Int J Software Informatics
International Journal of Software and Informatics, ISSN 1673-7288

E-mail: ijsi@iscas.ac.cn
http://www.ijsi.org
Tel: +86-10-62661040

Photomosaic Generation and Representative Photomosaic Selection for Photograph Browsing

Kiho Sakamoto, Takayuki Itoh

(Ochanomizu University, Tokyo 1690075, Japan)

Abstract Photomosaic arranges many small photographs to represent a large image. Our study applies the photomosaic to a photograph browser CAT. Our implementation displays photomosaic while zooming out, and individual photographs while zooming in. Here, many photograph browsing software displays a set of photographs in the order of their times. To maintain this order of photographs, our photomosaic-like image generation technique firstly arranges the given set of photographs in the order of their times, and then retouches so that the set of photographs forms a photomosaic-like scene. This paper presents our technique for photomosaic generation, a user evaluation to discuss what kinds of photographs are preferable to be applied, and an automatic photomosaic selection technique.

Key words: Photomosaic generation; Photograph browsing; Zooming interface; Representative photomosaic selection; Mean shift filtering; face recognition.

Authors (**Kiho Sakamoto, Takayuki Itoh**). **Photomosaic Generation and Representative Photomosaic Selection for Photomosaic Browsing**

Introduction

We often store a large number of photographs due to the digitalization and downsizing of cameras. Photograph browsing [1-7] has been an active research topic, which assists users to explore and browse the large number of photographs. We aimed the development of a new photograph browser featuring an artistic representation in addition to the all-in-one display of sets of photographs, because such features would make photograph browsing more enjoyable. Based on this discussion we focused on applying photomosaics to the photograph browser. Photomosaic is a technique to generate large images representing particular scenes, by arranging large numbers of small images. Photomosaic is very artistic and enjoyable, because it looks like a particular scene of impressionism arts while zooming out, or a set of well-arranged scenes while zooming in.

This paper presents our study on application of photomosaic to a photograph browser CAT [1], which features a level-of-detail control with a zooming interface. Supposing the given set of photographs is hierarchically clustered, CAT displays representative photographs of higher clusters while zooming out, or individual photographs in the clusters while zooming in. This feature effectively assists interactive exploration of large photograph collections. Here, our technique replaces the representative photographs by the photomosaic images. The new photograph browser displays photomosaic images while zooming out, or individual photographs while zooming in. This feature improves the smoothness of the photograph replacement between representative and individual photographs. The paper consists of three parts:

1. A new technique to generate photomosaic-like images and its application to CAT.
2. A user evaluation to discuss what types of photographs are suitable to be applied to the photomosaic generation.
3. A new technique to select adequate representative photomosaics of landmarks by reflecting the user evaluation, and the examples the photomosaic selection results.

2. Related Work

Since the paper presents a photograph browsing technique applying photomosaic, the most related studies include photograph browsing, photomosaic generation, and representative photograph selection techniques. This section briefly surveys these research fields.

Many photograph browsing software have been presented in these years. They can be divided into structured and unstructured techniques. Unstructured techniques scatter sets of images onto display spaces according to predefined rules, such as content similarity [3] or timestamps [4]. Recent sophisticated techniques can interactively select the rules to realize flexible layout of photographs [5]. Structured techniques construct trees or graphs to organize the photographs. Tree structure is especially well applied to the photograph browsing [1,2], because many people are familiar with the traversal of tree-based systems such as file systems. Early structured techniques just applied space-filling information visualization techniques, and then they are extended to control the arrangement based on various rules such as time, geography, and friendships [6,7].

This study applies a photograph browser CAT [1] which features a zooming user interface. We suppose that photographs are hierarchically clustered, and representative photographs are selected for each cluster. CAT places photographs in the rectangle subregions of the display space by applying a space-filling algorithm featured by a hierarchical data visualization technique "HeiankyoView"[8]. CAT also features a level-of-detail control technique with a zooming user interface. It displays representative photographs as shown in Figure 1(a)(b) while zooming out, and individual photographs as shown in Figure 1(c) while zooming in. Other existing photograph browsers such as PhotoMesa [2] also features automatic photograph placement algorithm and zooming user interface; however, CAT is better for our purpose because it displays representative photographs in the appropriately sized and shaped rectangular subregions while zooming out. Layout algorithms between Quantum Treemap (applied by PhotoMesa) and HeiankyoView (applied by CAT) have been experimentally compared in detail in Itoh et al. [8].

