Author(s): Malik, Aqdas & Marko, Nieminen

Title: Understanding the Usage and Requirements of the Photo Tagging System

Year: 2014

Version: Final published version

Please cite the original version:

Malik, Aqdas & Marko, Nieminen. 2014. Understanding the Usage and Requirements of the Photo Tagging System. Human IT. 12:3. P. 117-161. ISSN 1402-150X (electronic).

Understanding the Usage and Requirements of the Photo Tagging System

Aqdas Malik and Marko Nieminen, Aalto University

The amount of personal photos is massively increasing and managing them effectively and efficiently requires new approaches and practices. This study analyses the users' needs and their behaviour towards photo tagging in the context of a personal photo repository. Our results are from a qualitative study with 15 users. Study methods include a questionnaire and a task analysis approach in which we analyse and evaluate practices around semi-automated photo tagging. In the task analysis, we describe and use a photo tagging application called SmartImages for studying the actual tagging experience. The results from the study indicate that photo tagging in personal collections is rarely used as it is considered too laborious. The task analysis with SmartImages made users consider tagging worthwhile and beneficial. The results propose changes to the implementation of tagging functionality in photo management applications. We conclude that better visibility of the tagging feature and introducing social elements would improve the usage and benefit of tagging. Providing automated tag suggestions that are comprehensible, conceptually relevant and the relation between the displayed tags and the photo is clear would make users more willing to engage with tagging activities. Addressing the mentioned issues would help the users in managing the increasing number of personal photos.

Keywords: human–computer interaction, photographic practices, photo tagging, task analysis, user interface design

Malik, Aqdas & Marko Nieminen. "Understanding the Usage and Requirements of the Photo Tagging System." *HUMAN **IT*** 12.3(2014): 117–161.

With ever increasing adoption and popularity of mobile phones and digital cameras, people are taking more photos than before. Affordability of capturing devices, higher storage capabilities, better image quality, virtually no cost to capture photos, and multiple capturing devices in the household has fuelled the number of personal photo collections (Rodden & Wood 2003; Neustaedter & Fedorovskaya 2009; Whittaker, Bergman & Clough 2010; Sarvas & Frohlich 2011). As the number of photo libraries grow, the problem of managing, organizing, locating and interacting with photos also increase (Frohlich *et al*. 2002). Progressively it becomes difficult to manage and locate photos when needed. Organizing, browsing and finding photos becomes a challenging task for many users (Ren, Sarvas & Ćalić 2010; Whittaker, Bergman & Clough 2010). For effective organization and management of photos, human effort and time is required, something that many users would either postpone as a future task or either not be willing to get involved in (Frohlich *et al*. 2002; Whittaker, Bergman & Clough 2010).

Most of the commercial photo browsing and management applications still rely on chronological organization of photos (Gozali, Kan & Sundaram 2012). Though, some of the applications also offer users to tag their photos manually, but the feature has not been widely used or adopted by the users (Kirk *et al*. 2006). Still, most of the people rely on their cognitive or visual memory to locate their photos (Rodden & Wood 2003; Kirk *et al*. 2006). To locate a photo, the users try to remember the place or folder they saved a specific photo (Kirk *et al*. 2006). As people are increasingly using multiple devices for capturing photos, they also have to remember the device a specific photo was taken with.

Photo tagging can significantly help in overcoming the above mentioned issue by enhancing the usefulness of large set of photos (Ames & Naaman 2007). Automated and manual tagging systems, tag suggestions systems, similarity search systems, and face recognition and tagging systems are some of the popular genre of tagging systems (Kustanowitz & Shneiderman 2005). A number of experimental studies have incorporated

and tested the validity of various tagging techniques (Cui *et al*. 2007; Naaman & Nair 2008; Qin *et al*. 2011). These studies have mostly addressed the technical implementations and aspects such as introducing new algorithms or novel techniques for relevant tags. Research on tagging has mostly addressed motivational factors, tagging patterns, social images tagging, and testing various concepts with users (Bar-Ilan *et al*. 2008; Bar-Ilan *et al*. 2012; Golbeck, Koepfler & Emmerling 2011). A limited number of studies have been carried out on understanding the user needs and expectations as well as design requirements of tagging systems focusing on personal photos management (Wilhelm *et al*. 2004; Tsai & Hung 2008).

Our study bridges the gap between the tagging systems and users' needs and expectations from these systems. The study complements and extends the research conducted by Ames *et al*. (2010), Kirk *et al*. (2006) and Rodden and Wood (2003) on understanding the practices revolving around various phases of digital photography. In addition to general practices, Ames *et al*.'s (2010) longitudinal study on mobile photoware also investigated the patterns and usage of tags through Zonetag and Flickr. Meanwhile, their other study focused on the factors that motivate users in tagging photos (Ames & Naaman 2007). Studies by Kirk *et al*. (2006) and Rodden and Wood (2003) present findings on usage practices during various phases of photo-work including capturing, organization, browsing, searching, and tagging photos. As these studies cover multiple stages involved in digital photography, they offer limited understanding and input to users' tagging needs, expectations, and attitudes.

In this paper we present findings from the task analysis study with 15 participants by using a tagging application called SmartImages. The application generates and suggests various tags for photos and the users can append the suggested tags or create new tags for a specific photo. The application also allows face tagging, similarity search, and photo search features.

The overall research question of our study is:

What makes tagging meaningful, interesting and more adoptable for the users?

We achieve the results by studying real users using SmartImages and by analysing their behaviour and usage with the application. We limit the scope of our study by focusing mostly on tagging photos in the context of personal photos management.

Our study contributes to the body of knowledge in the following ways:

- Identifying the current practices in photo tagging with a focus on personal photos management.
- Reporting on the actual needs, expectations and attitudes of real users from a photo tagging system.
- Studying various interactions that can enhance the acceptance and adoptability of tagging systems.

Based on the study findings, we offer implications and practical recommendations that can be utilized by designers of digital photo management applications in personal as well as social contexts.

Related Work

Personal photography and more recently digital photography has attracted research interest from various disciplines including information processing and retrieval (Bar-Ilan *et al*. 2008; Bar-Ilan *et al*. 2012), machine learning (Tsai & Hung 2008; Xie *et al*. 2010), and human–computer interaction (Ames & Naaman 2007; Dong & Fu 2010; Nov & Ye 2010). With the ever increasing amount of digital photos, various solutions have been developed for managing them by the academic community as well as commercially. Besides many other solutions for managing photos, tagging is one of the integral features in many commercial[1] and experimental academic photo management applications (Cui *et al*. 2007;

Naaman & Nair 2008; Qin *et al*. 2011). Photo tagging or annotating is a process as well as an activity that lets the users or the machine assign terms or keywords to a photo. Recent efforts in the tagging domain has mainly focused on developing tools, techniques, and algorithms, and lately on various aspects of social tagging (Folksonomies) including usage patterns and motivations behind the activity (Cui *et al*. 2007; Naaman & Nair 2008; Qin *et al*. 2011; Strohmaier, Körner & Kern 2012).

Many researchers point out the multiple benefits of tagging photos. It provides new life and structure to unorderly scattered photos by converting them to valuable and useful repositories (Kustanowitz & Shneiderman 2005). Tagged photos can be easily organized, managed, searched and retrieved, hence minimizing the effort of locating them later on (Ames *et al*. 2010; Halaschek-Wiener *et al*. 2006; Sen *et al*. 2006; Wash & Rader 2007). With the convergence of personal photo management tools with sharing features, tagging can also enhance the social value of photos, reflect commitment and built reputation within the social network (Kustanowitz & Shneiderman 2005; Nov & Ye 2010).

Despite the number of benefits, tagging feature has not been widely adopted and used in personal photo repositories, and a majority of the users do not tag their photos (Frohlich *et al*. 2002; Rodden & Wood 2003; Kirk *et al*. 2006; Vennelakanti *et al*. 2011). Various popular and commonly used applications incorporate a tagging feature, but most of the users consider it demanding and difficult to use (Whittaker, Bergman & Clough 2010). Many users are even not aware of the tagging feature or the benefits it can bring to their content (Frohlich *et al*. 2002). One of the other main reasons for lack of interest and adoption of tagging is that the users need to spend effort as well as time (Kustanowitz & Shneiderman 2005; Vennelakanti *et al*. 2011). Many users treat tagging as a to-do project for the future, as they plan to organize and tag their photos at a later point, something that is never carried out by them (Frohlich *et al*. 2009). Most of the users still rely on remembering the saved location of the photo or simply browsing through them. These are the common

practices of retrieving the photos when needed (Whittaker, Bergman & Clough 2010).

