PhotoSurfing: A 3D Image Browser Assisting Association-Based Photograph Browsing

Hiromi Horibe Takayuki Itoh

Graduate school for Humanities and Sciences, Ochanomizu University

E-mail:{romiko,itot}@itolab.is.ocha.ac.jp

1. Overview

Recent evolution of computation performance and enlargement of memory space enabled us to store large number of photographs in our personal computers. Today, we can use or enjoy the photographs for various purposes on the computers. We will have more chances to create photograph albums on the computers and personally browse them or publish on the Internet.

We think there are two main operations to look a large number of photographs on the computers. The former operation is search for certain specific photographs, and the latter operation is appreciation of sets of photographs as looking at photo albums. We think we will have more and more chances to appreciate sets of photographs on the computer-based albums, due to the increase of digital photographs.

Let us discuss how the computer-based albums are more useful or enjoyable than paper-based albums, and how they will evolve in the near future. We suppose that most of browsers of paper-based albums look them starting from the page 1, and browse them turning over the pages. It means that they look photographs in the order defined by the owners of the albums, and the action does not bring surprising things. On the other hand, we can develop more flexible photograph browsers, which enable users to freely select photographs users want to look. We think this idea makes photograph browsing more enjoyable and surprising.

We think metadata is a good key to assist unexpected association and discover unexpected photographs. For example, we may want to look more photographs "taken in the same day", "taken in the same place", or "shooting the same object". We can gather such photographs if image files have metadata including date, place, and keywords. Such association will make users to enjoy more photographs, and discover unexpected photographs.

This paper presents "PhotoSurfing", a 3D image browser that assists association-based photograph browsing. Figure 1 shows the overview of PhotoSurfing. Initially PhotoSurfing places a set of images into a virtual 3D space, where three axes of the 3D space denote the values of three metadata.

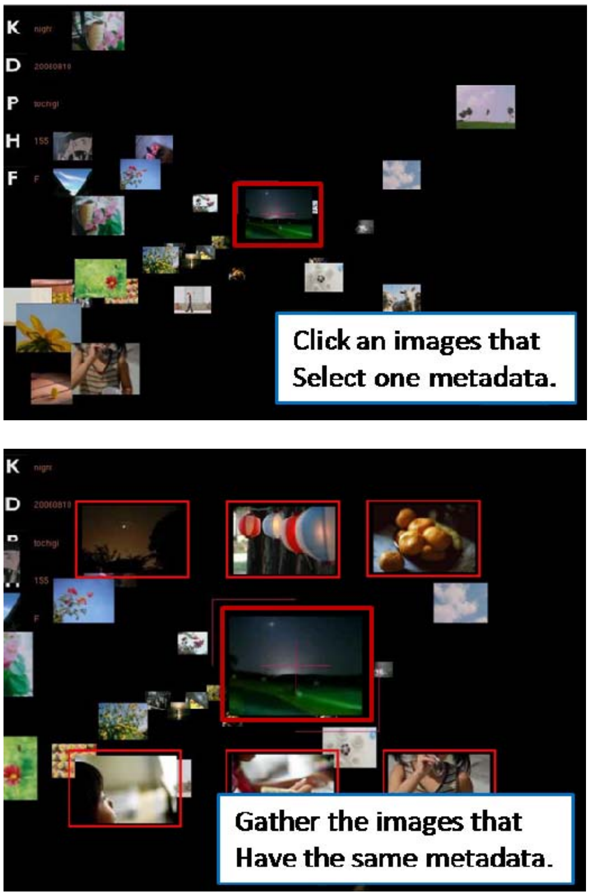

Figure 1. Overview of "PhotoSurfing". (Upper) A user selects an image shown in a virtual 3D space. (Lower) PhotoSurfing shows the images that have the same metadata as the selected image

Here, we need to implement suitable user interface to freely look at large number of photographs, and design the browser so that we can enjoy looking at photographs. PhotoSurfing provides a user interface for the walkthrough in the 3D space by mouse motion and selection of interested images by mouse clicking. This design enables us to look photographs as if we float on a 3D space, and to look at the photographs in the order we want to look. Also, when a user selects an image, PhotoSurfing shows a set of images that have same metadata as the selected image. When a user selects another image again, PhotoSurfing make the user jump to the position of the selected image in the 3D space. These functions make users help to look more photographs with metadata-based association, and discover unexpected photographs by jumping the 3D space.

