Tiling Slideshow: An Audiovisual Presentation Method for Consumer Photos

Wei-Ta Chu, Jun-Cheng Chen, and Ja-Ling Wu
National Taiwan University

The Tiling Slideshow system automatically organizes consumer photos and provides a novel audiovisual presentation. Displaying at the same pace as user-selected music, photos are elaborately manipulated and displayed to mold a novel browsing experience. In contrast to conventional photo slideshows, the proposed presentation provides tighter audiovisual coordination, and offers a more lively viewing experience.

Digital cameras have become an indispensable commodity for each family or individual in recent years. With the advance of digital storage technology, people can take pictures at will and have become accustomed to recording everything with photographs rather than text. Nevertheless, large amounts of photos without appropriate organization draw many potential problems in information access. People spend too much time browsing and often get lost in massive photo collections. Therefore, to help with organizing photos and to make the photo-viewing process more enjoyable, we developed some advanced analysis and presentation techniques that facilitate effective photo organization and efficient photo browsing.

We aren't the first researchers to try to tackle this issue. Some commercial photo browsers and research projects[1-3] have taken the approach of providing thumbnails or photo management functionalities. Although they do help users with managing and accessing image/photo collections, some critical problems still significantly impede users' browsing experience:

- Large amounts of disordered photos stuff the users' storage and make photo browsing and access tedious. One of the most popular ways to present photos is a slideshow. However, sequential presentation often takes much time and makes users tired.
- Consumer photos taken by amateurs often suffer from quality degradation. Techniques of quality estimation and photo filtering should play an important role in photo presentation and management.
- Conventional photo slideshows display photos one-by-one, according to alphabetical or temporal order. Therefore, photos taken in the same scene or having the same topic are separated into different slots, and the browsing experience is cut off.

We propose a new type of audiovisual presentation to address these shortages. With the help of image quality analysis, the proposed system automatically filters out low-quality photos. The remaining photos are then organized in terms of time and content characteristics so that similar photos can be displayed together. Finally, the proposed system elaborately manipulates photos and generates audiovisual slideshows that follow the music's pace. As Figure 1 shows, the time stamp of each photo's occurrence is determined by the beats of the music. As a strong beat occurs—for example, time instance (3) in Figure 1—the displaying content switches to another frame. Because it's likely that the final result will stick tiles of varying sizes on a wall, we call the proposed presentation a *tiling slideshow*.

The goal of the proposed tiling slideshow system is to generate a descriptive presentation via elaborate photo arrangements. According to the guidelines of technical writing, a solid paragraph contains a topic sentence that identifies the main idea and several supportive sentences that provide supportive materials. Many paragraphs are concatenated to convey the whole narration of an article.

Likewise, we advocate that a journey or an event can be reproduced by many *photographic paragraphs*, which are the frames in Figure 1. Each frame is composed of a larger-sized topic photo and several smaller-sized supportive photos. Tiling multiple photos into the same frame strengthens the atmosphere of the viewing experience. Photo presentation that's synchronous to music beats even improves the enjoyment of browsing.

