Sunrise: Towards Location Based Clustering For Assisted Photo Management

Kharsim Yousef
Department of Computer Science
University of Bath,
Bath, BA2 7AY.
kam26@bath.ac.uk

Eamonn O'Neill
Department of Computer Science
University of Bath,
Bath, BA2 7AY.
eamonn@cs.bath.ac.uk

ABSTRACT
We present an end to end system 'Sunrise' to assist users in using location based data to tag, cluster, and find digital photos. We suggest that location tagging of photos is a common desire of users but existing manual solutions are too time consuming and complex. Our 3-tier system offers a solution that integrates automated location tagging with existing manual approaches. The mobile component of our system captures GPS based data at user defined intervals to create trail logs of the user's movements. The desktop component automates location tagging of the photos and makes use of the GPS location data to provide assisted photo key-wording through a process of reverse geo-coding. The server component uses automated location clustering techniques to produce hierarchical location representations, to simplify photo navigation and generate temporal/spatial data combinations enabling assisted photo management.

Categories and Subject Descriptors
H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *User-centered design, Prototyping.*

General Terms
Design, Reliability, Experimentation, Theory, Verification.

Keywords
Geo-tagging, photo sharing, camera phone.

1. INTRODUCTION
Memories for life has become an important topic related to Life Caching[1], from capturing and collecting various information relating to a user's daily activities such as journal entries, photos and movies, to sharing, exchanging and interacting with others via blogs [2, 12]. Preserving such material is far from new; generations of people have kept diaries, scrapbooks and photo albums from the first permanent photograph in 1826 to modern digital albums. Today, digital photos have become increasingly easy to capture, and we are no longer constrained by the 24 or 36 exposures provided by traditional film cameras. A single memory card allows a user to capture hundreds of photos. Yet this very capacity has led to a problem of users storing many images with very similar obscure names, with consequent difficulties in managing the collection. We therefore need to look beyond the capture and storage capacities offered by current digital cameras, to ways in which we can simplify the process of managing the many thousands of images that an individual may take over potentially many years.

There has been much research into the importance of location in managing photos [17], the addition of metadata to images [1, 5, 11, 13], and methods to use GPS in automating the process of location tagging images [16, 4]. Much of this research is focused on the creation of map based interfaces for photo browsing. Our research focuses on GPS tagging as an enabler to provide enhanced features in addition to building map interfaces, such as automating street level key-wording of photos, enhancing search and enabling location based photo clustering.

In this paper we examine existing approaches to photo management and tagging, identify current limitations and new techniques that can be used to help us make the leap from the current approaches to solutions that can semi-automate photo management. In Section 2, we look at the process of manually associating information with existing photos, current photo management solutions and their limitations. Section 3 presents our implementation of an end to end system that enables the automated tagging of photos and integration with existing manual tagging applications. Section 4 presents an initial evaluation of our implementation and its use.

2. EXISTING APPROACHES
Many of us already have substantial collections of digital photos, many of them rejoicing in names as obscure as P1000806.jpg. This obscurity makes finding a particular photo very difficult even within a recently taken set of images, and approaching impossible in the very large collections that many people build up over time. Standard computer file management systems provide little help in this respect. Viewing either lists of obscure names or thumbnails of the images is little help to a human searching very large collections of images. Image recognition technology is also

[1] Life Caching: http://trendwatching.com/trends/life_caching.htm

ICMI '07 Workshop on Tagging, Mining and Retrieval of Human-Related Activity Information, November 15, 2007, Nagoya, Japan

of little use since it is typically impossible for the user to specify the target image in terms that could be used to enable an effective search. Users tend to describe photos, often vaguely, in terms of events and subjects rather than in terms of pixels.

Tagging, or the process of associating metadata to images, is becoming an increasingly common approach to addressing this problem. Tagging is essentially an exercise in library science; if the user is able to assign meaningful tags to each photo, it will be easier subsequently to find a particular photo. The potential effectiveness of tagging has made it the primary technique used by most current image management products. In this section we first report a survey that we conducted to better understand users' photo management habits and the needs of users. We then present an empirical study of a typical manual tagging process using existing applications. This was conducted to further our understanding of current limitations and to suggest requirements for future solutions.

2.1 Manual Photo Tagging Survey

Here we report the results of a user survey on Photo Management and Tagging which was completed by 50 participants, mainly university students and staff, with the majority aged between 20-29 and 30-49 respectively. The survey investigated users' experiences with managing digital photo collections, including the approximate number of photos maintained by survey participants, their experience of using photo management software and their use of various kinds of tags.

