Context-Aware Photo Selection for Promoting Photo Consumption on a Mobile Phone

Youngwoo Yoon, Yuri Ahn, Geehyuk Lee
Information and Communications University
119 Munjiro, Yuseong-gu
Daejeon, South Korea
+82-42-866-6161
{zeroyy, yurimong, geehyuk}@icu.ac.kr

Sungmoo Hong, Minjeong Kim
KTF
7-18 Shincheon-dong, Songpa-gu
Seoul, South Korea
+82-2-2010-9630
{sungmoo, kimi}@ktf.com

ABSTRACT
A mobile phone with a camera enabled people to capture moments to remember in the right time and place. Due to a limited user interface, however, a mobile phone is not yet a platform for enjoying the captured moments. We explored possible application scenarios for promoting utilization of user-created photos on a mobile phone. For the realization of the scenarios, we designed context-aware photo selection algorithms that take into consideration mobile phone contexts such as the current location and recent calls. A user study was conducted with a mobile phone prototype for the evaluation of the photo selection algorithms and also for user feedback about the photo consumption scenarios.

Categories and Subject Descriptors
H.5.2 [**Information interfaces and presentation**]: User Interfaces

General Terms
Algorithms, Design, Human Factors

Keywords
Mobile phone, media consumption, photographs, context-ware services, photo recommendation.

1. INTRODUCTION
A mobile phone these days can store a lot of photos that it creates, but utilization of the photos on a mobile phone is very limited. Users usually move photos from a mobile phone to a desktop computer or to a web site, and then organize them for better viewing and sharing. A mobile phone has not yet become a platform for 'consuming' photos produced by the phone. This may be due to the restricted user interface of a typical mobile phone that is characterized by a small-sized and low-resolution display and inconvenient controls. It is important to realize, however, that a mobile phone can provide people with unique photo consumption experience that other non-mobile computers cannot provide. As a mobile phone enabled people to capture important moments at right time and place, a mobile phone can enable people to remember and enjoy the captured moments at the right time and place. It is worth while pursuing application scenarios and supporting algorithms for helping people overcome the user interface limitations of a mobile phone and consume their photos better.

We started the current study with exploring ways to promote the consumption of photos on a mobile phone. We brainstormed for possible mobile phone application scenarios for the goal. While we discussed various application scenarios, we realized that algorithms for photo recommendation were needed and that it can do more than usual media recommendation algorithms do. Suppose that a mobile phone may want to show a different background image every time a user opens it. In this case a phone may choose one of the photos that it currently has at random, but certainly it can do it better than that. It may show a photo that was taken some years ago at the same place that a user is now visiting. It may show a photo that shows a friend when a user opens it in that friend's home. A mobile phone may bring a user a pleasant surprise if it is aware of a mobile context and use it in photo selection. The task of photo recommendation considering a mobile context differs from the same task on a non-mobile platform. We refer it by the term 'mobile-context-aware photo recommendation' to distinguish it from the same task in a non-mobile environment.

The GUI scenes of a mobile phone were explored in order to find appropriate places where photos can be utilized. Then, several application scenarios for photo consumption and the photo selection algorithms for the scenarios were developed. With a mobile phone prototype, we conducted a user study for the evaluation of the photo selection algorithms and also for user feedback about the photo consumption scenarios. We will describe in this paper the recommendation algorithms that we designed for the application scenarios in detail. The algorithms will make themselves examples that characterize the subject of context-aware photo recommendation for mobile applications. The concepts and algorithms introduced here are not restricted to a specific multimedia type, but we focused in the current study on photos since people do not yet create as many multimedia contents as they create photos.

MobileHCI 2008, September 2–5, 2008, Amsterdam, the Netherlands.

Section 2 summarizes related works and Section 3 defines the concept of context-aware photo selection for mobile applications. Section 4 presents the application scenarios that we developed for photo consumption on a mobile phone. Section 5 describes the context-aware photo selection algorithms and Section 6 summarizes user feedback that we received about the scenarios and the algorithms.

2. RELATED WORK

2.1 Photographs and Mobile Phones

As authoring and sharing multimedia content become easier, people use more photos and videos in an effort to provide themselves with informational and emotional information. Photo management tools such as Picasa [17] and iPhoto [1] show photos on a computer with simple and attractive interfaces. Flickr [5] is also useful service for viewing and sharing photos on the Internet. However, these services are not intended for proactive photo feeding, and they only display photos after explicit requests for photos by users.

Desktop widget services [6][22] currently support the ambient display of photos in a computer. The widgets located in a side of a display shows photos sporadically, and people notice the photos only occasionally. Loview [20], a digital picture frame, allows people to see their photos more easily with a familiar picture frame in contrast to a computer. In a recent study [19], the authors proposed a mobile phone as an ambient information display for information related to communication characteristics. In the present paper, mobile phones are utilized for feeding photos.

As well as developing services with photos, a number of studies have attempted to generate metadata for photos that can lead more intelligent and acceptable services. Davis et al. [3] inferred metadata from a current context. For example, a system infers a location name from physical location data and time information. Another study for metadata [12] inferred weather conditions from location, time and weather forecasting data. These studies are enabling context-aware photo selection algorithms, some of which are described in later sections.

