ASAP: Building Multi-Client Collaborated Environments Towards Synchronized Image Sharing

Zhigang Hua[2], Xing Xie[1], Hanqing Lu[2], Wei-Ying Ma[1]

[1]Microsoft Research Asia
5/F Sigma Center, No. 49 Zhichun Road
HaiDian District, 100080 Beijing, China
+86-10-62617711
{xingx, wyma}@microsoft.com

[2]Institute of Automation
Chinese Academy of Sciences
P.O. Box 2878, 100080 Beijing, China
+86-10-62542971
{zghua, luhq}@nlpr.ia.ac.cn

ABSTRACT
Digital images have become increasingly common and popular in mobile communications. However, due to the distribution of these images in various devices which are commonly available to users nowadays, there is a need to develop new technologies to facilitate the sharing of large amounts of photos across various devices. In this paper, we propose a novel approach called ASAP that builds multi-client collaborated environments to achieve synchronized image sharing. In the newly proposed ASAP, a user's interaction with any device automatically leads to a series of synchronized updates in other devices, resulting in a significant reduction of manual interventions from users. In this way, the ASAP presents multiple similar images simultaneously across devices in a way that allows automatic synchronization of the photos based on user interactions. Experimental evaluations indicate that it is effective and useful for image sharing in mobile environment. Finally, we will offer the discussion details on the new concerns and issues proposed by testers to guide our future work for improving the ASAP.

Categories and Subject Descriptors
I.4.9 [**Image Processing and Computer Vision**]: Applications

General Terms
Algorithms, Human Factors

Keywords
Adaptive content delivery, synchronized photo sharing, reflective UI, synchronized UI, mobile browsing experience, synchronized image viewer, synchronized update

1. INTRODUCTION
Mobile handheld devices with embedded digital camera are undergoing a considerable progress because of their portability and mobility currently. Accordingly, users acquire or gain access to an increasingly diverse range of mobile devices. People can easily capture and share personal images on these small-form-factor devices anywhere and anytime. Images play a more and more important role in sharing, expressing and exchanging information in our daily lives.

Users can use different ways to share their digital photos with their friends or relatives. The most common way for people is to transfer the photos captured in mobile devices to their desktop computers and then use the mature desktop applications to share their photos with others. The main problem is that desktop computers are not available for users all time, for instance, when people are on the move. Currently, people would like to upload their photos captured by mobile phones to the photo blogging sites that have been appearing worldwide such as FotoPages [5]. In this way, distant friends or relatives can view their live pictures in real time by connecting to such Web sites. However, the limited bandwidth in current mobile network is a bottleneck to cause the degradation of this sharing mode. Commonly, in our daily life, it is familiar that when friends come together, they would like to share their photos with each other. Usually, they may just let others manipulate their devices to view the photos inside. Furthermore, using the popular wireless connection techniques such as 802.11b [8] or Bluetooth [14], users can let their devices form into a network neighborhood to share their image resources. These sharing methods are mainly focused on offering accesses to let others visit the shared resources. However, when the number of photos and devices to be shared becomes increasingly large, it is a very hard task for users to browse such large amounts of photos across devices.

When photos distributed in various devices are shared together, the need to view multiple photos at one time often arises, such as the following scenarios:

Scenario 1: When friends come together, they wish to share photos in their devices which are taken during a specified time period to share the once memorable moments.

Scenario 2: Users want to find some visually similar photos across devices to compose a printed photobook for a permanent preservation.

Scenario 3: There are many photos captured during a travel in various devices owned by multiple persons. Users want to select some high-quality photos captured in a specified location and recommend them to friends from the large photo collections.

Comparative viewing of multiple photos across devices tends to be a tedious task for users since they must open and manipulate each photo in various devices. A user's scrolling/zooming/clicking interactions with these photos are independent of each other. Due to the small-form-factors of current mobile devices, users' interactions on these mobile devices are greatly constrained. For example, suppose that a group of people bring handheld devices embedded with camera and capture many photos through them during a travel. When a user browses a photo in one of the devices, he may want to view other photos which are visually similar with the photo. To compare all of the photos stored in various devices, the user will shift across various devices to open each photo to examine whether it is visually similar with the specified one.

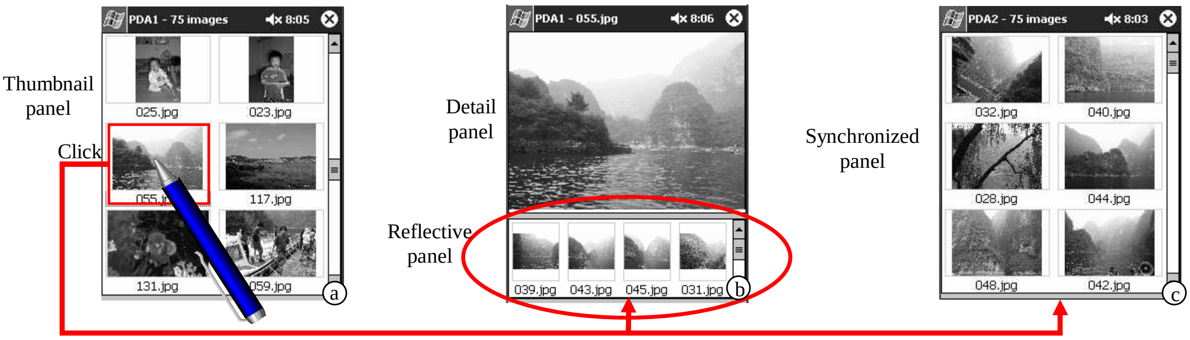


Figure 1. The synchronization between PDA1 and PDA2 when a photo was clicked by users on PDA1.