A number of manual, semi-automatic, and automated tagging techniques have been developed within academia as well as commercially. Content-based photo retrieval is a highly active research field and numerous advancements have been made in developing automated and semi-automated tagging systems. Most popular commercial applications yet rely on manual tagging compared to other techniques (Kustanowitz & Shneiderman 2005) as it offers better reliability and human touch. On the other hand, it is highly time consuming and require greater commitment from users especially for large sets of photos (Wang *et al*. 2012).

Though manual tagging is highly personalized data, it might not be well suited due to its limited relevancy and usefulness for other users. Wide acceptability and adoption of manual tagging is still limited as it is considered a laborious task that requires human effort and many are unwilling to involve themselves in this activity time (Vennelakanti *et al*. 2011). Due to limited use of photo tagging, the search feature performs poorly while locating specific photos. Hence, many users do not perform search to locate their photos as they are aware that the search would not yield the required results.

Meanwhile, semi-automated and automated tools require less human effort and time but the accuracy and validity cannot be guaranteed (Wilhelm *et al*. 2004; Cui *et al*. 2007; Qin *et al*. 2011). Automated tagging techniques still lack preciseness and human feeling as the tags are generated by the machine. Automated photo tagging is more prone to errors and offers limited personalization. On the other hand, it requires almost no human effort and large photo collections can be tagged in minimal time. Additionally, this technique keeps consistency across the tags, thus leading towards better organization and searching of pictures (Xie *et al*. 2010). In addition to the above mentioned techniques, some research work has also focused on engaging the users more with the tagging feature, in fun and playful manners. ESP game (Von Ahn & Dabbish

2004), and KissKissBan (Ho *et al*. 2009) offer novel techniques in the form of tagging as a game.

Despite various types of photo tagging applications and techniques, few studies have been carried out on understanding the users' needs, requirements, and attitudes as well as design issues of photo tagging systems within personal photo management context (Wilhelm *et al*. 2004; Tsai & Hung 2008). On the HCI front, the research community has mostly been involved in studying the users' motivations, perceptions, privacy concerns, and usage patterns related to photo tagging (Bar-Ilan *et al*. 2008; Besmer & Richter 2010; Kim & Rieh 2011). A number of researchers within the domain have designed and tested tagging applications, interfaces and concepts with real users (Cui *et al*. 2007; Qin *et al*. 2011; Bar-Ilan *et al*. 2012). Meanwhile, after the recent popularity of social media, some researchers have also focused on social tagging context (Heckner, Heilemann & Wolff 2009; Golbeck, Koepfler & Emmerling 2011). New techniques and interaction methods to enhance the usage and acceptability of tagging systems have also been proposed by some researchers (Kustanowitz & Shneiderman 2005; Matusiak 2006). On the cultural front, some work has also been conducted in understanding the cultural differences in tagging practices and adoptability (Dong & Fu 2010; Peesapati, Wang & Cosley 2010).

SmartImages

SmartImages is an application for viewing, browsing, searching, and tagging photo collections. The application suggests tags and allows the users to add photo and face tags in different photos as well as search for similar photos.

SmartImages was developed in a larger research project that had its focus on image analysis and retrieval. The authors together with PicSOM[2] and MUVIS[3] teams have been working closely in joint projects revolving around photo management and search solutions for the last two years. The research that is reported in this paper has had its focus on

user interface, usability, and user experience, meanwhile technical solutions have been developed by PicSOM and MUVIS team members. Our joint efforts resulted in SmartImages that was initially aimed at providing user feedback about the accuracy of the PicSOM and MUVIS image retrieval algorithms. The main motivation of using SmartImages in the current study has been to test the developed concept and its various features and functionalities with real users. By testing the concept with real users we also expected to obtain answers to our main research question, "What makes tagging meaningful, interesting and more adoptable for the users?" as well as a better understanding of users' practices, motivations, and attitudes towards photo tagging.

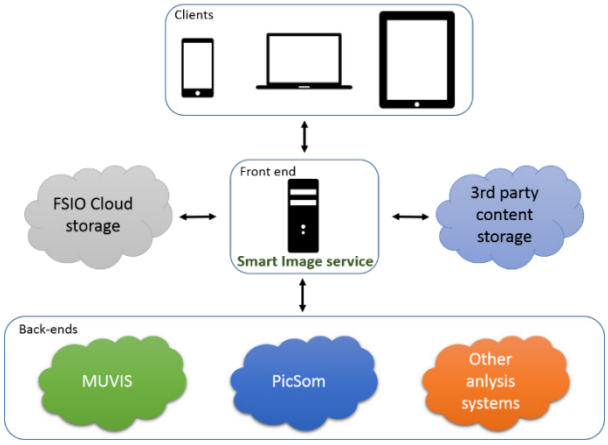

Figure 1: System architecture of SmartImages

The frontend of SmartImages is hosted by Packet Video[4] which provides user interface to the clients. REST API is used for the communication between the client and front-end as well as for the front-end and backend servers. The front-end delivers the web-based HTML5 user interface for the client devices. Additionally, the front-end communicates with the backend servers including PicSOM and MUVIS for the image

analysis. PicSOM is a multi-purpose content based image retrieval system that automatically generates relevant keywords based on colour, shape, texture and structure of a given image (Laaksonen *et al*. 2000). Meanwhile MUVIS, a framework for indexing, browsing, summarizing, and querying large image collections is responsible for providing content-based analysis service to the clients (Kiranyaz *et al*. 2003). The system architecture of SmartImages is presented in Figure 1.

SmartImages has been designed and implemented for tablet devices. The main view of the application displays a thumbnail view from the photos library (Figure 2). Thumbnail view contains nine photos and can be navigated through by swiping across or by tapping on the next or previous icons (Figure 4). Tapping on any thumbnail opens the photo in full view. In order to improve usability, the user interface provides screen tips that offer assistance to the users.

Search can be conducted by tapping on the search icon located at the top of the screen. Long tap on an image in thumbnail or full view initiates the similarity search feature. After analysing the content of the selected photo, the system displays those photos that are considered similar.

The users can also provide feedback on the generated results. This feedback is used as learning input by the system algorithms for customizing the future results accordingly (Laaksonen *et al*. 2000). The full view mode (Figure 4) allows the users to add and edit tags related to photo as well as faces. Tags can be added manually or by selecting the suggested tags generated by the system based on content analysis.

Data Gathering

Methodology

To gain a better understanding of the users' needs and requirements of various photo tagging related features, we utilized the task analysis method complemented by open-ended interviews. Knowing the users and

their needs by involving them in the design process is the core of user-centred design. Task analysis is one of the user-centred design techniques commonly used in the HCI domain for gaining insight into users, their usage practices and needs, and the tasks to be completed to achieve their goals in relevance to a system (Diaper 1989; Hackos & Redish 1998). Task analysis is considered a useful method in clarifying the objectives of each task within the application as well as pinpointing the significance of each task, e.g. tasks that are critical, unnecessary, missing or dependent on other tasks (Dumas & Redish 1999). This technique also enables the researchers to observe the difficulties and problems faced by the users while completing various tasks as well as getting an understanding of required skill level for completing the tasks efficiently. In task analysis method, a set of scenarios are developed beforehand and applied in relation to each task with the users. Scenarios are usually a set of narratives that describe what people do in particular activities (Carroll 1995). Using scenarios in task analysis method provide the users with close to real-life, concrete, and practical situations that can be easily comprehended by them (Benyon & Macaulay 2002). This method has been applied commonly for understanding the user requirement and goals by going through a set or sub-set of various tasks that are deemed necessary for using technology applications with real users (Freeman & Freeman 2011; Tappan *et al*. 2011).

The scenario based task analysis method allowed us to get insight into the participants' photo tagging goals and to analyse how they actually achieved the goals by completing the set of tasks. We were also able to get input from the participants on their preferences and importance of various photo tagging features. Using scenarios before evaluating each task also helped the users to comprehend the situation of feature use and gave an overall better understanding of the whole application. The open-ended questions that followed after each task, discussed various task related issues. Problems with using the feature, design issues and alternative solutions, use cases and usefulness of the feature in real life were some of

the themes covered by open-ended questions. Additionally, with open-ended questions the participants also managed to generate various ideas and features related to the topic.

All the sessions were video recorded on a video camera backed-up by audio recording on a mobile phone. After transferring the recordings to a PC, they were manually transcribed word by word for further analysis. Qualitative data generated from each session observations and open ended interviews was analysed by using the affinity diagram technique by one researcher. This technique is commonly used to organize individual notes and relevant data that is captured during interpretation sessions. With the help of Affinity diagram, this raw data can be hierarchically organized to reveal common themes, issues, and trends (Beyer & Holtzblatt 1997). In the scope of contextual design, this technique has been employed frequently by the practitioners to get insight into system requirements, limitations, and key problems faced by the users (Heikkinen *et al*. 2013; Olsson *et al*. 2013).

