PhotoGeo: a self-organizing system for personal photo collections

Yuri Almeida Lacerda, Hugo Feitosa de Figueirêdo, Cláudio de Souza Baptista and Marcus Costa Sampaio

Computer Science Department, University of Campina Grande. Av. Aprígio Veloso, 882, Bodocongó, Campina Grande – PB, Brazil. Tel.: +55 8333101122.
{yuri, hugoff, baptista, sampaio}@dsc.ufcg.edu.br

Abstract

Nowadays the photo-capturing devices are no longer limited to digital cameras but include mobile phones, PDAs and others. This is leading to a new problem: a very large number of digital photos captured and chaotically stored in multiple locations without being annotated. This paper presents a new system, called PhotoGeo, for self-organization of georeferenced photos. The proposed system uses metadata and external sources to organize the photos. These external sources include a map-based social network and user’s calendar. Furthermore, spatial clustering and temporal segmentation techniques are used to join the photos into clusters with similar features.

1. Introduction

Digital photography is definitively becoming very popular due to several advantages over the analog one which include easy processing, copying, sharing and cheap storing. The storage devices for digital photos (memory cards, hard-disks, DVD’s etc.) are getting cheaper and larger with every year so scores of digital photos can be stored on such devices. The photo-capturing devices are no longer limited to digital cameras but include mobile phones, PDAs and others. This in itself is leading to a whole new problem: a very large number of digital photos captured and chaotically stored in multiple locations without being adjusted.

There has been an increasing interest on developing software for storing and searching digital photos nowadays. Examples of commercial tools include Google - Picasa, Yahoo! - Flickr and Apple – iPhoto. Apart from that, many research proposals on this subject have also been published such as PhotoCompas [5], Photomap [9] and WMMX [8]. These systems aim to help users to organize their photo collections.

In order to become searchable, an image needs annotation at several levels: semantic, syntactic and structural. This annotation is known as metadata. Examples of such metadata include: author, date, place, subject, people involved, camera make and model, resolution, format, focus, exposition, aperture, and so on. There are at least three ways of annotating photos: automatically, semi-automatically, and manually. Obviously, users would like to have photo annotation with minor effort, thus as much automatic annotation may be done more satisfied will be the users.

Recent studies have shown that events and geographic locations are the most relevant features which users use to remember a given photo [4]. An event is characterized by both user and camera contextual situation during a photo capture. An example of such event would be an international conference held in London.

Apart from that, social networks have been investigated recently. These networks constitute online communities which enable information sharing. Some of such networks provide geographical information sharing. Basically, points of interest (POIs) such as restaurants, hotels, theaters, museum, and so on, are annotated and shared through digital maps among users. This information, available on the Internet, may help to identify location context.

Another important source of information for obtaining user context is related to the use of Web based calendars, in which users may annotate their appointments and share them with the network community.

This paper proposes a new system called PhotoGeo which provides self-organization of georeferenced photos by using spatiotemporal metadata; and data extracted from map-based social networks and Web based calendar. In order to validate the proposed ideas we built a proof of concept prototype. We used a data set of 2,412 photos.

The remaining of the paper is structured as follows. Section 2 discusses related work. Section 3 highlights the PhotoGeo architecture. Section 4 addresses the module that manages user photos. Section 5 presents
the module which is responsible for management of spatiotemporal data. Section 6 focuses on the automatic organization of the photos. Section 7 discusses the experiments and lastly, section 8 concludes the paper and highlights further work to be undertaken.

2. Related Work

The problem of automatic detection of events has been extensively addressed recently [1] [2] [3] [5] [6] [7]. The most used approaches proposed are: visual, temporal, and spatial. The visual approach is based on the content of the photo, i.e., visual features extracted using content-based information retrieval techniques [15]. Photos captured at the same events usually contain similar visual features. Examples of such features include color, texture, shape and structure. For instance, photos taken during a walking at the beach in a sunny day, will have great visual similarity.

On the order hand, the approach based on temporal metadata, use date and time extracted from photo metadata file (such as the EXIF format used in most digital cameras). Hence, this approach is based on the evidence that photos are taken sequentially in a given event. For instance, during a birthday party, people will take photos at different times of the party. There are several works which use this approach such as [1] [2] [3] [6]. Furthermore, there are works which combine visual techniques with temporal ones, like Ravichandran [7].

