

# Automatic Synchronized Browsing of Images across Multiple Devices

Zhigang Hua<sup>1\*</sup>, Xing Xie<sup>2</sup>, Wei-Ying Ma<sup>2</sup>, Hanqing Lu<sup>1</sup>

<sup>1</sup>National Laboratory of Pattern Recognition, Institute of Automation  
Chinese Academy of Sciences, Beijing, 100080, P.R. China  
{zghua, luhq}@nlpr.ia.ac.cn

<sup>2</sup>Microsoft Research Asia, 5/F Sigma Center  
No. 49, ZhiChun Road, Beijing, 100080, P.R. China  
{xingx, wyma}@microsoft.com

**Abstract.** Mobile devices are undergoing considerable progress during recent years. Using these portable devices, people can easily capture and share photos even when they are on the move. Doubtlessly, browsing a large number of images on such small-form-factor devices is still hard and time-consuming, especially when they are distributed across various devices. In this paper, we propose a novel synchronized approach to facilitate image browsing across multiple devices. In this approach, similar images across multiple devices can be simultaneously presented for users to make comparatively viewing or searching. Experimental results show that the synchronized approach is beneficial to improve users' browsing experience.

## 1 Introduction

Mobile devices are undergoing considerable progress in both hardware and software development during recent years. Using these portable devices, people can easily capture and share images even when they are on the move. For example, a recent popular trend is moblogging (mobile weblogging). It is through the use of a phone or other mobile devices for users to publish and share their resources on the Web in real time, whether that resource is text, image or other media. However, to make people really enjoy the ease of mobile communications, many hurdles still need to be crossed [4]. Among them, major crucial challenges include the limited accessing bandwidth and display sizes. While the bandwidth condition is expected to be greatly improved within the following years, however, in the foreseeable future, the display, i.e. the form factor, will continue to be the major constraint on small mobile devices.

---

\* This work was performed while the author was working as a visiting student at Microsoft Research Asia.

To deal with the display constraint, attention model [1][3] have been proposed recently to facilitate image browsing in small devices. In [1], it allows the delivery of more important regions instead of whole images to clients when their screen sizes are small. In [3], an image is decomposed into a set of spatial-temporal information elements which are displayed serially, each for a brief period of time. Although these approaches proved to be effective for browsing large images on devices with small displays, searching or comparing a large number of images across multiple mobile devices is still a hard task.

In this paper, we propose a synchronized approach to facilitate image browsing across multiple devices. In our approach, similar images on various devices can be simultaneously presented for comparatively viewing or searching. The ability to view multiple similar images across devices at one time is useful, such as when users want to make a comparison or search through image collections from various devices. In the novel approach, traditional image browsing interface constrained to one single device is extended to multiple devices. The rest of this paper is organized as follows. Section 2 introduces the system framework of our approach. Section 3 discusses in detail the communication protocol that is adopted in our synchronized approach. In Section 4, the image matching algorithm is presented. In Section 5, we give the experimental results in our use study. Finally, concluding remarks are described in Section 6.

## 2 Our System Framework

The details of our system framework will be described in this section, they are, user interface we adopt in use, the synchronization across devices implicitly indicated by implicit query of attention objects, and the system flow of our framework.

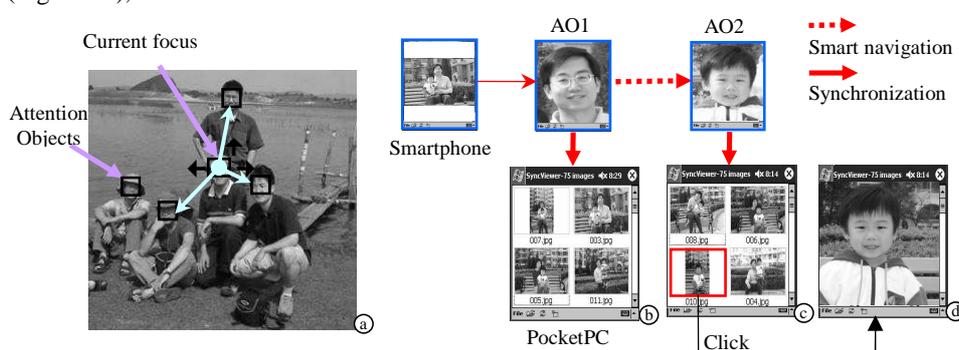
### 2.1 User Interface

An image attention model [1] is generated to reveal the informative regions in images as:

