

An Approach for a Personal Information Management System for Photos of a Lifetime by Exploiting Semantics

Khalid Latif, Khabib Mustofa, and A. Min Tjoa

Institute of Software Technology & Interactive Systems
Vienna University of Technology,
Favoritenstrasse 9-11/188,
A-1040 Vienna, Austria
{klatif, khabib, amin}@ifs.tuwien.ac.at

Abstract. Photo collections are one of the promising sources to tell story of life in this digital era. In this paper we present our work on organizing photos of a lifetime by exploiting semantic annotations. The complexity in using semantic technology is managed by introducing an annotation template corresponding to who, when, where, and what. Semantics of each dimension are semi-automatically annotated following the ontology for personal information. The story telling is done by exploiting these semantics to form trails of the photos. The notion of landmarks is used for this purpose which also ensures effective navigation in the lifetime photo-space.

1 Introduction

The size of personal photo collections has grown a great deal with the increasing use of digital cameras. On the other hand ever increasing storage capacity inspires the vision to accumulate information without the need to delete old photos. Organizing such a large number of photo collections requires effective use of the photo metadata. This metadata can be separated into two categories: the general photo characteristics and the photo contents. The later, although can be automated for low level feature description, is manually annotated to lower the “semantic gap” [4] [13]. Currently available personal photo management tools mostly exploit first type of metadata with support for free text comments [7] [15]. In contrast the semantic photo annotation tools are mostly developed for annotating digital photo archives such as for artwork [9]. There remains a need to bring together both worlds without flooding the user with the complexity of underlying semantic technology.

The primary goal of this research is to build a personal information management system for photos of a lifetime by exploiting their semantics. This is done through semi-automatically annotating photos about persons (who), time (when) or event, location (where), and other objects and actions of persons (what) using already available ontologies. Additionally user preferred photos

are declared landmarks, which serves two purposes. Firstly, while viewing an individual photo the landmarks being near to the photo in context are suggested as related photos. Secondly, the landmark photos are magnified when viewing whole photo collections. Thus the user is guided in navigating large photo space.

First we will briefly discuss existing approaches for personal photo management and semantic photo annotations. In the subsequent sections we will discuss the proposed annotation model, the use of landmarks for navigating lifetime photo collections and the user interaction for photo annotations. Finally we summarize the status of our work and an outlook of open issues.

2 Related Work

Photo collections, as one of promising sources to tell story of life in this digital era, has attracted many researchers to spend time for exploring its usage and proving its usefulness. Some solutions focus on personal photo organization while others bring together semantic annotations to photo collections in general. Due to the varying intentions of these tools we present them in different sections below:

2.1 Semantic Photo Annotation

Flickr¹ is an online application which allows to creating collections by dragging and dropping photos from timeline display. Annotations in Flickr are supported by tags. We have extended such tagging with ontologies and by connecting these tags with appropriate predicates from the ontology such as Person *playing* Football.

PhotoStuff [8] allows manual semantic annotation to images either on the web or local. It allows loading RDF-based ontologies for later annotating images with it. The taxonomy browser in SemanticLIFE is similar to their ontology tree view. Instead of exposing arbitrary property-value forms for encoding RDF annotations, our approach is based on the annotation template. The template based annotations guarantee sufficient metadata for describing images and the eas-of-use. Additionally instead of first drawing a region and then looking for the appropriate concept from the ontology for annotations we allow the user to select a concept and then draw a region for it.

Developed by Mindswap, SMORE [10] focuses more on semantic mark-up and ontology editing. It offers an extension for annotating text and images. Using these tools user can create metadata based on the loaded ontology. Furthermore, user can also define new concepts (class or property), to be added in the ontology.

Another promising multimedia annotation tool is M-Ontomat-Annotizer [4]. It supports metadata creation for an image or its selected region by associating it with a concept in the ontology.

¹ <http://www.flickr.com>

2.2 Personal Photo Organization

Among large number of other commercial photo management tools, Picassa² provides very good visualizations. It supports photo labelling and the photos having same label are considered one collection. Creation date is also used to automatically classify photos. Picassa provides good visualization for the timeline view of all collections. Similar to other commercial tools slideshow for collection is also supported.