The number of digital photos maintained by our survey participants ranged from participants who have fewer than 500 digital photos, whom we classed as beginners, to those who maintain more than 6,000 digital photos, whom we classed as advanced users. The resulting graph of user experience is positively skewed, with the majority of participants (68%) maintaining over a thousand photos, which we classed as moderately experienced.

Table 1. Tagging practices type/frequency of associated tags.

	Comments & Remarks	Place or Location	People's Name	Activity & Occasion	GPS Data
Always	0%	18%	4%	7.5%	0%
Mostly	15%	4%	15%	7.5%	0%
Neutral	15%	11%	4%	4%	0%
Sometimes	26%	11%	18%	22%	10%
Not at all	44%	56%	59%	59%	90%

Most (88%) of participants acknowledged the advantages of tagging their photos and some even suggested the need for additional tags not mentioned in our survey. User suggestions ranged from using a voice recorder capable of voice annotating photos at the time of capture to image analysis techniques that could automate the tagging of all their existing and future photos. All the participants acknowledged location as an important attribute of their photos and memories of events, but cited interaction complexity and time required as major disincentives to using photo management applications to tag photos with some form of locational data.

Table 2. Participants photo management software usage.

Always	Mostly	Neutral	Sometimes	Not at all
31.03%	13.79%	13.79%	10.34%	31.03%

Table 2 illustrates our survey participants' use of photo management applications. Despite the moderate experience levels of our survey participants and the survey results indicating that 77% of them used photo organizers that were fully capable of tagging photos, on average most participants added no tags whatsoever to their photos (see Table 1). The time taken to use existing manual tagging applications was frequently cited as the reason for this, although no measure of how it affected the various tagging stages could be derived from this survey. In the next section, we explore this issue through the process of tagging a large collection of photos using methods available to our survey participants.

2.2 Manual Photo Tagging Process

To better understand the process of manual photo tagging we selected a collection of 6,000 photos which we individually organized and tagged. The photo collection contained personal photo albums spanning a five year period and included photos taken during travel in Europe, America and Asia. The collection included 55 albums, each related to a particular location. The photo collection was particularly suitable to investigate our concerns due to the large number of images and their geographical dispersal.

Stage 1. Capturing photos: The initial stage, prior to our study, consisted of capturing the photos with a digital camera and transferring them to a computer's hard drive. At this stage, the images had only the default information associated with them by the digital camera and were not organized into any particular filing structure.

Stage 2. Folder organization: The second stage consisted of manually organizing each photo into individual folders based on a 'Continent/Country/City/Location' hierarchy. This approach was adopted to simplify the tagging process in the next stage. By starting with a location based hierarchy we could benefit from the photo organizer's automated 'import and tag' feature, in which each photo imported could be automatically tagged with its associated folder name, thereby simplifying the location tagging of each photo.

Stage 3. Photo organizer tagging: It is very difficult to predict how we might search for a given photo at an arbitrary time in the future. If we were to take a photo of a friend Sara at sunrise with an ocean view, we might tag that photo with (Sara, ocean view, sunrise). Then when we come to search for that photo in a few months, remembering "Sara", "ocean view" or "sunrise" will return a list of photos including the one we searched for. This form of subjective information is often an effective support to the human as "search engine" and typically forms the basis of manual tagging systems. To support this form of annotation in our collection, we imported all our photos into a dedicated photo organizer for manual tagging. Because spatial and temporal attributes are key indexical properties, we set out to tag these first. As the date is automatically embedded into most photos it required no additional work. Spatial tagging exploited the location based directory structure from the previous stage to aid the tagging process. Names of identified people such as friends, family members and colleagues, activities represented in photos (sports, events) and image descriptions were also added to each photograph to improve searching.

Stage 4. GPS tagging: In the final stage we imported the photos into a separate spatial tagging application to manually GPS tag each photo. The majority of applications that are targeted at managing and tagging photos, such as Photoshop Album, Corel Photo Album and Google's Picasa share the common user interface theme of a temporal approach to organizing and managing photos. Although there has been much research into the importance of geographic location in providing important contextual data to the user [e.g. 3, 8, 14], none of the current commercial applications have any spatial organizing capabilities, instead relying heavily on time based data and requiring users to manually create textual metadata tags to associate spatial or geographic properties with photos.

In comparison to true location tagging as GPS in Stage 4, textual tagging in Stage 3 does have its immediate advantages in simplicity. For example, it is easy to apply a tag such as "New York" to all photos relating to New York City, allowing users to easily search and find those photos in the future. But in the long term this represents merely a textual annotation rather than a physical location, and therefore is subject to linguistic, cultural and person-dependent attributes and interpretations that make it difficult to relate the textual identification to a physical world location [16].