$$\{AO_i\} = \{(ROI_i, AV_i, MPS_i, MPT_i)\}, \quad 1 \leq i \leq N \quad (1)$$

In addition to all previous automatic image browsing approaches making use of the image attention model [1][3], however, in this paper we provide a smart browsing mode named “smart navigation” to implement the image browsing. In this mode, pressing the directional buttons in mobile devices will result in a smooth scrolling to the closest attention object in the given direction. Figure 1a shows an example for this mode where the current focused attention object can be navigated to three closest objects by pressing left/right/up buttons. By using such a smart navigation, it can facilitate image browsing on small-form-factor devices with a substantive reduction of interactions.

In our approach, synchronized image browsing across devices is proposed according to the attention object that displays on the screen, which is assumed to represent a user's interest in the image. That's to say, the attention object that currently displays is used as an implicit query to maintain synchronization on other devices, by searching through their image libraries to find out images containing similar attention objects. Furthermore, a function is offered by us to facilitate the viewing of these search images, by automatically zooming into the matched objects with an appropriate display area when users click them (Figure 1d), which is to be described in details later.



**Fig. 1.** (a) Smart navigation. (b) Synchronization results of AO1. (c) Synchronization results of AO2. (d) Automatic zooming into the matched attention object.

## 2.2 Synchronization Based on Implicit Query

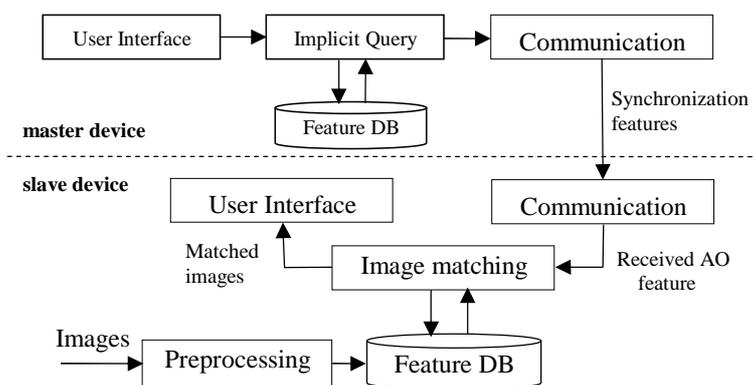
Since the synchronization across devices is implicitly indicated by the attention object that displays on the screen, we call this *implicit query* compared with *explicit query* where query task is directly specified by users such as web search engines. A synchronized process across devices can be summarized like this (in Figure 1, b and c respectively shows the synchronization results of two attention objects denoted AO1 and AO2):

1. A user selects one as a master device from multiple available devices, and begins to specify an image to browse through its attention objects using the smart browsing mode.
2. The feature of the current attention object is automatically extracted, and it is then delivered to the slave devices.
3. When a slave device receives such data containing feature of an attention object, it automatically starts to search through local image library to find out images containing similar features with the received one. The search images are sorted in an order according to the similarity and are then displayed on the screen.

4. If the user clicks an image in the search results on a slave device, the window automatically zooms into the matched attention object with appropriate display area.
5. Repeat steps 1 to 4 for images to be browsed across devices.

### 2.3 System Flow

In our synchronized approach, the feature of an attention object that currently displays on the master device is identified first, and is then automatically transferred to the slave devices for them to search for similar photos. When a user interacts with a master device's image using the smart browsing mode, the synchronized updates on other devices will be generated concurrently to display similar images. A complete system flow of our approach is shown in Figure 2.



**Fig. 2.** The framework of synchronized image browsing.

As shown in Figure 2, the framework mainly consists of three function modules:

1. The communication module is responsible for maintaining the synchronization relationship across various devices, that is, delivering feature data from a master device to its slave devices.
2. The image matching module is in charge of searching through local image library to find out images containing similar features with the received one.
3. The image preprocessing module is to save attention models and image features of each attention object as metadata into a local feature database for reuse. This is mainly due to the thin computing power in mobile devices such that the processing cost can be saved in real time.