Tim Kindberg et al. [11] described camera phone use and looked at six different kinds of camera phone images for social and personal purposes and affective and functional purposes. Nancy Van House et al. [8] also described social and personal use of camera phone images. This paper is proposing a system that shows user-created multimedia contents on their phones for themselves, and we are focusing on personal use of images rather than social use that is studied in [2][18].

2.2 Photo Recommendation Algorithms

The amount of accessible media is increasing dramatically as media authoring and storage devices improve; thus, viewing and searching through of a large collection of photos is now becoming problematic and intelligent retrieval, organization and recommendation issues are challenging for both users and developers. In recent studies [7][13][15][23], the authors use contextual information including the time and location to support more effective retrieval and organizations. Alexandar Jaffe et al. [10] also tried to summarize a large collection of photos using several representative photos.

Amy Hwang et al. [9] proposed Zurfer, a innovative mobile multimedia access system, and described a number of context-aware channels that make users access media according to spatial, social and topical dimensions. Antti Oulasvirta et al. [16] also studied on using contextual information on mobile phone and he proposed a renewed contact book providing mobile awareness and collaboration.

3. MOBILE-CONTEXT-AWARE PHOTO SELECTION

The contextual information that a mobile phone can use is much richer than that of non-mobile computers. A context for a mobile application may be represented by a high-dimensional composite variable whose components include time, location, interaction history, a user's schedule, and so on. Among others, location is one of the most unique components that distinguish a mobile context from a desktop context. Time is also an important component for a mobile context compared with a non-mobile context since a mobile phone is with a user almost all the time. In addition to the basic components, a mobile context contains social components since a user uses a mobile phone for communicating with other people. A recent call log and a recent message log are examples of the social components of a mobile context. In this study, we defined a mobile context to be the combination of environmental variables including time and location, social variables including call logs and message logs, and the current contents in a mobile phone like text messages and alarm events.

With this definition of a mobile context, we can define the concept of mobile-context-aware photo selection: selection of a photo that is most relevant to the current mobile context. Algorithms for mobile-context-aware photo selection are different from traditional photo recommendation algorithms in the following two important ways. First, proposed algorithms are aware of a mobile context such as a user's location and communication history. Second, they focus more on affective purposes than on informational purposes of photo uses. People tend to capture photos for affective purpose [11]. The algorithms are expected selects a photo that meets a user's needs to reminisce past events and scenes. The objective of the selection algorithms is to give affective satisfaction by selecting a photo relevant to the current mobile context.

4. SCENARIOS FOR PHOTO CONSUMPTION

The approach in this study is to encourage photo consumption without requiring an explicit user interaction for photo retrieval on a mobile phone. A spontaneous recommendation of photos allows users to have more chances to see photos but may be detrimental if it interrupts a user's tasks. For example, showing photos when a user is writing an important message is not acceptable. Therefore, before designing scenarios for photo consumption, we examined every scene of a mobile phone GUI in order to find opportune scenes which are appropriate places for photo uses.

4.1 Finding Opportune Scenes in a Mobile Phone GUI

21 major mobile phone functions were decomposed into user interface scene levels. The major functions included calling by searching a call log, writing a text message, listening to music, and taking a photo. Opportune scenes were selected according to the following four criteria:

Figure 1. Possible screen designs for the mBackground, mReminder, mMessage, and Ambient Photo

- Sufficient space for a photo on the screen
- Low cognitive load: a photo may not be welcome when a user is engaged in a complex task.
- Attention to the screen: showing a photo will not be effective if a user does not pay attention to the screen.
- Waiting time: a user will appreciate a photo recommendation when she must wait for a system response.

From dozens of scenes, the following six scenes were selected to be most opportune based on the above criteria.

Calling and receiving a call	***Alarming***
Power on / off	***Standby screen***
Writing / viewing messages	*Sending / posting messages*

We selected three out of these scenes that are most common during normal usage of a mobile phone, and developed application scenarios for them.

4.2 mBackground (Standby Screen)

A standby screen is the one that users see every time a phone is activated for use. Nevertheless, most mobile phone users set just one static image as a background image on this valuable place. mBackground service changes the background photo of a standby screen according to the current context, allowing users to see a variety of different photos on their mobile phones. The context in the mBackground service can include date, time, location, a call log, or the presence of nearby phones. The following scenario describes the mBackground service with examples.

Today is Mike's first day of work, and he arrived at the company earlier. When he opens his mobile phone to check the time in the office, the phone shows the current date and time, as well as a photo of the office that was taken on the day of his job interview.

After work, he goes downtown and can see the photo of his friend on the phone. His friend is also here, and Mike calls him to have dinner together.

The mBackground service allows users to consume their photos without distracting interactions. They only have to activate their phone to see a selected photo. The context-aware photo selection algorithm developed for mBackground is described in a later section.

4.3 mReminder (Alarm Window)

In most mobile phones, the alarm window has a static format and does not change with the context of an alarm. Some recent phones show a photo of a sender as an alarm for an incoming message, but they require explicit selection of a photo for each person in an address book.