Furthermore, textual tagging increases the complexity of searching algorithms required to correlate photos to spatial locations. For example, although most applications allow users to make "New York" a sub-tag of the Location tag set, it does not correspond to an actual physical location. If a search were performed to find pictures in the "United States", pictures that are tagged "New York" would not be found unless they were also tagged "United States". Hence we would need to rely on more complex textual analysis engines to make inferences, in turn affecting the precision and accuracy of search results. It is therefore important to have true location tags (e.g. lat/long) in addition to textual location tags (e.g. "Grandma's house") associated with photos, as they complement each other.

2.3 Manual Tagging Results and Implications

Our survey results indicated that people did little tagging of their photos, but the majority would have liked more tags. In our evaluation of current manual tagging techniques we found that a lot could be achieved with current solutions. Textual tags such as "Joe's girlfriend", "My Auntie's house" and "John's going away party" were well suited for manual tagging along with descriptions of people, pets, events and activity tags.

Textual tagging however becomes far more time consuming when physical location is involved due to the large number of tags and variants that can be used to describe a particular location (e.g. Continent, Country, State, City, Postcode, Street etc). GPS co-ordinate tagging further increases the time required due to the increased cognitive load to map photos precisely to real world locations. In our empirical study we quite quickly resorted to coarse grain GPS tagging (city scale 1-5km) in which, for example, photos that were taken across downtown Colorado were simply GPS tagged to a central point in downtown Colorado. Precise GPS tagging was found to be extremely time consuming and impractical when manually tagging large photo collections.

Users also indicated the need for automation, citing the time requirements associated with manual tagging as a limiting factor. In our empirical study we found that this varied widely, from 45mins per album to over three hours. This variation was influenced not only by the number of photos in a given album but also by the user's recollection of events, making it difficult to reliably measure the times needed.

Although it is not possible to provide the fully automated solutions many users would like, our survey and empirical study indicate both the importance of location tagging and the difficulties in undertaking it manually. With the rise of mobile devices and low cost GPS units, it is possible to develop at least partial solutions to assist users in location tagging their photos and making them easier to find and manage. In the following section we present an application that we have developed to support the automated GPS tagging of photos.

3. IMPLEMENTATION

In the previous section we looked at the manual tagging photo management process. Through our user survey and manual tagging study we showed the benefits of GPS tagging of photos but also the difficulties associated with the manual process of GPS tagging and commenting photos. An automated photo GPS tagging system has the potential to provide the benefits without the difficulties of a manual method. Here we present our implementation of such a system. Figure 1 presents a high level outline of our implementation architecture.

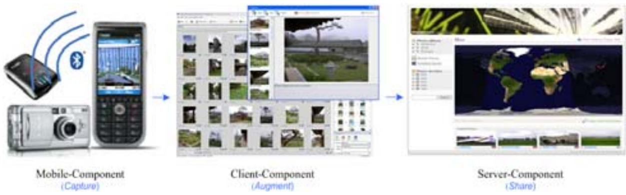

Figure 1. Automated clustering components: Mobile client, desktop and web components.

- Mobile Component. Provides a method to automate the process of GPS tagging newly captured photos. This component captures GPS based data at user defined intervals to create trail logs (GPX files) of the user's movements.
- Desktop Component. Automates location tagging of the photos and makes use of the GPS location data to provide assisted photo key-wording, e.g. "Paris trip photos", through a process of reverse geo-coding to obtain city/street level location data from the GPS lat/long co-ordinates supplied by the mobile component.
- Server Component. Makes use of the location data to run automated location clustering techniques to produce hierarchical location representations, with the aim of simplifying photo navigation and generating temporal/spatial data combinations for tagging.

3.1 Mobile GPS Data Capture Component

The ever increasing popularity of the mobile phone, with over two billion mobile subscribers worldwide, many of whom have their phone close to them most of the time, makes it an obvious platform for assisting users in capturing their locational data. Our GPS data capture component was built with the mobile phone in mind to perform the acquisition and logging functionality needed to enable features such as assisted key-wording and automated clustering in the desktop and server components. The primary task of the mobile application is to create a trail log of the user's movements using GPS data, for later association with the images captured through a user's own digital camera or built-in phone camera. There currently exist various commercial solutions for creating user trails based on GPS, from hardware based recreational GPS receivers from companies such as Garmin[2] and Magellan[3] to software based solutions available for smart phones and personal digital assistants. However due to our requirements for future flexibility and the relative simplicity of acquiring and logging GPS data we opted to develop our own solution.

The GPS trail logger was developed using C# NETCF2 for the Windows Mobile 5 platform, which simplified development and allowed us to deploy our application to our targeted phone and PDA form factors. The application was designed to allow the use of the device's built in GPS receiver if available or, currently more likely, to pair with an external GPS receiver over Bluetooth to create the user's movement log.