1070-986X/07/$25.00 © 2007 IEEE

Published by the IEEE Computer Society

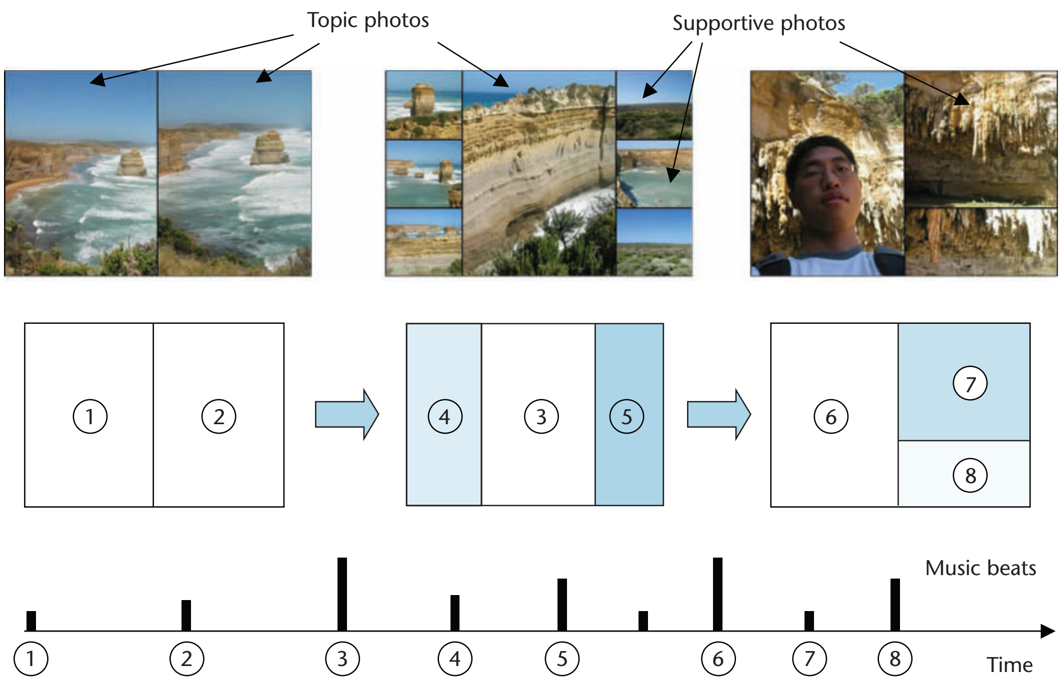

Figure 1. Example of a tiling slideshow.

On the basis of this idea, we propose a system that integrates visual and music analyses and automatically composes a vivid audiovisual presentation, as Figure 2 shows. The issues we discuss are summarized as follows:

- *Photo preprocessing.* We perform orientation correction and remove low-quality photos that are caused by blur and/or underexposure/overexposure.
- *Photo organization.* The proposed system automatically organizes photo collections by using time and content characteristics. We integrate them to perform finer clustering so that photos at the same scenic spots or presenting the same event are grouped together.
- *Music beat analysis.* We detect music beats and use them to drive the progress of presentation.
- *Temporal and spatial composition.* From the temporal perspective, photo presentation and frame switching are synchronous to music beats. From the spatial perspective, photos having similar characteristics are elaborately manipulated and arranged at the same frame.

Visual processing

In this section, we describe the steps necessary for visual processing in our system.

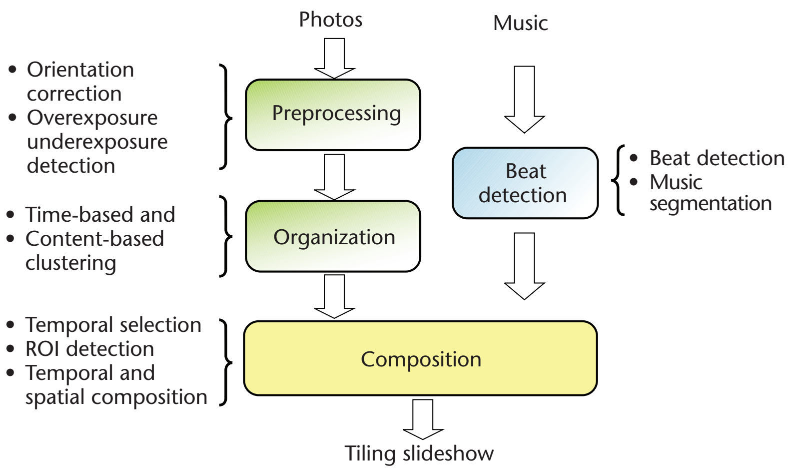

Figure 2. System flowchart of the proposed tiling slideshow.

Photo preprocessing

Before the system elaborately organizes the photos, it tries to correct photo orientation and filter out low-quality photos. These preprocesses prevent users from a great deal of tedious work.

Orientation correction. The orientation problem derives from the inconsistency between the user's intuition in browsing and the taken angles of photos. Recently, some studies[4,5] have been conducted for automatic orientation correction. However, the reported methods are often sophisticated and aren't computationally tractable. When it comes to aesthetically improving photo slideshows, though, it's important to perfectly

> **We can first utilize the time-based clustering to categorize photos by days, and then use the content-based approach to cluster photos having a similar visual appearance.**

correct orientation errors. Fortunately, more and more digital cameras are equipped with orientation sensors and simultaneously store orientation information as EXIF (exchangeable image file format) metadata[6] when shooting. Therefore, our system can easily and reliably correct photo orientation by checking EXIF information.

Blur detection. Most consumers aren't familiar with professional-level photography, so their photographs often suffer from unwanted defects. Blurred photos are often caused by hands shaking or the camera being out of focus. In this system, we adopt a wavelet-based method[7] to detect blur by checking edge characteristics in different image resolutions. With this information, the system can filter out photographs with severe blur degradation.