PhotoFinder [11] supports visualization of collections in several ways. Drag-n-Drop is supported throughout the application for different tasks. It also supports limited annotation for persons. We have extended their annotation strategy in two dimensions: (1) using image region for annotation and labelling and (2) annotating events and other photo contents using semantic web technologies. In addition the person information is reused and linked to the existing personal address book and contacts. This allows more semantic queries to be fulfilled from the photo collections such as searching for “photos of all friends in Salzburg”.

Girgensohn and colleagues investigated the performance issues in organizing large photo collections [7]. The tool support practically is not much different from the other commercial tools but comes with improved performance. Worth mentioning is a calendar view supported in their tool. In our prototype implementation we provide sorting and filtering based on time and date both for a particular collection and in lifetime view. We also provide sorting and filtering based on the concept taxonomy.

A different approach is presented in [3] for organizing large photo collections. It combines information from map data (GPS) with metadata of photos, by which a story of a trip will be more meaningfully constructed. Later on photos can be viewed on the location map and also based on the timeline. Additionally MyLifeBits provides an integrated view of the lifelong information items ranging from photos, documents, phone calls, emails, to web pages [6]. Our approach potentially differs from MyLifeBits as we exploit ontologies for managing semantics of these information items.

3 SemanticLIFE

The SemanticLIFE project is an attempt to realize the Vannevar Bush’s vision of the **Memex** “*a device in which an individual stores all his books, records, and communications. . . an enlarged intimate supplement to his memory*”; and **associations of thoughts** “*The human mind. . . operates by associations. With one item in its grasp, it snaps instantly to the next that is suggested by the association of thoughts, in accordance with some intricate web of trails.*” [5] The architecture of SemanticLIFE system is presented elsewhere [1]. The range of data sources starts from communication data (emails, phone calls, and chat sessions) to personal documents, photos, web-browsing sessions and calendar

² <http://picasa.google.com>

data. SemanticLIFE realizes the associations of items in personal digital diary by annotating contents of these information items with ontologies. In this paper we specifically focus on organizing and annotating photo collections by exploiting their semantics.

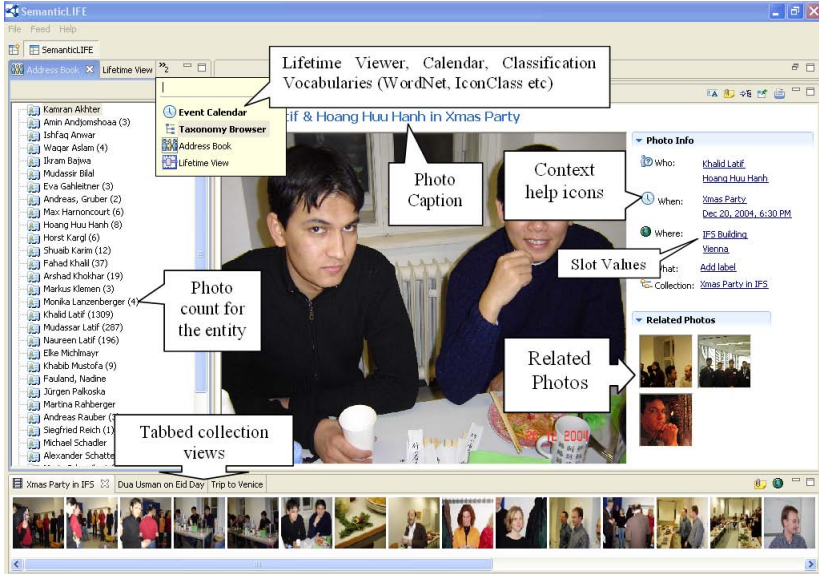


Fig. 1. Overview of photo annotation in SemanticLIFE

4 Annotation Model

The information of the photos can be separated into two categories: the general photo characteristics and the photo contents. The first category provides information about photo resolution, format, size, etc. Such information is present in the EXIF header of digital photos and is easily extracted. The second category describes what is depicted by the photo. The contents of personal digital photo vary largely, and may include a wide range of domains such as sports, entertainment, and sightseeing.