The output from our mobile component was in the form of a GPX[4] log file, based on the open standard XML data format for storing GPS data (waypoints, routes, and tracks). The GPX 1.1 schema was specifically selected because of its wide support in many free and commercial applications available today, allowing the trail data generated by our mobile component to be used by other applications outside of our implementation. Additionally it provided future proofing, by removing the need for the implementation as a whole to rely on the mobile component, as this could easily be substituted with a recreational GPS receiver. Their proprietary log outputs can be converted to the open GPX format through a number of freely available applications, making our mobile component independent by format from the functionality provided by both our desktop and server components.

The mobile application's interface and functionality is illustrated in Figure 2. The user is able to customize the logger by predefining the frequency of logging, which can either be based on timing (e.g. log every 5 seconds) or movement (e.g. log every 5 metres). Every entry logs a complete record of the current GPS data including latitude, longitude, elevation, heading, speed, and dilution of precision (DOP) data as outlined below.

3.1.1 GPX file format

```
<gpx version="1.0"><rte>
 <rtept lat="51.379102" lon="-2.328313">
  <ele>139.4</ele><time>2006-02-09T13:55:31Z</time>
  <course>22.0</course><speed>2.865455308</speed>
  <pdop>2.9</pdop></rtept>
</rte></gpx>
```

More frequent entries allow for improved trail accuracy but produce larger log files. Choice of frequency can therefore be determined by user needs and the device's memory limitations. The user interaction was designed with tourism applications in mind, whereby the user simply has to set a logging interval (optional) and press "Start". The application then runs in the background, allowing the phone to be used for all its normal functions.

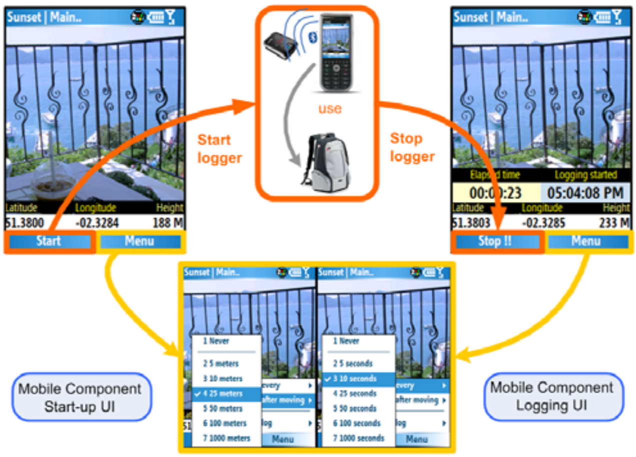

Figure 2. Mobile logger user interface, after the software is started, it will self manage itself.

The design itself was influenced by MAUI (Minimal Attention User Interfaces) [7]. MAUI focuses on reducing the number of user interactions that require direct user attention, in order to prioritize the user's engagement with real-world tasks. In a tourist scenario, this allows the user to walk around, enjoy her surroundings and the experience, and take photos, without having to think about the technology. When the user no longer requires the logging, simply pressing "Stop'" completes the GPX output and readies the application for another logging session.

3.2 Desktop GPS Translation Component

The desktop component makes up the middle tier of our implementation. Its purpose is to convert the user's movement trails (GPX files) collected by our mobile component or by third party solutions, to data that's more relevant to managing digital photos. It supports the following processes.

- GPS image tagging: Automates the process of GPS tagging each photo captured by the user, through cross-referencing the photos against the captured GPX data.
- Dynamic embedded tags: Extracts the data needed to use GPS beyond driving map interfaces, to enhance photo searching and clustering.
- Generating an online album: Combines all the generated data to enable the creation of an (XML) album metafile, which provides the data needed by our server component for importing, managing and organizing the digital images online.

---

[2] Garmin recreational GPS receivers: www.garmin.com

[3] Magellan recreational GPS receivers: www.magellangps.com

[4] GPX: the GPS Exchange Format: www.topografix.com/gpx.asp

3.2.1 GPS tagging images

The process of GPS tagging the images is performed after the user returns home, by initially transferring the images and GPX data to a desktop computer, then using our cross referencing application to match the timestamps of the captured images to the closest GPS positions at the time the photo was taken. The image capture times are extracted from the EXIF (Exchangeable Image File Format) data embedded in the photos; whereas the GPS time is extracted from the GPX file created by the phone. If the photos are taken with the phone's built-in camera, accuracy is assured. Our system also offers the flexibility of running the GPX logger on the phone while capturing the images with a separate digital camera, which typically is of higher quality than a phone camera. In this case, accuracy will be influenced by how well the phone's and camera's clocks are synchronized.

