For the long events the participants all had a trip: hiking, traveling, and a long roller skating trip. For the yearbook task whole set of images was used and no temporal or event restrictions were given. The participants selected about 10-20 in each task to be sent to friends, family, or people who were in the photographs. In the case of the yearbook, the particpants made the book mainly for themselves and planned to show it to friends and family.

The participants were satisfied on how well the photos they selected represented the event. In the long event task (task 2) they reported that they felt that they missed some photographs they would have liked to have. In the short event they felt that no photographs were missing, and in the yearbook task one participant reported that he got "almost" all of them, and another participant felt that she missed 5-6 photographs. As seen in Table 2, the participants were very happy with the photographs they had selected in tasks 1 and 3. In task 2 they thought they had missed some, but felt content anyway.

Overall, the FreeEye tool was scored high in our trials. As shown in Table 2, the overall average score for how well the tool helped the user in selecting photographs was 3.7 on a scale from 1-5 (1=terrible, 5=very good). Compared to the participants regular ways of selecting photographs for similar tasks it scored 3.9 on a scale of 1-5 where 3 was as good as their regular one and 5 was much better. All but one of the participants used Windows operating system's user interface to select their photographs, and the tool was considered better than Windows OS (average score of 4.1). The one participant used Picasa and he thought the tool was as good as Picasa (score of 3).

Table 2. User satisfaction results

	Task 1	Task 2	Task 3	All
How well the tool helped to select?	3.9	3.1	4.1	3.7
How well the selected photos reflected the event?	4.5	3.9	4.6	4.3
Compared to regular way of selection (3 = same, 5 = better)	4.2	3.3	4.1	3.9

Generally the tool was thought to be good in recollecting events and photographs taken. The way in which it showed forgotten photographs was mentioned as a positive thing. One of the main issues the participants had with the tool was that if they had a particular photograph in their mind, it was not always easily found. Especially Task 2 (long event) was considered harder to do than the other tasks because there were more pictures than in a short event and unlike the yearbook task, the long event was restricted in time. The quantitative data in Table 1 supports this: more time was spent per chosen photograph than in the other tasks, although the time spent between clicks was not significantly different.

3.2 Discussion

Our research interest is in building a user interface that leverages available information to facilitate the photo selection process, not to automate it. Selecting photographs from increasingly large personal collections is a common task for a variety of situations. For that reason we have built a tool where the user is in charge and does the final selection. In our tool we used only the visual similarity information to help the user select photos for emailing, uploading, or making a book. Surprisingly, the visual similarity was considered helpful and as the scores of our trial show the participants were quite happy with the tool and the selected photographs. In future research, we are adding other similarity measurements to the user interface: location, people, tags, and time. We are also planning to add controls for the user to change the importance of a parameter at any time (e.g., location similarity is more important than visual similarity).

4. CONCLUSIONS

The objective of the CeWe Challenge was to generate a photo story from a personal photo collection by automatic or semi-automatic means. Our approach was to develop an intuitive user interface tool that displays a number of photographs and their relation based on visual similarity. To test the tool we conducted user trials with five people using their personal photographs. The evaluation oucomes can be summarized as follows:

- The selected photographs reflected the events very well (4.3/5)
- The tool was considered helpful (3.7/5), and better or as good as their existing ones (3.9/5).
- The partcipants selected on average 10-23 photographs, and spent from 2-6 minutes in selecting the photographs.

The tool in its simplicity has potential as a general user interface for selecting media from a large collection. What we learned from our trial was that our tool seems to work well with personal collections: the participants knew their own photographs which helped them to feel in control. This became especially clear with one participant who had in her collection also photographs taken by someone else. This caused confusion and a feeling of being lost. The strength of our tool is that it is a general tool that is not coupled with any particular task or with any particular system. The other main strength is that according to our user trial, people found it useful and helpful.

7. REFERENCES

[1] Ren, K., Calic, J. FreeEye - Interactive Intuitive Interface for Large-scale Image Browsing. ACM International Conference on Multimedia, Beijing, China, October 19-24, 2009.

[2] www.youtube.com/watch?v=7v3PbuNLA-0

[3] Calic, J., Gibson, D. P., Campbell, N. W. Efficient Layout of Comic-Like Video Summaries. Circuits and Systems for Video Technology, IEEE Transactions on (2007) vol. 17 (7) pp. 931-936

[4] Felzenszwalb, P. F. and Huttenlocher, D. P. Efficient Graph-Based Image Segmentation. International Journal of Computer Vision, Vol. 59, No. 2, September 2004.