Utility of semantic annotations for describing such diverse photo contents is well established [4] [13]. But, comprehensibility in such an annotated lifetime photo space and usability of Semantic Web technologies from the user interaction point of view is still an open issue. Common users mostly want easy access to their photo collections for viewing, using in their homepage, creating presentations, or sharing with other people. It is difficult to provide a unified way for annotating personal photos with arbitrary RDF. Even simple and otherwise trivial annotations are complex and hard to grasp for non-experts, regardless of any simplification in the visualization. On the other hand a simplified annotation model can lead to pragmatic interfaces.

In our recent work [12] we presented a study of the LATCH³ model for organizing personal information items. The proposed annotation model corresponds to who, when, where, and what. The hypothesis is that such a structuring of annotation template on one hand provides adequate semantics to organize personal photo collections and on the other hand is easily comprehended by the user. The values for slots in the template are filled by creating semantic labels based on concepts in existing ontologies. Compared to keyword search such semantic annotations allow concept searching where users can specialize or generalize a query based on the concept hierarchy. The detail of the model is presented below:

Who: This axis describes the persons and other agents/actors depicted in the photo. For annotating persons the already available address book is presented to the user. Person names are also extracted from recently visited web pages and other user documents. The user is also suggested with these names during annotations. The detail of how we determine if a name should be suggested for a specific picture is explained in section 6.1. Digital photos are categorized as personal, professional photos, and art work [13]. Personal photo collections may include a photo of an art object. Such photos have a special actor, the *creator*, annotated as the original author of the object depicted in the photo. One such example is a painting depicted in a photo being annotated with its *artist* (where artist *isa* creator).

When & Where: These dimensions describe the time and location of the photos and the collections. Locations include concepts such as country, city, a region, and street address. For individual photos the time value describes the time the photo is taken which is extracted from the photo metadata in EXIF header. These dimensions can also be annotated with the scheduled events from existing calendar management applications like MS Outlook. An example case is explained in section 6.1. The time value is represented as an interval.

What: This axis has two dimensions: (1) *what actions* agents are doing such as a person eating sushi, playing football, and (2) what objects (other than agents described in *who*) are depicted in the photo such as palm tree, or a painting. The later can also be annotated in turn with context metadata, such as a painting within a photo maybe be annotated with the artist (who) information.

5 Photos as Landmarks

Humans make use of variety of practices for recollection. Use of *mnemonics* is one example of such practices. Originated with the ancient Greeks the idea is to associate parts of the information to well-known landmarks. In hypertext systems the opening web page is considered a landmark and every other web page in that

³ LATCH originally stands for *Location, Alphabets, Time, Category, and Hierarchy* [17]. We suggested replacing *Alphabets* with *Agent/Actor* for using it as organizing principle instead of ordering principle [12].

particular web application is linked with it [16]. Use of landmark events is also investigated for personal information space [14]. Traditionally personal photos are sorted into family album (the preferred ones) and “shoe-boxes”. We argue that, while keeping the distinction, the connection of the photos in family album with those in shoe-boxes could be established in digital archives. This is done by declaring important photos as landmarks within a collection. All other photos in the same collection are automatically associated with the landmark photos. Some personal photo management applications make use of ranking, such as a count from zero to ten in PhotoFinder, to weight importance of a photo. While declaring a photo as important is one aspect, more important is to associate other relevant photos with it and suggesting them to the user. Thus retrieval of photos is made efficient by forming trails of associations and the user is guided in exploring the large photo-space.

For declaring a photo a landmark users assign a numeric weight ($10 \geq w > 0$) to the photo. This weight is later used in determining the size of thumbnail in collection view and also in search results view. The landmark weight also contributes in determining the nearness of one landmark with the other. We have implemented different set of rules based on the weights using Semantic Web Rule Language (SWRL). For example a landmark photo LP_x having weight x will be considered near a photo P_0 if any of the following holds:

- o LP_x and P_0 are in the same photo collection
- o LP_x and P_0 are directly connected (through manual linking)
- o Both LP_x and P_0 are annotated with the same concept C from the ontology.
- o LP_x has an annotation of concept type C_1 ; P_0 has an annotation of concept type C_2 , and $SemanticDistance(C_1, C_2) < x$. The semantic distance is computed in several ways such as the manual associations, property-entity associations [2], topic similarity, and hierarchical concept distance. More detail of calculating semantic distance in nearness discovery of landmarks is presented in our prior work [12].

