Towards acquisition of semantics of places and events by multi-perspective analysis of geotagged photo collections

Slava Kisilevich, Daniel Keim
University of Konstanz, Konstanz, Germany
{slaks, keim}@dbvis.inf.uni-konstanz.de

Natalia Andrienko, Gennady Andrienko
Fraunhofer Institute IAIS
Sankt Augustin, Germany
{natalia,gennady}.andrienko@iais.fraunhofer.de

Abstract

Due to the pervasiveness of positioning technology combined with the proliferation of social-oriented web sites, community-contributed spatio-temporal data of people's historical positions are available today in large amounts. The analysis of these data is valuable to scientists and can provide important information about people's behavior, their movement, geographical places, and events. In this paper, we develop a conceptual framework and outline a methodology that allows to analyze events and places using geotagged photo collections shared by people from many countries. These data are often semantically annotated by titles and tags that are useful for learning facts about the geographical places and for detecting events occurring in these places. The knowledge obtained through our analysis carries an additional benefit. For example, it may also be utilized by local authorities, service providers, tourist agencies, in sociological and anthropological studies or for building user centric applications like tour recommender systems. We provide a conceptual foundation for the analysis of spatio-temporal data of places visited by people worldwide using community contributed geotagged photo collections. First, we define several types of spatio-temporal clusters of people's visits. Second, we discuss methods that can be used for analysis of these clusters. Third, we offer an analysis of tourist activities in Switzerland based on a case study.

1 Motivation

Ubiquity of location-aware devices, cheap storage and fast computing power has enabled collection and analysis of large amounts of spatio-temporal data. Different application domains like zoology, activity-based analysis or tourism in which data collection was a tedious and manual process (observation, surveys), benefit from the positioning technology and demand new analysis and techniques to cope with large quantities of these data.

Collections of geotagged photos have recently become available (Goodchild 2007) due to the availability of photo-sharing sites such as Flickr[1] and Panoramio[2], in which millions of users from all over the world upload their geo-referenced photos. The basic information provided by a person during photo upload is the location *where* the photo was taken, the time of the action, and the textual identifiers including title and tags. The photo may also be a member of some thematic group. A photo taken by a person can be regarded as an event, and collection of photos of a person can be considered as a trajectory. Such user-generated data have already been used in the analysis of attractive places (Crandall et al. 2009; Kisilevich et al. 2010a), movement behavior (Girardin et al. 2008a) and mobility (Andrienko et al. 2009a). The advantages of these data are (Girardin et al. 2008b): *unlike the automatic capturing of traces, the manual disclosure of location in the act of geotagging of photo provides additional qualities: positioning a photo on a map is not simply adding information about its location; it is also an act of communication which contains what people consider as relevant for themselves and others*.

Until now, these data were used as an alternative to the GPS-based data, mainly utilizing coordinates and timestamps. However, the title, tags, thematic group name as well as the photo itself, may reveal the context of the photo or describe the place where it was taken: some known event, a landmark or a person. Multimedia, computer vision and text mining communities realized the potential of geotagged data (Toyama et al. 2003) and proposed automatic approaches for such tasks as image summarization (Kennedy et al. 2007; Zhen et al. 2009a), landmark identification (Crandall et al. 2009), automatic event identification (Kennedy et al. 2007; Ahern et al. 2009, Becker et al. 2009), which includes clustering and retrieval of tag representatives. Information retrieval methods allowed automatic gazetteer creation using geotagged images (Popescu et al. 2008), Wikipedia, and web search engines and ontology induction from tagged images (Schmitz

[1] http://www.flickr.com
[2] http://www.panoramio.com

2006). However, pure automatic approaches of event or place exploration have several disadvantages that are important to draw attention to.

1. The automatic approaches usually utilize different constraints and assumptions that assure adequate performance. Such assumptions are for example the following:
    a. The representative tags or events are determined by the semantics of the textual information and not by the geographical constraints. This enables finding only one event per area (cluster). Clearly, if there are several events occurring in one place at different time or at the same time (overlapping events), only the most significant one will be selected. The significance of the events is purely algorithm dependent and can lead to loss of information about other events.
    b. The significance of the place is usually determined by the number of photos, users or other heuristics. Thus, the event with a *few* number of photos or people can be missed. For example, Jaffe et al. (2006) note that more than 1000 photos on a city scale are required in order to obtain meaningful results.
    c. The significance of the place is determined by the uniqueness of the textual semantics within the cluster. It means that in order to find a significant event in a cluster, other clusters, surrounding that cluster, should be analyzed and semantics extracted from them to be taken into consideration. This makes the algorithms non-scalable when large areas are used for exploration.
2. Experiments are performed using clean-room data samples, where class labels are manually prepared or taken from existing benchmark sources. Therefore, such issues as geographical errors, different languages or mistakes made in textual information are usually not raised.
3. Different representation models as well different algorithms produce different results.
4. Algorithm accuracy is reported with respect to the best-tuned parameters applicable to the training data. No real experiments were performed on arbitrary data.

In addition, all the mentioned approaches are user centric, aiming at providing solutions for exploration but not for analysis. Examples are: the representative tag viewer by Kennedy et al. (2007), tag maps by Jaffe at al. (2006) or tag mapping “world explorer” by Ahern et al. (2009). In contrast to these approaches, our paper aims at the *analysis* of places and events.

We claim, however, that combining the above mentioned techniques with geospatial visual analytics methods, GeoComputation, spatial and spatio-temporal data mining create new opportunities for the analysis of spatio-temporal data. The most important difference between the existing approaches and the methodology proposed by us, is the way in which event clusters are obtained. In contrast to the semantic-centric approach, we use spatio-temporal clustering based on geographical properties of the data as commonly used in geographical analytics. This allows us to apply different techniques like time-series, text or multimedia analysis on the same region in a chain of steps or to investigate different events that occurred at different time intervals or are overlapped.

For the sake of comparison to the existing approaches, we would like to mention the model used for semantic enrichment of trajectories. Furthermore, for the tasks in this paper that we regard as inapplicable, we will provide explanations respectively. Most of the spatio-temporal data is obtained by GPS devices and contains sequences of space-and-time referenced points measured at arbitrary chosen time intervals.