As the number of photos and devices to be compared rises, the task becomes increasingly tedious since the comparing task has to be manually done by the user. Definitely, the user's experience is greatly degraded by viewing all images one by one on the small displays with continuous shifting across the multiple various devices.

In this paper, we propose A Synchronized Approach for Photo sharing across devices (ASAP) to facilitate photo viewing across multiple devices. In [13], the paper presents a web browser called *Comparative Web Browser*, which simultaneously presents multiple web pages using multiple windows in a way that enables the content of the web pages to be automatically synchronized.

The ASAP can simultaneously present similar photos across multiple devices at the same time for comparatively viewing or searching. A user can interact with either of the available devices, and the user's interaction can automatically generate the synchronized updates in other devices. In this way, synchronized photo viewing across various devices is achieved. Moreover, the ASAP offers a variety of features to maintain the synchronization across devices ranging from low-level image features (e.g. color moment) to higher-level information (e.g. location and time).

In the case of this photo browsing synchronization, consider the aforementioned scenario where a user wants to view photos with visual similarity. In figure 1, there are two PDAs (Personal Digital Assistant) denoted PDA1 and PDA2 respectively, each stores a number of pictures which are captured by their embedded cameras. When a user clicks a specified photo in the *thumbnail panel* in PDA1 (Figure 1a), there are two steps to be automatically done simultaneously: 1) PDA1 itself searches through its local image library to display the visually similar photos on a *reflective panel* (Figure 1b); 2) at the same time, PDA1 sends the image feature to PDA2, and PDA2 will search out their local similar photos to display on a so-called *synchronized panel* (Figure 1c). Users can freely interact with photos belonging to a device by operations like clicking, zooming, scrolling or dragging. Based on users' behaviors, the ASAP automatically generates synchronized updates on the other devices. In this newly proposed approach, traditional photo browsing interface commonly constrained in one device is extended to multiple various devices.

The rest of this paper is organized as follows. In Section 2, related work is presented. Section 3 introduces the system framework of our approach. Section 4 discusses in detail the commonly available image features in mobile devices. Based on these features, we formulate the synchronized photo sharing problem into an image retrieval problem and present the corresponding algorithms in Section 5. In Section 6, we give the experimental results and demonstrate the usefulness of our proposed scheme. We will provide the discussion details on the concerns and issues proposed by testers in the use study in Section 7. Finally, concluding remarks are described in Section 8.

2. RELATED WORK

To let people really enjoy the ease from image sharing where a large amount of images are presented, there is a need to develop efficient methods for users to share them. Current methods are mainly designed to provide convenient interface for users to view the photos in a device.

Due to the increasingly large amounts of image collection, many approaches have been proposed to organize the collection for users to view conveniently. In [2] and [16], methods have been proposed to organize images according to their representative event or subject. Researchers have found that organizing photos by time improves users' performance in retrieval tasks [6][14]. In [19], the *World Wide Media eXchange* (WWMX) database was proposed to organize large collections of image media by location stamp.
When such a great collection of images in local libraries are presented, users need a convenient interface for them to directly navigate to the wanted images. Thus, many searching methods have been proposed to facilitate this direct navigation. They are mainly classified into two trends which are respectively based on low-level image features (e.g. color) and higher-level semantic info (e.g. event).

The visual similarity search function is based on a low-level image feature representation. In [6], an image feature is a high-dimensional vector with components comprising a smoothed color histogram, a color coherence vector, edge direction histogram and texture features. The bottleneck of the low-level feature based approach is its accuracy, which remains a general challenge in current CBIR technologies. Due to the inaccuracy in image search based on low-level image features, many approaches based on higher-level semantic info have been proposed to improve search performance.

Since we often describe photos by the events they document or the time they were taken, [1] proposed to adapt a similarity–based media segmentation algorithm to hierarchically cluster photographs with similar timestamps. Based on analyzing semantic info such as *who*, *when*, *where*, *what*, etc which are available from image headers or from users' manual annotations, the search results can rather consist with users' perception. In [6], it also provides a search function that uses features based on face detection and low-level color, texture and edge features combined with digital camera capture settings to provide high-quality search that is computed at the server but available from all other networked devices accessing the photo collection.

3. OUR SYSTEM FRAMEWORK

3.1 User Interface

We implement a synchronized photo viewer consisting of four panels which work cooperatively to achieve the synchronized photo browsing across devices. They are respectively named *thumbnail panel, detail panel, reflective panel* and *synchronized panel*. The details of the four panels are described as follows.

Thumbnail Panel

Thumbnail panel provides a thumbnail view for photos specified by users in a device, and each thumbnail contains a pointer to the corresponding full-resolution photo (see Figure 1a). Due to the thin computing capacity in mobile devices, thumbnail panel can help users to browse the photos quickly. When users select a thumbnail and click to view its detail, the window will be navigated to a detail panel where the detail view of the image is shown.

Detail Panel

Just as its name implies, the *detail panel* provides a detail view for images when users click (see Figure 1b). In this panel, users can freely zoom and scroll through a photo to view the details of each part within an image. On desktop computers, the detail view of an image can be quickly loaded from storage to display on the screen. While on current mobile devices, the time to load a detail image for display can not be ignored.

Reflective Panel

In this paper, we use *reflective panel* to aid user's photo browsing. Reflective panel locates at the bottom of the window and displays the local images which are similar with the currently browsed photo (see Figure 1b). Suppose that a user wishes to find a similar photo with the currently browsed photo in the local image library, it can be seen in the reflective panel. When a user operates an interaction which specifies a photo or a part of a photo to view, the similar photos in local image library will be automatically searched out to display on the reflective panel.

