Spaced Repetition Tool for Improving Long-term Memory Retention and Recall of Collected Personal Experiences

Norbert Györbíró  
University of Aizu  
ngyorbiro@acm.org

Henry Larkin  
University of Aizu  
research@logic.nu

Michael Cohen  
University of Aizu  
mcohen@u-aizu.ac.jp

ABSTRACT
A variety of electronic displays available for home use creates opportunities for intelligent applications. This paper presents a semi-passive photo reviewing tool for consolidating memories of experiences utilizing personal picture libraries. A form of spaced repetition algorithm is used to create visual journeys which link photos together around a user-chosen central theme. Systematically reviewing images from positive personal experiences can be useful to remember significant events, as well as to balance out stressful events in our lives. The design exploits existing digital home displays and aims to improve usage of media collections.

Categories and Subject Descriptors
H.1.2 [User/Machine Systems]: Human information processing; H.5.1 [Information interfaces and presentation (e.g., HCI)]: Multimedia Information Systems

Keywords
personal informatics, experience design, memory cue, spaced repetition, positive psychology

1. INTRODUCTION
To help remember significant information, we often take pictures and arrange them into collections. Photos can also be gathered and organized via personal lifelogs and social media websites, which may include contextual metadata such as location, participants, rating, and even emotional tags. However, memories and connections between places, events, and people can be difficult to recall. Memory recall in our brains can depend on several factors: emotional level, context, preservation or loss of information during encoding, etc. As time passes, memories are gradually forgotten or become altered, e.g., due to collision with newly encoded information [3].

Spaced repetition is a learning technique that reviews learning material at variable intervals of time, ideally just before such information would otherwise be forgotten, to maximize recollection and minimize learning time. It has been shown to enhance retention rate of information over longer spans of time. As such, it can be used to potentially boost long-term recollection of collected personal memories.

This paper is structured as follows. First, the research problem and outline of our approach based on photo journeys are introduced. In Section 3, a summary of photo collection attributes and spaced repetition systems is provided. Sections 4 and 5 present the algorithmic construction of photo journeys and the prototype implementation. Finally, Section 6 presents the conclusions and future direction of research.

2. MOTIVATION AND APPROACH
Automated lifelogging devices, such as the Microsoft SenseCam [5], a camera device worn around the neck which can capture a series of images based on movement or changes in light intensity, has been reported to be able to aid recollection of past events [9]. Sellen and Whittaker [8] state “However, the same study showed that the capacity for these images to help people recollect their past experience rapidly decreased after only three months, casting doubt on whether such devices can support long-term recollections,” concluding “we simply lack effective techniques for accessing digital archives.”

This paper presents a novel photo view-and-review personal information system to increase retention and recall of collected memories by utilizing the spacing effect in learning and presenting a context-aware selection of photos as the learning material. The main goal is recollection, i.e., thinking back in detail to past personal experiences, often called “episodic memories,” by automatically creating “visual photo journeys” according to one of the following:

Events (when) Events can consist of multiple sub-events and multiple participants. For example, a trip to Thailand may comprise several meals, day trips to the beach and forest, and with several people involved.

People (who) Our feelings regarding people in our lives are often drawn from shared experiences. By linking an image series for a person with experiences over varying periods of time and varying places, the connectivity associated with that person in our mind can be reinforced.

Places (where) A specific location or a wider area is defined. Some places—our hometown, childhood sports
field, a favorite café—can have many memories associated across varied people and events.

These dimensions were chosen based on usefulness ranking of retrieval cues. Psychological research on autobiographical memory [2] indicates that recollecting the past is dependent upon the kind of memory cue used, and suggests that places, events, and, people are stronger cues than time. A possible rationale could be that time is a more abstract concept and therefore does not have the same emotional or sensory stimulation capability. In a study on recalling and finding photographs from digital photo collections, Naaman et al [7] details that the most important cues are “who,” “where,” and “when,” in that order. However, regarding “when,” time of day ranks higher than the more general date.

Each visual journey contains a collection of pictures belonging to a user-selected slice through a multidimensional media library. Available media metadata is utilized to discover relations between photos and to vary the composition of journeys to increase the number and diversity of retrieval cues, thereby stimulating retention and recall of personal memories. Review times are scheduled for visual journeys. For effective learning, photos can be turned into memory cues by blurring or hiding parts of each image. Transformations are done automatically, e.g. by relying on the automatic face detection feature common in photo album applications and blurring the corresponding area. This prompts the user to actively think before each original, unmodified photo is revealed.

In general when using a spaced repetition system, the success of recall from retrieval cues is evaluated through user rating and used for scheduling subsequent review times. This would mean that the user devotes some time of the day to actively review scheduled parts of visual journeys. While this may be acceptable to some users, we are also considering a passive, background review mode for which user evaluation is optional. The system would automatically adapt based on a user’s physiological signals, such as heart rate, and utilize gaze tracking to confirm the user’s object of attention. Our system hopes to invoke a “curiosity effect,” through which users will naturally be curious as to “what picture will be next,” encouraging them to relate what they have seen with other memories.

