Department of
**Information Engineering
and Computer Science** DISI

UNIVERSITY
OF TRENTO - Italy

DISI - Via Sommarive 14 - 38123 Povo - Trento (Italy)
http://www.disi.unitn.it

PERSONAL PHOTO INDEXING

Ivan Tankoyeu, Julian Stöttinger, Javier
Paniagua and Fausto Giunchiglia

July 2012

Technical Report # DISI-12-028

Personal Photo Indexing

Ivan Tankoyeu, Julian Stöttinger, Javier Paniagua, and Fausto Giunchiglia
DISI, University of Trento
via Sommarive 14
38123 Povo, Trento, Italy
[tankoyeu | julian | paniagua | fausto]@disi.unitn.it

ABSTRACT

Sorting one's own private photo collection is a time consuming and tedious task. We demonstrate our event-centered approach to perform this task fully automatically. In the course of the demonstration, we either use our own photo collections, or invite the conference visitors to bring their own cameras and photos. We will sort the photos into a semantically meaningful hierarchy for the users within a couple of minutes. Events as a media aggregator allow a user to manage and annotate a photo collection in more convenient and natural to the human being way. Based on the recognized user behavior the application is able to reveal the nature of an event and build its hierarchy with a event/sub-event relationship. One important prerequisite of our approach is a precise GPS based spatial annotation of the photos. To accommodate for devices without GPS chips or temporary low GPS perception, we propose an approach to enrich the collection with automatically estimated GPS data by semantically interpolating possible routes of the user. We are positive that we can provide a well received service for the conference visitors, especially since the conference venue will trigger a lot of memorable photos. Large scale experimental validation showed that the approach is able to recreate a user's desired hierarchy with an F-measure of about 0.8.

Categories and Subject Descriptors

H.3.1 [**Information Storage and Retrieval**]: [Indexing methods]; J.4 [**Computer Applications**]: Social and Behavioural Sciences

General Terms

Algorithms, Performance, Human Factors

Keywords

Event-Based Indexing, Personal Photo Collection, Photo Context Analysis

.

1. INTRODUCTION

Digital photography has decreased the cost of creation and storing of personal photos. This often leads to a tremendous growth in the number of one's personal photo collection. In order to facilitate or even rid a user from the boring and tedious task of annotating and sorting personal photos reliable techniques to automate this process are required.

More than 70% of albums shared at Google Picasa[1] are related to a personal event. An event is defined as a phenomenon happening at known time and location. Therefore spatio-temporal metadata has a grand advantage over visual analysis for event-based indexing in personal photo collections. New camera devices are often equipped with GPS chips. The extraction of spatio-temporal information is computationally cheaper in comparison with the extraction of visual features. If the GPS information is not complete, we are able to (i) sync spatial information over different devices (e.g. GPS equipped smart phone and a non-GPS professional camera), and (ii) interpolate location data by using free online navigation services. All this leads to a new cost efficient and intelligible event-centered way to index one's personal photo collection.

We will provide a demonstrator that is able to manually annotate certain photos with GPS information (if needed), synchronizes GPS information over devices, interpolates missing GPS information, and provides a semantically meaningful hierarchy of events and sub-events based on the algorithm of [2, 1]. This may be carried out by the photos of conference visitors, or on our own example data-sets.