One of the widely used approaches in working with trajectories is based on extraction of significant places from a single trajectory using an object's stay time heuristic (Andrienko et al. 2007). This approach was later conceptualized by (Spaccapietra et al. 2008) by introducing a model in which trajectories are divided into sequences of *stops* (important places) and *moves* (movement **to** or **from** important place). Two main approaches are used to find important places in trajectories. The first considers only the characteristics of the trajectory (considerable time spent in a place). In the second approach, important places are obtained by intersecting the trajectory with the external application-specific geographical features provided by the user. In the first case, the obtained important places are still expressed in terms of geographic primitives and do not have any additional information, so that the analysis is usually performed by domain experts using visual analytics tools (Andrienko et al. 2007). In the second step, the obtained important places hold semantic information (id, location name) that can be used in the data mining process (Alvares et al. 2007a).

However, there are several problems with this approach: (1) An external database of geographic features should be available. But even if it is available, the algorithm can miss important places if the database is not complete. (2) The real context of a stop is not known. For example a person may be waiting in a traffic jam near a museum on the way to his/her work but the algorithm for finding important places may identify the person as visiting the museum by extracting a stay point (important place were a person spends considerable time) by intersecting the trajectory with the museum. (3) Since the data itself can have many contexts at different time intervals, the important place found may not correspond to the semantics that was attached to it (static semantics enrichment). (4) The extracted semantics describe only the data they are attached to, and cannot be used for other purposes.

Obviously, spatio-temporal data should contain more information to aid the analyst in understanding the context of the data. Since photo-collection data contains visual and textual information explaining the context of a photo, this data has invaluable potential for the analysis of the geographical places to which photos are geo-referenced, and the understanding of events that happen in the place where the photo was taken.

In this paper, we provide a conceptual foundation for the analysis of events and places using geotagged photo collections. We claim that a semantic enrichment of the spatio-temporal data should use additional components available in the data and take into account the temporal aspect. We define several types of semantic spatio-temporal clusters, and discuss methods for creation and analysis of these clusters.

2 Related work

2.1 Spatio-temporal clustering

Many methods were proposed to cluster spatio-temporal data. Trajectory patterns of moving objects were mined in (Giannotti et al. 2007) by finding regions-of-interest where many trajectories intersect with similar travel times. To find these places, the geographical space was divided into grids and the density of cells was computed. Then, a sequence mining algorithm was applied on these regions.

Palma et al. (2008) proposed a clustering approach based on DBSCAN (Ester et al. 1996) algorithm to find important places in trajectories. The original concept of point neighborhood used in DBSCAN was changed to allow finding important places in a single trajectory. According to a new definition, the important places are places where the speed of an object is considerably slower than in other parts of trajectory.

Zheng et al. (2009a) proposed a model to infer a user's travel experience and the interest of a location. In the first step, trajectories of people were divided into stops and moves. In the second step, density based clustering was applied on stops using different scales (neighborhood, city, country), by forming a tree-based hierarchical graph. For every level of the graph, the interestingness of the location could then be calculated.

Spatial generalization and aggregation of trajectories was proposed in (Andrienko and Andrienko 2010). The characteristic points (stops) of trajectories were discovered. Then, the points were grouped into clusters. The centroids of the cluster were used for building Voronoi tessellation (Okabe et al. 2000). The resulting Voronoi cells were used as splitting regions of the trajectory.

2.2 Place semantics

Semantic enrichment of movement data

(Alvares at al. 2007a; Alvares et al. 2007b) proposed a method of semantic enrichment of trajectories using the stop-and-move model. The method combines external geographical features and finds intersections between important places. Ontology-based semantic enrichment was proposed in (Baglioni et al. 2009) to interpret moving patterns.

Building gazetteers

Popescu et al. (2008) used different Internet sources like Wikipedia, Panoramio and web search engines to automatically collect, identify and categorize geographical names.

Working with photo collections

An algorithm for summarization of photo collections using textual attributes of a photo was presented in (Jaffe et al. 2006). The algorithm, based on Hungarian method (Kuhn 1955), first, performed hierarchical clustering of the region using cluster scoring as a heuristic for cluster creation. The score was composed from such components as *tag-distinguishability, photographer-distinguishability, cluster density, the sum of image qualities*. A visualization environment was proposed to visualize the representative tags for every cluster reflecting the tag's importance. The later work (Ahern et al. 2009) used k-means instead of the Hungarian clustering.

Kennedy et al. (2007) applied content and context based analysis for ranking clusters and finding representative images in a cluster. The cluster ranking was performed to assess how well the photos in a cluster are represented by a tag. They included such aspects as number of users, visual coherence, cluster connectivity, variability in dates. Following, an image analysis was used to select the best representative image from the high ranked clusters. Image organization and an engine for discovering landmark photos was proposed in (Zheng et al. 2009b). A worldwide landmark list was generated using geotagged images and articles from travel guides.

Becker et al. (2009) proposed an ensemble clustering approach (combining different features like titles, tags, keywords, description and content creation time) for event identification (concerts, music festivals, etc.) using photo collections. Different combinations of features were evaluated where the combination of all text features and tags alone achieved the highest performance.

3 Our previous work

This work is a continuation of a previous work on analysis of event-based movement data (Andrienko et al. 2009a), visualization of attractive areas using geotagged photos (Kisilevich et al. 2010a), and on semantic enrichment of visited places and pattern mining (Kisilevich et al. 2010b).

In (Andrienko et al. 2009a), five space and agent-centered analysis tasks for event and trajectory-based data were defined: *spatio-temporal aggre-*

gation of events, spatial clustering of events, spatio-temporal clustering of events, flow analysis and *interactions in space and time*.

The procedure for the visualization of attractive areas was proposed in (Kisilevich et al. 2010a). The process consists of applying density based clustering algorithm to the photo data and calculating the importance score of a photo using kernel density estimation. The importance score was then used for two purposes: (1) as a value for color generation and (2) an estimator for importance of the photo. The photo with the largest importance score was selected as a representative photo in a cluster.