Here we have a problem that photographs may occlude many other photographs, if we place the photographs strictly in the defined coordinate axis. PhotoSurfing supports a mechanism to reduce such occlusion.

2. Related Work

2.1 Image browsers

Graphical user interface and information visualization have been active research topics, because we think they make us easy to access various information. Image browser is a typical research topic that enables us to search for and enjoy photographs. For Example, many of large-scale image browser [1,2] shows images in 2D spaces. These techniques often suppose that users require specific images and therefore provide user interface so that users can easily zoom into the specific images. On the other hand, PhotoSurfing shows images in a virtual 3D space, and provides a user interface to fly the 3D space. This style brings users virtual reality to surf in the cloud of photographs, and purposelessly enjoy them.

Photosynth [3] and View Search Hokkaido [4] are widely used services on the Web, as examples of 3D photograph browsers. View Search Hokkaido applies so called “next generation retrieval technology (visual context search)”, which aims to retrieve the target image intuitively, in a word, using right brain of human. This system clusters images based on their features, not based on metadata. Photosynth [3] generates 3D models from photographs, and places the photographs based on the 3D geometry. These techniques are much different from PhotoSurfing on image layout policy and user interfaces; however, both techniques are good at sophisticated designs and functionalities, and reawake us the value of 3D image browsers.

2.2 Occlusion reduction

There have been several works [5,6] focusing on reduction of overlap of objects on the display spaces. Watanabe et al. [5] presented a technique to distribute icons without overlapping on a 2D display spaces. Miyazaki et al. [6] reported a technique for browsing the layered icons on a 3D space. These techniques construct the connection of the icons by applying Delaunay triangular mesh. PhotoSurfing also applies Delaunay triangular mesh based on the similar idea.

2.3 Metadata assignment

Existing photograph metadata assignment techniques can be roughly divided into the following two types: (1) automatic assignment, and (2) manual assignment.

There have been several automatic metadata assignment techniques. Recent most digital cameras can automatically record metadata (e.g. date, time, conditions) to image files. Moreover, several advanced studies on image analysis for automatic keyword assignment [7-9] have been presented, and they have been partially in actual use.

There have been also several software techniques for manual metadata assignment. For instance, we can easily search for images using iPhoto [10], according to manually predefined keywords.

3. Technical Components of PhotoSurfing

3.1 Metadata Setting

PhotoSurfing needs to assign metadata to every photograph as a preprocessing. We asked the following question to students in our university, to decide what metadata we assign:

> **[Question]** What is the important metadata to classify the image? (Multiple choices are allowed.)

Currently PhotoSurfing reads input data describing keywords, place, date, and preference level for each image, respecting to the result of the questionnaire survey. We require users to manually assign keywords, places, and preference level. Here, preference level is a numerical value during 0 to 5 which denotes a subjective assessment of photograph owner. We assume that place can be automatically converted into latitude and altitude values. Or, we can assume that future digital cameras will automatically record the latitude and altitude to image files. Also, PhotoSurfing calculates the feature values for each image. Current our implementation calculates the average hue value as a feature value.

PhotoSurfing needs to convert metadata into numeric values to specify the position of photographs on the screen. Especially, it needs to make a rule to calculate the numeric values from textual information. We think the following rules are effective:

[Position calculation 1:] If we calculate positions of photographs based on alphabetical order or Japanese syllabify order, it is effective to look for the photograph as if we use the index.

[Position calculation 2:] If we calculate positions of photographs based on distances of semantic meanings, it is effective to sequentially look at sets of semantically close photographs.

Now, we develop our system based on [Position calculation 1], which transforms the metadata into numerical values based on alphabetical order. This section calls such values as “metadata values”. Figure 2 shows the example of conversion of metadata into numeric values.

metadata	Before expressed numerically	After expressed numerically
Keyword	Jelly Beans	10
Date	20070512	23
Place	House	8
Hue	293	293
Fav.Rank	A	5

Figure 2. Example of expressing metadata numerically.

3.2 Image Placement onto a 3D space

PhotoSurfing places images onto a 3D space defined by an orthogonal coordinate system. PhotoSurfing calculates the positions of images from preselected three metadata values. This mechanism places semantically close images closer in the 3D space, and images which have same metadata onto a place in the 3D space.

Figure 3 shows an result of arranging the photographs onto a 3D space, where we assign “place” to X axis, “preference level” to Y axis, and “date” to Z axis, based on metadata values. Here, the photographs taken at the same place line up on a vertical line, because their X-coordinate values are equal. The higher the photographs are positioned, the higher that preference level is. In addition, we can intuitively understand date of photographs by the size of photograph, because newer photographs have smaller Z-coordinate values. If we assign different 3 metadata to 3 axes, we can get the quite different layout.