The mReminder service, which shows photos related to the current alarm content, is suggested in order to address this issue. mReminder can insert photos into alarm windows for events, messages, and calls by utilizing the content of an alarm. For instance, when an event alarm is activated by a scheduler, mReminder can attach to the window the photos of the people who will attend the event when the participants of the event are specified. An alarm window may become more informative with the use of relevant photos.

4.4 mMessage (Writing / Viewing Messages)

mMessage is a service that uses a photo related to a sender and contents of a text message as the background image of the text message. By placing a photo while writing and reading messages, communication between users becomes more informative and emotional [21]. Moreover, a user can use a selected photo for a background image of an outgoing message as people use a letter sheet with a photo.

Mike is writing a message for his friend, "Go for a beer? I'm buying today." As Mike types the words, the mobile phone shows a photo of beer as a background image. He likes the photo and sends it with the message.

4.5 Ambient Photo

The concept of 'Ambient Photo' was inspired by experiences that many people may unintentionally go through. It allows them to enjoy photos that may be under sheets and other items on their desks, for example. The ambient photo service shows a photo on the edges of a screen, as shown in Figure 1, and users can find and consume photos as they do with photo widgets of a desktop widget service [6][22]. The accidental but pleasant finding of valuable photos can bring secondary enjoyment to a user.

In the design of Ambient Photo, a mobile phone shows a small thumbnail of a photo (Figure 1) when users navigate menus or lists of entries in what is meant to be similar to finding something on a desk. This application design is not implemented in this study.

5. DESIGN OF THE MOBILE-CONTEXT-AWARE PHOTO SELECTION ALGORITHMS

We designed photo selection algorithms for the following services: mBackground, mReminder, and mMessage. From a collection of photos $P = \{p_i, 1 \leq i \leq n\}$, the photo selection algorithms of all services select a photo p_m where

$$m = \text{argmax}_i\big(S(p_i, C)\big)\ , 1 \leq i \leq n,$$

and the variable n is the total number of photos in a collection. The function S outputs scores given a photo p_i and the current context C. S and C are defined differently for the three services.

Datetime: Monday, January 23, 2006 10:20:06 AM
GPS coordinates: 36.3908, 127.4034
Location keywords: “laboratory”
Person keywords: “anonymous”
Event keywords: “”

Figure 2. An example of photo metadata. Datetime and GPS coordinates are mandatory, but all keywords are optional.

In this study, it was assumed that photo p is represented by metadata including date and time, GPS coordinates, and textural keywords. The textural keywords that are manually noted can include keywords of a location, people, and an event. Figure 2 shows example metadata for a photo. GPS coordinates are available on GPS-capable mobile phones that are becoming available in the market. A cell tower based or Wi-Fi based positioning may also become possible options.

5.1 Pilot User Study for mBackground

A pilot user study was conducted to determine the photos that are favorable to users of the mBackground service. Seven graduate students answered questions related to their preferences for the photos on their standby screen. According to the questionnaire, participants preferred photos that were taken in a similar context to their current state to photos taken recently or to those selected randomly. Participants were also asked about their preferences regarding the attributes of month, day, time, and location when selecting a photo for a given context. The participants favored photos having higher levels of similarity in terms of the month and location. They stated that the attribute of month can reflect the sense of the season. This may reflect the fact that their country has four distinctive seasons. In a question related to their preference for using call logs, most participants were skeptical about the effectiveness of call logs. They said they receive many calls from people with a business relationship and do not want to see them on the standby screen.

5.2 Photo Selection Algorithm for mBackground

The photo selection algorithm for the mBackground service selects a photo that is most similar to the current context C which includes the current date, time, GPS coordinates, and call logs. The similarities between the photos and the context are computed from the similarities of the five attributes below. The similarities of the five attributes output binary values.

- Month similarity: the month of the captured date of a photo and current month are compared. If two values are identical, the similarity is 1; if not, it is 0.
- Date similarity: if the date of a photo and current date are identical, the similarity is 1; if not, it is 0. Here, only the month and day information are compared, and only a case with identical dates for the current day is considered, as the exact current date was considered to be more meaningful compared to the day before or after.
- Time (of day) similarity: the raw time information of a photo is converted to the index of a cluster using a time clustering algorithm. If the indexes of clusters for the current context and a photo are identical, the time similarity is 1; if not, it is 0. The detail of the time clustering is described in the following sub-sections.
- Location similarity: similar to the time similarity, a clustering algorithm was used and the index of the clusters was compared. The clustering algorithm is also described also in the sub-sections below.
- Person similarity: only a person who communicated most frequently and recently according to call and message logs is considered. If the photo information of a person includes such a person, the similarity is 1; if not, it is 0.

The following sub-sections describe procedures to calculate time and location similarities. The priority and weighted sum methods that combine the five values of similarities are then introduced.

5.2.1 Time Similarity

People tend to perceive time of day as distinct period such as morning, afternoon, evening, and night. However, the partition of time is different for different person according to his or her lifestyle. Accordingly, a personalized partition of time instead of static groups was chosen. A k-means clustering algorithm was used to make clusters of photos based on their time information. The number of clusters was chosen to be five, and a distance metric is defined as time difference between photos. A difference in time does not exceed 12 hours, and the mean time of a group is calculated as a circular mean that consider circular characteristic of time of day.