Besides recollection, other application areas can include reminiscing and reflection. Referring to the definitions given in [8], reminiscing is a special case of recollection, where users re-live past experiences for emotional or sentimental reasons. Reflection can include considering patterns of past experiences, or looking at them from different angles or perspectives, for the purpose of seeing things anew and re-framing the past.

3. BACKGROUND

3.1 Personal photo collections

The proposed system accesses photos from a user’s personal collection and utilizes available metadata to infer relations among them. This section presents a brief overview of the characteristics of such collections and necessary metadata requirements for system applicability.

Compared to their paper photo album counterparts, digital photo collections contain a large number of items. This trend is driven by the popular usage of digital imaging devices for personal use (e.g. digital cameras, mobile phone cameras, and special lifelogging devices), and the rapid decrease in data storage costs. While only a few photos were taken to document an event in the “analog era,” digital recordings of, e.g., weddings now contain hundreds of images, especially if they are shared among event participants. Manually organizing photo collections has proven to be a burden that people can hardly cope with, as users have serious problems finding particular photos in their personal collections [10].

Automatic media management aims to assist photo search by deriving media metadata automatically (relying on hardware sensors and software algorithms) or with user assistance. Possible photo metadata can include:

- Location, which might be low level from a user’s point of view, e.g. GPS coordinates, cellular tower ID, or WiFi access point; or high level street address, city, region, country, etc.
- Date, day-of-week, time-of-day (dawn, morning, noon, afternoon, dusk, night, etc.), time, and light quality, as a useful memory cue [7], and if combined with location: season and weather information.
- Visual information, low level such as prominent colors and texture, high level, e.g. indoor/outdoor classification; face, object (perhaps linking to a corresponding Wikipedia entry), and scene category recognition. The latter two are considered to be difficult and are currently actively researched.
- Ratings or “starring” of photos, denoting subjective importance.
- Other including type of activity, personal emotional associations, user-provided keyword meta-tags, as commonly used, e.g., on social media websites.

Popular photo management software, such as Google’s Picasa [4] and Apple’s iPhoto [1], support basic date & time information, user-assisted face recognition, and location display & editing. Operations such as renaming pictures, attaching names to recognized faces, and adding location information can be applied en masse to aid organizing large photo collections. Media metadata are either stored within image files, e.g. through the Exchangeable Image File format (Exif), or in auxiliary files and databases. Social media sites, such as Facebook and Flickr, allow users to tag people on photos and add custom keyword meta-tags.

Since the proposed system relies on existing image metadata, certain assumptions are made about the photo archive: necessary face & location information is associated and photos are arranged into albums around events. Basically, image collections managed by one of the above-mentioned tools with additional user-assisted tagging satisfy these requirements.

3.2 Spaced Repetition Systems (SRS)

Paper or electronic flashcards can be used for memorizing factual knowledge. Each card poses or implies a question on its front, and provides the corresponding answer on its back. During study with flashcards, users choose a number of cards to review, ponder each question, and given the success or failure of each answer, group cards separately for later
review. Electronic flashcards automate and ease the management of cards, especially as the number of cards swells and a deck becomes unwieldy.

Research in psychology has shown that different intervals between reviews influence learning efficiency. The **spacings effect** phenomenon describes the fact that learning tasks are performed better if the repetitions are distributed sparsely in time as opposed to massed presentation (see e.g. [6]). Consequently, optimum learning intervals have been defined as the longest intervals that do not result in forgetting. The SuperMemo algorithm [11] supervises the learning process using a stochastic approach (since forgetting also has a stochastic nature). It can achieve results such that only around 10% of the to-be-remembered items are not remembered at the moment of repetition by adjusting the inter-repetition intervals based on user-assessment feedback. Numerous evolutionary versions of the algorithm have been developed and implemented in both commercial and open-source spaced repetition systems.

### 4. PHOTO JOURNEY CONSTRUCTION AND REVIEW

To create a new photo journey, the user chooses a person, event, or place. According to the selection, the user picks a representative key photo, which is used as a seed to discover related images in the photo library. Thus the key photo should have necessary metadata associated with it. In case of a person, the person’s face must be identifiable; in case of a place, location information should be present; and in case of an event, the key photo should belong to one or more albums containing related pictures.

Based on the theme, various heuristics are applied to find a set of relevant images. Finding photos with a particular person relies on facial metadata information, but dimensions of time and space can also be employed. A search method may link to one photograph from each encounter with the same person, or more from a particular meeting, e.g. from the recent past or from the first time together. Combined with location information, unique events can be singled out. For places, location can have different levels of granularity covering different sizes of geographical area. Summer vacation places might be handled at larger scales, encompassing a city or country, while frequently visited areas can be isolated from each other, emphasizing the hyperlocal nature of such places. Ratings, “starred” items, or user-edited photos can also be used as cues for importance. Furthermore, if the image metadata is obtained from a social media site, group ratings, comments and their statistical characteristics, as well as emotional tags can also be used. Additionally, metadata-derived images, e.g. maps or historical weather information, can also be included as journey items. Users can optionally control the compilation heuristics and the number of included pictures.