Underexposure and overexposure detection. Photos with a bad exposure are often due to incorrect camera setup parameters. We devised a simple detection method based on the intensity characteristics.[8] For example, when the number of darkness (or brightness) pixels in a photo is larger than a predefined threshold, the system indicates that this photo is underexposed (or overexposed).

Photo organization

To realize the idea of a photographic paragraph, we must organize photos so that photos in the same cluster are semantically related and are displayed in the same frame. Therefore, our system organizes photos based on time and content characteristics.

Time-based clustering. From the perspective of temporal context, photos taken within a certain time period usually share the same topic and record the same semantic events. Based on this idea, we adopted the time-based clustering algorithm proposed elsewhere.[2] Photos are first sorted by their shooting time. This algorithm dynamically detects noticeable time gaps by checking the time stamps of photos in a sliding window. These time gaps present changes of the shooting pace and reveal that photos are taken in different places.

Content-based clustering. From the content-based perspective, we exploit the dominant color and color layout descriptors defined in MPEG-7 as features. The dominant color represents an image's statistical color characteristics. The color layout represents the spatial distribution of colors and roughly describes the image's structure. The system uses the average of the normalized dominant color and color layout distance to measure the similarity between photos.

These two clustering approaches can be applied successively to organize photos at different granularities. For example, we can first utilize the time-based clustering to categorize photos by days, and then use the content-based approach to cluster photos having a similar visual appearance. We've shown the effectiveness of this clustering method elsewhere,[8] and we provide further evaluation in this article.

Region of interest determination. Because the frame space is smaller than that of multiple photos, it's inevitable that we'll need to shrink photos at some point to fit them in one frame. The simplest way is to directly resize photos according to their aspect ratios. However, blind resizing often causes significant information loss because the details of important objects would be rudely shrunk. To lose as little information as possible, we prefer cropping the region that retains the most important and attractive part.

To determine what is the most important or attractive in a photo, we use a user-attention model[9] to detect the region of interest (ROI). According to whether human faces exist in photos, the user attention is modeled by top-down or bottom-up approaches. Then the detected attentive region will be the unit for cropping and resizing. Chen et al. have provided additional details of ROI determination and related manipulation.[8]