The third approach uses both spatial and temporal metadata. This approach is based on the fact that a given event usually happens somewhere at sometime. Thus, a same event may contain photos from near locations.

WMMX uses spatialtemporal information to index the photos, but it does not implement automatic detection of events. PhotoCompas performs a classification of photos through location and events. The classification of photos is done using a new clustering algorithm. Moreover, automatic naming for events may be done based on placenames and temporal information.

PhotoMap enables semi-automatic annotation using spatial, temporal and social contexts of a photo [5]. Its metadata include nearby user’s friends, weather conditions, device address, and nearby important objects. PhotoMap provides a mechanism for organizing, sharing and retrieving the photo files. Nonetheless, PhotoMap does not organize the photos into events.

Among the commercial systems, iPhoto detects events automatically but using only temporal information. Picasa does not support automatic detection of events. Flickr enables the user to organize her photos based on spatial metadata and tags. Nonetheless, it does not provide any automatic organization.

For the best of our knowledge there is no related work which takes automatic detection of events based on the three approaches mentioned previously; together with spatially aware social networks and appointments automatically extracted from Web based calendars. These two latter issues constitute the main contribution of PhotoGeo.

3. The PhotoGeo Architecture

This section presents the PhotoGeo architecture which is depicted in Figure 1. PhotoGeo is a Web based photo organizer which uses spatiotemporal metadata. So that, user may insert and query photos in a simple way.

PhotoGeo is designed in a distributed architecture based on n tiers, which is compliant to the Model-View-Controller (MVC) design pattern, so that presentation, control and business model are in different tiers. In the data tier, there is a database server which is responsible for storing and querying the textual, spatial and temporal metadata. Part of these metadata come from external sources such as Wikimapia, which is an example of a spatially aware social network; and the Google Calendar, which is an example of a Web based calendar which implements the iCalendar standard.

The presentation tier uses a Web browser to generate content from JSP and HTML. The controller is implemented using the Struts framework.

The business model is composed of three modules: the photo manager, the data manager and the photo organizer. The photo manager is in charge of monitoring the photo storage structure; the data manager is responsible for spatiotemporal data from external sources. Lastly, the photo organizer automatically organizes the photos into events by using metadata and data extracted from the data collector. Also, this last module is responsible for naming these events. Following sections will further detail these modules.

4. Photo Manager

Apart from structuring the photos in directories, this module is also responsible for extracting metadata and storing them in a database server.

Figure 2 presents an Entity-Relationship diagram for the PhotoGeo schema. This schema contains the
entities: user, album, calendar, appointments, location, event, photo, country, state, city and place.

As it can be seen in the previous schema, PhotoGeo may contain several users, each one may have its calendar and many albums. The calendar stores information extracted from temporal data collector. Each album will have directories which store the many photos. Albums may have photos. Each photo may have a reference to an event, a place and a city. The georeferenced information extracted from the spatial data collector is stored in the tables city, state, country, and place.

Digital photos use a schema for annotating metadata, as for instance the EXIF Standard from Japan Electronics and Information Technology Industries Association – JEITA, which has been incorporated in a large number of digital cameras from different makers. These metadata are automatically extracted by PhotoGeo and inserted in the database server, so that they may be further queried.

5. Data Extractor

The data extractor is divided into two components: spatial and temporal data collectors. The spatial data collector obtains geographic information from a map-based social network, such as Wikimapia. The temporal data collector is responsible for obtaining user appointments from a Web-based calendar which implements the iCalendar standard, as for instance the Google calendar. In the following we detail these two collectors.

5.1. Spatial Data Extractor

The spatial data extractor obtains spatial information from external sources. PhotoGeo uses the Wikimapia, which is a map-based social network, in which people annotates in a World map, places of interest in a given location. PhotoGeo obtains such metadata through a Web-Service (http://wikimapia.org/d/), which generates a KML file with the spatial objects which are inside a given area of interest (also known as bounding box). KML is a XML based language to view geographic information in 2D and 3D in Web based maps. PhotoGeo extracts spatial data from the KML file and insert them in the underlying database.
Social networks enable the knowledge sharing in a given domain to a particular community of interest, forming a knowledge base. More recently, such networks have used maps to insert point of interests. An example of such network is Wikimapia, which aims to store information on points of interest such as shops, hotels, museums, restaurants, pubs, and so on.