The researcher watched the recorded session videos together with notes taken during each test session that lead to a detailed description of each test session. As we collected a large sum of qualitative data, affinity diagrams provided us a better and detailed understanding of the data. By using this technique we grouped responses of each participant in accordance with tasks and sub-tasks. Open ended answers by all the participants were also grouped to get a meaningful and broader understating of the user's needs as well as underlying problems and issues. By using the technique, we consolidated the data gathered from each session in accordance with various tasks. Consolidating the data lead us to categorize main themes emerging from each of the test task and open ended interviews. The main themes that emerged from data consolidation are: general imaging practices, actual usage and needs of a tagging system, user interface related findings, and new ideas generated by the participants. These main themes form the core of our results section.

After data consolidation, we categorized the data that helped us in clear understanding of various feature related issues, users' concerns, as well as task specific problems. We were also able to identify and group a number of generated ideas specific to each task.

Procedure and Practical Arrangements

The study was carried out in the usability lab of our university. Each session comprised of three phases, i.e. pre-tasks phase, task analysis phase, and post-tasks phase. In the pre-tasks phase, the participants were requested to fill in a questionnaire about their basic photography practices as well as background data. The pre-tasks questionnaire inquired about age, gender, occupation, camera devices, photos capturing frequency, organizing photos, storing photos, and tagging photos related practices and habits. Results of the background data is presented in Table 2, while the results of photographic practices of the participants is presented in Table 3.

After filling in the pre-tasks form, the researcher held brief discussion based on the respondents' answers. The aim of this discussion was to get a deeper insight into the photographic practices of the participants. Before initiating the task analysis phase, we introduced SmartImages to the participants. Various features of the application were briefly explained to stimulate the participants to envision the possibilities in relation to the photo tagging practice. After the brief walkthrough, the task analysis phase started by presenting scenarios for using the application as well as for each task to be completed with the application. In total, there were nine tasks in the task analysis session pertaining to navigation, searching and tagging. The first task was related to scrolling photos, three tasks were about adding and removing tags, two tasks addressed keywords similarity search, while the last three tasks focused on face tags and face similarity search. A detailed tasks list is presented in Table 1.

After completing each task, discussion about the task was carried out with the participants. Discussion in this phase primarily focused on

problems they faced, their liking and disliking about the task, interaction method, as well as their impressions and concerns with the task.

Task 1 *Navigational task*	"I would like to scroll through my already uploaded photos and find a photo of a table with coffee mug and water glasses"
Tasks 2–4 *Adding & removing tags tasks*	2. "I would like add some tags to this photo of 'a table with coffee mug and water glasses'" (user defined tag could be for example be 'coffee') 3. "I would like to search for the photo that I just tagged" (for example with a tag 'coffee') 4. "I would like to remove some of the irrelevant tags from this photo of 'a table with coffee mug and water glasses'"
Tasks 5–6 *Keywords & keywords similarity search tasks*	5. "I would like to search for the all the photos of Eiffel Tower" 6a. "I would like to search for all the photos that are similar to the Eiffel Tower photos" 6b. "I would like to remove all the non-related photos from the similarity search results of the Eiffel Tower"
Tasks 7–9 *Faces and faces similarity search tasks*	7. "I would like to search photos where both Mary and Peter are seen" 8. "I have noticed that in photo x Peter is not mentioned even though he is in the photo and I would like to add his tag in the photo" 9a. "I would select one of the photos where both Mary and Peter are seen, and search for similar photos" 9b. "I would like to remove all the non-related photos from the similarity search results of Mary and Peter"

Table 1: Tasks list carried out by each participant on SmartImages

In the post-tasks phase, an open-ended interview session was carried out with each participant for detailed discussions. Impressions about SmartImages and its various features such as similarity search, automatic tagging, and face recognition were discussed in detail. The session also addressed the liking and disliking of features, interaction methods, as well as their impressions with the application in general.

Before concluding the session we requested feedback, ideas, and suggestions related to the study from the participants. We managed to receive a number of excellent ideas and suggestions during the final phase of the session. Demographic information and photo capturing devices of study participants are presented in Table 2.

Study Participants

The study was advertised on the notice boards of the Computer Science building of our university. We set the following criteria for recruiting the eligible participants.

- The participant must own a digital camera or mobile phone equipped with camera.
- The participant must have taken at least five photos during the last month.
- The participant has at least 200 photos stored in her digital photo collection.
- The participant is not a professional photographer.
- The participant should have some experience with touchscreen devices.
- The participant must be fluent in English.

We also encouraged female participants to participate in the study (our aim was to recruit at least 25 per cent females for the study).

As this was the first phase of the SmartImages evaluation, we wanted to generate new ideas and suggestions for further development for our

application. Thus we opted to recruit test participants who are actively involved with taking photographs as we expect that this group of recruits will not only be engaged with digital photos capturing, but with their management, searching and sharing as well. Limiting the selection criteria to at least a moderate level of experience and engagement in various photographic activities was deemed necessary as almost all the above mentioned activities were covered during the evaluation and interview.

Results

Fifteen participants involved in capturing photos from their smartphone or digital camera were recruited for the study. Ten participants (67 per cent) were males while five (33 per cent) were females. Eleven participants were students, two were engaged in entrepreneurship, one working as researcher and one as an engineer. Nine participants were 18–24 years old, five were between 25 and 34 years old, and one was 39 years old. Five participants owned a Nokia smartphone, five were using HTC, three had Samsung, while two were using Apple phones. Twelve out of fifteen participants owned a digital camera. Six of them owned a Canon camera, three had Nikon, two had Sony, and one had a Samsung digital camera. Eight participants also owned a tablet device. Most common tablet device was Apple Ipad which was owned by six participants, followed by Samsung and Asus tablets owned by one each. PC was the most common storage place for the photos mentioned by twelve participants. Online storage service Dropbox and phone memory is used by four participants each. Three use external hard drive for storing their photos. Meanwhile Skydrive, USB and Kuvat.fi is also used by one participant each.

	Age group	Gender	Occupation	Current phone model	Current digital camera model	Current tablet model
P1	18–24	Male	Student	Nokia Lumia	Nikon D3100	ipad 2
P2	25–34	Male	Student	Nokia N9	-	-
P3	25–34	Male	Student	Nokia N8	-	-
P4	25–34	Female	Entrepreneur	iPhone 4	Canon 1000D	iPad 1
P5	25–34	Male	S/W Eng.	HTC Desire	Canon S90	iPad 2
P6	18–24	Female	Student	iPhone 3GS	Canon 400D	iPad
P7	18–24	Female	Student	Nokia	Canon D500	Galaxy Note
P8	18–24	Male	Student	HTC Desire	Nikon D40	-
P9	25–34	Male	Entrepreneur	HTC Chacha	Sony DSC-V1	Asus Nexus 7
P10	18–24	Female	Student	Nokia Lumia	Canon G12	-
P11	18–24	Male	Student	Galaxy S3	Nikon D60	iPad
P12	18–24	Female	Student	HTC	Sony	-
P13	18–24	Male	Student	HTC desire	-	-
P14	18–24	Male	Student	Samsung	Canon	-
P15	35–39	Male	Researcher	Samsung S2	Samsung	iPad

Table 2: Demographic information and photo capturing devices of study participants

In the following sections we will report the findings from our study. In the first section, we present our findings and results from the pre-test questionnaire and interviews. The second and third sections will highlight the findings from the task analysis phase as well as the post-test interviews.

Imaging Practices

Capturing Photos

All participants (n=15) answered that the mobile handset was the preferred device for capturing photos. Four subjects captured photos from their handsets on a daily basis, nine captured at least once a week, while the remaining two took photos at least once a month. Out of eight participants owning a tablet device, five mentioned that they never use the tablet for capturing photos, while only three seldom used it for capturing photos. The main reason for not capturing photos with a tablet device is due to the nature of the device as well as the setting in which the device is usually used.