Music analysis

Music plays an important role in multimedia

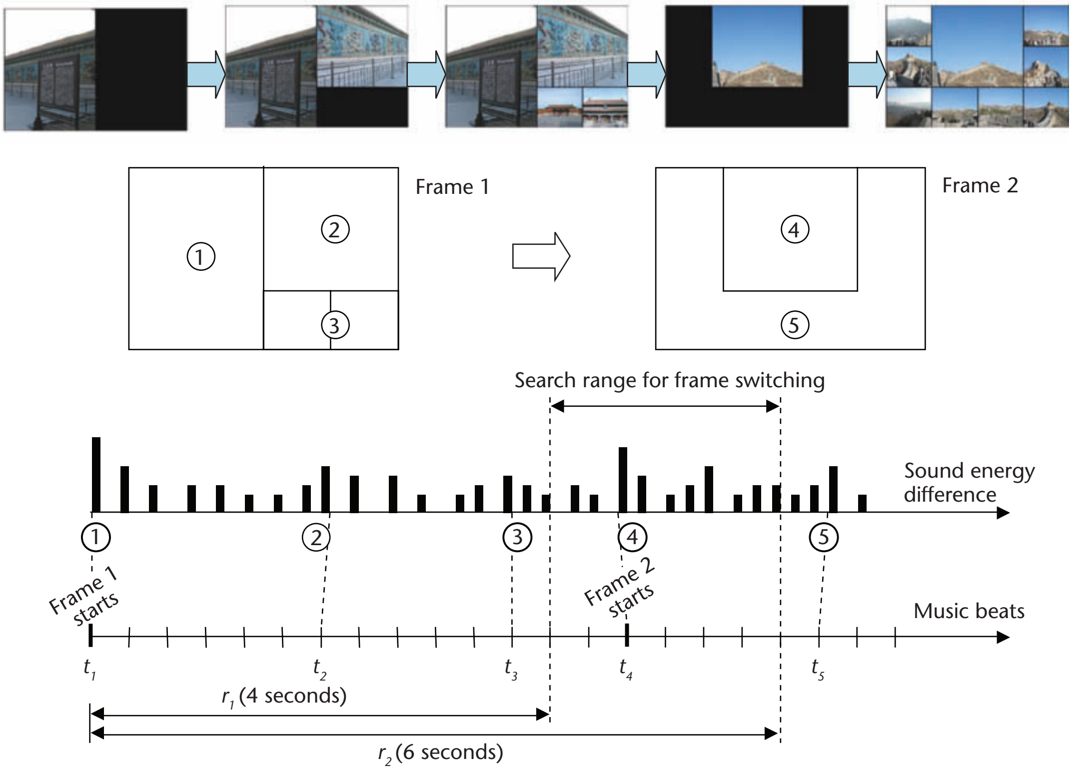

Figure 3. An example of determining the timing for frame switching and photo display.

presentation. Keeping with the pace of the music, a tiling slideshow can then help construct this notion of solid photographic paragraphs, while also putting effort into concatenating these paragraphs into an affective photographic story. To achieve this goal, we detect music beats based on the algorithm proposed in Scheirer.[10] This method analyzes music signals in different frequency bands and estimates beat information therein. Beat information serves as the timer for photo presentation. In the following, we describe how the timing for frame switching and photo displaying are determined.

Timing for frame switching

In addition to music beats, we also consider sound energy differences between adjacent audio frames for frame switching. In the example of Figure 3, if the starting time of Frame 1 is t_1, we check the sound energy differences in the range from ($t_1 + r_1$) to ($t_1 + r_2$), and detect the largest energy difference in this range. To guarantee the coordination between visual and aural media, the time stamp of the nearest beat to the largest energy difference is set as the timing for frame switching, like time stamp t_4 in Figure 3. In our implementation, r_1 and r_2 can be adjusted to control the displaying speed and meet different people's preferences.

Timing for photo display at a frame

At each frame, we have to determine each photo's occurrence time stamp. Many approaches exist for achieving this task. One of them is to unequally dispatch display times according to each photo's allocated area. In our implementation, we prefer to average and distribute the displaying duration (for example, from t_1 to t_4) to each photo. We find the time stamps of the music beats that are nearest to the average distributed points. With this elaborate design, the proposed scheme synchronizes the visual display with the music's pace.

Tiling slideshow composition

After completing the aforementioned processes, we try to put photos from the same cluster in the same frame. At the composition stage, we have to face several challenges:

- *Challenge 1*. Given a time-limited music clip, we often have to select a subset of photo clusters for displaying. If a frame lasts for 4–6 seconds (as assigned by users), the slideshow can only afford at most 60 clusters of photos if the user selects a 4-minute music clip. The importance of a photo cluster is, therefore, defined to be the metric for cluster selection.

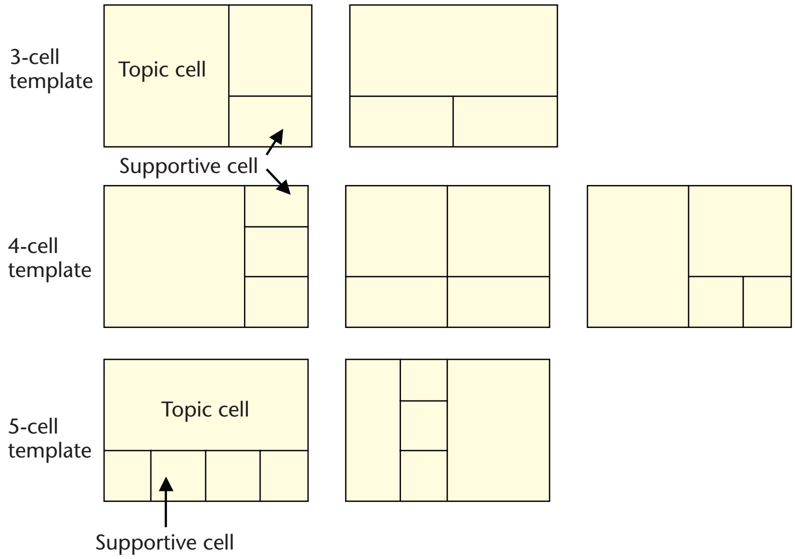

Figure 4. Examples of different kinds of templates.

- *Challenge 2.* Given a cluster of photos, we try to reasonably manipulate them so that more important or attractive photos occupy a larger space, and photos having similar characteristics are located nearby. Based on predefined templates, we have to devise a method to select the most appropriate one to be the display platform.

- *Challenge 3.* Once the system determines the match between photos and displaying templates, we have to elaborately resize or crop the original photos to fit in with the limited region.