The gathered images form a visual photo journey, a persistent collection of cards from the SRS perspective. To facilitate active review, each card of the deck will have a memory- cued photo on the obverse and the original, unmodified one on the reverse. Utilizing a bounding rectangle, faces can be disguised as shown in Fig. 1. Certain aspects can be emphasized by e.g. employing the “Ken Burns effect” (after the technique mastered and popularized by the documentary filmmaker): zooming into or panning over an area, where perhaps the speed is controlled by the recorded arousal level of the user when the particular photo was taken. For photos lacking description or hints about their contents, more general and global transformations such as mosaic effects, distortions, tiling, and scrambling could be applied.

Photo journeys are handled according to the SRS algorithm. At first, all unseen items scheduled to be reviewed. Reviews are time-limited, which parameters can be set by the user. After some reviews, the deck might contain examined cards and some unexamined cards. Unseen items are mixed in with scheduled seen items to ensure that all items are reviewed eventually. Card scheduling is influenced by user ratings during review, chosen based on the success or failure of recalling the original photo from the associated memory cue. The user can select among “Not remembered,” “Hardly remembered,” and “Well remembered” responses. An option to exclude certain photos (irrelevant or boring matches) is provided. Users can cancel or reschedule sessions if they want. Finally, there is an option for deleting a visual journey.

During review, visualizations are used between each card to relate the items back to the chosen theme and key photo as part of an animation cycle and to display the user’s progress through review session. The recurring visualization emphasizes the connections among the photos and with the theme that is being recalled. The animations can, e.g., show the key photo in the center with spokes radially pointing to linked images, showing the progress as a cycle, as depicted on Fig. 2.

As reviews are scheduled, the user devotes a certain amount of time each day, depending on the number of cards, their maturity, and the number of existing visual journeys. Each visual journey can last from a few seconds to a few minutes, putting one’s mind in a state of reliving and recalling the event, location, or people. To increase the immersion, users can also associate selected music with journeys that is played back during reviews. The contents of journeys can be automatically updated if new photos are added or photos are removed from the collection. If the experience is not remembered well, the number of photos included in a journey can be increased automatically, adding more memory stimuli.

#### 4.1 Passive Review

Passive review is an optional background review mode for visual journeys which can be used with digital photo frames. In this mode, items are shown according to precompiled scheduling but without user interaction. The photos are evenly spaced in time based on the session length.

An SRS typically needs some form of feedback to assess what was and what wasn’t remembered well. As this mode is largely passive, e.g. a digital photo frame in the background of one’s desk, there may also be a need to determine if the user is paying attention or busy with some other task. There are several approaches for providing ratings. The simplest algorithmically is for the user to provide direct input feedback whenever they feel like it. If the system assumes the user is familiar with the scene (since if it was unfamiliar, they would spend longer looking at it and be more inclined to provide feedback), then the user need only interact when a memory wasn’t remembered well, recording a “Not remembered” signal, thereby increasing the recurring frequency of that item.
A more complicated form of feedback would involve the use of eye-tracking to determine when a user is focusing on an image longer. However, reasons for such fixation could be various: either it has a positive effect, such as an inspirational photo, or the memory recall is taking some time, as the user recalls the event and circumstances surrounding it, “playing it back” within their minds.

5. PROTOTYPE IMPLEMENTATION

Several current and upcoming software platforms provide opportunities for display implementation. JavaFX is expanding to new devices such as TV sets. AJAX-based HTML websites also have potential, with web browsing interfaces available on many devices that can be used for background “picture frame” interfaces (as seen in Fig. 3), such as TV gaming consoles, the Apple iPad, etc.

A prototype application using HTML5 is currently being implemented for devices that support web browsers. The picture-serving backend is implemented in Java using an open SRS algorithm. Photos and metadata are fetched from either Picasa or iPhoto collections.

6. CONCLUSIONS AND FUTURE WORK

The proposed system leverages existing learning algorithms and applies them in the form of an assistance tool for remembering personal experiences. Complementary to still photos, audio clips, music, and video media including stereoscopic (“3D”) stills and videos may also be used as memory stimuli. The system could be applied to mobile phones, laptops & desktop computers, digital picture frames, and TV gaming and media systems, including online social media. Such devices and services can then be linked to a single database via computer network.

Memory triggering also has other functions in daily life.
Due to the emotion-arousing ability of memories, it also becomes possible to evoke moods, e.g. to attain a positive effect from certain triggers such as inspirational photos capturing moments when one may have performed his or her personal best.

7. REFERENCES