The photo viewer uses these rules to find landmark photos near the currently selected photo. User can set a threshold value (default to 4) for the number of relevant photos to show in the photo viewer. The priority is given to the photos with higher landmark weight. While viewing one photo from a collection the user is provided with photos which are semantically near the photo in context. Thus the whole photo collection turned out to be a web of trails.

6 User Interaction

Other than the taxonomy browser (classification view) which is an integral part of SemanticLIFE, three views are provided to the user for navigation and annotation of photos: (1) lifetime photos view, (2) collection view, and (3) the photo view. In the lifetime view representative thumbnails of all categories are displayed along with their titles and event information. The thumbnails are generated from combination of the landmark photos in the collection similar to “My Photos”



Fig. 2. Photos arranged on a map by the user. Magnification of a photo thumbnail depends on its landmark weight.

view in Windows XP in which first four photos are used to generate thumbnail of a folder. The collections in this view are sorted based on the timeline.

The collection view by default uses date/time for sorting the photos. Photos can also be sorted and filtered based on the concepts in the taxonomy hierarchy. A scattered plot mode with a background location map is also supported (see figure 2). Users can freely place the photos on a background map. The settings are preserved and could be seen anytime by selecting the location map mode in the collection view toolbar. Lastly the application is developed as an Eclipse rich client platform (RCP) which allows the user to arrange the views at the position of their choice.

6.1 Connecting Life Items

SemanticLIFE's repository is fed with different desktop information such as calendar entries/appointments, web browsing cache, emails, and address book. The photo annotation is an integral part of the system, so it utilizes and reuses the existing information by far. Most of the personal photos come from planned events, such as birthday party or a conference. Information about such events (if present) is fed by Outlook and Sunbird adaptors, and is stored in the repository after appropriate transformation to RDF (c.f. figure 3 showing scheduled event in Sunbird, the event website as visited by the user, and photo taken in that event.) Such existing items are an added help to the user in photo annotations. Items in the same date/time range are suggested to the user for their possible reuse during photo annotation. We also apply ANNIE⁴ to the recently

⁴ <http://www.gate.ac.uk/annie/>

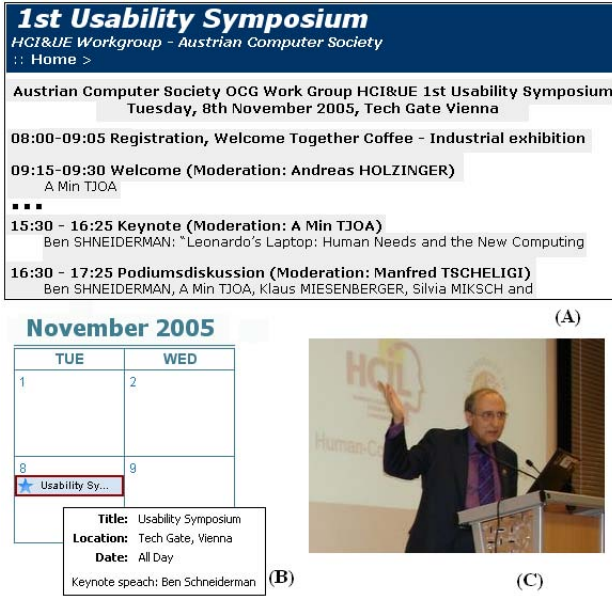


Fig. 3. (A) Part of the website showing program of the event, (B) scheduled event in Mozilla Sunbird, and (C) a picture taken in that event

fed life-items (such as web pages, emails, and documents) for extracting named-entities such as person names. Clicking the *context help icon* in photo view (c.f. figure 1) displays a list of relevant entities being extracted by ANNIE. Any of these items/entities could be dragged and dropped on the photo or whole collection. Depending on the item type and its meta-data the appropriate slot is filled, thus users do not have to re-type.