Cluster selection (for challenge 1)

A cluster of photos are expected to be displayed within a music segment. Nevertheless, it's often the case that thousands of photos (and therefore hundreds of photo clusters) can't be completely displayed because a time-limited music clip (for example, 4 minutes) can only be divided to tens of segments. To solve this problem, photo clusters are sorted based on their cluster-based importance in descending order, and the first *N* clusters are picked for presentation if only *N* music segments are available.

We estimate the importance of each cluster based on two features: photos per minute (PPM) and photo conformance (PC). PPM denotes the shooting frequency of photos in a cluster, while PC denotes the content-based similarities between photos in the same cluster. These two features are fused together to describe the importance of a given cluster. Details of the cluster-based importance measurement are discussed elsewhere.[8]

Template design and determination (for challenge 2)

Once the clusters to be displayed are determined, the problem now is how to select appropriate layouts for presentation. According to the guidelines of a publication layout,[11] we design several templates for showing different numbers of photos:

- *Showing limited content in a limited space.* Presenting too many photos in the same frame confuses viewers and obscures the presentation. Therefore, the number of photos in a frame is limited to no more than 12 in this work. A region that displays one photo is called a *cell*, and we design various templates containing between 1 to 12 cells.

- *Enlarging important photos to drive visual perception.* Photos at the same frame are elaborately scaled into different sizes to show their relative importance. Therefore, the areas of different cells in a template should be differentiated to show variations.

- *Designing layouts by adjusting uniform subunits.* This not only enriches the spatial arrangement but also maintains the regularity of presentation. In our implementation, we divide a frame into 12 equal-sized, basic units. To construct a template that consists of four cells, for example, one construction method is to merge the left nine regions into a large cell and leave the remaining three regions as three small cells. Figure 4 shows some of our templates for showing three, four, and five photos.

Intuitively, if the number of photos in a selected cluster is four, for example, we just take the templates with four cells for presentation. As we mentioned, we designed several templates with a variety of cell numbers to enrich layouts. Then the problem is to determine which template is appropriate for showing the given photo cluster. Conceptually, more representative or important photos should be allocated a larger space. Therefore, we first define template- and photo-based importance values to be the metrics for template determination.

- *Template-based importance.* As Figure 4 shows, each tiling template consists of at least one topic cell and several supportive cells. Based on the ratio of the cell's area over the whole frame, we calculate each cell's importance. The system then sorts the cells' importance values in the same template in descending order.

- *Photo-based importance.* The system calculates the photo-based importance based on the face region and attention value. The system applies a linear weighting method to combine these two features and derive the photo-based importance.[8] It also sorts the calculated photo-based importance in descending order.

On the basis of these importance values, the system determines the best match between templates and photos by finding the pair that has the most similar importance distribution. In this system, we try to accomplish this work from two perspectives: by vector angle and relative entropy.

In the vector angle approach, we respectively pack template- and photo-based importance values into vectors. Given a set of k-cell templates, $\Gamma = \{T_1, T_2, \ldots, T_s\}$, a cluster that contains k photos should be mapped to one of these s templates. The included angle between the photo-based importance vector **PV** and a template importance vector $\mathbf{TV}_i$, $i = 1, 2, \ldots, s$, is defined as the metric of template determination:

$$i^* = \arg\min_{i=1,2,\ldots,s} \operatorname{acos}\left(\frac{\mathbf{PV} \cdot \mathbf{TV}_i}{\|\mathbf{PV}\|\|\mathbf{TV}_i\|}\right) \quad (1)$$

where $\mathbf{TV}_i$ is the corresponding template-based importance vector of the template T_i. The minimum included angle between two vectors denotes the best match between photos and templates. Because both importance vectors are sorted in descending order, this process also determines which photo should be put into which cell.

On the other hand, we can construct probability mass functions to describe the distribution of template importance (P) and photo-based importance (Q). From this viewpoint, the symmetric Kullback-Leibler (KL) distance between the photo-based importance distribution and the template importance distribution can also be defined as the metric of template determination:

$$i^* = \arg\min_{i=1,2,\ldots,s} \left(D\left(P_i \| Q\right) + D\left(Q \| P_i\right)\right) \quad (2)$$

where P_i is the distribution of the ith template, and $D(\cdot\|\cdot)$ is the KL distance.

These two approaches both determine templates based on the concept of importance distributions. Because the results based on these two approaches don't pose a significant difference, we mainly use the vector angle approach in our current implementation.