2. SYSTEM OVERVIEW

This section provides a list of functionalities of the system.
Semantic geo-interpolation and extrapolation - assigns geo information for images with missing GPS coordinates.
Event Detection - separates photo collection to a set of albums related to detected events.
Spatial routine detection - detects "routine" behavior of a user based on spatio-temporal information.
Event Hierarchy Building - builds the hierarchy of events for non-routine behavior.
Momentum of Attraction detection - detects the most interesting moment related to the image either within a photo collection or within the given event. The image can be used as a cover of a photo collection or an album related to an event.
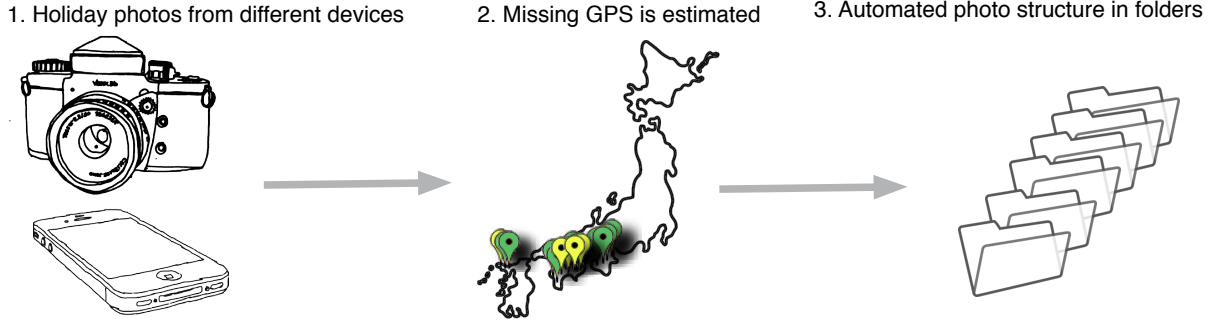
---

[1]https://picasaweb.google.com

Figure 1: Demonstration workflow: Conference visitors copy their unsorted images. Optionally they can annotate a few pictures with their GPS location using Google maps. Then, the missing GPS information is estimated for the remaining pictures. Finally, the photo events are estimated and put into a folder structure.

3. SEMANTIC ESTIMATION OF GPS DATA

We present the novel approach for estimating missing coordinates for images with absent geo information. First, the system splits a photo collection on a set of event-related clusters ($e_1 - e_4$) based on temporal information (Δt) only. The result is visualized in Figure 2, markers with letters annotate photos with GPS data, dark ones photos without GPS data. Considering the position of the image in accordance to temporal boundaries there are two possible cases for assigning missing data points:

1.Extrapolation (Figure 2(2)) is the task of extending a known sequence of values A_{e_1} or C_{e_4} .

2.Interpolation (Figure 2(3)) is the task of estimation of a unknown sequence of samples within two known data points A_{e_2} and D_{e_2}.

The system is able to make linear interpolation and semantic interpolation. Semantic interpolation allows to the system to understand the travel mode of a user (walk, bike, car) based on time differences between two points. After choosing the appropriate travel mode it interpolates images with the given timestamps querying an online navigator (currently Google). The semantic analysis is done based on suggestions of travel routes using the Google Maps API. If no route is provided, the locations are linearly interpolated.

Figure 2: Examples of extrapolation (2) and interpolation (3).

4. DEMONSTRATION

The demonstrator can be summarized as follows (compare Figure 1):

1. Conference visitor uploads his personal images taken during the conference visit, if needed some "key" photos are manually spatially annotated.
2. Missing GPS information is estimated.
3. Event hierarchy is built.
4. Final folder structure is event hierarchy and copied back to phone, camera or USB stick.

A user uploads photos to the system from different camera enabled devices. The system extracts spatial and temporal information from images and creates a chronologically ordered collection. Based on time differences between images, the system assigns images in event-related clusters. For images with absent geo-information the GPS coordinates are estimated automatically. After the step of the enrichment the given set with geo-data the system runs the algorithm for event-detection described in detail in [2]. The output of the system contains the folder structure which reflects the event hierarchy Figure 1 . Finally, the system presents the most active moments of his photo-related activity, the "momentum of attraction" as favorite images.

5. CONCLUSION

We demonstrate an event-based personal photo indexing framework. This framework analyzes the spatio-temporal context of personal photo collection to build a meaningful hierarchy of photos. In large-scale experiments we could show that we meet the user's desired hierarchy on average with an F-measure of 0.8.

6. REFERENCES

[1] Javier Paniagua, Ivan Tankoyeu, Julian Stöttinger, and Fausto Giunchiglia. Indexing media by personal events. In *ACM ICMR*, 2012, to appear.
[2] I Tankoyeu, J Paniagua, J Stöttinger, and F Giunchiglia. Event detection and scene attraction by very simple contextual cues. In *ACM MM J-MRE Workshop*, 2011.