A few products, such as the Ricoh Pro G3, combine high quality camera functionality with built in GPS tagging of photos. This removes the need to check clock synchronization. However, this approach limits the user to the use of an expensive high end camera, instead of his existing devices of choice. In addition, it simply tags individual photos and does not support the collection of the trail data that enables the creation of memory maps by our server component [Figure 6].

Accuracy is also influenced by the logging frequency set by the user. For example, logging every two seconds will provide better accuracy than logging every 30 seconds. As noted above, there is a trade-off to be made here between the user's requirement for accuracy and the size of the GPX file generated.

We consider the GPS lat/long data to be a core embedded tag, along with important tags such as time-date, shutter speed, exposure and user defined tags such as "Nana's House" etc. These tags form an integral part of the photo ecosystem and allow for improved search ability. Dynamic tags on the other hand are more fluid in nature and can be created on the fly or periodically, based on embedded data contained in the images themselves and their embedded tags. For example, although GPS co-ordinates are important, they are often not as useful for "human powered" searching as textual data such as "Trafalgar Square" or "Oxford Street". Dynamic tags may be generated to convert the embedded GPS lat/long data to textual data (e.g. country, city, street name) relevant to the GPS position, thereby further improving search ability.

The major differentiator between embedded and dynamic tags is that dynamic tags are primarily derived from embedded tags. Therefore, embedded tags would be given priority in search results. Also, due to the external sources used in dynamic tags, there is a need to keep their contents fresh (even if only once in a while) to ensure that dynamic tags have up-to-date results to further improve their accuracy. For example, if we considered the image pixels as embedded data then an image analysis engine could produce various dynamic tags that could be associated with the photo. Unlike the static embedded tags, the dynamic tags could benefit from being updated as the engines used to generate their data are improved.

Dynamic tags can benefit greatly from the numerous Mashup web applications available today to provide the additional data needed to assist in tagging. For our implementation, we needed to translate the GPS data into more meaningful (to human users) address, place, city, state, and location information. This conversion process is known as "Reverse Geo-coding", as opposed to "Geo-coding" that translates a place name into lat/long location. In order to achieve this, we use the MLBS online reverse geo-coding service to combine GPS data from our images with street level data to produce our dynamic location tags. Each query to the server returned locational data consisting of the street name closest to the GPS co-ordinates as well as several surrounding street names, with associated city, province, post code and country data.

While such services are improving all the time, as databases are created and up-dated, they are not yet comprehensive. For example, reverse geo-coding conducted on GPS co-ordinates located inside the MGM Grand Las Vegas (Latitude: -115.170751, Longitude: 36.102004) produces data relating to surrounding street names, city data, country info etc, but does not actually return the MGM Grand itself as part of the resulting data, even though the GPS coordinates are based on a point inside.

```
Reverse Geo-coding Results (MGM): "65 E
Tropicana Ave, Las Vegas, NV 89109, United
States: 36.1009580279132, -115.170759067782;
Las Vegas Blvd S, Las Vegas, NV 89109, United
States: 36.100787273184, -115.170759067782; SR-
593, Las Vegas, NV 89109, United States:
36.100787273184, -115.170759067782; Island Way,
Las Vegas, NV 89109, United States:
36.100910216589, -115.16962525638;"
```

3.2.2 Generating the online album.

Having automated the process of associating GPS co-ordinates and other data with each image, our system also offers the user the opportunity to combine our automated data with any manual tags that the user has already created. Through a drag and drop process from within the user's photo organizing application or file explorer, our desktop component is able to extract embedded tags from within the user's images to augment our automated data, and is therefore complementary to existing applications [Figure 3].

Finally after all the data is collated, the system generates a concatenated album XML file, containing both data collected from the user's manual tags and automated system entries.

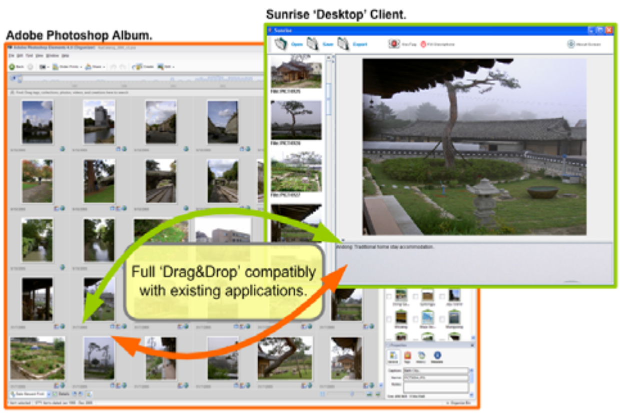

Figure 3. Roundtrip integration with existing software.