Spatial composition (for challenge 3)

The final task to generate a tiling frame is to put photos into the designated cells. However, the aspect ratio of the targeted cell is often different from that of the original photo. Moreover, the resolution of photos taken by current digital cameras is at least 2 million pixels (about $1{,}600 \times 1{,}200$), which is significantly larger than the targeted resolution (720×480). Therefore, it's unavoidable that we need to resize and/or crop photos to fit in with the template.

So as not to largely distort the content of each photo, we want to find a region that has the same aspect ratio as the targeted cell and possesses the largest attractive content. As we previously noted, we can find the ROI through top-down or bottom-up approaches. In top-down cases, we first find the centroid of the largest face region. Starting from this position, we expand the region in four directions (top, down, left, and right) according to the targeted cell's aspect ratio. The expansion stops when at least two boundaries of this region reach the photo's boundaries. The selected region is then resized to stick on the targeted cell. Likewise, the only difference in bottom-up cases is that we start expansion from the centroid of the salience-based ROI.

Through the processes of cluster selection, template determination, and spatial composition, we can construct photographic paragraphs. After determining the timing for displaying photos and switching frames, the system concatenates photographic paragraphs and generates a tiling slideshow. To facilitate more gorgeous presentations, we also include transition effects such as fade-in and fade-out. We released the executable program of this system and provide some examples at http://www.cmlab.csie.ntu.edu.tw/~wtchu/TilingSlideshow.

Evaluation

We considered both objective and subjective perspectives in evaluating our methods. We used five photo sets taken by different amateur photographers for evaluation, as listed in Table 1 (next page). Two of them were taken while traveling, two were taken in special events such as wedding and graduation ceremonies, and one of them contains only landscapes. Table 1 also shows the

Table 1. Data from the photo set evaluation.

Set	Type	Number of Photos	Descriptions	Length of Music (in Minutes and Seconds)
1	Travel	780	Traveling in Japan—including people, cityscapes, and landscapes.	3:31
2	Travel	522	Traveling in Australia—including people, cityscapes, and landscapes.	4:38
3	Wedding	388	A Chinese wedding ceremony where people are the primary targets.	3:49
4	Graduation ceremony	227	A graduation ceremony where people are the primary targets.	4:38
5	Landscape	133	Pure landscape in Taiwan, including mountains, rivers, and forests.	4:38

Table 2. Clustering performance evaluation.

Slideshow	Number of Frames	Number of Photos	Number of Frames with a Clustering Error	Minimum Photos in a Frame	Maximum Photos in a Frame
1	37	127	1	2	8
2	48	172	1	2	8
3	38	155	0	2	9
4	48	212	0	1	9
5	48	131	8	1	7

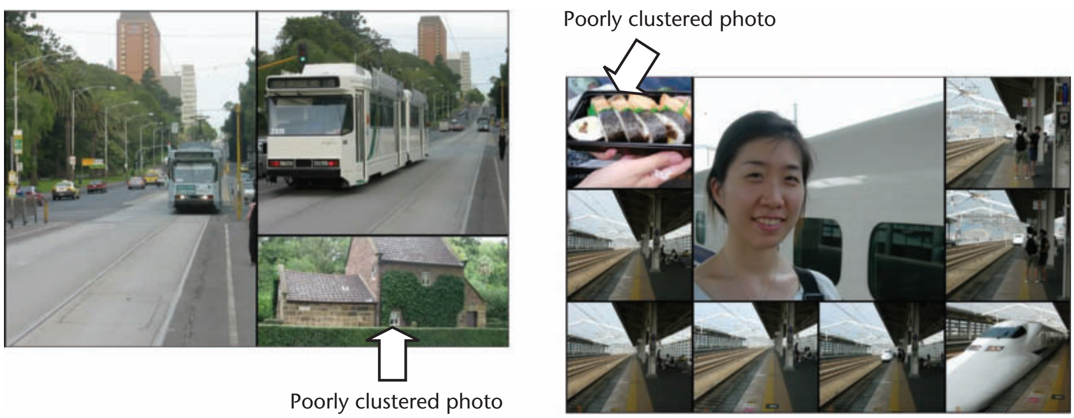

Figure 5. Examples of poorly clustered photos.

final presentation's music length. These music clips are all pop music. We use the same music clip for photo sets 2, 4, and 5 to show that different results would occur from visual diversity.