	How often do you take photos with your phone camera?	How often do you take photos with your digital camera?	How often do you take photos with your tablet?	Where do you save/store photos?	How often do you organize your photos?
P1	Everyday	Seldom, when I need better quality photos	Never	Skydrive, PC	Never
P2	Few times a week		-	Laptop	Around once a year
P3	5-10/ week	Rarely	-	Phone, PC	Never
P4	Every week	Few times a year	Never	Laptop, Dropbox, Kuvat.fi	Never

P5	Daily	Monthly, on bigger occasions	Never	Picasa, PC	Around once a year
P6	Few every week	Around once every month	Never	Laptop, Phone	At least once every month
P7	Twice every week	Weekly	Rarely	Dropbox, Computer	Around once a year
P8	Once a week	Once a month, on trips, take many photos	-	Multiple hard disks, Dropbox	Once every month
P9	Few times a month	Few times a year	Couple of times a year	Laptop, External HD as backup	Around once a year
P10	Almost every-day	Once a month	-	External hard drive, PC	Never
P11	Weekly	Monthly	Few times a year	PC, Laptop	Around once a year
P12	Once a month	Couple of times in a year	-	Computer	Around once a year
P13	Few time a week	-	-	PC	Around once a year
P14	Almost every-day	Maybe once a year	-	Phone, Laptop	Around once a year
P15	Few times a month	Few times a year	Never	Phone, USB, Dropbox	Around once a year

Table 3: Photographic practices of study participants

Meanwhile, four of the participants expressed in the interviews that digital cameras are meant for taking photos mostly on scheduled and special events such as birthdays, picnics, and trips abroad. Out of twelve participants owning a digital camera, only one of them used it at least once a week, three of them use the device at least once a month, while remaining eight stated that they use the camera a few times a year. Details

about the capturing practices by each study participants are presented above in Table 3.

Renaming Photos and Folders

Based on the discussions held after filling in the pre-test form, on average between 1 000 to 10 000 photos are stored in different devices including mobile phone, tablet device, PC, laptop, as well as on cloud services. Only two participants (N=2/15) mentioned that they rename some of the most important photos after transferring them to their PC or laptops. The rest of the participants (N=13/15) stated that they do not rename their photos as they considered renaming photos a daunting task. Additionally it is quite impossible and time consuming to rename every photo. "Usually, I transfer at least a hundred photos in one sitting, and renaming each photo is close to impossible task for me" (P#5). This task is also disliked as coming up with good names of the photos is quite difficult and may require a lot of thinking from the users. On the other hand, all the users except two mentioned that renaming is mostly limited to the photos folders. This renaming helps them in locating photos later on. None of the participants stated renaming the photos in their mobile or tablet device. Half of the participants also stated that the search feature for locating photos is never used as they lack trust and confidence in search and they are sure that using search to locate their photos would not yield any results. "I don't remember the time when I used local search for photos, as I am pretty sure I won't find anything through it…" (P#2). Therefore, to locate their photos they usually browse through their photos or try to recall the exact location of the photo.

Organizing Photos

Only one participant (N=1/15) stated that she actively organizes her photos on a regular basis, roughly once a month. Six participants said that they do it less often and probably around a couple of times a year, while the remaining eight stated that they do photo organization but on

a limited scale and rarely. Most of the organization is limited to creating new folders and naming them with some descriptive details that can be year, month, event name, person(s) name or a combination. After creating and naming folders, the photos are dragged and dropped or copy-pasted from the default folders to the newly created folders. Three participants also mentioned that during the process of organization, they delete duplicate photos by choosing the best one of the multiple photos taken of an instance. Two of the participants also mentioned that organizing photos can be fun, exciting, and sometimes surprising as they get a chance to explore the forgotten photos or to view photos with a fresh perspective. Opportunity to re-visit the old memories after a while is considered a secondary factor for organizing their photos: "… [it is] always a nice feeling to find some photos not seen for a long time… (P#12)".

Tagging of Pictures

All the participants (N=15) in our study knew about tagging photos. Tagging within the context of personal photos management is not actively practiced by the studied users. Even though all the participants were aware of the benefits of tagging, almost all the participants were reluctant to engage in tagging their photos. A majority of the participants stated that they do not get involved in tagging their photos as it is a time consuming activity. Furthermore, it seems daunting for the users to manually tag photos one by one as there are thousands of photos located in different devices. Fourteen participants mentioned that they have not used any system that automatically suggests tags. One of the participants who had used an automatic tagging system could not recall the application name as it was used couple of years ago. She tried the system out of curiosity to see the capability of automatic tagging. The user abandoned the application after using it for a while as the system did not perform according to her expectations.

Some respondents mentioned that the tagging feature is either not available or they have not actively tried to find out in their photo management application what restricts them from using the feature. One of the users stated "I think it [tagging] must be hidden somewhere inside the menu, should be there as it is one of the most commonly used [photo management] application (P#9)". Meanwhile, another participant commented "I know that my [photo management] application has this feature and [I] remember seeing it few times by mistake [laugh]…" (P#14). Another common reason pointed out by three participants was coming up with suitable tags. The participants expressed that thinking and producing good and relevant tags for each photo seems very difficult. Moreover, we noted that the participants were concerned about remembering the exact tags later on while searching for the photos. Additional reasons mentioned by one to three respondents for not tagging photos were laziness, not knowing how to perform tagging, and no perceived benefits of the activity. Thirteen participants stated that they never tag or give keywords to their photos. Only two participants engage themselves in tagging their photos but scarcely. Twelve participants said that they would tag their photos more often if relevant tag suggestions are provided to them. The remaining three participants stated that they would not like to tag their photos even if tag suggestions are offered. According to them, the tag suggestion provided by the system would not be relevant enough in organization and later retrieval of their photos.

Usage and Needs of the Photo Tagging System

Photo Tagging Feature

The tagging tasks were easily completed by all the participants (n=15/15), though for three of them it took slightly longer time to complete the tasks successfully due to limited familiarity with tagging practice. Most of the users mentioned that the current implementation of tagging feature makes the screen area too busy. Even though the appended and suggested

tags were visually separated in different sections, the feature was considered confusing. Four of the participants pondered the difference between appended and suggested tags. During the tasks, we also noticed that after adding a new tag or appending a suggested tag the participants expected feedback from the system. The current implementation did not display any feedback to the user after adding a tag. Missing feedback after appending the tag to a photo kept the users wondering whether the tag was added or not. On the visual appearance of tagging feature, around half of the respondents considered the visual icons representing "add suggested tag" and "remove tag" as boring, unattractive, and dull. This feedback from the participants points out the importance of using eye-catching, nicely designed and illustrative icons set in applications.

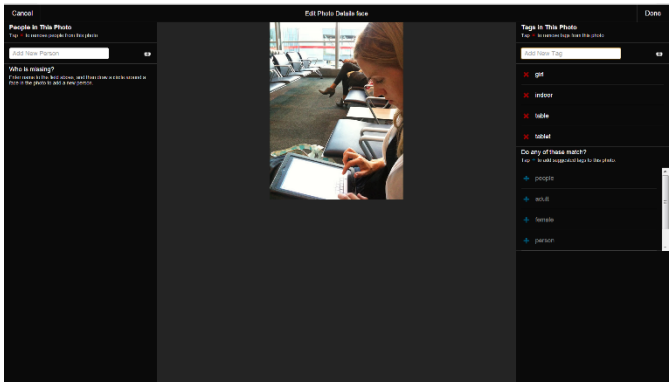

Figure 2: Appended tags and suggested tags on the right panels

After completing the tagging tasks, many participants expressed that the current implementation of tagging feature should be more playful and interactive. They regarded playfulness and interactivity as one of the core aspects that attracts them towards a tablet device. Participants also stressed that adding playfulness and interactivity would make the tagging feature, as well as the whole application, more appealing and fun to use. One of the users stated, "It [the current implementation] feels so dull, and do not give a feel of using on a tablet" (P#7). Another participant

commented, "It would be more enjoyable if I can drag and drop the tags on top of a specific object or person or to the picture…" (P#9).

Tags Suggestions Feature

Tags suggestion is a feature that suggests machine generated tags after analysing the contents of the image. At the time of testing, the system for tag generation was still in its infancy and sometimes it generated irrelevant or meaningless tags. Even though the feature did not always generate highly relevant and meaningful tags, the idea of tags suggestions was considered highly valuable by all the users. Two participants (n=2/15) were doubtful about the suggested tags as they felt that the automatically generated tags cannot comprehend the actual content of the picture. Additionally, three participants also pointed out concerns over the privacy of their photos. As already mentioned in the previous section, the users suggested that the system should automatically append the tags that are highly relevant and later request for confirmation. In case the users have already provided or confirmed tags to a photo, the system should automatically append these to highly matching photos or objects and request for users' confirmation later on. One of the participant stated "… it wouldn't harm if the system automatically tags my images and then ask my confirmation" (P#4).