3.3 Server Sharing Component

The server component builds upon data generated by the mobile and desktop components and exploits the automated locational data to demonstrate the usefulness of physical space as a more prominent and usable feature of the photo management interface. The server component consists of a process to extract the data from the album XML file, populating a database and generating an online album to present the photos [Figure 4].

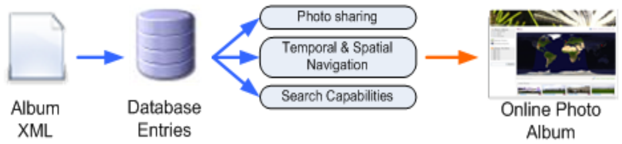

Figure 4. Online album creation process, XML data extraction and database population.

Having database entries consisting of images and associated metadata improves search ability but does little to affect the overall photo navigation experience. Search is typically used to find specific items, as opposed to general photo navigation, which is conducted through structured browsing and hierarchical menus [15]. Due to the availability of image timestamps from the very first digital cameras, building up temporal hierarchies for photo navigation has always been an available option. But other automated hierarchical solutions have not had such an impact due to the limitations of the collected data.

With the introduction of our automated approach we can demonstrate that this is no longer the case. Building on the automatically collected spatial data, it is possible to use the natural hierarchical structures associated with location data to create automated spatial hierarchies to enhance navigation. Furthermore, we can have automated clustering approaches that make use of combining both temporal and spatial data to produce new hierarchical structures. Figure 5 illustrates temporal, geographic and geo-temporal clustering of photos. We have fully implemented both the temporal and spatial navigation approaches and they are fully automated. We have also developed the geo-temporal "holidays" (New Years and Christmas in Figure 5) prototype to demonstrate one possible approach to using the combined data.

Figure 5. Example of automated clustering.

4. EVALUATION

In evaluating the system we focused on two main aspects: the accuracy of the captured data and the usefulness of the derived data in achieving our overall goal to simplify photo management. To test the accuracy of the mobile logging component, we equipped a number of participants with a Windows Mobile 5 phone (i-mate SP5m[5]) that was pre-loaded with our logging client. The logging frequency was preset to 5 seconds. We chose this setting to provide a high degree of accuracy while avoiding generating overly large files, as noted in Section 3.1. The participants were briefed on using the phone's functionality and the logger's interface. They were then asked to follow a pre-planned route through the city, in which they were encouraged to take any photos they wanted at any point on the route. The route itself took approximately 30 minutes to complete. Upon completion, the log data and the images were collected and put through the automated tagging process.

We developed a cross referencing web client that took the GPX output from Section 3.1 and the generated album XML file from Section 3.2 and combined them with the Google Maps API[6] to produce a map of the users' walking trails and image capture locations overlaid on a map of the city. Figure 6 shows the trail (red line) and image capture locations (black camera icons).

We found the GPS trail produced to be generally accurate, following both street layouts and the user's actual path. The GPS position did suffer slight inaccuracies such as two obvious moments in Figure 6 showing the users apparently walking through the river or jumping in for a swim, though this GPS inaccuracy is expected in a city environment.

Figure 6. Example of auto-generated memory map.

There was also a noticeable deviation mid way through the route [Figure 6: mid right]. When the users were shown the resulting map, they immediately commented on the deviation to be a point where they simply "lost their way". When asked if they preferred that trail section to be removed they wanted to retain it for its "comic value" Although the deviation was unexpected as the participants were city residents, it did highlight the need to

[5] i-mate: http://www.clubimate.com

[6] Google Maps API: http://www.google.com/apis/maps/

consider limiting or omitting trail data or parts of it when sharing or even storing such information.

Having evaluated the mobile logging and desktop photo/trail cross-referencing component for accuracy, we evaluated the usefulness of the resulting data in simplifying the process of sharing and managing photos. To do so, we ran a user-based evaluation in which we asked ten participants to use our prototype online album system [Figure 7] to perform a number of navigation tasks that involved finding photos by date, location and expressed user preference for either time or location-based navigation.

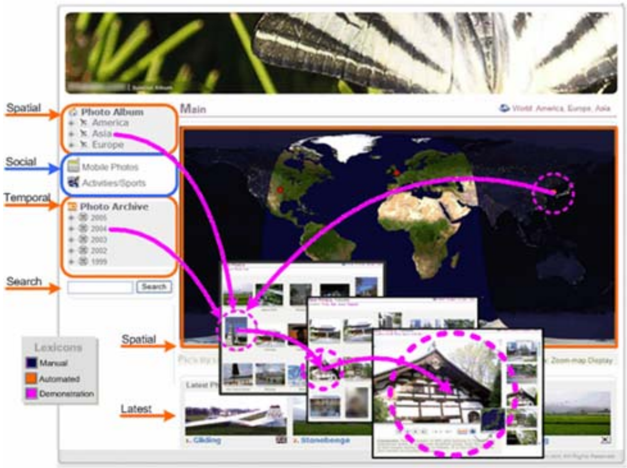

Figure 7. Clean clustered navigation interface.