A four-step process was proposed in (Kisilevich et al. 2010b) to extract movement sequence patterns using semantic enrichment process. In the first step, every photo was semantically annotated by a nearest point of interest (POI) using an external database of POIs. The photos that were assigned to the same POI created a semantic cluster with the POI being a representative of the cluster. For example, if the POI is a *train station*, the question can be asked: *Are there people who take photos near a train station* or *how many people take photos near a train station?*

In the second step, photos that were not semantically annotated due to the absence of POI in the neighborhood, were clustered into regions. The obtained regions were considered as new unknown POIs. In the third step, a movement sequence was generated for every user, using the POI identifiers assigned to her photos. In the fourth step, a sequence mining algorithm was applied to the sequences in order to find frequent patterns. As a consequence, the pattern of type A → B could be interpreted like this: *people who visit the area A also visit the area B* and pattern of type A → * → B could be interpreted like: *people who visit the area A may continue to any other place and from any other place come to B.*

The current paper extends the previous works in several aspects:

(1) To reflect the importance of time in cluster creation and analysis, we provide a taxonomy of possible types of spatio-temporal clusters.

(2) The semantics enrichment process is discussed with respect to time.

(3) We discuss the possible external data sources that can facilitate extraction of semantics.

(4) The methods supporting semantics extraction are outlined.

4 Importance of time in understanding space

In Section 1 we argued that time is important for understanding spatial patterns. In this section, we provide a taxonomy of spatio-temporal clusters and present possible data sources of semantics knowledge. Additionally, we discuss methods that facilitate semantic extraction and understanding of spatio-temporal clusters.

4.1 Types of spatio-temporal clusters

Vasiliev (1997) in her “Mapping Time” monograph defines five time categories that are used by geographers: *moments* (moment in time, single instance, dating of an event, no duration), *duration* (intervals, continuance of an event in space), *structured time* (sequences, ordering of events, organization of space by time), *time as distance* (time as a measurement of distance) and *space as clock* (space as a measure of time). These categories define different time interpretations and representations on the map. We derive our taxonomy using the basic definitions of these categories where *moments*, *duration* and *structured time* are the most important for our definition of types of spatio-temporal clusters.
Note: our definition of *events* is similar to that of Becker et al. (2009) – *an event is something that occurs in a certain place at a certain time and characterized by some photo activity of people.* For the more general definitions of *events* in geospatial domain we refer the reader to (Worboys and Hornsby 2004; Cole and Hornsby 2005; Hall and Hornsby 2005).

1. Stationary (moments) - the cluster is called stationary when the subject of a photography does not change in time. Landmarks like monument, museum, airport are good candidates to be found in such a cluster.
2. Reappearing (duration, intervals) - clusters can be reappearing when the photographic activity in the area increases in one time period and decreases in another. Two types of reappearing clusters can be expected:
    a. Regular - clusters in which some periodic events take place. Such events attract people at regular periods. For example: new-year fireworks on the main square of the city or Oktoberfest in Munich.
    b. Irregular - clusters in which aperiodic events take place. Such events attract people at irregular periods.
    For example: a concert or football game.
3. Occasional (moments) - clusters in which some events happen occasionally, e.g. a traffic accident.
4. Regular moving events (structured time) - clusters that represent the same event taking place in different places, e.g. olympic games (taking place on a regular basis in different places) or scientific conferences. This is probably the most complex type of a cluster

since, for each particular place, it is irregular or occasional, and can only be discovered by inter-place comparisons.

In addition to the types of spatio-temporal clusters, we would like to differentiate between two types of semantics: (1) semantics of places and (2) semantics of user's behavior. Semantics of places and users are interconnected and one can enrich the other. Knowing the semantics of a place, we can infer the semantics of users who take photos in these places. For example, the place that has semantics of sport (stadium, football) will indicate that people who take photos there like sport. Likewise, characteristics of a user can have implication on the analysis of the event, user's behavior or a place. Using the profile of the person, we can interpret the profile of the cluster. For example, if we know that a person often takes photos of sport events and nature, then occurrences of his photos in a cluster may help us to identify the semantics of a cluster.

4.2 Potential sources of semantic data

The following is the list of potential data that can be used to extract semantic information about the places or events. The primary source of information is the photo collection data that include all the relevant information like coordinates, tags and titles. Additional source of information are Wikipedia encyclopedic pages and GeoNames database.

Geo-referenced photo collections

Panoramio contains millions of geotagged photos. It is used by Google Maps and Google Earth as one of the visualization layers. Its publically available API allows downloading a photo metadata by providing a bounding box of the desired area. The following is the most important information provided by the API: *photo id* and *coordinates, owner id* and *name, photo url* and *title*.

Another source of geotagged photos is Flickr. Flickr has larger user database and its API allows for receiving more meta information than Panoramio, such as *thematic photo groups, contacts (favorites) of users* and *user information including place of residence (filled by 13% of users)*. The Flickr API does not allow downloading metadata by specifying exact boundaries of the area of interest. Therefore, we used an approach similar to Web crawling. We downloaded all the photo metadata of arbitrarily selected subjects and obtained the list of their contacts as well as the list of groups their photos belong to. The same procedure was iteratively applied on other retrieved users. We began collecting the data from the beginning of June, 2009. By the end of March 2010, we collected 87,665,970 entries from 7,449,723 users and 394,830 thematic photo groups. This amount of data allows us to analyze virtually every place on the Earth if it was previously attended by photographers.

We are aware that user-generated data like photo collections can include incorrect spatial and temporal information. For example, 10,117 photos do not include the date and 55,176 photos are dated after 2010. These photos have to be excluded from the temporal analysis. However, there are cases in which it is difficult or impossible to detect incorrect entry: adjustment of the camera clock to the local time (in most cases adjusted manually by the person) or correct geotagging during the upload process (if the camera was not equipped with GPS). Still, not all of these problems are critical. Spatial aggregation does not require timestamps. Aggregation level in space and time may be larger than position or time reference errors.

Points of interest

Wikipedia database[3] can be used as a source of POI data. This database is an on-going community project aimed at applying geographic annotation to articles describing interesting sites around the world. The database that is currently available contains 815,085 entries of various sites such as cities, landmarks, monuments, buildings, towers annotated with coordinates and titles.