Most of the respondents expressed that the evaluated system might not be intelligent enough to detect the photo contents, and in most of the cases irrelevant and useless tags will be generated. During the study, we also discussed the possible instances for confirming the tag suggestions. Slightly less than half of the participants responded positively about confirming or adding tags while the photos are being transferred to another device or uploaded to an online service. Tagging in those specific phases was deemed a far better alternative than tagging while in the browsing mode. One of the participant stated, "… suggesting tags while my photos get transferred might be a good option to start tagging" (P#15). We consider that the tagging activity during the uploading or

transferring phase was considered more acceptable as the users are already involved in a photo management activity and in that phase it might require less effort of them. Additionally, one of the participants also queried about the possibility of suggesting tags while the photo is being captured. "… I am not really sure whether it is possible to suggest tags as I am taking a picture… I believe this should be possible with mobile phone [camera]…" (P#3). Based on the discussion with the study participants, we believe that suggesting tags while the photo is being captured, "live tags", or instances when the photos are being uploaded or transferred, can be highly useful and effective ways to engage them in the tagging activity.

Keyword Search Feature

All the participants completed the keyword search related tasks with ease. Some participants noticed the variation in search views as currently two slightly different search user interfaces (UI) were implemented. In the basic search view the user inputs a term in the search box, while in the other UI the user can pick keywords from the displayed terms. Two different search alternatives in different view confused the participants. Our findings suggest that uniform interaction methods and design should be implemented across the application and explicitly for a specific feature. At least three participants queried about the auto suggest terms feature in the search box. As the feature is currently used commonly in various applications, we believe that the users would expect this feature in most of the cases where search is implemented. Another suggestion pointed out by one of the participants was that the system should suggest alternative or similar terms in case of no matches or search results. For instance, if the user searches for "van" in case of no results the system can suggest "wagon", "automobile", or "car" photos. Displaying closely matching results or relevant suggestions will boost users' confidence with search and will also help in adopting the system.

Similarity Search Feature

The similarity search feature enables the users to find photos that are similar to a selected one. The similarity search feature considers various features such as colour, shape and face, as well as other criteria. After objects extraction and analysis, the system displays a number of photos that are considered similar or matches closely the selected photo. All the participants were well aware of the screen tips feature when our discussion reached similarity search tasks. Though the similarity search feature was not self-evident, all the participants managed to complete the tasks by reading the displayed screen tip. The participants pointed out that similarity search is not an obvious or clearly evident feature, and they would have not been able to complete the tasks without getting assistance from the screen tip. Although the feature was not considered self-evident, all the participants regarded it easy to learn once used.

The reaction to the similarity search feature was mixed among the participants. Some of the participants stated that this feature can be very useful for locating all the photos of a similar place, person(s), or an object. This feature was also considered useful for detecting multiple shots usually taken as a backup, thus enabling them to easily delete multiple shots of the same photo and choose only the best one. Though similarity search was considered a useful feature by some, the majority were concerned about the relevancy of the similar photos. The participants stated that the feature might be useless if the results are not relevant enough.

During the sessions, we continuously realized issues due to improper or lacking feedback by the system. As noted earlier on, the participants expected a feedback while they carried out the similarity search tasks. More than half of the participants tapped the screen for a very long time, in some cases even up till ten seconds. The users expected and kept on waiting for some sort of initiation by the system. The current system is designed such that long press (2 seconds) and then release initiates the similarity search feature. On the basis of feedback and user interaction, the feature should provide instant feedback upon long press by the user.

This feedback should inform the user that the search process has been initiated. We also noticed issues with the thumbs up/down icons used for providing relevancy feedback to the system. Some of the participants understood and related the icons with the actual context, while most of the participants associated the icons to like/dislike or favourite photo. One of the participants stated that, "…. It [thumbs up icon] reminds me of favourite images and not relevant images" (P#3).

To sum up, we believe that the expectation of the similarity search varied across different participants. Many of the users expected or would like to view similar photos based on the place, person, or objects in the photos. Meanwhile, some of the users expected to see all the other photos taken in the surrounding area or from the same trip, event, or occasion. It became evident from the responses that similarity search results simply based on texture and colour was not considered important at all. Some of the participants also pointed out the possibility to carry out similarity search based on a specific object or person in the photo. Extending this feature to multiple objects and persons was considered more practical and useful. As already witnessed in other tasks, the participants expected more interactivity such as moving photos to different positions according to their relevancy. Lastly, high relevancy of the similarity search results was considered as the key for adopting this feature.

Face Tagging Feature

SmartImages also includes a face tagging feature that allows the users to tag faces in their photos. All the respondents except two users faced some level of difficulty in completing the task, even after reading the screen tip multiple times. More than half of the participants were assisted by the moderator in completing the face tagging task. After completing the face tagging tasks it became evident that the main face tagging task (task 8) was not straightforward and the task required much effort from the participants. Despite the effort required from the participants, most of them liked the feature and considered it a "must have". In the current design,

face tags are located on the left panel of the displayed photos, while the photo tags and tag suggestions are located on the right panel. Some participants were puzzled by separate locations of face tags and photo tags and suggested to have the face and photo tags together. They argued that combining face and photo tags would make the UI simpler and create less screen clutter. Another argument for combining them was that the face tags space was considered useless in cases where there are no faces in the photo. On the other hand, a majority of the participants favoured separate face and photo tags sections. They felt that separate locations for face and photo tags keeps the clarity between different types of tags.

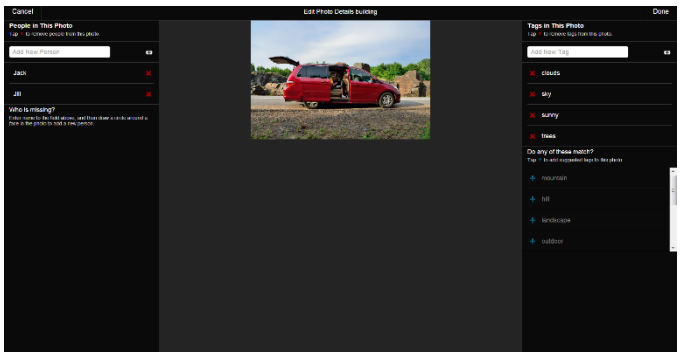

Figure 3: Face tags on the right panel

During our discussions, we explored various ideas related to face tagging. Most of the ideas are novel and if properly designed and implemented can change the way of photo search. For instance, one of the suggestions was to display all the photos of a particular person in the search results even if only one photo is face tagged. By implementing this feature, the users can easily confirm all the face tags for the displayed photos quickly and with ease. Another suggestion was the semi-automatic tagging of photos of famous people that can be helpful in some of the cases. For instance if the user has a photo with a famous actor, the system should automatically suggest the tags for that renowned

person. After suggesting the tags, the system should confirm those tags. This semi-automatic tagging can also be implemented for common objects, buildings, and places that will minimize the need of manual tagging by the users. Minimizing the effort required by the users to tag will motivate them to adopt and use the system with confidence.

General Findings about SmartImages

Main View

The main view of SmartImages consists of nine thumbnail photos with left/right swiping to move across a set of photos. All the participants (N=15) expressed contentment over this interface, as the photos are large enough to be viewed properly and in a glimpse they are able to scan through multiple photos.

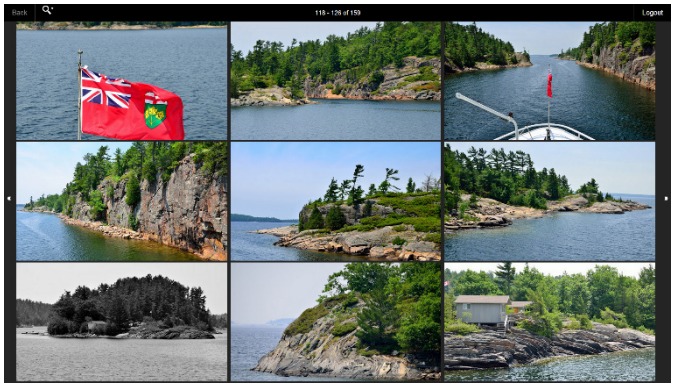

Figure 4: Main view of SmartImages

Displaying nine thumbnails as compared to having fewer photos in this view minimized scrolling through the photos and required less effort from the participants to locate a specific photo. Three users tried to zoom in/zoom out the main view, as they felt that it is quite natural to have this functionality in touch screen devices. They stated that zooming out to a certain extent will help them in viewing more photos in the

same view and thus will minimize the photo scrolling even more to locate photos. One-fourth of the participants noticed that the image thumbnails are being cut-off, and in some of the cases a very important aspect of the photo was not visible.