CAT had a problem that switch of displayed photographs may look very sudden. We expect applying photomosaic as representative images of the clusters would solve this problem.

(a) (b) (c)

Figure 1: Example of zooming operation of CAT. (a)(b) Display of representative images while zooming out. (c) Display of individual images when zoomed in.

Automatic photomosaic generation is an active research topic. AndreaMosaic [9] is an orthodox technique, which selects small block images based on color matching and places them in a reticular pattern. Gianpiero et al. [10] presented a unique technique, which generates randomly edged small block images based on the edges of the original image, so that the generated photomosaics well preserve the edges and shapes in the original images. Several recent techniques further improved the quality and performance [11,12]. These photomosaic generation techniques arrange block images based on local similarity of colors and edges. In other words, they do not consider the semantics and meta information of photographs during their arrangement processes. Many photograph browsing software arranges the given photographs based on their semantics or meta information, such as timestanps, photogenic subjects, and positional information. We used to search for the particular photographs with such hints; and therefore it would be often difficult to search for the photographs if orthodox photomosaic generation algorithms arrange them. To solve the problem, this paper proposes a technique to generate photomosaic-like images from the sets of block images arranged in the order of timestamps. Computational complexity is another problem of the orthodox photomosaic generation algorithms. The next section discusses our photomosaic-like image generation technique requires smaller computation time.

Representative photograph selection is also an active research topic. VisualRank [13] is one of the most robust approaches to determine the ranks of individual photographs; the ranks can be applied to the selection of representative photographs. Several techniques for representative selection focusing on personal photograph collections have been recently presented [14,15,16]. On the other hand, there are few studies for representative photograph selection techniques for photomosaic generation in our survey. Thus, we discuss that criteria representative selection for photomosaic is different from those for ordinaly photograph collections in this paper. Also, we present a new approach to automatically select satisfactory representative photomosaics.

3. Photomosaic Generation and Zooming Interface

This section presents a new photograph browser, which displays photomosaic as representatives of clusters of photographs. This strategy makes interactive exploration of large photograph collections smoother. This section calls the reference image for photomosaic generation "*representative image*", and a set of arranged small images "*block images*".

3.1 Photomosaic-like Image Generation

The new technique for photomosaic-like image generation presented in this paper firstly arranges the given set of block images in the order of their timestamps, because it makes easier for users to look for particular images. The technique then retouches the arranged block images so that they look like the particular scene in the representative image. It repeats the arrangement of block images, if the number of blocks is larger than the number of block images.

The technique applies the HSB color system for block image retouch. The HSB color system describes colors by three variables: hue, saturation, and brightness. This technique calculates the RGB values of the each pixel of the final image, from the HSB values of representative and block images. Let the average HSB value of a block image (h_1, s_1, b_1), the average HSB value of the corresponding block in the representative image (h_2, s_2, b_2), and the ratio of saturation and brightness between the former and latter average values, $s_{12}=s_2/s_1$ and $b_{12}=b_2/b_1$. This technique retouches the HSB value (h, s, b) of a particular pixel as (h', s', b') by the following equations:

$$h'=h_2, \quad s'=s_{12}s, \quad b'=b_{12}b, \qquad (1)$$

The above formulation substitutes the average hue of representative image to the hue of all pixels of the block. They also multiply the ratio of average saturation and brightness to the saturation and brightness of each pixel of the block image. This formulation preserves the silhouette of the scene of the block images while retouching the hue.