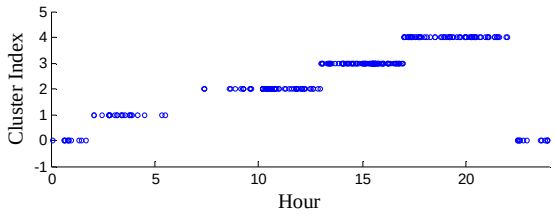

Figure 3. A possible time clustering result.

Figure 3 depicts a sample result for a collection of photos. Each user or a photo collection results in different clustering results.

5.2.2 Location Similarity

For location similarity, a photo collection was classified into several clusters according to their GPS coordinates. Unlike time similarity, each collection of photos has different number of clusters for locations. Some people who travel to many places may have a larger number of clusters. In addition, the size of a cluster should not be uniform. People recognize a country as a cluster if the country is not familiar but recognize a building as a cluster if the location is highly familiar to them.

There are the following three steps for location clustering and another step for location similarity calculation that uses clustering results.

1. Event Segmentation

The initial event segmentation for a set of photos follows the method proposed by Naaman et al. [13]. According to the timestamps of photos, the algorithm divides a set of photos into several event segments. Here, 'story line' is assumed [13] in which all photos were taken by one mobile phone or camera. In a list of photos sorted by time, the algorithm creates segments by dividing photos at a point in which the difference in the time between two neighbor photos is larger than h_{over} . In the prototypes here, h_{over} was 13 hours.

An additional segmentation step then occurs to create segments at a fine level. The algorithm separates a group at a point between two photos if the differences in the GPS coordinates and time between the photos are outliers over all differences between two adjoining photos. The method used by Naaman et al. [13] was adopted to determine the outlier values.

2. Segment Grouping

Segments of events are grouped to make clusters of location; a cluster contains one or more segments. In this step, a representative location keyword of a segment that is most frequently found for all photos in the segment was chosen. If two or more keywords of the same frequency are present, the keyword that has the lower frequency for all photos on the mobile phone is selected, as the keyword is representative of the segment and has uniqueness for the segment over a collection of photos. Following this, one or more clusters with the same representative location keywords are grouped into a cluster.

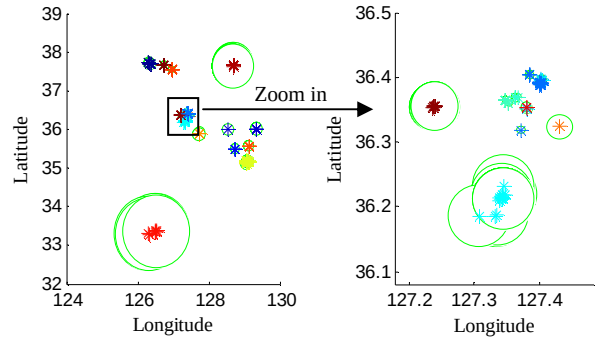

Figure 4. A possible location clustering result. Circles are making boundaries of a cluster.

The grouping mainly relies on the location keywords created by the users, as the keywords contain familiarity with the place. It was assumed that people tend to state location keywords of finer levels of detail in proportion to the level of familiarity.

3. Calculate Boundary

To be able to decide the current location is covered by which clusters, it was necessary to set boundaries for clusters. Each cluster has a value of d as is determined by its distance from a nearby cluster:

$$d = \text{EuclideanDistace(center of current cluster, center of nearest cluster)} \ / \ 3.0.$$

A center of a cluster is the mean of all points in the cluster. All photos in the cluster create a circle with a radius of d, and the region of all circles defines a cluster and its boundary. By determining d according to the distance from the nearest cluster, an isolated cluster covers a larger area, but clusters in a complex region cover only a small area. As shown in Figure 4, numerous clusters exist where the user lives, and the clusters have relatively small boundaries.

4. Location Similarity Calculation

When the current GPS coordinates are given, all clusters that cover the coordinates are activated and all photos in the activated cluster have a value of 1 as the location similarity.

5.2.3 Priority Method

From the similarities of the five attributes, the priority method determines the scores of the photos by considering the attributes individually. The five attributes are ordered with early attributes dominating later attributes. For instance, a photo with a similarity of 1 for the first attribute ranks higher than photos with similarities of 1 for all attributes except the first attribute. The most significant parameter in the priority method is the order of the attributes. The attributes are in two ways based on the properties of the attributes and based on the results of a pilot user study. The first order of attributes is

Date – Location – Month – Time – People.

It was considered that the date similarity should be considered initially as uncommon when an exact date match occurs. If the date similarity is considered in the second or third place, there would likely be no photos that satisfy this criterion, and users would be less likely to enjoy such photos as those taken on the last birthday of a friend or at Christmas parties. It was also considered that the location information is more visible in photos

compared to the month and time information. According to negative comments on user call logs, the people similarity factor was considered to be insignificant. Another order of the attributes was developed according to the pilot interview results;

Month – Location – Date – Time – People.