Navigation

Sideway scrolling is implemented for scrolling or navigating through the thumbnails as this approach is more aligned with the tablet devices interaction. Eleven participants (N=11/15) initiated the navigation by sideways scrolling, while the remaining four tried navigating down-to-upwards scrolling. Four participants who started with down-to-upward scrolling quickly realized the actual orientation for navigating through the thumbnails. The arrow icon on the right side of the screen also assisted them in realizing the actual navigation orientation. After using the sideways navigation for a while, all the respondents mentioned that it would be more natural to have sideways navigation in landscape mode, while down-to-upwards navigation will feel more natural in portrait mode. We also noticed that most of the participants that initially navigated down-to-upwards have limited experience with tablet devices or do not own a tablet device and they were more aligned with the smartphone interactions.

While navigating through the thumbnails important issues were raised by many respondents. These issues were not given due importance while designing the application as it was deemed for testing purposes. Swiping effects were not implemented in the application and the transitions between different views were not smooth enough. Additionally, some users also showed their dissatisfaction over lack of clarity in various visual indicators, and sluggish system responsiveness. The developers had considered these issues as minor and they had been mostly neglected. After analysing the participants' interaction and discussions with them, the significance of these issues became more important.

Screen Tips

In SmartImages, we implemented the screen tips feature in various views. The main aim of the screen tips was to introduce and familiarize the users with different features and functionalities by providing simple instructions. These instructions were meant to guide and assist the users in achieving the tasks effectively. Although this feature was noticed by most participants, a few paid close attention and read through the instructions. Three participants completely missed the screen tips feature. These participants expressed that the main reason for missing out the screen tips feature was its small size, low visibility due to poor contrast, and users' focus on the actual tasks.

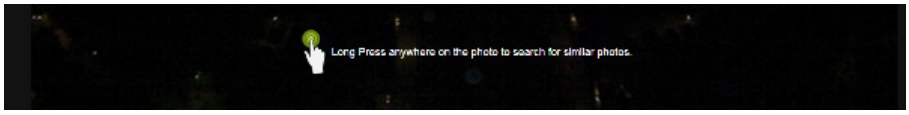

Figure 5: Screen tip visible at the bottom of the full view image

After completing various tasks and getting familiar with the screen tips feature, all the participants considered the feature highly useful and helpful in getting to know about various functionalities of the system as well as in easily completing the tasks.

Discussion and Implications

Results from our study strongly support numerous findings from previous research on personal photography practices. The results from the background questionnaire confirm the findings from literature that the number of photos captured and stored by the users has grown remarkably (Kirk *et al*. 2006; Neustaedter & Fedorovskaya 2009; Whittaker, Bergman & Clough 2010; Sarvas and Frohlich 2011) and almost all of them face difficulties in managing their photos. It is evident that the number of personal digital photos is increasing at such a pace that the users cannot easily organize, search, and locate their photos without

spending substantial effort and time. Multiple capturing devices, online and offline locations where the photos are saved, and the number of sources a typical user receives photos from are also adding further complexity. Similarly, all the participants agreed that in the near future, managing their photos will get even more difficult due to the growing number of photos captured, shared, and saved on multiple devices and platforms. This fact strongly backs up the findings by other researchers (Frohlich *et al*. 2002; Ren, Sarvas & Ćalić 2010; Whittaker, Bergman & Clough 2010).

Based on the discussions with our study participants, we can clearly see the preference of using mobile phones as a primary capturing device. Immense integration of mobile handsets in society, higher quality of photos, increasing storage, capabilities of easy and quick sharing with family and friends, as well as on social network services, appear as reasons behind this preference over other capturing devices (Ames & Naaman 2007). The "always with me" nature of mobile phones also complement the usage as a main capturing device. Mobile phones also make it possible to view the captured photos anywhere and anytime as well as show to others in collocated settings (Van House 2009).

Meanwhile, the participants expressed that digital cameras are meant for taking photos mostly on scheduled and special events such as birthdays, picnics, and trips abroad. Digital cameras are mostly considered a device for serious photography when compared to mobile cameras and tablet devices. Due to high picture quality, advanced features, and better device handling digital cameras were favoured over the other two devices. Interestingly, the tablet device was deemed as the least favoured device for capturing photos. The main reason for not capturing photos with a tablet device is due to the nature of the device as well the setting in which the device is usually used. The participants mostly use their tablet devices while they are at home, commuting in public transport, or in a library or classroom. In the above mentioned settings they usually do not take that many photos. If they have to capture a photo in those settings,

they still preferred using mobile camera or digital camera over the tablet device as one of the users stated, "… it [the tablet device] is not designed for taking photos" (P#9). Additionally, as compared to mobile phone and digital cameras, tablet devices offer inferior picture quality and device handling while taking the photos.

The current study seeks to fill the existing gap pointed out by researchers that emphasises the need of user-centric approach in understanding users, their needs and requirements from the personal photography practices (Van House *et al*. 2005; Kirk *et al*. 2006) with a specific focus on tagging and related features (Wilhelm *et al*. 2004; Tsai & Hung 2008). We believe that tagging photos is one of the important features that offer solutions which can potentially solve many issues related to the organization, searching, and location of photos (Frohlich *et al*. 2002; Kustanowitz & Shneiderman 2005). However, the subjects of our study still considered tagging as an extra task in taking and managing photos. Therefore, we expect that the solution for smooth organising of digital photos is a more elaborate combination of automated image feature extraction and manual tagging.

In accordance with prior research, our results also indicate that even though people are aware of the potential benefits of tagging photos in personal photo collection (Kustanowitz & Shneiderman 2005), the activity is not practiced widely as most of our study participants do not engage with photo tagging activities (Kirk *et al*. 2006; Vennelakanti *et al*. 2011). Even though the benefits of tagging are considered obvious by many participants, the feature has not been fully utilized by the software vendors or adopted by the users. Based on the findings, we propose strongly to develop photo management solutions that consider the issues such as better visibility of the tagging feature and seamless integration to the services. The benefits and value of tagging needs to be highlighted clearly by the software vendors as some of the users are currently unaware of the tagging feature and its potential benefits.

Though none of the participant completely favoured fully-automatic tagging, they expected and wished for automatic tagging to a certain extent together with collaborative tagging as proposed by some of the previously developed tagging solutions including Zonetag and EasyAlbum (Cui *et al*. 2007; Naaman & Nair 2008). For instance, one of the participants coined the idea of "… an option that allows the system to automatically tag the same objects and persons in other photos" (P#3). Based on the discussion, we tend to agree that automating the tagging feature to a certain extent can minimize the effort required by the users (Xie *et al*. 2010) and at the same time it can enhance the confidence and adoptability of the system.

Our findings propose that the suggested tags should not give a feeling that the terms have been generated by the machine. A feeling of human touch is considered as the core for adopting and willingness to use the suggested tags feature. The respondents stated that a tags suggestion feature will motivate them to get involved and use the tagging feature. Interestingly, we also noted that many participants pointed out the importance and acceptability of social aspects of the tagging feature e.g. tag suggestions generated by other people (Naaman & Nair 2008; Nov & Ye 2010). Participants were more willing to adopt the tags suggestions provided by other people instead of tags generated by the system. Users' own tagging pattern and their tagging history were also considered crucial as potential sources for tag suggestions. Furthermore, tag suggestions based on the location of photo capturing was also considered important by many participants (Cui *et al*. 2007; Wang *et al*. 2012). Participants were more willing to get tag suggestions of exact locations such as "Alepa grocery store", "Aalto University Library", or "Otaniemi gym" instead of general terms such as building, place, or people. One participant also suggested that tag suggestions based on the activity or an event can also be important in some cases. For instance, suggesting tags of an event or activity taking place where the photo is captured, for example, "Helsinki sailing event 2013", "night of the arts 2013", or "Espoo day 2013".

Regarding the priority of tags suggestion sources, we believe that a combination of tag suggestions based on users' own pattern, location, and other people in the service as well as the system can provide more meaningful and relevant tags. As an alternative, offering the users an option to choose the source of tag suggestion can also be considered. The system should analyse the image contents and base the tag suggestion according to the users' preference of various sources.

We also realized that the tagging activity during the uploading or transferring phase was considered more acceptable as the users are already involved in a photo management activity and in that phase it might require less effort from them. Based on the discussion with the study participants, we believe that suggesting tags while the photo is being captured as "live tags" or instances when the photos are being uploaded or transferred can be one of the highly effective and interesting options to engage them in the tagging activity.