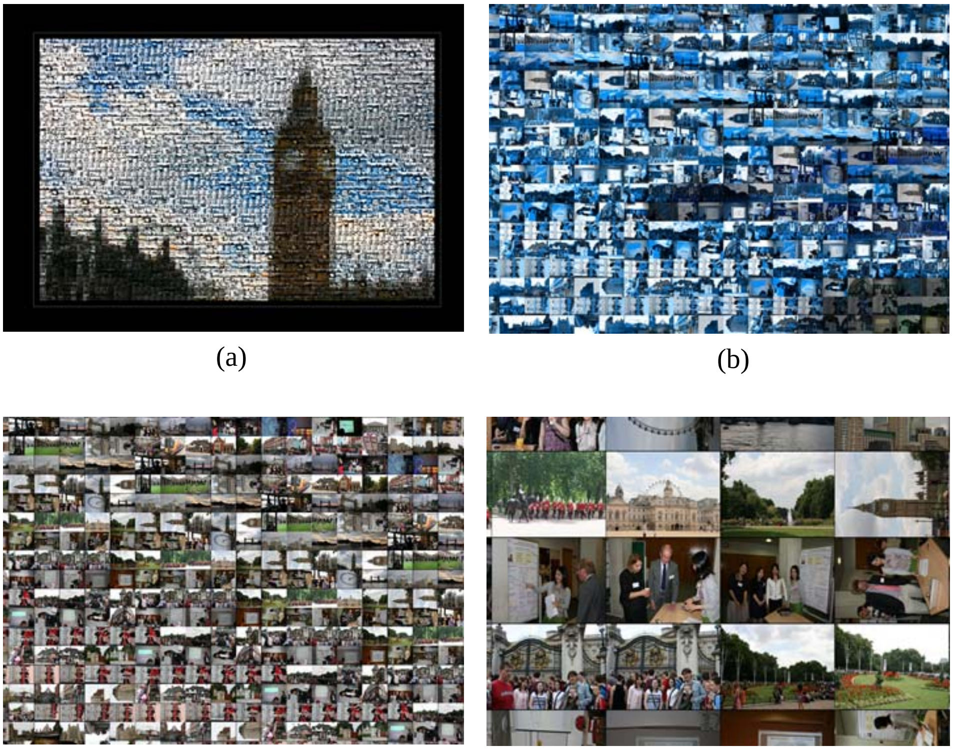
Computational time is another advantage of the presented technique. Suppose the number of original photographs n_P, the total number of pixels in the photomosaic n_M, the number of pixels in a block n_B, and the number of blocks m, where $n_M = n_B m$. Conventional photomosaic generation technique requires $O(n_B n_P)$ computation time to select a corresponding photograph of a block. As a result, it takes $O(m n_B n_P) = O(n_M n_P)$ computation time to complete the photomosaic generation. On the other hand, the photomosaic-like image generation technique presented in this paper requires much smaller computation time

$O(n_M)+O(n_B n_P)$. The technique firstly requires $O(n_P)$ computation time to calculate the average HSB value (h_2, s_2, b_2) of a block. Totally it requires $O(n_B n_P)$ to calculate for all the blocks. The technique then requires $O(n_M)$ computation time to calculate HSB values by applying equation (1) for each pixel of the representative image.

3.2 Photomosaic with zooming user interface

We present the photograph browsing applying CAT [1] and photomosaic. It firstly displays a photomosaic as the representative image of the given photograph set, and then switches to display individual photographs according to the zooming operations.

Figure 2 shows an example of a photomosaic generated by our technique. Figure 2(a) is the photomosaic generated from 174 photographs used as 60 by 45 pixels of block images. The representative image is divided into 5,256 blocks, and therefore the block image arrangement process is repeated 30 times. Figure 2(b) is a partial zoom-up of the photomosaic images. This zoom-up view shows that buildings or trees are taken in the block images. Their colors are much different from the colors of real buildings or trees, but we can recognize them from their silhouette in the block images. Then the browser switches the displayed images from the photomosaic images to the input images as shown in Figure 2(c). This switch looks very smooth, because it just changes in the hue of the images. Figure 2(d) is a partial zoom-up of the input images.



(a) (b)

(c) (d)

Figure 2: Zooming operation with a photomosaic image. (a) Photomosaic generated by our technique. (b) Partial close-up. (c) Switched to the original photographs. (d) Close-up to some of the original photographs.

Also, our implementation optinally switches the photomosaic images while the repetation of zooming in and out opetations. The browser displays a set of photographs instead of the photomosaic, when a user zooms in the particular portion of the display space. When the user zooms out after the zoom in operation, it displays another photomosaic. Even if the browser displays the different photomosaic, it can seamlessly switch the display from the set of photograph to the photomosaic. The browser provides an enjoyable show by randomly selecting different photomosaics during zoom in and out operations.