The interviewees indicated that the month information has a high value as they may want to be able to sense the season or the weather. They were considering photos of a beach for summer and photos of snow for winter. The priority methods with two orders of the attributes were compared to the weighted sum method and the random method in the user study.

5.2.4 Weighted Sum Method

As the name implies, this method computes the overall similarity between a photo and its context information by summing the five values of similarities using weights. When $s_j(p_i, C)$ denotes the similarity of attribute j for photo p_i and current context C, the combined similarity is given by

$$S(p_i, C) = \sum_{j \in A} s_j(p_i, C) \cdot \mathrm{pref}_j \cdot \mathrm{ipf}_j{}^{w'_j}$$

$$A = \{Month, Date, Time, Location, People\}$$

where

$$\mathrm{ipf}_j = \log\left(\frac{|P|}{|\{p \in P : s_j(p, C) = 1\}|}\right)$$

is the inverse photo frequency most similar to the inverse document frequency in the information retrieval. The weighted sum method emphasizes the rareness of relevant photos – with similarity of 1 – for attributes by introducing what is termed ipf. For example, the location attribute will have more weight if the user is in a remote location where the number of photos of the location is low. Similarly, photos captured at dawn or midnight will have a greater weight for the time attribute. However, ipf for the month and date attributes were not counted, as the number of relevant photos for these attributes cannot reflect how much less familiar the users are with the context. Unlike the location and time, the instances (or values) for the month and date attributes are equally available to the users. Thus, a low number of photos for these attributes may indicate a lack of interest. The date attribute would be emphasized when there are many relevant photos on a special day, however. Therefore,

$$w'_j = 1, \text{if } j \in \{\text{Time, Location, People}\}$$
$$w'_j = 0, \text{if } j \in \{\text{Month, Date}\}.$$

Pref was also used to introduce subjective weight for the attributes. In the user study, however, pref was 1 for all attributes. Pref and the factors that complement the differences between the attributes should be investigated in a future study.

5.3 Photo Selection Algorithm for mReminder and mMessage

The scope of mReminder was limited to an event alarm in the scheduler application. The photo selection algorithm of mReminder selects a photo that is relevant to the alarm event including the event name, location, and the people involved. The algorithm of mMessage selects a relevant photo for given message sender/receiver as well as the content. Both algorithms use the cosine similarity in the vector space model. The vector for a photo includes all textural keywords of the location, people and the event as index terms. The vector for an event includes the keywords of the event name, location, and the people involved. The vector for a message includes the keywords of the sender/receiver and all of the terms in a message. The weights of term t_j for the i-th photo is defined by

$$w_{i,j} = \mathrm{tf}_{i,j} \times \log\left(\frac{|P|}{|\{p \in P : t_j \in T(p)\}|}\right)$$

where $T(p)$ is a set of keywords of a photo and $\mathrm{tf}_{i,j}$ is a term frequency of term i within a document (photo, event, or message) j. The weights for the event and message vectors were defined similarly, and the similarity is defined as

$$S(p_i, C) = \sum_j \mathrm{w}_{C,j} w_{i,j}$$

where C includes keywords of event, location, and participants in the mReminder or keywords of sender or receiver and message content in the mMessage.

6. SERVICE IMPLEMENTATION

The three services of mBackground, mReminder, and mMessage were implemented on a Samsung M4500 PDA phone which runs the Windows Mobile 5.0 operating system and the .NET Compact Framework 2.0. The emulated services run as independent applications that are not coupled with existing scheduler or text-message applications. The implemented application obtains information regarding the current context for the three services from remote computers or via manual input by the user since the phone does not have GPS capabilities. For given context information, the application runs photo selection algorithms and shows appropriate photos with customized UIs. All of the photos are in the phone, and the metadata of the photos are stored in a separate XML file. For a query, in addition, the total run time for selecting and displaying a photo is less than one second.

The mMessage service is for writing and reading messages. When a user writes a message, the system monitors the inputted characters and runs the photo selection algorithm in the background. It then displays the selected photo on the screen. In this implementation, the photos are different for every word and users can affix a photo to send it with the message.

6.1 Considerations on mBackground

The mBackground service displays partial metadata related to the selected photo to increase the acceptance rate for the photo. When a photo is selected due to a date similarity factor, the user may not recognize the reason for the selection without an explanation. The description of a selected photo includes the year, date and location information. The difference between the current year and the year a photo was taken is shown, and date difference is also displayed at a proper level of detail. As shown in Figure 5, “Two years ago September, ABC Mart” and “Last year Today, Han River” are possible outcomes. The keywords of people in photos were not considered since the photo itself shows subjects.

Figure 5. Implemented applications of the mBackground, mReminder, and mMessage: The first two are figures of the mBackground.

In the mBackground service, it was assumed that people encounter the standby screen many times per day with the same context. The restoration time idea was developed in order to show different photos depending on the context. If a photo is selected as a background of the standby screen, the photo should not be shown for one or two hours of restoration. In this mechanism, users will see a variety of photos in the order of the score.

7. USER STUDY

User studies of the three application services and the photo selection algorithm for the mBackground service were conducted.