Geographical features

GeoNames[4], a geographical database contains over eight million geographical names and consists of seven million unique features, 2.6 million populated places and 2.8 million alternate names. All features are categorized into one out of nine feature classes, and are further subcategorized into one out of 645 feature codes (mountains, lakes, monuments). The elements of the dataset are organized into *isA* (conceptual inheritance) or *partOf* (spatial inclusion) relations. This dataset is freely available for download or accessible through web services.

---

3 http://de.wikipedia.org/wiki/Wikipedia:WikiProjekt_Georeferenzierung/Wikipedia-World/en#Static_layer

4 http://www.geonames.org

4.3 Methods for semantic enrichment

In this section we describe six main methods for semantic enrichment. Table 1 briefly summarizes the proposed methods.

Table 1. Methods for semantic enrichment

Method	Achieves
Spatial and spatio-temporal clusters	Grouping of the photos into clusters using distance metrics and time-stamps
Text analysis	Grouping of the photos into contexts using textual information (title, tags)
Content-based analysis	Grouping of the photos into contexts using visual similarity
Analysis of events (times series)	Temporal characteristics of the cluster. Frequency of events
POI database and entity relations	Retrieving photo topics (nature, landmarks) and hierarchies of concepts
User profiling	Profiling of users in a cluster using photo semantics

Detection of spatial and spatio-temporal clusters
Clustering can serve as a primary tool for organizing the collection of photos into groups. Among the possible methods used for spatial clustering are: grid based (Girardin et al. 2008a), density based (Adrienko et al. 2009a; Kisilevich et al. 2010a) and hierarchical clustering (Zheng et al. 2009a).

Clustering based on grids is data independent and does not take into consideration the distribution of points. The number of cells should be known in advance and many trial and errors are required to find the suitable number and size of the cells. Density based clustering is based on the neighborhood density (min points) and minimum distance between points using (usually Euclidean) distance function. The method produces an arbitrary number of clusters based on the selected parameters. In general, variations of density based clustering can be applied where the time component is taken into consideration (Andrienko and Andrienko 2009b). For example, the spatio-temporal cluster will be formed if there are more than 5 people that took photos within the range of 1 hour and the distance between them is no more than 100 meters. Such approach would create event-centered temporal clusters.

Hierarchical clustering can be applied to form clusters at different scale levels. Clusters on every level can be analyzed separately and different semantics can be applied at different scales. For example, the tag that identifies the name of a city can be assigned to the cluster on the city scale, while tags that identify names of neighborhoods will be assigned to clusters at the neighborhood scale. Spatial clusters can be produced by bounding the data with time limits. In this way, only the data that falls into time interval will be clustered whereas the clustering algorithm will cluster points without taking the temporal aspect explicitly.

Text analysis

Text analysis of titles and tags can be used for finding events that happen in a cluster or in different parts of the world at the same time or at different times. For this, the representative tags and titles can be obtained for several clusters and matched for similarity. Examples of such events are new-year celebrations that take place at the same time in different parts of the country or the world. At the global scale, difference in time zones should be taken into account by clustering every region separately with adjusted time intervals.

Content-based analysis

Similarity between images in a cluster can facilitate finding different contexts. For example a photo may not have title and tags, or its title is meaningless for analysis (written in a language not known to the analyst or does not represent any event or place). The visual similarity can be still found between other images in a cluster. Thus, if we know that a photo is visually similar to a nearby photo, then these two photos can be grouped together. If needed, the representative photo can then be found. Content-based analysis can also reveal the heterogeneity of the cluster. If people take photos of a single point of interest, then large amount of photos will be similar to each other.

Analysis of time series of counts of events

Time series analysis helps to understand the type of the cluster in terms of its temporal characteristics. For example, we can count the number of people who take photos in the area every day within one year period. The shape of time series indicates the type of a cluster. If the number of events does not differ significantly, the cluster can be classified as stationary. If there are bursts in activity at several intervals, the cluster can be classified as reappearing. If there is a single high peak, the cluster can be classified as occasional.

Additionally, the change of contexts can be detected by using time series analysis. Following the context detection using text and content based analysis, we can classify clusters to any of the spatio-temporal types that we already defined in terms of number and variety of events occurring in a cluster during some time period. Let us consider an example: There is a cluster in which sport events are held in winter. In the summer time, most of the photographers take pictures of the nature. In this case, we can classify the cluster as reappearing if the context (winter sport, nature photography) appears in the cluster several times. If the context appears only once, it is probable that the cluster is occasional.

POI database and entity relations

POI database is a valuable source of semantics. The methods used in automatic gazetteers construction may be employed in retrieving topic of the photo (the title on the photo relates to a landmark or nature) or build an ontology of types from thematic photo groups. In this case, the photos that are in the group *birds* can be classified in general as of type *nature*.

User profiling

As was shown above, cluster semantics may help to determine the semantics of user behavior and vice versa. The user profiling can be performed statically, by using information from all the photos that belong to the user or dynamically, for every cluster or for different time periods.

5 Case study: using time-series analysis and text clustering for extracting semantics of events and places

In this section, we demonstrate analysis of temporal patterns and semantic acquisition using combination of time series, text analysis and external data sources presented in Section 4.

5.1 General scenario

Let us briefly consider a possible scenario by employing the methods presented in Section 4 in the analysis of a geographic region.

1. We apply a clustering algorithm to outline areas of people's visits. Although, the cluster and its size reveal spatial information, it explains neither the dynamics of the interest nor why the place was interesting to the photographers. Therefore, additional introspection should be performed.
2. We apply time-series analysis to investigate peaks of activity. The dynamics of the subject of interest can change over time and the same cluster can encompass different events that also change over time. The temporal component of the semantic enrichment will change the way we analyze spatio-temporal processes and as a result, different patterns of spatio-temporal clusters will appear. The number of taken photos or number of people can be used as dependent variable. At this level, we can already infer the spatio-temporal type of the cluster according to the selected dependent variables.
   While peaks of activity can point to some interesting time periods, we still cannot deduct what was the reason of such activity.
3. We apply clustering techniques for extracting significant keywords using photo tags and/or titles that can show the photographers' intended subjects and point of interests when taking the photos. In fact, text clustering techniques can be applied on all the photos in a cluster or separately on photos for each time interval. This approach can reveal changing trends of place interestingness.
4. We can use external POI databases like Wikipedia to acquire additional information about the cluster if there are points of interest in the area. This information can be matched against the topics obtained from the text clustering step.
5. We can apply image clustering to find representatives that visually highlight the place or in cases where the text clustering does not provide meaningful categories.