Discussion with the study participants also points out the importance of the face tagging feature. Various ideas and suggestion related to this feature were also explored in detail. Implementing these ideas and suggestion can make the face tagging feature more appealing and usable for the users. Considering the feedback from the participants, it also became evident that the screen tips feature can be highly helpful in assisting the users. Screen tips can introduce the concepts that are new and difficult to understand. The feature can also make users confident and comfortable while trying new application features. Based on the users' feedback, we would recommend that the screen tips feature should be utilized by the applications having novel and uncommon features as well as interaction methods. The feature should be prominent and stand out more to seek users' attention. We also noted that the continuous appearance of screen tips should be avoided, as excessive appearance might divert the users' attention from the actual task. We would recommend that the screen tips feature should automatically hide once the user gets familiar with the

task in hand, or by providing an option to enable/disable the feature through the application settings.

We also found out a few issues specific to tablet device interactions. Applications designed for tablet devices should be promptly responsive to users input and provide appropriate feedback to the users. For instance, while performing the similarity search many users tapped the photo for a very long time and expected output. In another case of adding tags to a photo, no feedback was provided to the users and they kept on waiting for the feedback. We also realized that the users expected playfulness and interactivity as one of the core aspects that attracts them towards a tablet device. While using an application on a tablet device they look forward to novel interactions and design elements. Guiding clues and visual indicators such as scroll bars and loading/progress bar should be designed carefully. Clear and simple clues assist the users in completing their tasks successfully. In our task analysis study, some of the users were puzzled or not able to complete their tasks successfully due to unclear visual indicators. During our study, most of the respondents complained about the delayed response from the application. In addition, some of the respondents also pointed out lack of smoothness and effects during the screen/view transitions.

Adhering to interface style and reliability are few among the other important implications that can be highly beneficial for designers and developers irrespective of the application domain. For instance, minor issues such as swiping effects, smooth transitions, visual indicators clarity and fitting system responsiveness should perform accurately. We acknowledge that these issues should be given due importance even while designing applications for testing purposes as they play a vital role in guiding and assisting the users as well as promoting the adoptability of the system. Addressing these properly will ensure the test user satisfaction and confidence with the system. This is an important learning specifically from the HCI perspective for the application developers and designers

to consider while working toward concept evaluations of their applications.

Conclusions

In this paper, we introduced a tagging application SmartImages. The application allows the users to add photo and face tags manually, append tags from the list of suggested tags, and search for similar photos. Evaluating the application with 15 participants enabled us to understand the users' needs from a tagging application. During the evaluations, we managed to get detailed input about various features and functionalities that are essential for tagging applications. In addition to the answer to the research question, the evaluation method used in the research facilitated the discovery of many new ideas related to tagging. In general, the presented implications can also be utilized by tagging systems, imaging and other multimedia applications. Some contributions can be adopted by the touch screen applications as well.

It became evident that the users expect high level of interactivity and playfulness from the touch screens. These elements are given high importance and valued by the users. Many users feel quite natural to experiment with various interaction techniques in touch based applications. The elements of interactivity and playfulness is more needed in multimedia applications such as imaging, audio, and videos as these applications are mostly used for leisure and fun. Incorporating these elements into multimedia applications may enhance users' engagement and devotion. We also acknowledge that factors including tasks simplicity, engaging in a fun way, and automation to some extent can contribute to the acceptance and adoptability of photo tagging systems. The adoption of a tagging system lies in high level of accuracy and relevancy of the results. On the basis of our interactions with the participants, we conclude that the tagging system should require minimal effort from users and optimally the users should interact with the system to confirm the suggestions.

Relevancy of the results is crucial for applications that utilize search in any form. If the search results are not relevant, it is highly likely that the users will avoid using the search. This also holds true if the users think that their search term would not yield any results. During the similarity search tasks, the users pointed out that they would like to see targeted and to the point search results even if they are few. Instead of displaying hundreds of photos as search results it would be far better to show them a couple of results that match perfectly to their search criteria. On the other hand, if the search yields no results then displaying closely associated results can keep the users motivated with the search feature. For instance, if the user inputs "Playing" as the search term, and there are no photos tagged with the term, displaying "Football" photos as closely associated results can be helpful for the users.

As the users have become more experienced and comfortable with technologies as well as their experience with various devices and platforms have grown, they expect features and functions that are refreshing. They also expect that new features and functions can provide them with positive surprises. For instance, in SmartImages' case, the users were positively surprised with the similarity search feature and the interaction method used to conduct the similarity search.

Despite the number of feature and functionalities in SmartImages, we would like to point out some of the limitations of the application as well as the testing setup. Most importantly the participants used SmartImages with preloaded photos generated by the system, so the fact that the participants did not interact with their own photos as they normally would have could have influenced their opinions and reactions. Secondly, at the time of testing, a few of the generated tags seemed irrelevant or meaningless due to the fact that the system being was still in its infancy. This might have contributed negatively to the participants' opinions even though during the session we stated this issue to all of them. During the session, we also found out a few minor UI bugs that were not detected beforehand; these issues were reported instantly after the test session to

the development team and were fixed immediately before the next test session.

We believe that there is still much to investigate in improving the user interfaces as well as the overall user experience and adoption of photo tagging systems. Future work should investigate various tagging interfaces, comparing their acceptability and potential benefits with the users. Investigating the new ideas and concepts presented above can also be an interesting avenue for further research. Developing design guidelines for tagging systems in general and more specifically for photo tagging systems can also be highly beneficial for developers and designers of these systems.

Aqdas Malik is a doctoral candidate at Aalto University, School of Science, Department of Computer Science and Engineering and currently working with the Strategic Usability Group (STRATUS).

Marko Nieminen D.Sc. (Tech.) is Professor of Usability and User Interfaces at Aalto University, School of Science, Department of Computer Science and Engineering and heading the Strategic Usability Group (STRATUS).

Contact: aqdas.malik@aalto.fi
Web: cse.aalto.fi/en/research/groups/strategic_usability

Notes

1. Examples: Flickr, Google Picasa, Corel Digital Studio, Adobe Photoshop Album.
2. http://picsom.ics.aalto.fi/picsom/databaselist
3. http://muvis.cs.tut.fi
4. http://www.pv.com

References

AMES, MORGAN & MOR NAAMAN (2007). "Why We Tag: Motivations for Annotation in Mobile and Online Media." *CHI '07: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, April 28–May 3, 2007*. New York: ACM. 971-980.

AMES, MORGAN ET AL. (2010). "Requirements for Mobile Photoware." *Personal and Ubiquitous Computing* 14.2: 95–109.

BAR-ILAN, JUDIT ET AL. (2008). "Structured Versus Unstructured Tagging: A Case Study." *Online Information Review* 32.5: 635–647.

BAR-ILAN, JUDIT ET AL. (2012). "Tag-Based Retrieval of Images Through Different Interfaces: A User Study." *Online Information Review* 36.5: 739–757.

BENYON, DAVID & CATRIONA MACAULAY (2002). "Scenarios and the HCI-SE Design Problem." *Interacting with Computers* 14.4: 397–405.

BESMER, ANDREW & HEATHER RICHTER LIPFORD (2010). "Moving Beyond Untagging: Photo Privacy in a Tagged World." *CHI '10: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, April 10–15, 2010*. New York: ACM. 1563–1572.

BEYER, HUGH & KAREN HOLTZBLATT (1997). *Contextual Design: Defining Customer-Centered Systems*. San Francisco, CA: Morgan Kaufmann.

BLAŽICA, BOJAN, DANIEL VLADUŠIČ & DUNJA MLADENIĆ (2013). "A Personal Perspective on Photowork: Implicit Human–Computer Interaction for Photo Collection Management." *Personal and Ubiquitous Computing* 17.8: 1787–1795.

CARROLL, JOHN M., ED. (1995). *Scenario-Based Design: Envisioning Work and Technology in System Development*. New York: Wiley

CUI, JINGYU ET AL. (2007). "EasyAlbum: An Interactive Photo Annotation System Based on Face Clustering and Re-ranking." *CHI '07: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, April 28–May 3, 2007*. New York: ACM. 367–376.

DIAPER, DAN. (1989). "Task Analysis for Knowledge Descriptions (TAKD): The Method and an Example." *Task Analysis for HumanComputer Interaction*. Ed. Dan Diaper. Chichester: Horwood. 108–159.

DONG, WEI & WAI-TAT FU (2010). "Cultural Difference in Image Tagging." *CHI '10: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, April 10–15, 2010.* 981–984.

DUMAS, JOSEPH S. & JANICE REDISH (1999). *A Practical Guide to Usability Testing*. Rev. ed. Exeter: Intellect Books.

FREEMAN, MARK & ALISON FREEMAN (2011). "Online Grocery Systems Design through Task Analysis." *Journal of Enterprise Information Management* 24.5: 440–454.

FROHLICH, DAVID ET AL. (2002). "Requirements for Photoware." *CSCW '02: Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work, New Orleans, LA, USA, November 16–20.* New York: ACM. 166–175.