7.1 Photo Collection

The mBackground service assumes the use of personal photos; only the owner of the photos can evaluate them. Consequently, six paid university students were recruited and collections of photos were made that included the metadata for each subject. Table 1 describes the participants and the corresponding photo collections. The participants were paid the equivalent of $50 and they were asked to collect approximately five hundred of photos taken by them. The application scenarios assumed the use of a camera phone, but it was difficult to collect photos captured by camera phones. For this reason, photos captured by digital cameras were also acceptable. However, photos that were not recognizable on the low-resolution screen of a mobile phone were manually excluded. The participants were also instructed to collect photos from different contexts.

Table 1. Information of the participants and photo collections in the user study. (M = male, F = female)

Participant (Age, Gender)	No. of photos	Avg. no. of keywords	No. of location clusters
A (27, M)	319	3.51	115
B (23, M)	668	3.19	221
C (21, M)	483	2.76	89
D (22, F)	493	2.56	61
E (21, F)	474	4.04	119
F (22, F)	382	3.11	97

All photo sets have time spans of more than one year. The metadata of the photos included the date, time, and GPS coordinates. The photos could include (this was not mandatory) textural keywords regarding the location, people, or the event. The date and time information was extracted using the EXIF data [4] of the photos, and the textural keywords were manually inputted by the subjects with a tagging tool made for this study. The GPS coordinates were also manually specified with a customized map service [14].

In designing the location clustering algorithm, we assumed that users input location keywords in different granularities according to the familiarity for the places. The algorithm relies on this assumed pattern of location keywords, so we observed and confirmed the users really input keywords as we assumed. In a database of geo-tagged photos with location keywords, we manually tagged granularity for all location keywords in five levels. The Table 2 summarizes granularities of location keywords.

Table 2. Five different granularities of the location keywords.

Granularity	Coverage	Examples
1	Building or more smaller places	my home, my lab
2	Group of buildings	the university, park
3	Town or street	East Village
4	City	LasVegas, Amsterdam
5	State or more larger places	Nevada, Korea

The database contains photos of six users and all users currently reside in a city, Daejeon, so we considered Daejeon as a home place for all users. The photo database showed no one typed only keywords of 3 or higher granularity for the photos captured in the home place. Mean location keyword granularities for the home and the other places were 1.4 and 2.4. The users inputted keywords in finer levels for the home place but coarse keywords for the other places. Figure 6 depicts distributions of location keywords of different granularities for the home and the other places.

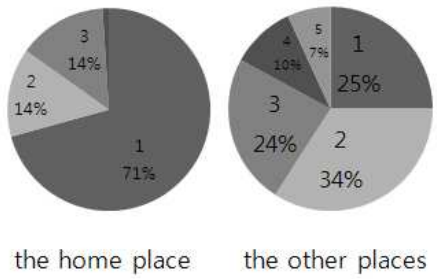

Figure 6. Distribution of location keywords of different granularities for the home and the other places.

7.2 User Study of the Context-aware Photo Selection Algorithm for the mBackground

A user study of the context-aware photo selection algorithm for the mBackground was conducted in order to obtain feedback from users and for the guidance of future developments of the algorithm. In the user study for mBackground, the priority, weighted sum, and random methods are compared. The random method selects a photo purely at random. For the priority method, two different orders of attributes were used as described in the previous section.

Priority1: Date – Location – Month – Time – People

Priority2: Month – Location – Date – Time – People

To evaluate the algorithms, it was necessary to generate several different contexts including different call logs. Initially, 20 names of people who were likely to call most frequently or who had called recently were gathered regardless of the existences of their name in the collection of the photos from the subjects. The evaluation tool then generated 50 artificial contexts automatically by extracting the date and time information for a photo selected randomly along with the location information from another photo. For an artificial context, two persons were also randomly selected from the list of 20 names. The selected contexts follow a probabilistic density of contexts in the photo collections, but it was ensured that the contexts of home or workspace occupied 20% of total contexts. For each context, a short textural description was written and some photos from the Internet (not from the photo set) were prepared to describe the context. With the description and prepared photos, the users assimilated the context in the test by reading the description and seeing the photos. Figure 7 (left) shows an example of the description and the context describing the photos shown to the subjects.

In this study, the participants rated each photo for a given context at five levels. They were unaware of the photo selection algorithms being tested. Each context was iterated three times and the photos selected previously did not appear within a context and photo selection algorithm, as the restoration time was considered when designing the mBackground service. In total, a subject rated 600 photos (50 contexts * 3 iterations * 4 algorithms) and required nearly one hour to complete the test. A desktop computer with a 19-inch monitor displayed the photos, as shown in Figure 7 (right), and the participants evaluated the photos in a dimmed room.

Figure 7. Screens shown to the users when evaluating the context-aware photo selection algorithm: The left figure shows an artificial context, and the users rate the selected photos with the right figure.