Synchronized Panel

When a user's interaction in the active device was transformed into the synchronized updates in other synchronized devices, the similar photos are automatically searched out by the ASAP to display on the *synchronized panel* (see Figure 1c). In a synchronized panel, users can view the similar images without redundant interactions in these devices which are inevitable in traditional ways.

3.2 Basic Concepts of the ASAP

In the ASAP, the interactions with any device in the group automatically generate the synchronized updates on other devices. Several notations in the ASAP are introduced as follows:

Active and Synchronized Devices

We assume that a user wishes to browse photos across multiple available devices. The user interacts with one device at a time, which we call the *active device*. The interaction in the active device will automatically lead to synchronized updates on other devices. We call the other devices the *synchronized device*s.

Synchronized Group

When various devices begin to keep synchronization with each other, they form into a *synchronized group*. Each device in a synchronized group serves as an equivalent peer. This means that user's interactions with any device will be transformed to other devices to maintain the browsing synchronization.

Synchronization Features

The *synchronization features* are transformed from users' interactions with the active device and are then automatically delivered to other devices by the ASAP. When the features reach other devices, they are used to search for similar photos from local image libraries.

Image-Level and Region-Level Synchronization

In this paper, we provide a hierarchical synchronization scheme including two-level synchronization according to the operated object of a user's interactions. When a users' behavior is handled on a specified photo, the image-level synchronization is employed. Otherwise, when a user's behavior is manipulated on a part within a photo, the region-level synchronization is used instead.

3.3 Implicit Query Based on User Interactions

Because synchronized updates across devices in the ASAP are implicitly indicated by users' interactions with devices, we call this *implicit query*. Figure 1 and Figure 2 show the examples illustrating the concept of synchronized photo browsing across two PDAs.

Thumbnail Clicking

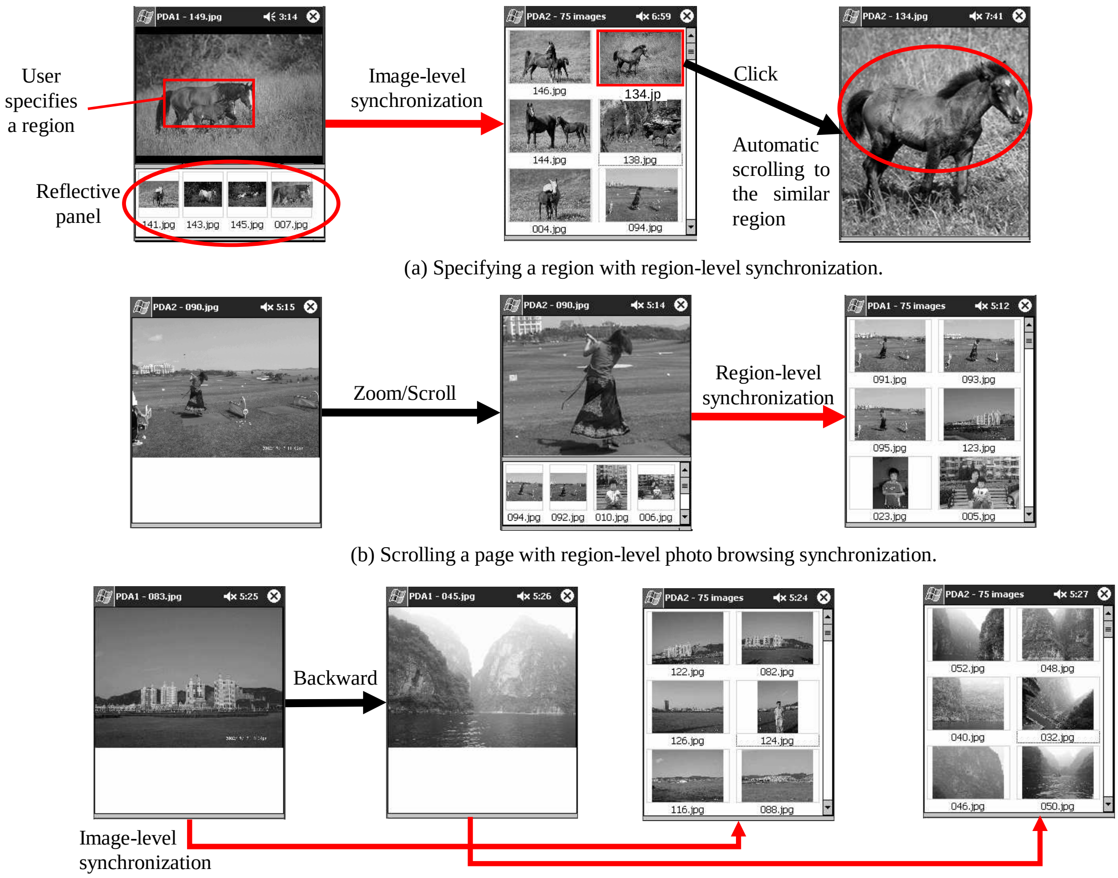
Figure 1 shows two PDAs, midst them PDA1 works as an active device and PDA2 acts as a synchronized device. When users click a photo in PDA1, the most similar photos on the synchronized device PDA2 are automatically shown on the synchronized panel, and the top most similar photos are shown in the reflective panel in PDA1. This transformation is *image-level synchronization*.