6. Methods used in gazetteer creation supported by the domain expert can be employed in building hierarchies of concepts for the given cluster, e.g. a photo of an animal will be classified as nature. Retrieval of other places with similar events can be performed by searching for areas with similar semantics.

5.2 Spatial clustering

We used a subset of photos referring to the territory of Switzerland[5]. For discretizing the space we use a method (Andrienko and Andrienko 2010) that divides the territory to non-overlapping polygons of given size in a way that reflects the distribution of points. In brief, the generalization method groups points into spatial clusters and uses the centroids of the clusters as generating points for Voronoi tessellation (Okabe et al. 2000) of the territory. We applied spatial clustering to the positions of the photos and built Voronoi cells (1183 in total) with average diameter of 2km around the obtained clusters. The whole operation took between three to five seconds using sampling approach with about 20,000 points. The general steps of the algorithm are described below:

Algorithm 1: Territory tessellation
Given: Sequence of positions of points P = {xi, yi} and desired radius *r*
Output: A set of Voronoi cells V
Description of the algorithm:
1. Group the points of P in spatial clusters with desired radius r
   S = SpatialClusters(P; r)
2. Compute the centroids of the spatial clusters C = Centroids(S);
3. Generate Voronoi cells around the centroids
   V = Voronoi Tessellation(D);

---

[5] We don't present results of a complete analysis of the photos on the territory of Switzerland but only provide several examples as illustrations of what can be detected. It is clear that many more events occurred, and the challenge is to develop such methods that will find as many of them as the available data permit.

5.3 Time series analysis

For every cluster, we calculated frequencies of people's visits and the number of taken photos aggregated by month and built time series graph spanning 5 years (2005-2009). Figures 1 and 2 show a part of Switzerland with examples of different temporal patterns of events for selected regions denoted as A, B, C, D, E. Figure 1 shows frequencies of people's visits while Figure 2 displays frequencies of taken photos.

According to Figure 1, people visit the region labeled A in all seasons. In total, 71 people visited this region and took 1721 photos. We can observe a sharp increase in the number of photos taken in every year in January. A possible explanation is that some local repeated event takes place during the winter. Figure 2 shows that the event pattern corresponds to *regularly reappearing* type of the spatio-temporal cluster.

A photo activity in the area near Muensingen (labeled as B) starts in May 2008 showing a smooth increase until July 2008. In total, 38 people (1576 photos) visited the area. Starting in August 2008, there is a steep increase in the number of photos until October 2008 followed by steep decrease until January 2009. Afterwards, there is little activity in April. According to the activity pattern, this cluster can be classified as *occassional*.

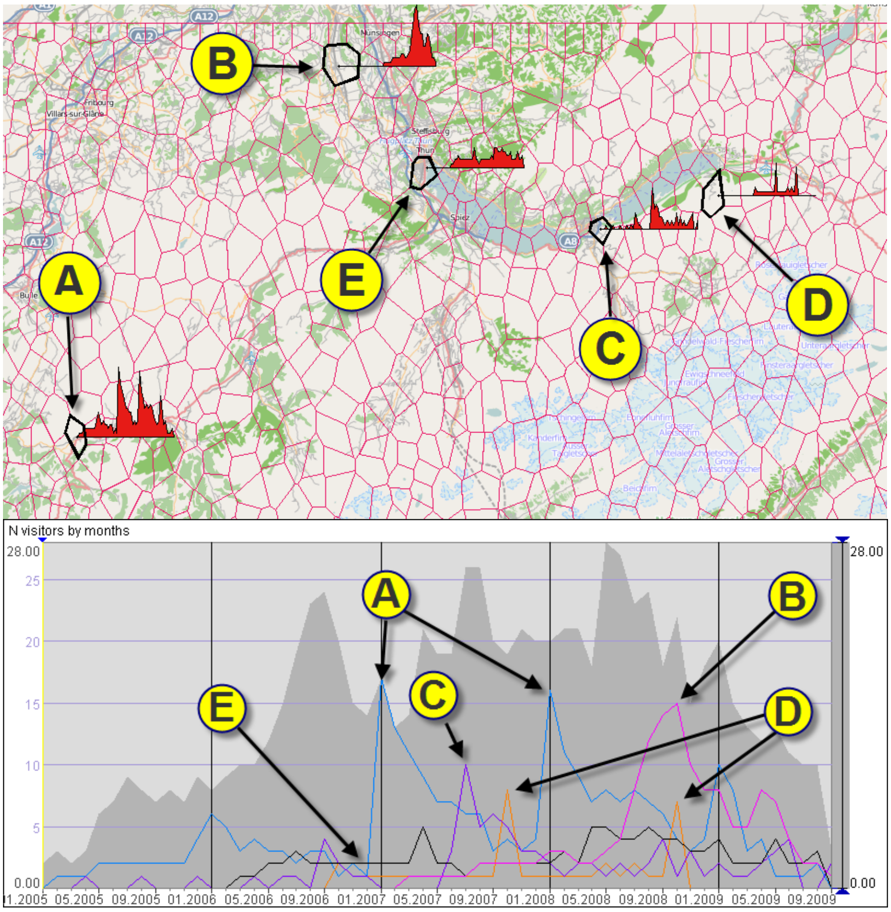

Figure 1. Switzerland. Time series (top) and graph (bottom) of frequencies of people's visits

Regions labeled C (45 visitors, 1637 photos) and D (17 visitors, 1071 photos) have two high peeks in the number of taken photos in July, 2007 (C) and October, 2007 (D) according to Figure 2. For the region C, the number of photos taken in July 2007 constitutes 87% of all photos taken from 2005, while in D almost 90% of photos were taken on October, 2007. According to Figure 1, there are two peaks of visits in October 2007 and 2008. This fact allows us to conclude that while region C and D can be classified as *occasional* in terms of taken photos, cluster D can be classified also as *irregular* with respect to visits of people.