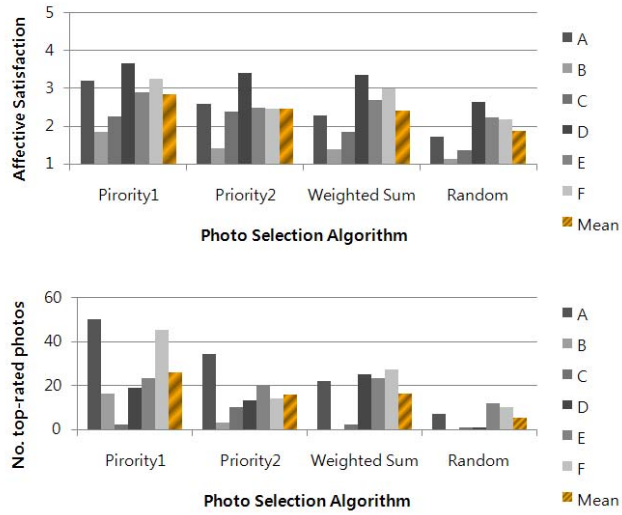

Figure 8. Evaluation results of the context-aware photo selection algorithm.

Figure 8 summarizes the results of the test excluding the results of subjects A and B. The priority method performed best and the random method has the poorest performance. Here, we conducted paired t-test to see significant differences. In terms of the mean ratings, the priority 1 method had higher ratings than the random method and the weighted sum method ($p < 0.01$). The priority 2 method also showed significantly better ratings than the random method ($p < 0.01$). The numbers of top-rated photos were also significantly different among the methods. In particular, the number of top-rated photos in the priority 1 was 5 times that of the random method.

Table 3. The mean ratings for the photos having a positive and zero similarities of month, location, time, date, and people attributes. Here, 'matched photo' means photo that have a positive similarity for a certain attribute.

Attr.	Avg. No. matched photos	Avg. No. not-matched photos	Mean rating of matched photos	Mean rating of not-matched photos
Month	21	127	2.0	1.8
Date	1.2	149	1.8	1.9
Time	38	111	1.8	1.9
Location	19	131	2.8	1.8
People	9	141	1.7	1.9

In the random method, we observed the mean ratings for the photos having positive similarities of month, location, time, date, and people attributes. We tried to find importance of attributes by comparing mean ratings between the photos having a positive similarity and the photos having a zero similarity. Table 3 summarizes the results. It means that the location attribute have more power in selecting photos. The mean ratings between location matched cases and not-matched cases were significantly different ($p < 0.05$). In interviews after the test, the participants also stated that they mainly rated the photos based on the information of location. The captured dates of photos were hard to be recognized and matched with the context.

All participants stated that viewing photos similar to the current context is good for affective satisfaction. They said those photos arouse reminisces and the photos were proper to share with nearby people. The participants answered the location is the most remarkable attribute, but they said the location attribute is less remarkable in home places since they don't reminisce home and work places. We also found different ratings for familiar or unfamiliar places. All subjects were familiar to the university, so we compared mean ratings for the photos of the university and the other places. Mean rating for the photos of the university and the other places was 2.77 and 3.26 respectively (significantly different, $p < 0.05$). This comparison is based on the priority 1 method and location matched photos.

7.3 Experiencing Services

To evaluate and improve the application scenarios and the photo selection algorithm, another user study was conducted in which subjects experienced with the three services and were interviewed. In this user study, six possible scenarios for the mBackground, mReminder, and mMessage were created. Table 4 shows the scenarios briefly; in the actual user study, the scenarios were more elaborate than the description in Table 4.

Context information was created for the mBackground, and event schedules for the mReminder and text messages for mMessage were created according to the scenarios. The photo selection was completed by the algorithm on the mobile phone. The author's photo collection was used in this test.

Table 4. Six different scenarios in the experiencing services

Scenario	Brief description
mBackground scenario #1	He is now climbing the mountain where he visited before. When he activated the phone, a photo captured at the mountain is shown.
mBackground scenario #2	He is now studying in his laboratory and his phone shows a photo of a campus festival that was captured exactly three years ago.
mReminder scenario #1	There is an alarm for a meeting of a club and the phone shows a photo captured at the previous meeting.
mReminder scenario #2	There is an alarm for a dinner appointment with a friend and the phone shows a photo of the friend.
mMessage scenario #1	He receives a text message from his friend asking him, "What goggles do you use?" The phone shows a photo of a set of goggles. If he is using the goggles, he also may reply with the a photo.
mMessage scenario #2	He receives a message from an old friend and the phone shows a photo of the friend.

Figure 9. The phone used in experiencing the services. The figures are for the mBackground and mMessage scenarios.

15 unpaid students participated in this study. The mean age of the participants was 24, and three of them were female. After experiencing the services with the scenarios, they answered a questionnaire and were interviewed about their answers. First, all participants highly accepted this concept of the services. More specifically, 12 subjects responded that they want to use the mBackground services, but the other subjects thought that varying background images of the standby screen might be annoying and they wanted to use their favorite photos. For the mReminder and mMeesage service, all participants expressed positive responses. However, a few subjects worried about excessive changing of a background photo in writing a message as they type words. About half participants also commented that they would like to be able to adjust the photo selection algorithms and have feedback mechanism so that photos of, for example, ex-lovers or embarrassing events are not shown. In addition, the half participants preferred "2005" not "two years ago" in presenting the information of the selected photo in the mBackground service. They said "2005" is more informative and recognizable, but the

people in the other side said “two years ago” is more emotional and sentimental.