4. Evaluation of Photomosaic toward the Automatic Selection

4.1 Subjective evaluation

It is important to reduce manual operations to make photograph browsers easier and convenient. Therefore, automatic representative photomosaic selection is an important problem for the photograph browsers featuring the zooming user interfaces. We aim to develop a technique to automatically select preferable photographs before generating photomosaics from them, and call this technique “automatic photomosaic selection” in this paper. Representative photograph selection itself is a very difficult problem, and therefore many studies [14-16] have been already presented. Here, we found it is important to discuss the difference what kinds of photographs are preferable as representatives between general photographs and photomosaic. Therefore, we conducted a user evaluation of the preferences of photomosaic. We showed photomosaic generated by our technique to 28 subjects, and asked to answer the questions regarding the preference for photomosaic; for example, “which representative image do you think the best to be the photomosaic?” as five-point Likert scale evaluations. We prepared photograph collections of abroad trips, and generated several photomosaics from each of the collections. We randomly selected four to eight pieces of photomosaics for each question.

Figure 3(a) shows an evaluation result of photomosaic where famous landmarks are taken in the original photographs. Big Ben and London Tower had higher ratings. We suppose the main reason of the higher rating is the contrast between the buildings and background, which makes the recognition of the landmarks easier, as well as that these landmarks are famous. We suppose this knowledge may be useful for the automatic representative photograph selection, because it is easy to calculate the contrast between the buildings and backgrounds. This strategy is also good to avoid selecting too dark or single colored photographs as representatives.

Figure 3(b) shows an evaluation result of photomosaic where one or more persons are taken in the original photographs. The result denotes that photographs which take too distant or close persons (e.g. close-up and group photographs) had relatively low rating. We suppose that photomosaic may get bad impressions if the original photographs are focused only on human faces, because it is often difficult to identify where they are. We therefore think it should be careful to select photographs focusing only on human faces as representatives. At the same time, we found that subjects joined to the trips or parties could identify the persons taken in the photomosaic, while it was impossible to identify them for the other subjects who do know them.

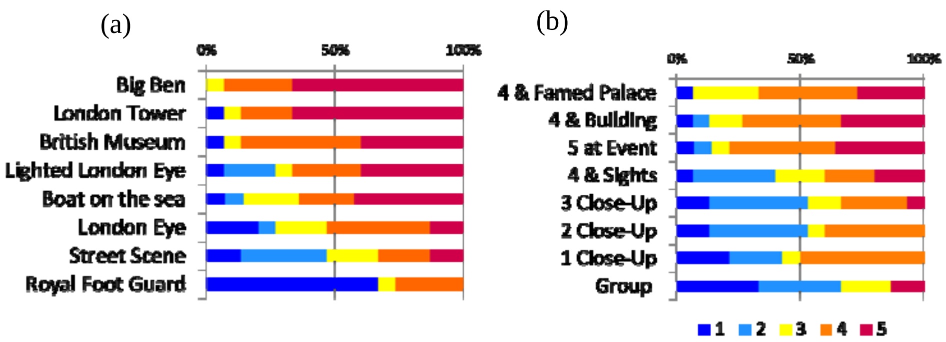


Figure 3: Result of user evaluation. (a) Landmarks in the photomosaic. (b) Persons in the photomosaic.

4.2 Discussion for Automatic Representative Photograph Selection

Here we discuss the criteria for automatic representative photograph selection for the two types of photographs: ones which take landmarks, and other ones which take human faces. We suppose criteria to select representative images are different between ordinary photographs and photomosaic. When we are to select a representative image from a set of photograph, many people will prefer to select the photographs which famous landmarks or persons are taken. This preference can be also applied to photomosaic selection, according to the user evaluation presented in the previous section. On the other hand, we found several differences of characteristic preferences between photomosaic of landmarks and persons. The following sections discuss the conditions for preferable representative image selections based on the evaluation results introduced in the previous section.

4.2.1 Conditions for the Photomosaic of Landscapes

We conducted the following hypothesis from the evaluation results regarding the photomosaic of landmarks. First, the preference of photomosaic is related to the contrast between foreground objects and background. If the contrast is large as shown in Figure 4(a), it is easy to clarify and recognize the objects from background. Otherwise, the color of photomosaic is monotonic as shown in Figure 4(b). In addition, photomosaics tend to be preferable if they take large foreground objects placed around the center of the input photographs, as shown in Figure 5. We suppose these points are good criteria for the selection of preferable representative images of landmarks.