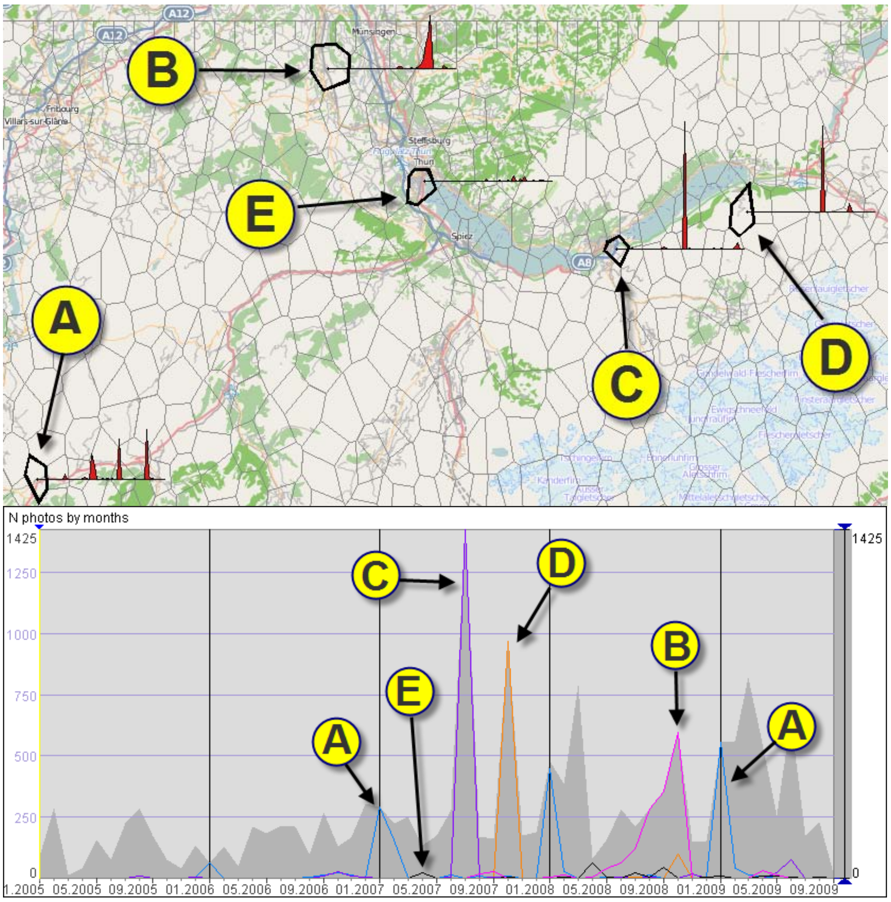

Figure 2. Switzerland. Time series (top) and graph (bottom) of frequencies of taken photos

The inspection of visits and photo activities in region E (25 visitors, 224 photos) shows that there is no variance in the number of photos (30 photos on average) and number of visits (2-3 people on average visit the place each month). This fact suggests that the cluster is of *stationary* type. The area that the cluster occupies corresponds to the town of Thun. The possible reason of this stable activity is that the city attracts photographers during all seasons and there are no important local events taking place during a specific time that could attract more people than during any other period.

5.4 Text clustering

To understand the observed temporal patterns, we extracted tags from photos taken in places A, B, C, D, and E and created two model representations. In one representation, a photo was treated as a separate document (all tags of a photo were saved as one document) in another model the owner of all photos in a cluster was treated as a document, so all unique tags from photos of the owner were collected and saved as one document. We applied two clustering algorithms (the operation took about 1 sec per algorithm) on these models: Lingo (Osiriski et al. 2004a; Osiriski and Weiss 2004b) and STC (Suffix Tree Clustering) with default parameters[6]. These algorithms use different clustering approaches (term-document matrix (Lingo) vs. suffix tree clustering (STC)) and produce different cluster quality (high cluster diversity (Lingo) vs. low cluster diversity (STC)). However, they create overlapping cluster categories. This is an advantage over the methods for automatic representative tag and event extraction proposed in the literature (see Section 2.2), since the photo can have different tags that may describe several categories like (trees, sun, summer). In addition to the understanding of the observed temporal patterns, our goal is to show how results may differ due to model representations, clustering algorithms, language differences or mistakes made during tagging, and stress the importance of visual analytics. Table 1 and 2 present 10 most frequent categories extracted from region A and B using two model representations (owners and photos) and two clustering algorithms applied on them (Lingo and STC). The number of occurrences of every category in documents is given on the right side of each category in parentheses. The tag syntax is preserved.

Let us inspect the obtained cluster categories. The quick look on the categories suggest that people use four languages to tag their photos (Table 2, Lingo owner): English (Snow, 9), Spanish (Suiza, 10), French (Suisse Vaud, 1), German (Schweiz, 3). At least three different contexts can be extracted from the categories: *places* (Vaud – the Swiss canton, Gstaad – small village, Chateux Doex – municipality), *events* (Balloon, Montgolfiere, Festival), *season* (Snow). *Balloon* is the most frequently used term but different variations are used like *hotairabolloon*, *ballon*, *ballons* that are treated as different entities by the clustering algorithms. Similarly, categories of region B are expressed in different languages (Table 3, Lingo owner): German (Autofriedhof, 1), English (Carwreck, 8) and French (Suisse, 5, Lingo photos). Several contexts can be extracted: *places* (Gürbetal, Bern, Kaufdorf), *cars* (Volkswagen Beetle, Ford Zephyr, Vw,

[6] Part of the Carrot2 workbench, http://project.carrot2.org

Fiat), *objects' state* (Abandoned, Cemetery, Carwreck, Rost, Old, Oldtimer, Junkyard, Scrapyard), *nature* (Forest).

5.5 External POI database

We applied a spatial query on external dataset of geotagged Wiki pages using coordinates of the area A and B. One page was retrieved for region A: *Pays-d'Enhaut* (a district in the canton of Vaud in Switzerland). Three municipalities are located within the district: Château-d'Œx, Rossinière, and Rougemont) and three pages for region B: *Gürbetal* (Gürbe Valley), *Historical Carcemetery Gürbetal* (the page exists only in German) and *Rümligen Castle* (the page exists only in German). Surprisingly, the Rümligen Castle was not among the attractions of the region B. However, it turned out that this castle and its adjacent territory is a private area. Pays-d'Enhaut was not tagged presumably because it represents high level of abstraction of the area or the majority of the visitors does not know the name of the district.