Currently, Wikimapia contains approximately 7,137,841 places around the world inserted by users. This data set is constantly updated. Although data obtained from open social networks such as Wikimapia, are not so trustable, there are auditing mechanisms which aim to reduce inconsistency in the data sets.

The information collected from a social network may be used to separate and describe events, to describe photos and help users to annotate them with geographic locations.

5.2. Temporal Data Extractor

The temporal data extractor is responsible for acquiring information related to user appointments. Online Web calendars have been largely used nowadays. There are very complex calendars with a rich set of functions which can be accessed from anywhere and shared with other people.

Hence, when users introduce their appointments in a calendar, this information may be used for automatically organize photos, as they denote the user context in a given time of the day. These appointments describe information on what, when, where and who is taking part in a given appointment.

The iCalendar Standard is implemented by several of such calendars, as for instance the Google Calendar. PhotoGeo accesses the Google Calendar through a URL that the user provides to generate a file using the iCalendar standard. This file is read using the iCal4J framework, and after that the data is inserted into the database server.

6. PhotoGeo Organizer

As it can be viewed in the PhotoGeo architecture (Figure 1), in the business logic layer, there is a module for the photo organization. This organization is done automatically for the user. This organization may be also called non-supervised classification as it is known in the data mining field.

PhotoGeo Organizer has the following requirements:

- the photos should have in their metadata, latitude, longitude, date and time; and
- users must use their Web based calendar, which implements the iCalendar standard, in order to set their events.

The aim of the self-organization is to reduce the effort of users during the organization and annotation of their photo collections.

The algorithm for self-organization is divided into four steps: spatial classification, temporal classification, improvement and self-naming. After executing these four steps there will be a set of photos separated into events.

In the two first steps of the algorithm – spatial and temporal classification-, the photos are grouped according to similar properties. Thus, in the first step, clusters of photos are formed based on geographic proximity. In the second step, each cluster is divided into subgroups which are close in time.

After the separation of the photos in clusters, some inconsistency may arise, for instance photos from a same event that are in different clusters or different events in the same cluster. Thus, it is necessary to execute a further step to correct these errors. This step is known as improvement and it enables to either join or divide clusters, remove elements from a cluster and add them in another one. In order to do that, some comparisons with photo metadata are performed. For instance, if two photos were into different groups due to the large spatial distance between them, but concerning time they are close to each other, then probably they should be in the same group.

The last step is the naming of the extracted groups. Hence, this step matches the photo metadata to spatial and temporal information from the database, so that it may name the events.

The organizer module is based on photo classification algorithm, thus it is possible to have wrong classifications in given events or even the creation of inexistent events. When these exists such inconsistencies the system enables the user to correct them manually. That is, the system also enables the manual organization of photos into events.

In the following we further discuss each of the four steps of PhotoGeo algorithm.

6.1. First step: Spatial clustering

The first step aims to form clusters of photos which are near geographically. Hence, it is possible to use the geographic metadata of the photos to proceed in this step. The location is represented as a 2-D point (latitude, longitude) extracted from the georeferenced photos.
As the location is expressed in two dimensions, it is not possible to use a strategy based on segmentation, which is used in the next step. Thus, a possible approach would be to use the political division of the countries. For instance, the photos would be grouped by cities where they were taken from. The problem with this solution is that some photos may be in a same event but taken from different cities.

Other possible solution to the problem of spatial clustering of photos would be the use of a non-supervised algorithm used in data mining. In this case, the photos taken from nearby places, even if they are from different cities, may be put in the same cluster.

We have adopted this latter approach in our algorithm. In order to do that, we tested two different approaches in order to compare them and choose the best one for the problem we are addressing.

The first approach is the non-supervised classification based on partitioning, using the K-means algorithm. The second approach is the non-supervised classification based on density, using the DBSCAN algorithm.

To test these two algorithms, we submitted a set of georeferenced photos to both. The results obtained were similar using both algorithms, then we decided to use DBSCAN as it is easier to set the input parameters than K-means.