(a) (b)

Figure 4: Examples of photomosaics containing (a) large or (b) small contrast.

Figure 5: Examples of preferable photographs, which take large landmarks as indicated by blue circles.

4.2.2 Conditions for the Photomosaic of Persons

We conducted the following hypothesis from the evaluation results regarding the photomosaic of persons. First, the preference of photomosaic is related to the sizes of their faces in the photographs. We cannot enjoy looking at the background if the sizes of faces are too large, as shown in Figure 6(a). Or, we cannot recognize the characteristics or expression the human faces if their sizes are too small in the photomosaic, as shown in Figure 6(b).

(a) (b)

Figure 6: Examples of photomosaics containing (a) large or (b) small sizes of faces.

As a result, we conducted that moderately sized faces bring well-balanced composition to generate photomosaics. In addition, large contrast between human faces and background will bring preferable photomosaic, as same as the photomosaic of landmarks. We suppose these points are good criteria for the selection of preferable representative images of persons.

5. Automatic Representative Photomosaic Selection for Landmarks

Our evaluation concluded that photographs, which have larger contrast between landmark and background, are usually preferable as photomosaics. Also, we concluded that photographs are preferable if landmarks in representative images are large and placed at the center of a photograph, as shown in Figure 5.

Based on the above discussion, we developed and tested the following technique for the selection of photographs taking landmarks. Our implementation firstly applies a segmentation process to divide the image space into multiple parts by the Mean Shift Filtering provided by OpenCV. It then calculates the areas and centers of the segments, and selects the segment which is sufficiently large and close to the center. It then calculates the contrast between the average colors of the specified segment and other segments. Our implementation repeats the above process for all photographs, and calculates the value v applying the following equation to determine the adequateness of the photographs as representative photomosaics:

$$v=p/a_{ls}+qx_{ls}+r/d_{col} \qquad (2)$$

Where,

- a_{ls} is the number of pixels of the specified segment.
- x_{ls} is the difference of the number of pixels between the x-coordinate of the center of the image and the x-coordinate of the center of the specified segment.
- d_{col} is the difference of the average color values between the specified segment and the other segments. Our current implementation simply calculates the differences of their averages of RGB values.
- p, q, and r are user-specified constant values. Our current implementation applies p=3.0x10^7, q=2.0x10^1, and r=5.0x10^3 for photographs containing 3888x2592 pixels. We may need to adjust these constant values for different resolutions of photographs.

Figure 7 shows examples of photographs which are highly or poorly evaluated by the equation (2). Upper and central four photographs in this figure are highly evaluated; where these photographs actually had relatively large foreground objects close to the center of the photographs, and these objects had relatively clear contrast against the background. On the other hand, lower two photographs in this figure are poorly evaluated; where they had too fine objects and low contrast.

v=783

v=1276

v=1696

v=1796

v=614596

v=73590

Figure 7. Examples of highly or poorly evaluated photographs. The upper and central four photographs are highly evaluated, while the lower two photographs are poorly evaluated.

We conducted the user evaluation to validate the effectiveness of the presented photograph evaluation method. We showed the 12 photomosaics to 22 participants, and asked to answer the evaluation of the each of the photomosaics in the five-point Likert scale, where 5 is the best and 1 is the worst. At the same time, we calculated the value *v* for each of the photograghs corresponding to the photomosaics. Figure 8 shows the scatterplot illustrating the relations between the values *v* (assigned to the x-axis) and evaluations of participants (assigned to the y-axis). Pearson correlation coefficient between them was -0.428, where -1 is the best result.

Figure 8 illustrates that the value *v* had preferable correlation with the evaluation of participants, except two photograph (pointed by a red circle in Figure 8) had poor evaluations though the values *v* were relatively small.