Clustering performance

In this experiment, we invited the owners of the evaluated photos and asked them to judge whether the photos at the same frame belong to the same event or scenic spot. After hearing the content owners' thoughts, we learned that only a few frames consisted of poorly clustered photos, as we show in Table 2.

The result of the fifth photo set is slightly worse than others because of significantly diverse luminance conditions. In this instance, the content-based clustering method could erroneously separate photos that should be clustered together. Even so, the system often placed poorly clustered photos in supportive cells rather than the conspicuous topic cell. We also noticed that because of the different photographers' picture-taking habits, the average number of photos at the same frame (cluster) varied for different data sets. Figure 5 shows two poorly clustered examples.

Cropping performance

To evaluate the cropping performance, we checked each photo in frames and judged whether it was badly cropped. Table 3 shows the experimental results, where we can see few photos that are badly cropped. We also checked whether badly cropped photos were placed in the topic cell, because bad cropping would provide viewers with a more negative aesthetic experience. Generally, people are sensitive to seeing erroneously cropped human faces. Therefore, there's a higher probability of sensing a cropping error in photo sets 3 and 4, in which people are the primary targets.

Figure 6 shows two examples of badly cropped photos. The reason for bad cropping could be a result of nonrobust face detection, as is the case with the side view face in Figure 6a. Furthermore, significantly different aspect ratios between the targeted cell and the important object could also cause crude cropping to the original photo, as is the case in Figure 6b.

Table 3. Cropping performance evaluation.

Slideshow	Number of Photos	Number of Badly Cropped Photos	Number of Badly Cropped Photos in the Topic Cell
1	127	5	1
2	172	5	0
3	155	8	6
4	212	9	5
5	131	2	0

Figure 6. Examples of badly cropped photos.

User study

For subjective evaluation, we compared users' satisfaction of the slideshows generated by the proposed system to ACDSee (see http://www.acdsystems.com) and Photo Story (see http://www.microsoft.com). ACDSee generates conventional slideshows by sequentially displaying photos one by one. It has no ability to accompany the slideshow with music. On the other hand, Photo Story generates camera motion effects on single photos and sequentially switches photos as well. Photo Story also has the option of providing music to accompany the slideshow.

We invited 27 evaluators to join the user study. We asked them to judge different results according to their subjective perception. Generally, they found the tiling slideshow much more satisfactory than the other options. They felt that the tiling slideshow provided a more impressive presentation and easily helped them experience the content conveyed by the photos. We provide more detailed experimental settings and results elsewhere.[8]

Conclusion and discussion

The evaluators' comments let us know that we have definitely made some strides in creating a better photo presentation slideshow for consumers. However, on the basis of the proposed idea and from the feedback we received, we're investigating various issues and extensions. In the following paragraphs, we detail some of this discussion from different perspectives.

Influence of user intervention

One of the major factors that affected subjective satisfaction is the user's preference. For instance, one person might always prefer to put a specific person's photo on the topic cell. In addition, some photos might be significantly valuable to someone even if they're blurred. Our system didn't take these kinds of subjective preferences or specific considerations into account in the fully automatic process. We're currently working on providing different profiles so that users can select whether to perform photo filtering, select the granularity of photo organization, and control the pace of frame switching. This flexibility somewhat provides a method for users to generate personalized tiling slideshows.

Photo organization

On the basis of photo characteristics, we might want to have different clustering requirements at the organization stage. For the photos that have a strict temporal order, such as for travel, a wedding, and a graduation ceremony, we suggest that time-based clustering should be applied first to find social event boundaries. Content-based clustering could then be applied to each time-based cluster for finer organization. In this way, the final presentation would follow the timeline and reproduce the progress of the targeted photo set. On the other hand, if the targeted data are photos in daily life or are simply landscapes, we could apply content-based clustering first to collect photos with similar appearances.

Impact from high-level concept detection

Some erroneous results, such as badly cropped or poorly clustered cases, derive from the semantic gap between low-level features and high-level concepts. Because we only exploit content-based features in organization and manipulation, main objects or side view faces could be crudely cropped, and photos that should be clustered together in a human sense would be separated. To generate more elaborate results, we can appeal to semantic concept detection or image annotations that are widely used in photo sharing communities.

High-level concept detection also affects the arrangement of a frame. For example, a photo with the sky on the top would be more favorably considered than the one without it. A photo with an apparently vertical structure, such as a high-rise building, should also be placed in a vertical-bar cell rather than horizontal-bar cell. Taking these considerations into account would enhance visual coordination.