At the end of this first step, there will be a set of clusters which will contain photos close geographically. Figure 3 presents the result of spatial clustering of four events which contain photos distributed spatially (small rectangles represent individual photos). It is interesting to notice that using only spatial clustering it is not possible to detect the events appropriately, as a given cluster may contain more than one event. For instance, two events occur in the city of João Pessoa – Paraíba, Brasil –, during the Brazilian Database Symposium and during the Summer vacation. These events occur at different times, but using only spatial clustering they will be put in the same cluster. Thus, it is necessary to refine the algorithm in order to better segment the photos into events. Next step will introduce the temporal clustering which may improve the segmentation.

6.2. Second Step: Temporal segmentation

According to Cooper et al. [1] and Graham et al. [3] personal photos have an explosive behavior. For instance, people usually take several photos in special events such as birthday party, wedding party, and so on. Thus, the time interval between two consecutive photos taken at two different events tends to be greater than the time interval between two consecutive photos taken at the same event.

The temporal segmentation aims firstly to order the photos based on temporal metadata. Then it is calculated the time interval between consecutive photos. The average and standard deviation of the time intervals are also computed. Lastly, consecutive photos are separated into different clusters if the time interval is greater than a given threshold \( t \).

The value of \( t \) is given by the sum of the average, standard deviation, and a constant \( k \). In our prototype we have chosen \( k = 1 \) hour. The constant \( k \) is used to adjust the cases in which the average and standard deviation have values too small. At the end of this step each cluster will have a set of events. This step may cause at least two kinds of inconsistency:

- several events that are joined in a unique one;
- a unique event is segmented into several events.

Figure 3. Example of spatial clustering of four events using a map extracted from Google Maps.

6.3. Third step: Improvement

This step aims to reduce the inconsistencies from the previous step. In this step the algorithm extracts external information aiming to improve the quality of segmentation. There are at least two types of information that can be extracted externally: spatial and temporal.

To obtain spatial data externally, the algorithm requests them from the SpatialDataCollector module. It asks for places which are close to the cluster generated in step 1. The input parameter is the bounding box of each cluster and a buffer of \( N \) meters. In our test we used \( N = 500 \).

Using the geographic coordinates of each photo it is possible to find out the place name. This information is...
extracted from the Wikimapia social network, which has places from all over the world. Concerning the temporal clustering, the external information extracted comes from the TemporalDataCollector which extracts information from user’s calendar.

We take two hypotheses to solve temporal inconsistencies:

- If a given event has photos which are taken from different user appointments, then this event will be divided.; and
- If there are two or more consecutive events and they are in the same appointment in user’s calendar, then these events should be unified into a unique one.

6.4. Fourth step: Self-naming

Finally we reach the fourth and last step of the algorithm. This step is responsible for naming the events obtained from previous steps. We use the spatial and temporal information to name the events.

The following format was adopted to name an event: “Event name + Place name + City + State + Country + Begin data + Event duration”. From the user calendar we obtain the event name which matches with the date the photo was taken. If there are in a unique event photos from more than one place, they will be annotated with spatial information related to city, state and country. If there are more than one city, the spatial information will be restricted to state and country, and then successively.

Information on date and event duration are extracted from the database. An example of an event would be: “ACM Symposium on Applied Computing – Vila Galé Hotel – Fortaleza – Ceará – Brazil – 05/16/2008 - 5 days”.

7. Experiments

PhotoGeo outputs a set of photos clustered into events. In our experiments we used five collections of 2,412 georeferenced photos (http://lsi.dsc.ufcg.edu.br/collection.zip). Also, we used the calendar of the owners of the collections.

Most of photos were captured from Brazilian cities. Collection A has 596 photos which are distributed in approximately one year and eight months (see Figure 4). During this time span there are months with zero to three events.

Collection B contains 162 photos and has a time span of eleven months. Collection C contains 238 photos and has a time span of approximately eleven months. Collection D contains 541 photos and has a life span of one year. Lastly, collection E contains 812 photos taken during two years and a half. This collection has photos from Brazil and Canada.

In the collections A and B almost all photos have events registered in the user’s calendar. In the collection D, only 25% of photos have appointments registered in the calendar. Collections C and E have no events described in users calendars.