Region Specifying

Users can freely use a pen in PDA1 to drag a region for highlighting (e.g. Figure 2a). Our ASAP is capable of *region-level synchronization*. That is, for the photo region that is specified by users, our system will automatically find photos containing similar regions with the user-specified region across devices. The search results are shown on the reflective panel in PDA1 and on the synchronized panel in PDA2. Furthermore, when users click such photos searched, the ASAP will automatically scroll through that photo so that the similar regions can directly appear before users.

Zooming/Scrolling a Photo

When browsing through a detail view of a large photo on the small-screen mobile devices, users usually zoom in the photo to scroll up or down to view each part of interests. In the ASAP, for the photo part that currently appears in the detail panel of the active device when users are scrolling, the ASAP will automatically search for the most similar parts across devices (e.g. Figure 2b).



(a) Specifying a region with region-level synchronization.

(b) Scrolling a page with region-level photo browsing synchronization.

(c) Forward and backward navigation with image-level synchronization.

Figure 2. Examples for the synchronization across two PDAs in the ASAP.

When the photos that contain the most similar part are searched out, they are shown on the reflective panel in PDA2 and on the synchronized panel in PDA1. Similarly, when users click such photos, the ASAP automatically scrolls through that photo to display the similar regions.

Forward and Backward Navigation

When users navigate forward or backward in the active device, the ASAP displays the next (or previous) photo of the users' viewing history, and also shows the similar photos on the synchronized panel in the synchronized devices (e.g. Figure 2c).

3.4 System Flow

The ASAP identifies synchronization features of the photo that is currently browsed based on the transformation of user's behaviors, and then automatically transfer the features to other synchronized devices for them to search for similar photos. When a user interacts with an active device's photo, the ASAP will generate automatic synchronized updates in other devices concurrently. A complete system flow of our approach is shown in Figure 3. The ASAP system overview is as follows:

1. As shown in Figure 3a, our system automatically segments the input photos (we adopt the JSEG algorithm [4]) to partition each photo into a series of regions. Based on the attention model proposed in [12], these segmented regions are then grouped to generate a set of larger regions. The low-level features of these grouped regions are then extracted. Together the semantic features like time and location inserted in image headers which are available for current imaging devices, they are saved to a local image feature database.
2. The user selects one as an active device from multiple available devices, and begins to specify one or several folders that contain images in the device. The ASAP automatically shows the thumbnails for photos included by the selected folders on the *thumbnail panel*.
3. The user scrolls through the detail panel and selects to click a thumbnail to browse its detail view.
4. The ASAP automatically fetches the corresponding feature of the specified photo from its local image feature database. Then it multicasts the fetched feature to other synchronized devices for them to search for similar photos and display the search results on their synchronized panels. At the same time, it also searches for similar photos and shows the top most similar photos on the reflective panel.
5. On the detail panel, the user zooms/scrolls to navigate or manually drags to specify a region within the photo.
6. The ASAP identifies the area that currently displays in the detail panel, and it fetches the features implicitly indicated by

the specified area and multicast them to the synchronized devices for them to search for the photos containing similar regions. At the same time, the active device searches for the similar photos concurrently. Furthermore, when users click these photos searched by the ASAP, they will be automatically scrolled through the similar regions.

7. Repeat steps (2) to (6) for photos to be browsed.

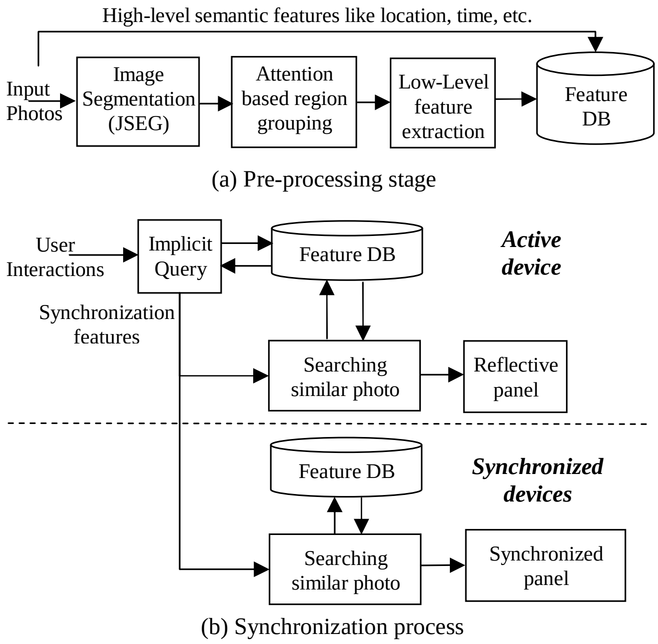


Figure 3. The system flow of the ASAP.

4. IMAGE FEATURES TO MAINTAIN SYNCHRONIZATION ACROSS DEVICES

We classify the available features in current imaging devices into two types, namely low-level features and semantic features.

4.1 Low-Level Image Features

Surely, the visual similarity search function is based on a low level image feature representation. In our approach, images in a database are first segmented into series of regions. The segmentation method we use is the so-called JSEG algorithm [4]. And these segmented regions are then clustered into a small number of groups based on the attention model proposed in [12].

We use three properties to describe a region within a digital photo. They are respectively the low-level feature extracted from the region, its importance weight and the position of its centroid. The former two properties are used in the similarity measure between two images, and the last property is used for the generation of automatic scrolling through the region within a photo. For the first property, many kinds of low-level features may be used. In the current implementation, we use color moment, which is shown to be robust and effective [17]. We extract the first two moments from each channel of CIE-LUV color space. For the second property, we currently use the percentage of region area. The only requirement is that the sum of importance weights for an image should be equal to 1. For the last property, we get it by averaging the coordinates of all of the included pixels within a region.