We then tried to find relevant information in the Web by supplying the extracted categories to a public Web search service. In this way, we managed to obtain the following information:

A. (Chateau-d'Oex): International Hot-air Balloon Festival takes place every January.
B. (Kaufdorf): a car cemetery with many old-time cars. It is probable that the car cemetery existed before summer 2008 but did not attract much attention. On 5th of July, 2008 an article about it was published by the administrator of the Switzerland group of Flickr users on his web site. This article, evidently, attracted many Flickr users to visit this place in summer and autumn 2008.
C. (Interlaken): Red Bull Air Races, July 14, 2007.
D. (Axalp): Air Show, October 10-11, 2007.
E. (Thun) : Touristic city with several attractions (that do not change over time) like: the twelfth century Castle, sixteenth century town hall, view of the Alps.

5.6 Summary

In the use case, we demonstrated the feasibility of the proposed methodology by applying several approaches of spatio-temporal analysis using geotagged photos combined with time series analysis, semantic interpretation based on text clustering, and extraction of additional knowledge from external sources. We succeeded in finding and explaining the events occurred in five selected regions of Switzerland without any prior background knowledge. The overall runtime performance suggests that the most time consuming operations like spatio-temporal and text clustering can be applied on the countrywide territory without much time overhead.

Table 2. Region A. Cluster categories using 2 model representations and 2 clustering algorithms

#	Lingo (owner) 25 clusters	Lingo (photos) 77 clusters	STC (owner) 16 clusters	STC (photos) 16 clusters
1	Suisse Vaud (12)	Châteaudoex Ballon (962)	Switzerland (39)	Châteaudoex (884)
2	Hotairballoon (10)	Montgolfière Châteaudoex (507)	Châteaudoex (25)	Ballons Châteaudoex, Hotairballoons Montgolfière Paysdenhaut (612)
3	Schweiz Switzer-land (9)	Switzerland (485)	Suisse (24)	Switzerland (485)
4	Châteaudoex Bal-lon (8)	Balloon (440)	Chateaudoex (20)	Balloon, Festiv-al (461)
5	Balloon Chateau-doex (6)	Festival (289)	Balloon (18)	Ballon (324)
6	Montgolfiere (5)	Chateaudoex Switzerland (276)	Vaud (13)	Châteaudoex Ballon (166)
7	Chateux Doex (4)	Ballons Châteaudoex Hotairballoons Montgolfière Paysdenhaut (215)	Ballon (12)	Chateau Doex Balloon, Doex Balloon Fiesta Festival (163)
8	Gstaad (4)	Switzerland Suisse (164)	Schweiz (9)	Balloon Châteaudoex Color Coleur (129)
9	Snow (4)	Ballon Switzer-land (142)	Chateau Doex (8)	Hot Airballon Chateaudoex Switzerland (96)
10	Switzerland Suiza (4)	Chateau Doex Balloon (140)	Ballons (8)	Hotairballoons Châteaudoex Ballons Mont-golfière (67)

Table 3. Region B. Cluster categories using 2 model representations and 2 clustering algorithms

#	**Lingo (owner) 19 clusters**	**Lingo (photos) 76 clusters**	**STC (owner) 16 clusters**	**STC (photos) 16 clusters**
1	Autofriedhof (27)	Gürbetal Autofriedhof Junkyard Scrapyard Schweiz Switzerland (775)	Autofriedhof, Kaufdorf, Switzerland (32)	Autofriedhof, Schweiz Switzerland, Schrott Schrottplatz (1444)
2	Switzerland (21)	Oldtimer Old (229)	Old, oldtimer (13)	Car, Auto (610)
3	Volkswagen Beetle (5)	Rusty (281)	Gürbetal (13)	Decay, Classic, Old Oldtimer (520)
4	Carcemetery (4)	Bern Kaufdorf (218)	Bern (12)	Kaufdorf Gürbetal Schweiz Switzerland Autofriedhof (451)
5	Automobile (3)	Schweiz Suisse (206)	Cars (11)	Schrott Schrottplatz Junkyard Scrapyard (209)
6	Lomolca (3)	Archiv2008 Abandoned (193)	Abandoned, Rost, Schrott (11)	Kaufdorf Bern, Switzerland Schweiz, Carcemetery (206)
7	Schrott Schrottplatz (3)	Cars Cemetery (160)	Beetle, Mercedes, Käfer (10)	Schweiz Suisse Switzerland (206)
8	Carwreck (2)	Car Abandoned (153)	Vw, Fiat (10)	Gürbetal Schweiz Switzerland Autofriedhof (151)
9	Ford Zephyr (2)	Bern Gürbetal (118)	Suisse (8)	Autofriedhof Schweiz Switzerland Kaufdorf (116)
10	Forest (2)	Auto Autofriedhof Autoverwertung (113)	Auto Autofriedhof (5)	Schrottplatz Kaufdorf Gürbetal Autofriedhof (112)

6 Conclusion and future work

In this paper, we proposed a conceptual framework and methodology that would allow analysis of events and places using geotagged photo collections. We defined four main types of spatio-temporal clusters that can be classified by time series analysis and refined by semantic enrichment process using temporal component: *stationary*, *reappearing*, *occasional* and *regular moving*. We discussed methods for analysis of these types of clusters and identified publicly available datasets that can help in the semantics enrichment process. We stressed the importance of the interactively driven analysis that can overcome weaknesses of the purely automatic approaches and help extracting new knowledge from the data. With the example from the selected regions in Switzerland, we showed how time-series analysis, text clustering and additional contextual information can be applied to extract semantics and interpret the region under investigation.

Currently, parts of the framework are implemented as separate services. In our next work, we will build a visual analytics framework integrating the methods proposed in this paper. It will facilitate the discovery and interpretation of the types of spatio-temporal clusters that we defined.

Acknowledgements

This work was partially funded by the German Research Society (DFG) under grant GK-1042 (Research Training Group Explorative Analysis and Visualization of Large Information Spaces"), and by the Priority Program (SPP) 1335 (“Visual Spatio-temporal Pattern Analysis of Movement and Event Data”).