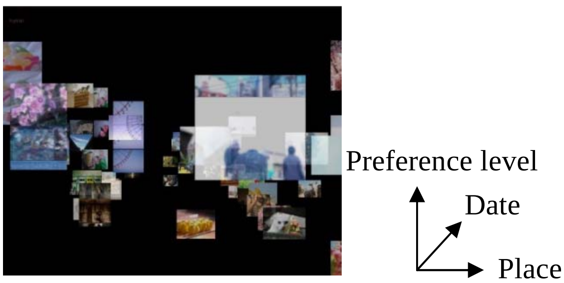

Figure 3. Example of photograph placement on a 3D space.

3.3 Image Selection

PhotoSurfing provides a user interface to select images in the 3D space. When a user selects an image, PhotoSurfing shows a set of images which have the same metadata, as shown in Figure 1. For example, when a user clicks an image, and selects a metadata “date”, PhotoSurfing shows the images taken on the same day. If the user selects different metadata such as “place”, PhotoSurfing shows the different set of the images taken at the same place. Or, the user can select another image and jump to the image in the 3D space.

3.4 Occlusion Reduction

Figure 4 is a partially enlarged display of Figure 3. Here, we can observe occlusion, where multiple photographs overlap on the display while they are placed in a 3D space.

Figure 4. Appearance of occlusion.

Occlusion among 3D objects is a general problem of 3D visualization techniques. PhotoSurfing supports an occlusion reduction algorithm shown in Figure 5. The algorithm firstly constructs a Delaunay triangular mesh connecting the center points of images in a 2D display space. It then relocates the images by stretching short edges of the triangular mesh.

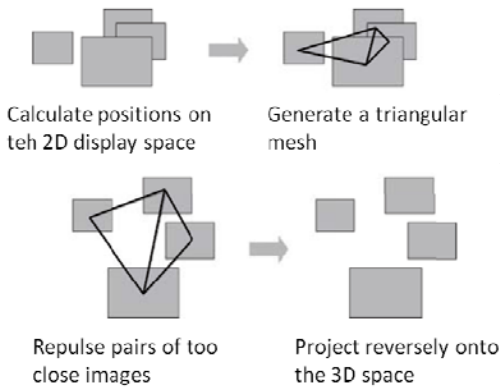

Figure 5. Algorithm for occlusion reduction.

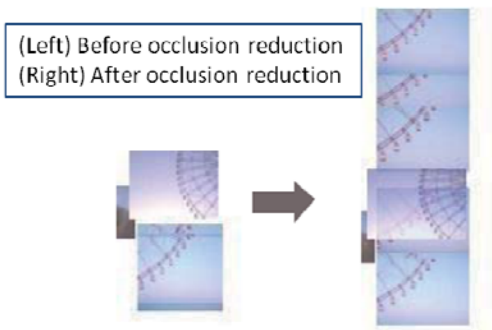

Figure 6. Result of occlusion reduction.

Figure 6 shows the result of the above algorithm. The result shows that the algorithm successfully reduced the occlusion among images.

Table 1 shows the total number of the images, and the number of the images those movements were more than width of images in our experiment. As shown in the table, number of largely moved images rapidly increases due to the increase of the total number of the images. We afraid that the result does not adequately denote the meaning of photographs, if many images largely move due to the occlusion reduction process. Moreover, the process is not effective, if the images place at the completely same position. In this case we may want to frequently switch images which place at almost same positions, as shown in Figure 7. We need to improve the process to solve the above problems.

Table 1 Distance of image by occlusion processing

Total number of images	Number of largely moved images
10	0
50	4
100	10
150	27

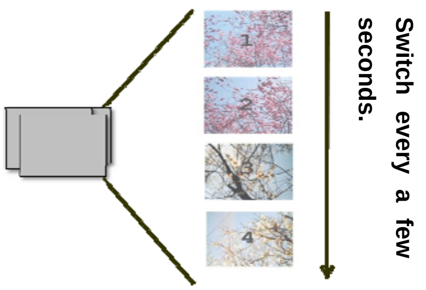

Figure 7. Switch of images which place at almost same positions.

3.5 User Interface

We support the following user interface.