### 3 Communication Protocol

To make the synchronized image browsing approach work effectively, the communication module plays an important role in maintaining the synchronization relationship across mobile devices. In the protocol, the devices are classified into two categories according to their usages or functions: *master device* is defined as the device that is operated by users; *slave device* is defined as the device which works synchronously with the master device.

Several mobile devices can form a synchronization group by using wireless connection techniques like Bluetooth or 802.11b. Each group can only have only one master device at a time. Instead of directly delivering the whole image content, communication module only transmits the feature of an attention object, which is defined as:

$$F_{trans} = \{AO, Type, Feature\} \quad (2)$$

In the equation, three properties are assigned to the attention object to be transmitted.  $F_{trans}$  represents the feature data of an attention object to be delivered from a master device to its slave devices.  $AO$  stands for the properties of an attention object, which is previously described in Equation 1.  $Type$  is to indicate the category an attention object, which is classified into *saliency*, *face* and *text* by [1]. This acts as a matching criterion in the image matching module discussed later.  $Feature$  is the low-level feature extracted from the current attention object, such as color, texture, moment, etc.

### 4 Image Matching

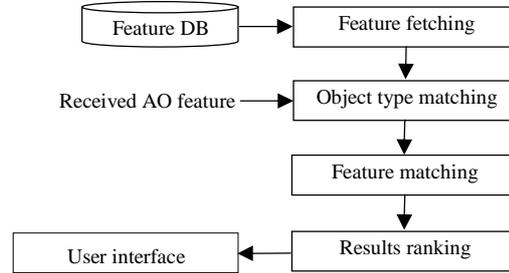
To achieve synchronized image browsing across devices, it is necessary for us to develop a method to search out synchronization images from an image collection with a large amount on various devices. In this section, our method is described in three aspects including feature extraction from images, image matching algorithm and a function to facilitate the viewing of synchronization images.

**Features to Maintain Synchronization.** In order to save the computational cost consumed in image processing, the feature extraction from images is done in the preprocessing module. In our approach, images are first analyzed to extract attention objects using the attention model proposed in [1]. In addition to the properties defined in Equation 1, in this paper, we extract low level image feature from each attention object for the use in the image matching. Many kinds of low-level features can be extracted from an attention object, however, we adopt color moment as the feature in our current implementation, which is shown to be robust and effective [2]. We extract the first two moments from each channel of CIE-LUV color space. In future work, we are considering using some semantic image features such as location, time, subject and event to extend the low-level feature that is currently used.

**Image Matching Algorithm.** More detailedly, assume there is one attention object  $AO_o$  within an image  $I_0$  being browsed on the master device. Suppose an image collection in a slave device be  $I = \{I_1, I_2, \dots, I_N\}$  with  $N$  image, and  $I_i$  ( $1 \leq i \leq N$ ) is defined as the set of its included attention objects:  $I_i = \{AO_{ik} | 1 \leq k \leq n_i\}$  (where  $n_i$  is the number of objects in  $I_i$ ). In our implementation, the similarity between two attention objects is measured using the Euclidean distance. Figure 3 gives the process to make image matching. Here is the step-by-step description of the algorithm to find out similar objects with  $AO_o$  from  $I$ :

1. For each image  $I_i$  ( $1 \leq i \leq N$ ) within the image collection  $I$ , do the step 2.
2. Examine the type of each attention object  $AO_{ik}$  ( $1 \leq k \leq n_i$ ) within  $I_i$ . Group those attention objects containing the same type with  $AO_o$  into a set, let it be  $I'_i$  (which includes  $n'_i$  members, and it satisfies:  $n'_i \leq n_i$ ). Do the steps 3-4.
3. Calculate the Euclidean distance  $d_{ik}$  between the low-level feature of  $AO_o$  and each attention object  $AO'_{ik}$  within  $I'_i$ . Select  $AO'_{it}$  ( $1 \leq t \leq n'_i$ ) with the least distance  $d_{it}$ , which is used to measure the similarity between  $AO_o$  and  $I_i$ .
4. Identify whether the distance  $d_{it}$  is under an assigned threshold ( $\gamma$ ). If the distance is under this threshold, the corresponding  $AO'_{it}$  and  $d_{it}$  is inserted into a set denoted  $\zeta$ .

By this way,  $\zeta$  is gradually formed. The search results are finally displayed on the slave devices in an order according to the similarity calculated by the above steps.