The results did not indicate a significant user preference for time-based over location-based navigation. Participants preferred having both available, which may be attributed to the ways in which users' recollections of different events can vary. The additional tags and spatial structures provided users with more flexibility in finding photos. The user's recollection of a particular event influenced her choice of the most appropriate method (time, space, mapping or search functionality). Based on this evaluation, we can tentatively suggest that the addition of the location-based navigation features does aid in enhancing the user's experience.

5. DISCUSSION

Whilst we have managed to build an end to end system for, automating location tagging of photos and integrating with users existing tagging solutions, there are still many issues that need to be addressed. In designing the mobile component, we relied mainly upon GPS for obtaining the location data. However, given our choice of the mobile phone as the data capture platform, it is also possible to use cell-ID data to achieve similar results. This approach is less accurate than GPS but could nevertheless support the creation of the hierarchical spatial menus illustrated in Figure 5 as they were largely city scale. The main disadvantage would be the loss of precise locational information for map display and searching, but this approach would remove the need for the GPS hardware.

Although GPS and cell-ID logging have the advantages of being reasonably accurate and very unobtrusive to the user (when running in the background on a phone), they are merely an indication of where the picture was taken from. The combined picture/location represents the data from the viewpoint of the user taking the photo rather than the location of the picture's contents. Thus, inaccuracy is introduced in relation to the distance from the camera to the subject e.g. a picture taken of Madison Square Garden from a distance would not be tagged "Madison Square Garden" but by the location from which the picture was taken.

In addition, as seen in Section 3.2, limitations in the way the reverse geo-coding service worked further impacted the tagging, so that a picture acquired whilst inside Madison Square Garden would not be reverse geo-coded to "Madison Square Garden" but to the closest street: "West 31st St". This anomaly is due to the in-car navigation heritages of most of the reverse geo-coding services available.

As part of the desktop component, we managed to integrate with existing commercial products through a "drag and drop" approach to combining our automated tags with existing user tags [Figure 3]. This process could have been made more intuitive as it required the user to launch another application (Sunrise Desktop) in addition to the user's existing photo management application. Better integration through the use of application plug-ins may have helped streamline this process. Also, a better end user experience may be provided by the use of a wizard to guide users step by step in tagging newly imported photos and in acquiring required files such as the GPX trail and the album XML for automated tagging.

In our initial implementation, the online album interface was entirely self-generated from the uploaded XML files. This however constrained users to the richness of their tagged data and the logical structures for automation. On the other hand, user-created clusters such as "Me, my friends and Pizza" that are not necessarily logically structured or do not necessarily contain every photo relevant to the cluster, could benefit users by providing the means to create their own clusters of choice. In turn, such user-defined clusters could be used by the system to provide "training data" that is fed back into the system to create additional on the fly hierarchies. For example, if the user were browsing the "Me, my friends and Pizza" cluster, the system would create automated branched off clusters that allowed the user to find similar results (e.g. "Me, my friends and Food"), to drill down into results ("Me, my friends and Pizza, on Saturdays"), and to pull outwards ("Me, my friends").

6. CONCLUSION AND FUTURE WORK

Through a user survey and our own tagging studies, we found that location tagging has many potential benefits to end users in enabling assisted management of their photo collections; furthermore we identified the need for solutions that augment rather than replace current manual tagging practices.

In this paper we presented our end-to-end solution that enables the automation of GPS data collections, co-ordinate transformation into more meaningful (to humans) information for assisted photo key-wording and a flexibly modular photo navigation interface, that allowed users to find photos by auto-generated clusters, natural location data, keywords and map based navigation methods.

In understanding Photowork [9] the authors noted that participants in their study used simple organising principles for their filing of photos, mainly driven by time and events, and relied heavily on these in searching and browsing photos. In this paper we have shown that this merely creates textual annotations and therefore is subject to linguistic, cultural, person-dependent attributes and interpretations. However in combination with assisted metadata it is possible to utilise both resources to enrich sharing and browsing practices among users.

In the ever growing network and social environments of the Web and so called Web 2.0 phenomena, it has become more important for users not only to be able to search through photos they have taken themselves and are familiar with, but also through large archives of their friends' and family's photos, organized through multiple clustering techniques to simplify finding specific photos even if they remember only the location, time or event represented.