Definition 1: The low-level feature of a digital photo is defined as the collection of its included segmented regions:

$$Rgn_i = \{(F_i, W_i, Cen_i)\}(1 \le i \le N) \qquad (1)$$

where

Rgn_i, the i^{th} region within the photo
F_i, the low-level feature of Rgn_i
W_i, the importance weight of Rgn_i
Cen_i, the centroid position of Rgn_i
N, the number of the regions within the photo

and, the sum of *weight*s satisfies: $\sum_{i=1}^{N} W_i = 1.0$.

4.2 Semantic Image Features

Most digital cameras follow the PIMA/DIG/IIIA EXIF standard. This standard defines the format of header fields in JPEG images saved on to storage cards in digital cameras. It allows for the insertion of various types of metadata by the camera processor, such as basic image parameters (e.g. height, width and compression), location/time of capture, and capture settings (e.g. flash usage, focus distance and exposure time).

Studies show that users associate their personal photos with event, location, subject, and time [10]. Among them, event are frequently tied to time and location, subject is often defined by combinations of who, what, when, and where. According to the status of the current imagining devices, we choose image *location* and *time* as the metadata to identify the semantic features of a digital photo. The geographic location where the photographs were taken provides critical context that transcends language, culture, and user-dependent taxonomies [19]. For example, knowing that a photo was shot at a spot says a lot about the photo even before a single pixel is viewed. The time when the pictures were taken is helpful for locating particular photos or events, because it is usually easier to remember and associate when an event occurred relative to time [14].

Currently, the majority of the images would be accompanied by time stamps which indicate the time being captured. In [19], the paper provides detailed ways on how to acquire location info for digital photos. Mobile phones, for example, are beginning to use a combination of GPS and cell-tower triangulation to determine their location represented by longitude/latitude, and this location info can be inserted into the headers of images captured by their embedded cameras.

4.3 Annotation Features

In general, annotation methods have been applied in CBIR to extend the available features to semantic texts or other types of information. For example, Search engines such as Google's Image Search make good use of surrounding keywords of images within Web pages as the annotation features, and it search images via keywords instead of low-level features [7].

The technically simplest solution for acquiring feature metadata from digital photos is to let users manually annotate them via a convenient UI. In our case, by default, clients use plain text as the annotation, which offers an interface allowing the texts to be linked with the wanted semantic features such as *who, where* (BTW: appending the time stamp available in imaging devices, a subject an image indicates can be determined). Users can add feature annotation to past photos, or correct incorrect annotations. The clear disadvantage is the tedious labor required.

5. SYNCHRONIZED BROWSING OF PHOTOS ACROSS DEVICES

The hierarchical two-level synchronization are discussed in this section, namely image-level and region-level synchronization.

5.1 Image-Level Synchronization

As for image-level synchronization, since various users may have different demand for the synchronization results, we provide an option interface for users to select the wanted feature. For example, if a user wishes to find photos with similar time to the currently browsed photo, he can select the *time* feature to maintain the synchronization. According to the category of synchronization features specified by users, there is a need to develop different similarity measures to evaluate the similarity distance between two photos according to various types of used features.

Case 1: Similarity Measure for Low-Level Image Features

As for low level features, in our approach, the distance between two images is measured using the Earth Mover's Distance (EMD) which has a rigorous probabilistic interpretation [15]. EMD is suitable for region-based image similarity measure for it matches perceptual similarity well and can operate on variable-length representations of the distributions. In this case, the ground distance is an equally weighted Euclidean distance between the features of two regions, namely color moments currently used in the ASAP. EMD incorporates the properties of all the segmented regions so that information about an image can be fully utilized. EMD is robust to inaccurate segmentation by allowing many-to-many relationship of the regions to be valid.

Case 2: Similarity Measure for Semantic Image Features

If the *time* feature is selected to maintain the synchronization for a photo, it is quite simple to measure the similarity between two photos in various devices. For two given photos, P_0 and P_1 stored in an active device and a synchronized device, respectively with a time stamp t_0 and t_1, the similarity can be described as follows:

$$Sim_t(P_0, P_1) = \|(t_0 - t_1) - O_{01}\| \qquad (2)$$

The variable O_{01} represent time offset of the active device with respect to the synchronized device's clock. This time offset info can be acquired when the synchronization across devices is initialized.

When *location* info represented by longitude/latitude is used to be the synchronization feature, the longitude/latitude of a photo is mapped to an index with a single 8-byte in [19]. Using this identified index, the similarity distance in the *location* feature between two photos can be evaluated. As for other higher-level semantic features such as annotation texts which probably identify a specified aspect of a photo such as person name, the similarity on this text-level can be referred to the current popular text-centric search algorithms.

5.2 Region-Level Synchronization

As for region-level synchronization, which occurs when a user zoom/scrolls or specifies a region within a photo, currently our approach offers only low level features of the selected regions to maintain this level synchronization. Unlike the image-level synchronization, there is a need to search similar regions instead of whole photos to achieve region-level synchronization across devices. The summary of the region-level synchronization can be described as follows:

1. Get the regions included by the specified area. We achieve this by examining whether the size overlapped by the specified area and a region exceeds an assigned threshold λ (e.g. 100 pixels).
2. The low-level features of these included regions (gained in the step 1) are then fetched from local feature database as the synchronization features. They are multicast to synchronized devices to search for photos containing similar regions, and the active device itself also searches concurrently.