References

Ahern S., Naaman M., Nair R., Yang J. (2009) World explorer: Visualizing aggregate data from unstructured text in geo-referenced collections. Proceedings of the 7th ACM/IEEE joint conference on digital libraries 1-10

Alvares L.O., Bogorny V., Kuijpers B., Macedo J.A., Moelans B., Vaisman A. (2007a) A model for enriching trajectories with semantic geographical information. Proceedings of the 15th annual ACM international symposium on advances in geographic information systems 1-8

Alvares L.O., Bogorny V., Macedo J.A., Moelans B., Spaccapietra S. (2007b) Dynamic modeling of trajectory patterns using data mining and reverse engineering. Tutorials, posters, panels and industrial contributions at the 26th international conference on conceptual modeling 83:149-154

Andrienko G., Andrienko N., Wrobel S. (2007) Visual analytics tools for analysis of movement data. ACM SIGKDD Explorations Newsletter 9(2):38-46

Andrienko G., Andrienko N., Bak P., Kisilevich S., Keim D. (2009a) Analysis of community-contributed space- and time-referenced data (example of Panoramio photos). Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems 540-541

Andrienko G., Andrienko N. (2009b) Interactive spatio-temporal cluster analysis of vast challenge 2008 datasets. Proceedings of the ACM SIGKDD workshop on visual analytics and knowledge discovery

Andrienko G., Andrienko N. (2010) Spatial generalization and aggregation of massive movement data. IEEE transactions on visualization and computer graphics

Baglioni M., Macedo A.F., Renso C., Trasarti R., Wachowicz M. (2009) Towards semantic interpretation of movement behavior. Advances in GIScience 271-288

Becker, H., Naaman M., Gravano L. (2009) Event identification in social media. Proceedings of the ACM SIGMOD Workshop on the Web and Databases (WebDB'09)

Crandall D., Backstrom L., Huttenlocker D., Kleinberg J. (2009) Mapping the worlds photos. . Proceedings of the 18th international conference on world wide web.

Cole S., Hornsby K. (2005) Modeling noteworthy events in a geospatial domain. GeoSpatial Semantics 3799:77-89

Ester M., Kriegel H.P., Sander J., Xu X. (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. Data mining and knowledge discovery 226-231

Jaffe A., Naaman M., Tassa T., Davis M. (2006) Generating summaries and visualization for large collections of georeferenced photographs. Proceedings of the 8th ACM international workshop on multimedia information retrieval 89-98

Giannotti F., Nanni M., Pinelli F., Pedreschi D. (2007) Trajectory pattern mining. Proceedings of the 13th ACM SIGKDD international conference on knowledge discovery and data mining 330-339

Girardin F., Fiore F.D., Ratti C., Blat J. (2008a) Leveraging explicitly disclosed location information to understand tourist dynamics: a case study. Journal of Location Based Services 2(1):41-56

Girardin F., Calabrese F., Fiorre F.D., Biderman A., Ratti C., Blat J. (2008b) Uncovering the presence and movements of tourists from user-generated content. 9th International Forum on Tourism Statistics

Goodchild M.F. (2007) Citizens as voluntary sensors: spatial data infrastructure in the world of Web 2.0. International journal of spatial data infrastructures research 2:24-32

Hall S., Hornsby K. (2005) Ordering events for dynamic geospatial domains. Spatial Information Theory 3693:330-346

Kenedy L., Naaman M., Ahern S., Nair R., Rattenbury T. (2007) How flickr helps us make sense of the world: Context and content in community-contributed media collections. Proceedings of the 15th International conference on Multimedia

Kisilevich S., Mansmann F., Bak P., Tchaikin A., Keim D. (2010a) Where would you go on your next vacation? A framework for visual exploration of attractive places. GeoProcessing 2010

Kisilevich S., Keim D., Rokach L. (2010b) A novel approach to mining travel sequences using collections of geotagged photos. 13th AGILE international conference on geographic information science

Kuhn H. (1955) The Hungarian method for the assignment problem. Naval Research Logistic Quarterly 2:83-97

Okabe A., Boots B., Sugihara K., Chiu S. (2000) Spatial tessellations – Concepts and applications of Voronoi diagrams. John Wiley

Osiriski S, Stefanowski J, Weiss D (2004a) Lingo: Search results clustering algorithm based on singular value decomposition. Intelligent information processing and web mining: proceedings of the International IIS: IIPWM'04 359-368

Osiriski S., Weiss D. (2004b) Conceptual clustering using lingo algorithm: Evaluation on open directory project data. Intelligent information processing and web mining: proceedings of the International IIS: IIPWM'04 369-378

Palma A.T., Bogorny V., Kuijpers B., Alvares L.O. (2008) A clustering-based approach for discovering interesting places in trajectories. Proceedings of the 2008 ACM symposium on applied computing 863-868

Popescu A., Grefenstette G., Moellic P.A. (2005) Gazetiki: automatic creation of a geographical gazetteer. Proceedings of the 8th ACM/IEEE-CS joint conference on digital libraries 85-93

Schmitz, P. (2006) Inducing ontology from flickr tags. Collaborative Web Tagging Workshop at WWW2006 210-214

Spaccapietra S., Parent C., Damiani M.L., Macedo J.A., Porto F., Vangenot C. (2008) A conceptual view on trajectories. Data Knowledge Engineering 65(1):126-146

Toyama K., Logan R., Roseway A. (2003) Geographic location tags on digital images. Proceedings of the 11th International Conference on Multimedia (2003)

Vasiliev I.R. (1997) Mapping time. Cartographica: The international journal for geographic information and geovisualization 34(2):1-51

Worboys M., Hornsby K. (2004) From objects to events: GEM, the geospatial event model. Geographic Information Science 327-343

Zheng Y., Zhang L., Xie X., Ma W.Y. (2009a) Mining interesting locations and travel sequences from gps trajectories. Proceedings of the 18th international conference on world wide web 791-800

Zheng Y.T., Zhao M., Song Y., Adam H., Buddemeier U., Bissacco A., Brucher F., Chua T.S., Neven H. (2009b) Tour the world: building a web-scale landmark recognition engine. Proceedings of ICCV