We have shown that automating the acquisition and tagging of locational data is feasible and does add value to the photo organizing process in time saved and simplicity for the user. More fundamentally, we have shown that we need to consider a shift from the current techniques in which photos are tagged after capture (having returned home) to a more on the spot approach to tagging.

The mobile client created as part of this research has since been freely released[i] to the public, supporting both Smartphone and Pocket PC devices. To date and within a short period of being released, it has achieved over nine hundred downloads and positive user and industrial feedback.

7. ACKNOWLEDGMENTS

Thanks to Chun Ho, Jonathan Cruickshanks, Keng-Ming Chang, Kieron Macknight and Tor Erik Linnerud for their contributions. We also thank Vodafone Group R&D for their support of this research.

8. REFERENCES

[1] Bederson, B. PhotoMesa: A Zoomable Image Browser Using Quantum Treemaps and BubbleMaps. Proceedings of the 14th annual ACM symposium on User interface software and technology (UIST 2001), pp.71-80.

[2] Bryant, L. 2005. Smarter, simpler, social: An introduction to online social software methodology 2005), *Retrieved April 5th*.

[3] Diomidis, D. S., Position-annotated photographs: a geotemporal web. IEEE Pervasive Computing. 2(2):72-79, 2003.

[4] Gemmell, J., Lueder, R., and Bell, G. 2003. The MyLifeBits lifetime store. In Proceedings of the 2003 ACM SIGMM Workshop on Experiential Telepresence (Berkeley, California). ETP '03. ACM Press, New York, NY, 80-83

[5] Girgensohn, A., Adcock, J., Cooper, M., Foote, J., and Wilcox, L. Simplifying the Management of Large Photo Collections. Human-Computer Interaction INTERACT'03, IOS Press, 2003; pp.196-203.

[6] Hammer LR. Further hoarding preferences in hamsters. psychon Sci 1972; 26: 139–40.

[7] J. Pascoe, N. Ryan, and D. Morse, "Using While Moving: HCI Issues in Fieldwork Environments," ACM Transactions on Computer Human Interaction, vol. 7, pp. 417-437, 2000.

[8] Kimber, D., Foote, J. and Lertsithichai, S. 2001. FlyAbout: Spatially Indexed Panoramic Video 2001), ACM Multimedia, 339–347.

[9] Kirk, D., Sellen, A., Rother, C. and Wood, K. 2006. Understanding photowork 2006), Proceedings of the SIGCHI conference on Human factors in computing systems, 761-770.

[10] Kjeldskov, J. 2002. "Just-in-Place" Information for Mobile Device Interfaces. In Proceedings of the 4th international Symposium on Mobile Human-Computer inter-action (September 18 - 20, 2002). F. Paternò, Ed. Lecture Notes In Computer Sci-ence, vol. 2411. Springer-Verlag, London, 271-275.

[11] Kuchinsky, A., Pering, C., Creech, M.L., Freeze, D., Serra, B., and Gwizdka, J. FotoFile: A Consumer Multimedia Organization and Retrieval System. In Proceedings of the SIGCHI conference on Human factors in computing systems (CHI 99). (Pittsburgh, Pennsylvania, 1999), ACM Press; pp. 496-503.

[12] Nardi, B. A., Schiano, D. J., Gumbrecht, M., and Swartz, L. 2004. Why we blog. Commun. ACM 47, 12 (Dec. 2004), 41-46.

[13] Rodden, K. and Wood, K. R. 2003. How do people manage their digital photographs? 2003), Proceedings of the SIGCHI conference on Human factors in computing systems, 409-416.

[14] Smith, T.R., Andresen, D., Carver, L., Dolin, R, Fischer, C., Frew, J., Goodchild, M., Ibarra, O., Kemp, R.B., Kothuri, R., Larsgaard, M., Manjunath, B.S., Nebert, D., Simpson, J.,Wells, A., Yang, T., Zheng, Q. A digital library for geographically referenced materials, IEEE Computer 29(5): 54-60, 1996.

[15] T-S. Lai, J. Tait and S. McDonald. "Image Browsing and Navigation Using Hierarchical Classification" Challenge of Image Retrieval, Electronic Workshops in Computing (eWiC) Newcastle upon Tyne, UK. 25-26 February 1999

[16] Toyama, K., Logan, R., and Roseway, A. 2003. Geographic location tags on digital images. In Proceedings of the Eleventh ACM international Conference on Multimedia (Berkeley, CA, USA, November 02 - 08, 2003). MULTIMEDIA '03. ACM Press, New York, NY, 156-166.

[17] Wagenaar, W. A. 1986. My memory: a study of autobiographical memory over six years 2 1986), *Cognitive psychology(Print)*, 18, 225-252.

---

[i] http://www.vodafonebetavine.com/web/SunsetGPSLogger/