Similarity Measure

More detailedly, assume that there are n regions $R_{01}, R_{02}\ldots R_{0n}$ being selected within the specified area in a photo P_0 in the active device. For the sake of simplicity, we take a photo P_1 in a device as an example to measure its similarity with the selected regions within P_0. Suppose P_1 contains m regions, which are respectively labeled $R_{11}, R_{12\ldots}R_{1m}$. Here, the distance between two regions is measured using the Euclidean distance.

Definition 2: The similar region pair between photos P_0 and P_1 is defined as a set of region pairs and their Euclidean distance:

$$SimRgnSet_k = \{(R_{0n_k}, R_{1m_k}, Dist_k)\} \quad (1 \le k \le S) \qquad (3)$$

where

$SimRgnSet_k$,	the k^{th} similar region pair within the set
R_{0n_k},	the region belonging to P_0
R_{1m_k},	the region belonging to P_1
$Dist_k$,	Euclidean distance between R_{0n_k} and R_{1m_k}
S,	the number of the members in the set ($S \le n$)

Here is the step-by-step description of the algorithm to find similar region pairs between $R_{01}, R_{02}\ldots R_{0n}$ and $R_{11}, R_{12\ldots}R_{1m}$:

1. For each region $R_{01}, R_{02\ldots}R_{0n}$ within P_0, do the step 2-3.
2. Calculate its Euclidean distance with each region within P_1 such as $R_{11}, R_{12\ldots}R_{1m}$. Select the region within P_1 with the least distance, and let it be labeled R_{1t} ($1 \le t \le m$).
3. Examine whether the distance is under a assigned threshold γ. If the distance satisfies the condition, the corresponding region R_{1t} is inserted into a similar region pair set denoted ζ (initialized as null).

By this way, ζ is gradually formed. Suppose ζ finally consists of n' members ($0 \le n' \le n$). Depending on whether n' is larger than 0, the final similarity distance between two photos is various:

- If n' is equal to 0, P_1 is thought to have no matched region with P_0. The similarity distance is assumed to be $\propto$.
- If n' is larger than 0, The overall similarity distance of P_1 compared with the selected regions of P_0 is defined like this:

$$Sim_r(P_0, P_1) = \frac{\sum_{i=1}^{n'} Dist_i \times R_{1i}.W_i}{n'} \qquad (4)$$

Correspondingly, for other photos to be compared, their similarities with P_0 can be measured like P_1. According to the similarities measured by the above method, the search results in each device are sorted in a priority order. They will be displayed on the synchronized panels in synchronized devices and on the reflective panel in active device.

Region-Level Synchronization Results

As for the region-level synchronization, we provide a further method to facilitate photo viewing on the search results. That is, when users specify one of photos midst the search results to click, the window will be automatically scrolled to navigate to the matched regions. Specially, when the number of the matched regions is more than 1, we adopt an automatic scrolling approach to automatically scroll through these similar regions (this can be implemented by using the centroid property of each region, i.e. *Cen* in Definition 1).

The generation of automatic scrolling through similar regions is referred to the method of the generation of automatic browsing path proposed in [11]. In this case, the user's click into these photos results in a smooth scrolling through the similar regions included in ζ with each being displayed serially for a brief period of time. This mode can facilitate synchronized photo browsing, because users can view the interested regions which can be automatically scrolled through without excessive user interactions such as zooming/scrolling.

6. EVALUATION RESULTS

As shown in Figure 4, our evaluations were conducted with two PDAs (Compaq iPAQ 3670 Pocket PC) with 64M memory, a resolution of 320x240 and PocketPC 2002 as the operating system. Both of the PDAs are equipped with an 802.11b wireless network card to connect with each other. The two PDAs respectively stores 75 images ranging from a variety of types including family, indoor, outdoor, group. They are mainly collected from personal albums offered by the colleagues. We asked ten computer science undergraduate students to participate with our use study, including six males and four females. They are familiar with the operations in mobile devices like PDAs, and they never have any knowledge of the ASAP before.

The ten subjects were firstly asked to view an easy instruction on how to use the ASAP. Afterwards, the common interactions in the ASAP were each conducted five times for each user over a course of fifteen minutes: specify a photo in thumbnail panel to click, drag to specify a region within a photo, zoom/scroll through a photo, and click into the synchronized photos to experience the automatic scrolling through similar regions for these photos. By the assigned exercise, the users mastered the operations used in the ASAP.

Figure 4. Two Pocket PCs connected by wireless 802.11b.

6.1 Task-Based Evaluations

After the users got familiar with the interactions in the ASAP, we asked them to do evaluations which accurately reflected the common use of the ASAP. They were asked to finish three tasks to evaluate the performance of our approach as follows:

Task 1: *Select 25 photos randomly from the photo collections on the two PDAs, and use the three features like the low-level feature, time and location to examine the synchronization results in the ASAP. Rate the three features in a preference ordery.*

Task 2: *Select 6 random photos from the photo collections, and use either zooming/scrolling or dragging to specify regions within these photos to examine the region-level synchronization results.*

Task 3: *Find related photos on the two PDAs for five assigned pictures, and the number of search results for each assigned picture should not be less than 5.*

In the task-based evaluations, the subjects counted whether the synchronization results were relevant with regard to the photo specified by them. The definition of relevancy was left to the user, but two scenarios were recommended to the testers:

- Whether the found images would be reasonable replacements for the original if it were lost or deleted?
- Whether they would group the found photos together with the original to compose a themed subject?