6.2 Annotation of Individual Photo

Other than entering the metadata by hand, we support two strategies for photos annotations: (1) Existing information items such as persons in address book, event entries in calendar, and ontology contents from a vocabulary are dragged and dropped on the photo. This associates the prevalent concept to the whole photo. (2) A specific region of the photo could be annotated by first selecting the target concept from existing vocabulary (such as *Gondola* taken from WordNet). The taxonomy browser loads the ontology vocabularies in a tree structure for this purpose. The annotation marker on the photo view tool bar is then selected and a rectangle is drawn on the photo. This annotates the image region with currently selected concept in classification view. An RDF listing of the annotation is presented in figure 5. The rectangular region is hidden in the photo view unless the target concept is selected from the photo information which highlights the region (cf. figure 4).

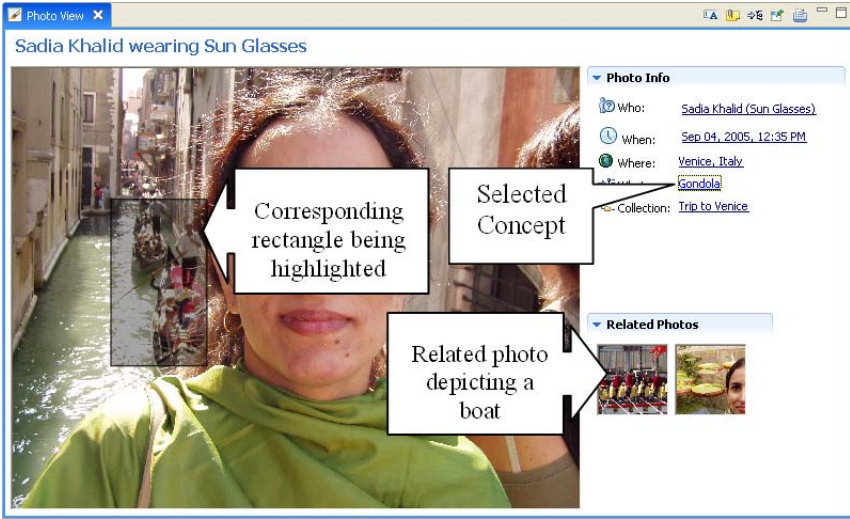


Fig. 4. The photo viewer with concept and region highlighting support. The selection of concept *Gondola* has highlighted the associated region.

```
<rdf:RDF xml:base="http://www.ifs.tuwien.ac.at/slife-core.owl"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:reg="http://www.w3.org/2004/02/image-regions#">

  <foaf:Image rdf:ID="slph_1004285">
    <reg:hasRegion>
      <reg:Rectangle rdf:ID="slph_1004285_r01">
        <reg:coords>...</imreg:coords>
        <reg:regionDepicts rdf:resource=".../wordnet/1.6/Gondola-1"/>
      </reg:Rectangle>
    </reg:hasRegion>
  </foaf:Image>
</rdf:RDF>
```

Fig. 5. Abridged RDF listing of an annotation showing a rectangle within a photo which depicts a concept *Gondola* from WordNet

6.3 Annotating Collections

Collections can be created either (1) manually by dragging and dropping the selecting photos or (2) suggested by the system for un-sorted photos. The later task examines the EXIF header for possible match in date/time and other available characteristics. Similar to photo annotations, whole collections can also be tagged with ontology concepts or linked to other personal information items such

as an event from the calendar data. Associating the metadata with the collection replicates the semantics to all member photos. Moreover a collection can become a part of another collection.

7 Conclusion and Future Work

We have presented our work on managing personal photo collections using semantic annotations. The proposed annotation model and use of landmarks have made the large photo collections a web of connected photos in which user can navigate from one collection to photos belonging to specific concepts and then to other semantically related photos. This way we tried to close the gaps in bringing together semantic photo annotations and personal photo management.

In near future we intend to extend our work for semantic relation discovery by exploiting more the semantics of the ontology concepts compared to the syntactical concept distance. Additionally we are working on semi-automatically developing a unified view of the concept hierarchies which otherwise belong to different domain ontologies. In DynamOnt project we are investigating an approach for dynamically building ontologies by reusing existing concept hierarchies through their mapping with the foundational ontology. We hope to enhance the work presented in this paper by benefiting from the findings of that project.