8. CONCLUSION AND FUTURE DIRECTIONS

This study explored new scenarios for promoting photo consumption on a mobile phone. Also the paper introduced context-aware photo selection algorithms for the realization of the scenarios. User studies showed that people are highly accepting of the services and that the photo selection algorithm was quite effective. In the proposed application, users preferred our mobile context-aware photo selection algorithm than random method used in lots of photo feeding systems. Moreover, Location was found to be a predominant attribute in the mobile context-aware photo selection algorithm.

In the current study we had to use photos taken by digital cameras instead of those taken by a mobile phone. Also, we had to rely on an artificial experimental setup in a user study to evaluate the scenarios and photo selection algorithms. The next step of this research will be to conduct a long-term field study that will overcome these limitations of the current study. In the field study, it is expected that the proposed services may change the behavior of users when capturing photos and inputting metadata. In addition, the context-aware photo recommendation function could benefit from machine-learning algorithms. Personalized preferences of the attributes of photos and user preference patterns concerning location and time of the photos would be learnable.

9. REFERENCES

[1] Apple iPhoto. http://www.apple.com/ilife/iphoto/.

[2] Counts, S. and Fellheimer, E. Supporting social presence through lightweight photo sharing on and off the desktop. In Proc. CHI 2004, ACM Press (2004), 599-606.

[3] Davis, M., King, S., Good, N., and Sarvas, R. From context to content: leveraging context to infer media metadata. In *Proc. Multimedia 2004*, ACM Press (2004), 188-195.

[4] EXIF data. http://www.exif.org/.

[5] Flickr.com. http://www.flickr.com/.

[6] Google Desktop. http://desktop.google.com/

[7] Graham, A., Garcia-Molina, H., Paepcke, A., and Winograd, T. Time as essence for photo browsing through personal digital libraries. In *Proc. JCDL 2002*, ACM Press (2002), 326-335.

[8] House, N. V. and Davis, M. The Social Life of Cameraphone Images, In Proc. Pervasive Image Capture and Sharing Workshop, (2006).

[9] Hwang, A., Ahern, S., King, S., Naaman, M., Nair, R., and Yang, J. Zurfer: mobile multimedia access in spatial, social and topical context. In Proc. Multimedia 2007, ACM Press (2007), 557-560.

[10] Jaffe, A., Naaman, M., Tassa, T., and Davis, M. Generating summaries and visualization for large collections of geo-referenced photographs. In *Proc. MIR 2006*, ACM Press (2006), 89-98.

[11] Kindberg, T., Spasojevic, M., Fleck, R., and Sellen, A. The Ubiquitous Camera: An In-Depth Study of Camera Phone Use. *IEEE Pervasive Computing*, 4, 2 (2005), 42-50.

[12] Naaman, M., Harada, S., Wang, Q., Garcia-Molina, H., and Paepcke, A. Context data in geo-referenced digital photo collections. In *Proc. Multimedia 2004*, ACM Press (2004), 196-203.

[13] Naaman, M., Song, Y. J., Paepcke, A., and Garcia-Molina, H. Automatic organization for digital photographs with geographic coordinates. In *Proc. JCDL 2004*, ACM Press (2004), 53-62.

[14] Naver OpenAPI http://openapi.naver.com/.

[15] O'Hare, N., Gurrin, C., Jones, G. J. F., and Smeaton, A. F. Combination of content analysis and context features for digital photograph retrieval. In *2nd IEE European Workshop on the Integration of Knowledge, Semantic and Digital Media Technologies*, (2005).

[16] Oulasvirta, A., Raento, M., and Tiitta, S. ContextContacts: redesigning SmartPhone’s contact book to support mobile awareness and collaboration. In Proc. MobileHCI 2005, ACM Press (2005), 167-174.

[17] Picasa http://picasa.google.com/.

[18] Sarvas, R., Oulasvirta, A., and Jacucci, G. Building social discourse around mobile photos: a systemic perspective. In Proc. MobileHCI 2005, ACM Press (2005), 31-38.

[19] Schmidt, A., Häkkilä, J., Atterer, R., Rukzio, E., and Holleis, P. Utilizing mobile phones as ambient information displays. *Ext. Abstracts CHI 2006*, ACM Press (2006), 1295-1300.

[20] SK Telecom Loview. http://www.loview.co.kr/, http://wireless.engadget.com/2005/12/09/sk-telekom-loview-digital-picture-frame-receives-photos-via/.

[21] Voida, A. and Mynatt, E. D. Six themes of the communicative appropriation of photographic images. In *Proc. CHI 2005*, ACM Press (2005), 171-180.

[22] Yahoo! Widgets. http://widgets.yahoo.com/.

[23] Yu, Z., Zhou, X., Zhang, D., Chin, C., Wang, X., and Men, J. Supporting Context-Aware Media Recommendations for Smart Phones. *IEEE Pervasive Computing*, 5, 3 (2006), 68-75.