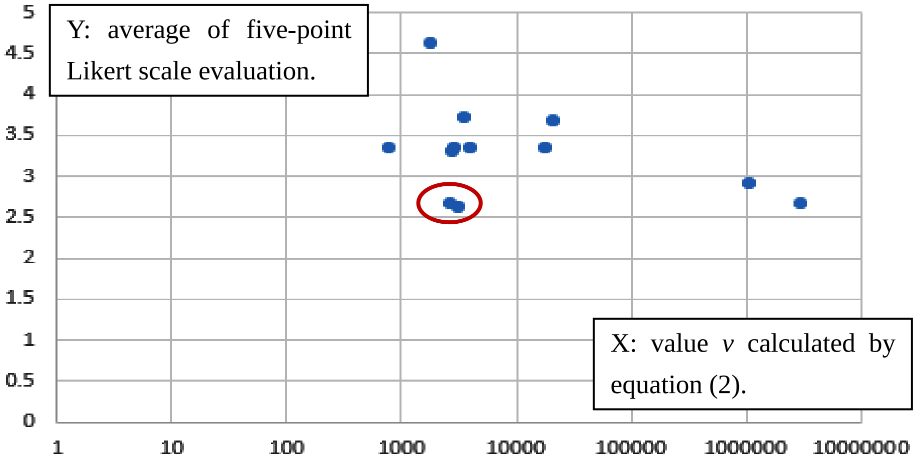


Figure 8. Correlation between the calculated evaluation and answered evaluation.

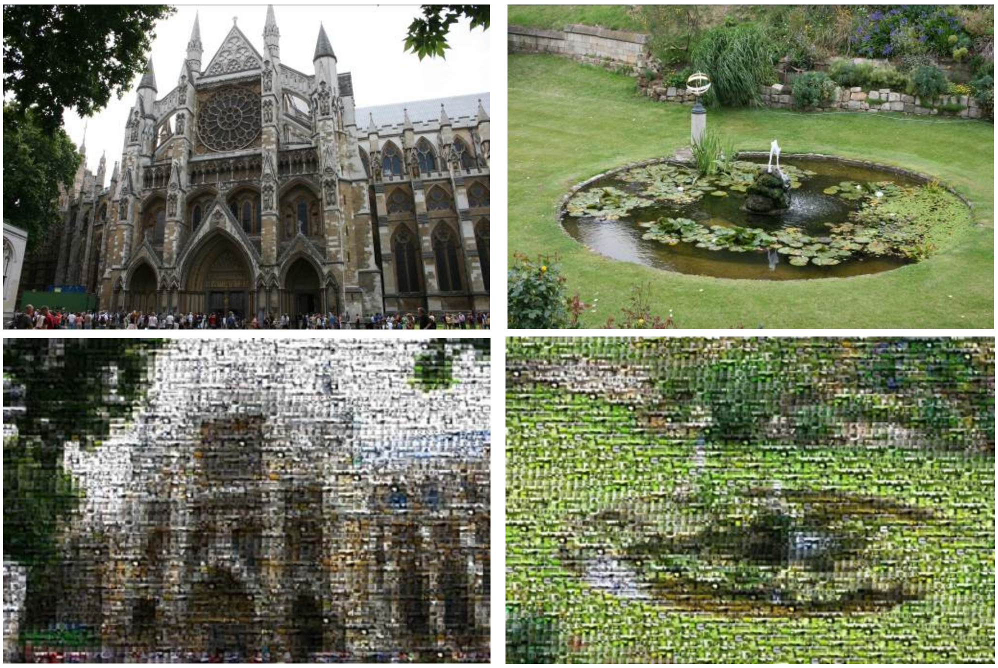


Figure 9. Photographs which our method highly evaluated though evaluation of participants was not high.

Figure 9 shows the two photographs which our method highly evaluated as representative photomoisaics though evaluation of participants was not high. We found that foreground objects and background in these photographs had large contrast but close hue. It may often happen that our photomosaic-like image genetation technique generates unclearly-looking photomosaics from such photographs, since the technique just preserves hues of the original photographs as described in equation (1). We may need to revise the equation (2) so that we can poorly evaluate such photographs shown in Figure 9.

6. Conclusions and Future Work

This paper presented a photograph browsing technique applying photomosaics. This technique displays photomosaic as representative images on a photograph browser CAT. It realizes smooth switch of displayed images by applying photomosaic rather than original photographs. This paper then introduced a user evaluation to determine which kinds of photomosaic are preferable for representative images of the photograph browsing. We found that photomosaic of landmarks got higher ratings if contrast between buildings and backgrounds are clearer. Also, we found that photomosaics focused only on human faces got relatively lower ratings, because it is generally difficult to identify where they are taken. However, these images even got higher ratings from subjects who know the taken persons. We supposed that preference of the representative photograph selection depend on the possibility of identification of taken persons, according to the free comments of the subjects. Finally, this paper described techniques for automatic selection of representative photographs taking landmarks and persons for photomosaic generation, and introduced the results with photographs taking landscapes.