![Figure 4. Photo Temporal Distribution](image)

We used these photos to validate our PhotoGeo algorithm. After running the tests, the users evaluated the effectiveness of the proposed algorithm, by answering a questionnaire with the following questions:

- Question 1: How many events were correctly detected?
- Question 2: How many names have well represented the events?
- Question 3: Try to find five photos without regarding their events. The proposed storage structure has enabled you to find them: ( ) slow ( ) normal ( ) fast?
- Question 4: What is the importance that you give to self-organizing your photos by using the metadata from the Web calendar? ( ) low ( ) medium ( ) high
- Question 5: In general, how do you define the self-naming given to the events? ( ) very bad ( ) bad ( ) normal ( ) good ( ) excellent.
- Question 6: Do you have any suggestion for the storage structure of PhotoGeo?

The main contribution of this research is on the use of external sources in the step 3 of our algorithm to better separate the photos into events. Hence, we have done two kinds of analysis: full and partial.

The full analysis was done by executing all steps of the proposed algorithm. We have used the following metrics:

- the quality of classification;
- the quality of the self-naming of events; and
• the importance of the calendar.

On the other hand, the partial analysis has taken into account only the steps 1, 2 and 4. Thus, the influence of the step 3 in the results is evaluated.

7.1. Full analysis

According to the results obtained from the execution of the entire algorithm, users from collections A and B have considered the classification and naming of the events excellent, as can be seen in Figure 5. This result was achieved as almost all events were annotated in users’ calendar.

In collection C the PhotoGeo results were evaluated as good obtaining 83% of user acceptance. The user mentioned that there was an event that was divided into three others. In this particular event, the photos were captured into three consecutive days. The user liked the use of a calendar and mentioned that the PhotoGeo system is very useful to classify automatically the photos in a first step, and then the user in a second step may manually refine the annotations.

Figure 5. Result of the event clustering

Collection D had two inconsistencies during event separation. In the first one, an appointment was set to a given day (e.g. a birthday party), but this event went from midnight. Hence, as the time interval was greater than the average, these photos were allocated into different events. The second inconsistency was related to a user trip which took two days and should be put in a unique event, but it was considered two events. Nonetheless, this trip was not annotated in the user calendar.

Finally, in the collection E 84% of events were separated correctly. Only 20% of the events had a correct naming. This happen due to the fact that in the PhotoGeo database there was no information about Canadian cities, so that photos taken in that country could not have the spatial dimension properly set, but only the temporal one. Figure 6 presents the result of the applied questionnaire.

The first metric that we are going to analyze is the correct separation of the events. Based on the results of question 1 (Figure 6(a)) there was an average of 87% of approval of users. Another important result is related to Figure 6(c), which presents that the majority of users had rapid access to their photos.

Figure 6. Results from the questionnaire
With regard to the quality of the self-naming of events we use the answers to questions 2 and 5. Figure 6(b) presents the result for the question question 2. The collection E was not very well evaluated. This happen because PhotoGeo does not have data on Canada. This problem could be solved using world wide gazetteers. The results for question 5 (Figure 6(d)) vary between average and excellent.

7.2. Partial Analysis

The results of the classification without the step 3 of the algorithm may be visualized in Figure 5. In this analysis it was verified whether the event separation is according to user requirements.

One could notice that in all collections the results using the full algorithm were better. By not using the step 3, the average of user acceptance drops from 87% to 60%. These results may prove the importance of the inconsistency correction which done in this step.

8. Conclusion

This paper presents a system for self-organizing personal georeferenced photo collections using map-based Wiki social interaction and users calendar. The proposed algorithm is based on four steps: spatial clustering, temporal segmentation, improvement and self-naming. The output of the algorithm contains a set of photos separated into events, each one of these may have self-naming.

In order to validate the proposed ideas we built a prototype and used a set of georeferenced photos. Then, we made a user analysis of the efficacy of the prototype. The results obtained from this evaluation were satisfactory, and show the importance of the user calendar. The self-naming had a 77% of approval. Lastly, the third step – improvement – has proven to be very useful as it improved the results in 27% on average.

As further work we intend to use other algorithms for spatial clustering. Also, we intend to use gazetteers to improve the quality of the self-naming process.

9. Acknowledgements

This work was partly supported by CAPES (Brazilian Committee for Graduate Studies). The authors would like to thank the people who participated in the evaluation of the PhotoGeo system.

10. References