Task 1 Observation

As for the evaluation results on rating the synchronization features, see Table 1 above. 70% of users rate the synchronization features generally as this: *Location*, *Low-Level*, *Time*. Totally, 80% of the participants rate *location* as the prior selection to be the synchronization feature, and the remaining 20% tend to use low-level image feature as the prior synchronization feature.

Table 1. The rating results of the synchronization features.

Synchronization features rating	Percentage
Location, Low-Level, Time	70%
Low-Level, Location, Time	20%
Location, Time, Low-Level	10%

The testers mostly think that the synchronization results specified by the *location* feature consist most with their perception. The low-level features was mainly rated as the second, while two users rate it as the prior and only one user rate it as the last. This indicates that the synchronization based on the analysis of low-level features is accepted by users, but it needs technical improvements, which is a common challenge for the current Content-Based Image Retrieval (CBIR) techniques. There are no participants to select *time* as the prior synchronization feature. The main reason is due to that different devices may capture irrelated photos at one time.

Task 2 Observation

Each evaluation is put into one of the following three categories: perfect, good and bad. "*Perfect*" means that the synchronization result is perfectly consistent with users' perception. "*Good*" means that the synchronization result is correct but there are minor errors and those errors do not affect the overall results. "*Bad*" means that the synchronization result contradicts with user's perception.

Figure 5 shows the distribution of our performance evaluation of image-level and region-level synchronization based on the three categories. More than 60% of the results are assumed to be prefect. And about 88% of the results are either perfect or good. In both evaluations, the "bad" category is mainly due to the synchronization results inconsistent with user's visual perception. The results indicate that the region-level synchronization in the ASAP is commonly accepted by users.

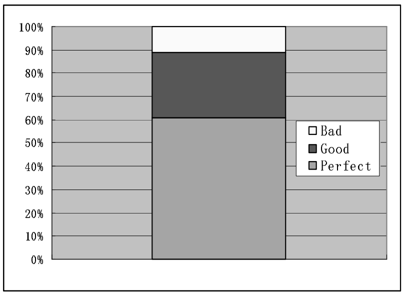


Figure 5. The evaluation results of region-level synchronization.

Task 3 Observation

In this task, all of the ten testers were halved into two groups. Midst the two groups, one group was asked to use our ASAP approach, and the other group was requested to use the traditional photo sharing approach. Each member in the two groups performed this task on the two PDAs.

Figure 6 shows the results of the average search time for each assigned photo. We can see there is a significant reduction of search time using our approach compared to the traditional sharing method. We can see that for each photo, we can achieve about 5~8 fold time reduction. Specially, we record the features being selected to maintain the synchronization for the group using our synchronized approach. In this evaluation, we set the low-level feature as the default feature to be used. If users feel unsatisfied with the low-level synchronization results, they can freely select other semantic features to maintain the synchronization. Totally, 60% satisfy the synchronization results of the default low-level features initially assigned, and 32% continue using *location* to adjust the synchronization, and the remainder 8% users select *time* as the synchronization features. From this test, it further supports that users tend to use *location* as a more accurate synchronization feature.

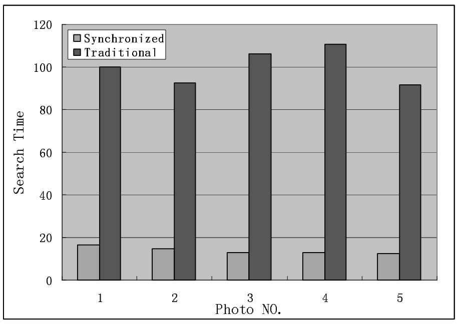


Figure 6. The search time for the five assigned photos.

6.2 Overall Evaluation

After users finished the assigned tasks above, we ask users to evaluate the effectiveness and ease of use in the ASAP with *YES* or *NO*. The results are encouraging that each user indicated that the ASAP is effective for them to browse photos across multiple devices. Each of them thinks that our approach is an interesting and useful work because they commonly have a need to search or compare related photos across devices.

The users claimed that when they are staying with their friends or families during traveling or get-together, they would like to share their photos together synchronously using our proposed ASAP. The users indicated its perceived usefulness in outdoor situations where one can't be in front of his or her computer, for they commonly transfer the images captured in imaging devices to their desktop computers for storing or processing. On the other hand, users seemed to think it useful in "quickly glimpsing over photos across devices", for they think our approach can save their interactions with various devices.

7. DISCUSSIONS

In the end, we handed out a survey to each of the participants. We tried to learn user's feedback to guide our future work to improve the ASAP. The survey involved two problems as follows:

1) What concerns you have toward the ASAP?
2) Due to the obvious privacy issues in photo sharing, what means you would like to make out to protect your privacy?

7.1 User Concerns

Users almost seemed to raise a common concern that the ASAP needs to provide more features to maintain synchronization across devices. They thought features should be extended beyond *the low-level image feature*, *time* and *location* that are currently offered in the ASAP. For instance, they proposed the name of a person within photos, the subject or event of photos, the flower breed in a flower photo and so on. Actually, these proposed semantic features are beyond the current techniques to be automatically extracted from images. After we explained the technical difficulties to them, they suggested that we could provide an annotation interface for them to identify their wanted features to maintain synchronization across devices. They also argued that in some cases several types of features need to be combined together to maintain the synchronization. For instance, they illustrated a scenario where they want to browse visually similar photos with a specified location. In this case, they propose to combine the low-level image feature and semantic feature *location* together to build the synchronization feature.