We are now developing an automatic representative selection technique for photomosaics of human faces as an on-going work. Our evaluation concluded that photographs with moderately sized human faces are better for representative images to be shown as photomosaics. Our current development is based on this discussion. Our implementation firstly recognizes human faces, and gets their sizes. It then selects the photographs, which take moderately sized human faces. Finally, it applies the processes developed for the selection of landmark photographs to select small number of representative photographs. We would like to test the availability of this implementation, and discuss if the technique really selects preferable representative photographs.

Following are our other future issues. We would like to find out other conditions for preferable photomosaic generation. Especially, we will need to address the selection of the best photographs from sets of them taking the same objects or scenes. Near-duplicate detection [15] is one of the useful concepts to solve the problem. We expect more various user evaluations will bring wide aspect of conditions. For example, different lighting of landscapes, different races of human faces will bring different impression of photomosaics. We would like to widely explore which factors bring preferable photomosaics.

References

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

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

[3] J. Yang, J. Fan, D. Hubball, Y. Gao, H. Luo, W. Ribarsky, Semantic Image Browser: Bridging Information Visualization with Auto- mated Intelligent Image Analysis, IEEE Visual Analytics in Science and Technology, 191-198, 2006.

[4] P. Brivio, M. Tarini, P. Cignoni, Browsing Large Image Datasets through Voronoi Diagrams, IEEE Transactions on Visualization and Computer Graphics, 16(6), 1261-1270, 2010.

[5] C. T. Vi, K. Takashima, H. Yokoyama, G. Liu, Y. Itoh, S. Subramanian, Y. Kitamura, D-FLIP: Dynamic and Flexible Interactive PhotoShow, Advanced in Computer Entertainment, Lecture Notes in Computer Science, 8253, 415-427, 2013.

[6] A. Gomi, T. Itoh, MIAOW: A 3D Image Browser Applying a Location- and Time-Based Hierarchical Data Visualization Technique, Advanced Visual Interface (AVI10), 225-232, 2010.

[7] A. Gomi, T. Itoh, A Personal Photograph Browser for Life Log Analysis based on Location, Time, and Person, ACM Symposium on Applied Computing, Multimedia Visualization Track, 1250-1257, 2011.

[8] T. Itoh, Y. Yamaguchi, Y. Ikehata, and Y. Kajinaga, Hierarchical Data Visualization Using a Fast Rectangle-Packing Algorithm, IEEE Transactions on Visualization and Computer Graphics, 10(3), 302-313, 2004.

[9] AbdreaPlanet, AndreaMosaic, http://www.andreaplanet.com

[10] G. DiBlasi, G. Gallo, M. Petralia, Fast Techniques for Mosaic Rendering, Computational Aesthetics in Graphics, Visualization and Imaging, 29-39, 2005.

[11] H. Narasimhan, S. Satheesh, A Randomized Iterative Improvement Algorithm for Photomosaic Generation, Nature and Biologically Inspired Computing, 777-781, 2009.

[12] S. B. M. Sah, V. Ciesielski, D. D'Souza, Refinment Techniques for Animated Evolutionary Pohotomosaic Using Limited Tile Collections, Applications of Evolutionary Computation, Lecture Notes in Computer Science, 6025, 281-290, 2010.

[13] Y. Jing, S. Baluja, VisualRank: Applying PageRank to Large-Scale Image Search, IEEE Transactions on Parrern Analysici and Machine Intelligence, 30(11), 1877-1890, 2008.

[14] C..-H. Yeh, Y.-C. Ho, B. A. Barsky, M. Ouhyong, Personalized Photograph Ranking and Selection System, ACM Multimedia, 211-220, 2010.

[15] W.-T. Chu, C.-H. Lin, Automatic Selection of Representative Photo and Smart Thumbnailing Using Near-Duplicate Detection, ACM Multimedia, 829-832, 2008.

[16] C. Li, A. Gallagher, A. C. Loui, T. Chen, Aesthetic Quality Assesment of Consumer Photos with Faces, IEEE International Conference on Image Processing, 3221-3224, 2010.