Acknowledgement

This work as part of the SemanticLIFE project is generously supported by ASEA-UNINET (ASEAN-EU Academic University Network) and HEC (Higher Education Commission of Pakistan). This research is also supported by the DynamOnt project which is funded by the Austrian Governments FIT-IT Research Initiative on Semantic Web.

References

1. Ahmad, M., Hoang, H.H., Karim, S., Khusro, S., Lanzenberger, M., Latif, K., Michlmayr, E., Mustofa, K., Nguyen, H.T., Rauber, A., Schatten, A., Tho, M.N., Tjoa, A.M.: Semanticlife - a framework for managing information of a human lifetime. In: Proceedings of 6th International Conference on Information Integration and Web-based Applications & Services, Jakarta, Indonesia, OCG Press (2004)
2. Aleman-Meza, B., Halaschek-Wiener, C., Arpinar, B., Ramakrishnan, C., Sheth, A.: Ranking complex relationships on the semantic web. *IEEE Internet Computing* **9**(3) (2005) 37–44
3. Aris, A., Gemmell, J., , Lueder, R.: Exploiting location and time for photo search and storytelling in mylifebits. Technical Report MSR-TR-2004-102, Microsoft Research (2004)
4. Bloehdorn, S., Petridis, K., Saathoff, C., Simou, N., Tzouvaras, V., Avrithis, Y., Handschuh, S., Kompatsiaris, Y., Staab, S., Strintzis, M.: Semantic annotation of images and videos for multimedia analysis. In: Proceedings of 2nd European Semantic Web Conference, Heraklion, Greece, Springer Verlag (2005)

5. Bush, V.: As we may think. *The Atlantic Monthly* (1945)
6. Gemmell, J., Bell, G., Lueder, R., Drucker, S., Wong, C.: Mylifebits: Fulfilling the memex vision. In: *Proceedings of ACM Multimedia*. (2002)
7. Girgensohn, A., Adcock, J., Cooper, M., Foote, J., Wilcox, L.: Simplifying the management of large photo collections. In: *Proceedings of Human-Computer Interaction*, IOS Press (2003)
8. Halaschek-Wiener, C., Schain, A., Golbeck, J., Grove, M., Parsia, B., Hendler, J.: A flexible approach for managing digital images on the semantic web. In: *Proceedings of 5th International Workshop on Knowledge Markup and Semantic Annotation*, Galway, Ireland (2005)
9. Hollink, L., Schreiber, G., Wielemaker, J., Wielinga, B.: Semantic annotation of image collections. In: *Proceedings of Workshop on Knowledge Markup and Semantic Annotation*. (2003)
10. Kalyanpur, A., Hendler, J., Parsia, B., Golbeck, J.: (Smore semantic markup, ontology, and rdf editor) <http://citeseer.ist.psu.edu/555327.html>.
11. Kang, H., Shneiderman, B.: Visualization methods for personal photo collections: Browsing and searching in the photofinder. In: *Proceedings of IEEE International Conference on Multimedia and Expo*. (2000) 1539–1542
12. Latif, K., Tjoa, A.M.: Combining context ontology and landmarks for personal information management. In: *Proceedings of IEEE International Conference on Computing & Informatics*, Kuala Lumpur, Malaysia (2006)
13. van Ossenbruggen, J., Troncy, R., Stamou, G., Pan, J.: Image annotation on the semantic web. Technical report, W3C Semantic Web Best Practices Working Group (2006) http://www.w3.org/2001/sw/BestPractices/MM/image_annotation.html.
14. Ringel, M., Cutrell, E., Dumais, S., Horvitz, E.: Milestones in time: The value of landmarks in retrieving information from personal stores. In: *Proceedings of 9th International Conference on Human-Computer Interaction*, Zurich (2003)
15. Schneiderman, B., Kang, H.: Direct annotation: A drag-and-drop strategy for labeling photos. In: *Proceedings International Conference on Information Visualisation*, London, England (2000)
16. Sorrows, M.E.: Recall of Landmarks in Information Space. Phd dissertation, School of Information Sciences, University of Pittsburgh (2004)
17. Wurman, R.S., Sume, D., Leifer, L.: Information Anxiety 2. Que (2000)