**Fig. 3.** The process to make image matching.

**Automatic Zooming into the Matched Attention Object.** We provide a further function to facilitate image viewing on the search results. That is, when users specify one of images in the search results to click, the window will automatically zoom into the appropriate attention object using a display area no less than its property MPS (seen in Equation 1).

This can be implemented by using the ROI property (seen in Equation 1), which represents the region information of an attention object within an image. The generation of automatic zooming into the matched attention object is referred to the method of the generation of automatic browsing path proposed in [3]. This function can let users view interesting regions without too much interactions like zooming/scrolling.

## 5 Experimental Results

Our evaluations were conducted with a Smartphone (Dopod 515) and a PDA (Compaq iPAQ 3670 Pocket PC). Each of them stores 75 images ranging from a variety of types including family, indoor, outdoor, group. We asked eight computer science graduate students to participate with our use study, including four males and four females. They are familiar with the operations in mobile devices like PDA and Smartphone, and they never have any knowledge of our synchronized approach before. The eight subjects were firstly asked to view an easy instruction on how to operate in our approach. After users get familiar with it, user evaluation was then performed to measure the performance.

**Task-Based Evaluations.** The participants were first asked to finish two tasks as follows:

- 1) Find out all images on the two devices that are relevant to ten assigned attention objects distributed in six images.
- 2) Test at least 10 images freely to examine the synchronization results across devices. After the test is finished, finish a questionnaire to rank the effectiveness of our approach to facilitate image browsing across devices.

**Task 1 Observation.** In this task, the subjects were divided into two groups, and each group consisted of four members with two females and two males. One group was assigned to use conventional browsing, and the other was left to use our synchronized way.

Here, we adopt the average time reduction and recall as the measures to evaluate the performance. Recall is to represent the ratio of the number of correctly found images to the total number of target images. Notice the reason that we choose recall is to manifest that users commonly failed to find all of their wanted images for the inconvenience of interactions in small devices. The test results are listed in Table 1. As can be seen, the search time has been greatly reduced by about 64.6% in our approach, and the recall has also been improved to 100%. Obviously, testers can use synchronized image browsing to facilitate the search task prior to examining them one by one on the two devices. Furthermore, through the observation of the search results by users, it's found that the testers failed to seek all matched images, hence resulting in a lower recall.

**Task 2 Observation.** The subjects were first asked to select at least 10 images freely to examine the synchronization results. After such a task is finished, they were then asked to finish a questionnaire. The questionnaire asked for general information about our approach. The items are listed in Table 2. These questions were answered on a Likert scale, where 1=strongly disagree and 5=strongly agree. The table shows that users were overwhelmingly positive about our approach, claiming that such a service should be an essential functionality for any device. Note also that the rating of user satisfaction from the synchronization results is below 4, which remains to be improved in our future work. In general, users commonly thought ours is effective to improve the browsing experience.

**Table 1.** Average search time and recall.

<b>Approach</b>	<b>Search Time (sec)</b>	<b>Recall</b>
Conventional	132.7	75.9%
Synchronized	46.3	100%

**Table 2.** 1=strongly disagree, 5=strongly agree. Here, “it” refers to synchronized approach proposed in the paper.

<b>Question</b>	<b>Rating</b>
It greatly reduces the interactions across various devices when browsing images on them.	4.68
It should be an essential functionality in any device.	4.46
By using it, I find similar images across devices easier.	4.22
I would like to use it when I stay with my families, friends or colleagues to share images together.	4.70
I feel satisfactory with the synchronization results it gives.	3.89
It improves the browsing experience across devices.	4.57

## 6 Conclusions

In this paper, we proposed a novel synchronized approach to facilitate image browsing and searching across devices. The details of the synchronization across various devices are offered, including the system framework and the algorithm. With the satisfactory results from our use study, we are planning to extend our work to other media such as video.

## 7 References

- [1] L.Q. Chen, X. Xie, X. Fan, etc. A visual attention model for adapting images on small displays. ACM Multimedia Systems Journal, Vol. 9, No. 4, 2003, pp353-364.
- [2] 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.
- [3] 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.
- [4] W.Y. Ma, I. Bedner, G. Chang, etc. A framework for adaptive content delivery in heterogeneous network environments. MMCN 2000, Jan. 2000, San Jose, USA.