PhotoSurfing supports various viewing functions, including zoom in, zoom out, and translation, for the walkthrough of 3D space, as well as 3D flight simulator software. As a result, users can look at images with a virtual reality as if they freely float in the 3D space. By this sense, we can enjoy photographs in the flexible order, contrasting looking at photographs of booklet albums in predefined page order.

We would like to develop PhotoSurfing so that everyone can intuitively operate. We are designing button operations and menu manipulations as easy as possible.

4. Conclusion

This paper presented PhotoSurfing, a 3D image browser that assists association-based photograph browsing. Our on-going works include high-performance implementation for large-scale collections that contain thousands or tens of thousands of images, and improvement of user interface for smooth surfing of the 3D space.

Also, we would like to advance the development of PhotoSurfing by solving the following problems in the future.

We would like to calculate metadata values of keywords not only based on the alphabetical order, but also based on the semantics, as discussed in Section 3.1. There have been several tools to retrieve semantic distances among keywords (e.g. WordNet Similarity [11]). We can order photographs based on the distances of their keywords, by applying a dendrogram. Similarly, we can order photographs based on the physical distances.

Currently we only apply hue as a color-based feature value; however, we would like to apply more various feature values. Currently, we are discussing the adoption of color correlograms [12] used for the color distribution analysis in the image.

Our current biggest interest is an analysis of trajectory of users while using PhotoSurfing. We would like to observe how users gather images that have same metadata according to their association, and jumps to unexpected positions, by tracking users' behavior as shown in Figure 8. Such observation is useful not only for improvement of PhotoSurfing, but also design of photograph albums or new image browsers.

Finally, improvement of occlusion reduction process is also an important future work, as discussed in Section 3.4.

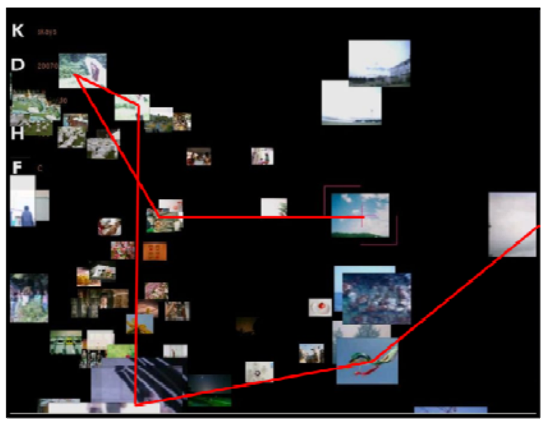

Figure 8. Tracking of users' behavior.

References

[1] Bederson, B. B., PhotoMesa: A Zoomable Image Browser Using Quantum Treemaps and Bubblemaps, ACM Symposium on User Interface Software and Technology (UIST 2001), pp. 71-80, 2001.

[2] Gomi A., Miyazaki R., Itoh T., Li J., CAT: A Hierarchical Image Browser Using a Rectangle Packing Technique, 12th International Conference on Information Visualization (IV08), pp. 82-87, 2008.

[3] Photosynth, http://labs.live.com/photosynth/

[4] View Search Hokkaido, http://www.view-hokkaido.jp/

[5] Watanabe N., Washida M., Igarashi T., Bubble clusters: an interface for manipulating spatial aggregation of graphical objects, UIST '07: Proceedings of the 20th annual ACM symposium on User interface software and technology, pp. 173-182, 2007.

[6] Miyazaki R., Itoh T., An Occlusion-Reduced 3D Hierarchical Data Visualization Technique, 13th International Conference on Information Visualization (IV09), 2009.

[7] Jane G., Kristina S., Abe C., Functionalities for automatic metadata generation applications: a survey of metadata experts' opinions, International Journal of Metadata, Semantics and Ontologies, pp. 3-20, 2006.

[8] Greenberg J., Metadata Extraction and Harvesting: A Comparison of Two Automatic Metadata Generation Applications, Journal of Internet Cataloging, Vol. 6, No. 4, pp. 59-82, 2004.

[9] Liddy E. D., Allen E., Automatic metadata generation & evaluation, Annual ACM Conference on Research and Development in Information Retrieval, pp. 401-402, 2002.

[10] iPhoto, http://www.apple.com/jp/ilife/iphoto/

[11] WordNet,Similarity,
http://www.d.umn.edu/tpederse/similarity.html

[12] Huang J., Kumar S. R., Mitra M., Zhu W. J., Zabih R., Image indexing using color correlograms .Proc. IEEE Conf. on Computer Vision and Pattern Recognition '97, pp. 762–768, 1997.