GOLBECK, JENNIFER, JES KOEPFLER & BETH EMMERLING (2011). "An Experimental Study of Social Tagging Behavior and Image Content." *Journal of the American Society for Information Science and Technology* 62.9: 1750–1760.

GOZALI, JESSE PRABAWA, MIN-YEN KAN & HARI SUNDARAM (2012). "How Do People Organize Their Photos in Each Event and How Does it Affect Storytelling, Searching and Interpretation Tasks?" *JCDL '12: Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries, Washington, DC, USA, June 10–14.* New York: ACM. 315–324.

HACKOS, JOANN T. & JANICE REDISH (1998). *User and Task Analysis for Interface Design.* New York: John Wiley & Sons.

HALASCHEK-WIENER, CHRISTIAN ET AL. (2006). "A Flexible Approach for Managing Digital Images on the Semantic Web." *SemmAnnot 2005: Proceedings of the 5th International Workshop on Knowledge Markup and Semantic Annotation, Galway, Ireland, November 7.* 49–58.

HECKNER, MARKUS, MICHAEL HEILEMANN & CHRISTIAN WOLFF (2009). "Personal Information Management vs. Resource Sharing: Towards a Model of Information Behavior in Social Tagging Systems." *ICWSM '09: Proceedings of the Third International Conference on Weblogs and Social Media, San Jose, CA, USA, May 17–20.* Menlo Park, CA: AAAI Press. 42–49.

HEIKKINEN, JANI ET AL. (2013). "Mobile Devices as Infotainment User Interfaces in the Car: Contextual Study and Design Implications." *MobileHCI '13: Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services, Munich, Germany, August 27–30.* New York: ACM. 137–146.

HO, CHIEN-JU ET AL. (2009). "KissKissBan: A Competitive Human Computation Game for Image Annotation." *HCOMP '09: Proceedings of the ACM SIGKDD Workshop on Human Computation, Paris, France, June 28, 2009.* New York: ACM. 11–14.

KIM, YONG-MI & SOO YOUNG RIEH (2011). "User Perceptions of the Role and Value of Tags." In *CHI '11: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, USA, May 7–12.* New York: ACM. 671–674.

KIRANYAZ, SERKAN ET AL. (2003). "MUVIS: A Content-Based Multimedia Indexing and Retrieval Framework." *ISSPA 2003: Proceedings, Seventh International Symposium on Signal Processing and Its Applications, Paris, France, July 4, 2003.* New York: IEEE. Vol. 1: 1–8.

KIRK, DAVID ET AL. (2006). "Understanding Photowork." *CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Montréal, Canada, April 22–27.* New York: ACM. 761–770.

KUSTANOWITZ, JACK & BEN SHNEIDERMAN (2005). "Motivating Annotation for Personal Digital Photo Libraries: Lowering Barriers While Raising Incentives." College Park, MD: University of Maryland.

LAAKSONEN, JORMA ET AL. (2000). "PicSOM–Content-Based Image Retrieval with Self-organizing Maps." *Pattern Recognition Letters* 21.13: 1199–1207.

LEE, CHEI SIAN ET AL. (2009). "Tagging, Sharing and the Influence of Personal Experience." *Journal of Digital Information* 10.1: 1–15.

MATUSIAK, KRYSTYNA K. (2006). "Towards User-Centered Indexing in Digital Image Collections." *OCLC Systems and Services* 22.4: 283–298.

NAAMAN, MOR & RAHUL NAIR (2008). "ZoneTag's Collaborative Tag Suggestions: What Is This Person Doing in My phone?" *IEEE Multimedia* 15.3: 34–40.

NEUSTAEDTER, CARMAN & ELENA FEDOROVSKAYA (2009). "Capturing and Sharing Memories in a Virtual World." *CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, April 4–9.* New York: ACM. 1161–1170.

NOV, ODED & CHEN YE (2010). "Why Do People Tag? Motivations for Photo Tagging." *Communications of the ACM* 53.7: 128–131.

OLSSON, THOMAS ET AL. (2013). "Expected User Experience of Mobile Augmented Reality Services: A User Study in the Context of Shopping Centres." *Personal and Ubiquitous Computing* 17.2: 287–304.

PEESAPATI, S. TEJASWI, HAO-CHUAN WANG & DAN COSLEY (2010). "Intercultural Human-Photo Encounters: How Cultural Similarity Affects Perceiving and Tagging Photographs." *ICIC '10: Proceedings of the 3rd International Conference on Intercultural Collaboration,* New York: ACM. 203–206.

QIN, CHUAN ET AL. (2011). "Tagsense: A Smartphone-Based Approach to Automatic Image Tagging." *MobiSys '11: Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services, Washington, DC, USA, June 28–July 1*. New York: ACM. 1–14.

REN, KAN, RISTO SARVAS & JANKO ĆALIĆ (2010). "Interactive Search and Browsing Interface for Large-Scale Visual Repositories." *Multimedia Tools and Applications*, *49*.3: 513–528.

RODDEN, KERRY & KENNETH R. WOOD (2003). "How Do People Manage Their Digital Photographs?" *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems, Fort Lauderdale, FL, USA, April 5–10.* New York: ACM. 409–416.

SARVAS, RISTO & DAVID M. FROHLICH (2011). "Digital Photo Adoption." *From Snapshots to Social Media-The Changing Picture of Domestic Photography*. London: Springer. Ch. 6: 103–137.

SEN, SHILAD ET AL. (2006). "Tagging, Communities, Vocabulary, Evolution." *CSCW '06: Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work, Banff, AB, Canada,* November 4–8. New York: ACM. 181–190.

STROHMAIER, MARKUS, CHRISTIAN KÖRNER & ROMAN KERN (2012). "Understanding Why Users Tag: A Survey of Tagging Motivation Literature and Results from an Empirical Study." *Web Semantics: Science, Services and Agents on the World Wide Web* 17: 1–11.

TAPPAN, JACQUELINE ET AL. (2011). "Mobile Application for Utility Domains." *MobileHCI '11: Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, Toronto, Canada, September 23–26.* New York: ACM. 521–524.

TSAI, CHIH-FONG & CHIHLI HUNG (2008). "Automatically Annotating Images with Keywords: A Review of Image Annotation Systems." *Recent Patents on Computer Science* 1.1: 55–68.

VAN HOUSE, NANCY A. (2009). "Collocated Photo Sharing, Story-telling, and the Performance of Self." *International Journal of Human-Computer Studies* 67.12: 1073–1086.

VAN HOUSE, NANCY ET AL. (2005). "The Uses of Personal Networked Digital Imaging: An Empirical Study of Cameraphone Photos and Sharing." *CHI EA '05: Extended Abstracts on Human Factors in Computing Systems; The SIGCHI Conference on Human Factors in Computing Systems, Portland, OR, USA, April 2–7.* New York. 1853–1856.

VENNELAKANTI, RAMADEVI ET AL. (2011). "The Picture Says it All! Multimodal Interactions and Interaction Metadata." *ICMI '11: Proceedings of the 13th International Conference on Multimodal Interfaces, Alicante, Spain, November 14–18.* New York: ACM. 89–96.

VON AHN, LUIS & LAURA DABBISH (2004). "Labeling Images with a Computer Game." *CHI '04: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vienna, Austria, April 24–29*. New York: ACM. 319–326.

WANG, MENG ET AL. (2012). "Assistive Tagging: A Survey of Multimedia Tagging with Human-Computer Joint Exploration." *ACM Computing Surveys (CSUR)* 44.4: art. 25.

WASH, RICK & EMILEE RADER (2007). "Public Bookmarks and Private Benefits: An Analysis of Incentives in Social Computing." *Proceedings of the American Society for Information Science and Technology* 44.1: 1–13.

WHITTAKER, STEVE, OFER BERGMAN & PAUL CLOUGH (2010). "Easy On That Trigger Dad: A Study of Long Term Family Photo Retrieval." *Personal and Ubiquitous Computing* 14.1: 31–43.

WILHELM, ANITA ET AL. (2004). "Photo Annotation on a Camera Phone." *CHI '04 Extended Abstracts on Human Factors in Computing Systems; The SIGCHI Conference on Human Factors in Computing Systems, Vienna, Austria, April 24–29*. New York: ACM. 1403–1406.

XIE, LEXING ET AL. (2010). "The Accuracy and Value of Machine-Generated Image Tags: Design and User Evaluation of an End-To-End Image Tagging System." *CIVR '10: Proceedings of the ACM International Conference on Image and Video Retrieval, Xi'an, China, July 5–7*. New York: ACM. 58–65.