Another concern proposed by users is that the delay caused by searching for synchronization results across various devices. Due to the thin computing capacity in current mobile devices, it's challenging to quickly search out the synchronization results from a large collection of photos. We are considering using two plausible strategies to solve this matter in our future work. One is proposed in [3], where the summary of features for all images in each device will be multicast to other devices when connection across devices is initialized. In this case, the target devices which contain the wanted resources can be quickly confirmed. The other is based on the history records, that is, the active device will cache the synchronization results. When the photo needs to be synchronized again, the device will fetch the cached synchronization results and transform them to the target devices to prevent them searching renewedly through a large photo collection.

7.2 How to Protect Privacy?

Privacy is the key problem users are concerned when they share resources with others. Almost each of them argued that the shared photos need privacy protection. The users tended to authorize various access rights to different users when they share their photos. For instance, they wish to share some specified picture with their buddies, while others can not access such pictures.

A rather creative suggestion from a user is that users can freely set photo sharing level for different users such as the *public*, *protected* and *private*. He claimed that *public* means that everyone can access even without authentication, while *protected* needs authentication to access and *private* can not be accessed by anyone else. In this case, an active device can only access public photos in synchronized devices if they haven't been authorized. There are also two users claiming that they should have an access to specify whether some of his interactions should be transformed to the synchronized updates to other users when his device acts as an active device, because he felt his interactions may be traced or monitored by others.

We leave questions relating to privacy for image sharing as open problems for future research.

8. CONCLUDING REMARKS

Currently, the predominant methods for sharing photos across devices are simply letting each other to access the photo resources using wired or wireless connection techniques. The independence of user's behaviors in various devices results in tedious browsing experience due to the excessive user interactions with frequent shifts across devices.

In this paper, we have described a synchronized approach for photo sharing across multiple devices (ASAP). The ASAP simultaneously presents similar photos across multiple devices in a way that allows automatic synchronization of the photos based on user interactions. In the ASAP, a users' interaction in a device is transformed into synchronized updates on other devices to achieve the synchronized photo sharing. The ASAP provides a hierarchical synchronization scheme including image-level and region-level synchronization. Experimental evaluations indicated that our approach has significantly improved users' browsing experiences across multiple devices. From the users' feedback, we will continue to improve the ASAP in future work.

9. REFERENCES

[1] J. Adcock, M. D. Cooper, J. Doherty, J. Foote and A. Girgensohn. Lynn Wilcox: Managing digital memories with the FXPAL photo application. ACM Multimedia 2003 Conference, Nov, 2003, Berkeley, California, USA.

[2] M. Balabanovie, L.L. Chu, G.J. Wolff. Storytelling with digital photographs. CHI 2000, April, 2000, Amsterdam, Netherlands.

[3] W. Chang, G. Sheikholeslami, J. Wang, and A. Zhang. Data Resource Selection in Distributed Visual Information Systems. IEEE Transactions on Knowledge and Data Engineering, Volume 10, Number 6, February 1998.

[4] Y. Deng, B. S. Manjunath, and H. Shin. "Color Image Segmentation", in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR '99, Fort Collins, CO, vol.2, pp.446-51, June 1999.

[5] FotoPages. http://www.fotopages.com

[6] U. Gargi, Y. Deng, and D.R. Tretter. Managing and Searching Personal Photo Collections. Proc. SPIE Storage and Retrieval for Media Databases, pp. 13-21, January, 2003.

[7] Google Image Search. http://images.google.com.

[8] IEEE 802.11, the Working Group for Wireless LANs Standards: http://grouper.ieee.org/groups/802/11/

[9] F. Jing, M.J. Li, H.J. Zhang and B. Zhang. An effective region-based image retrieval framework. ACM Multimedia 2002 Conference, Dec. 2002, Nice, French.

[10] H. Lieberman, H. Liu. Adaptive Linking between Text and Photos Using Common Sense Reasoning. In *AdaptiveHypermedia and Adaptive Web-Based Systems*, P. deBra, P. Brusilovsky, R. Conejo (eds.), Springer-Verlag, Berlin, pp. 2-11, 2002.

[11] H. Liu, X. Xie, W.Y. Ma and H.J. Zhang. Automatic browsing of large pictures on mobile devices. ACM Multimedia 2003 Conference, Nov. 2003, Berkeley, California, USA.

[12] Y.F. Ma and H.J. Zhang. Contrast-based Image Attention Analysis by Using Fuzzy Growing. ACM Multimedia 2003 Conference, Nov, 2003, Berkeley, California, USA.

[13] A. Nadamoto and K. Tanaka. A comparative web browser (CWB) for browsing and comparing web pages. WWW 2003 Conference, May 2003, Budapest, Hungary.

[14] K. Rodden and Kenneth R. Wood. How Do People Manage Their Digital Photographs? CHI 2003, April, 2003, Florida, USA.

[15] Y. Rubner, C. Tomasi, and L. Guibas. "A Metric for Distributions with Applications to Image Databases." Proceedings of the 1998 IEEE International Conference on Computer Vision, January 1998.

[16] C. Shen, N.B. Lesh, F. Vernier, C. Forlines and J. Frost. Sharing and Building Digital Group Histories. CSCW'02, November, 2002, New Orleans, Louisiana, USA.

[17] M. Stricker, and M. Orengo. "Similarity of Color Images", in Storage and Retrieval for Image and Video Databases, Proc. SPIE 2420, pp 381-392, 1995.

[18] The official Bluetooth wireless info site: http://www.bluetooth.com/

[19] K. Toyama, R. Logan, A. Roseway and P. Anandan. Geographic Location Tags on Digital Images. ACM Multimedia 2003 Conference, Nov. 2003, Berkeley, California, USA.