Extensions

Our tiling slideshow can be adopted to build many interesting applications. For example, given a traveler's schedule and his or her corresponding photos, the traveler can make a photo-based tour. Similar ideas can also be applied to generate customized auto guidance or electronic lecturing. Moreover, many research issues still exist if we take textual information into account. We can allocate some space for text and coordinate more diverse media (such as text, photos, and music). Finally, although our tiling slideshow is currently presented as a video clip, we can come up with different kinds of visualizations in the future. For instance, we can just output a text-based script that describes the involved photos and the time stamps and locations for presentation. This script could then be read by a multimedia player such as Adobe Flash and then it could be visualized. **MM**

Acknowledgments

This work was partially supported by the National Science Council of the Republic of China under grants NSC 95-2622-E-002-018, NSC 95-2752-E-002-006-PAE, and NSC 9-2221-E-002-332. It was also supported by National Taiwan University under grant 95R0062-AE00-02. We would like to thank Jin-Hau Kuo and Chun-Yi Weng for their valuable discussions, and the anonymous reviewers for their valuable comments.

References

1. J. Geigel and A. Loui, "Using Genetic Algorithms for Album Page Layouts," *IEEE MultiMedia*, vol. 10, no. 4, 2003, pp. 16-27.
2. J.C. Platt, M. Czerwinski, and B.A. Field, "PhotoTOC: Automating Clustering for Browsing Personal Photographs," *Proc. IEEE Pacific Rim Conf. Multimedia*, IEEE Press, 2003, pp. 6-10.
3. J. Kustanowitz and B. Shneiderman, "Hierarchical Layouts for Photo Libraries," *IEEE MultiMedia*, vol. 13, no. 4, 2006, pp. 62-72.
4. A. Vailaya et al., "Automatic Image Orientation Detection," *IEEE Trans. Image Processing*, vol. 11, no. 7, 2002, pp. 746-755.
5. J. Luo and M. Boutell, "Automatic Image Orientation Detection via Confidence-Based Integration of Low-Level and Semantic Cues," *IEEE Trans. Pattern Recognition and Machine Intelligence*, vol. 27, no. 5, 2005, pp. 715-726.
6. Japan Electronic Industry Development Assoc., *Digital Still Camera Image File Format Standard*, standard JEITZ CP-3451, 1998.
7. H. Tong et al., "Blur Detection for Digital Images Using Wavelet Transform," *Proc. IEEE Int'l Conf. Multimedia & Expo*, IEEE Press, 2004, pp. 17-20.
8. J.-C. Chen et al., "Tiling Slideshow," *Proc. ACM Multimedia Conf.*, ACM Press, pp. 25-35, 2006.
9. Y.-F. Ma et al., "A Generic Framework of User Attention Model and Its Application in Video Summarization," *IEEE Trans. Multimedia*, vol. 7, no. 5, 2005, pp. 907-919.
10. E.D. Scheirer, "Tempo and Beat Analysis of Acoustic Musical Signals," *J. Acoustical Soc. of Am.*, vol. 103, no. 1, 1998, pp. 588-601.
11. J.V. White, *Editing by Design: A Guide to Effective Word and Picture Communication for Editors and Designers*, R.R. Bowker, 1982.

Wei-Ta Chu is a postdoctoral research fellow at National Taiwan University (NTU). His research interests include content analysis and multimedia indexing. Chu received a PhD in computer science from NTU in 2006. He was awarded the best technical full paper at the 2006 ACM Multimedia Conference and received the best PhD thesis award from the Institute of Information and Computing Machinery.

Jun-Cheng Chen received an MS in computer science from NTU in 2006. His research interests include video processing and multimedia systems. He was awarded the best technical full paper at the 2006 ACM Multimedia Conference.

Ja-Ling Wu is the head of the Graduate Institute of Networking and Multimedia in NTU. His research interests include digital signal processing, image and video compression, content analysis, digital watermarking, and digital rights management systems. Wu received a PhD in electrical engineering from the Tatung Institute of Technology in Taipei, Taiwan. He received the lifetime distinguished professor award from NTU in October 2006.

Readers may contact the authors at wtchu@cmlab.csie.ntu.edu.tw, pullpull@cmlab.csie.ntu.edu.tw, and wjl@cmlab.csie.ntu.edu.tw, respectively.

For further information on this or any other computing topic, please visit our Digital Library at http://www.computer.org/publications/dlib.