



Sharing, Discovering and Browsing Geotagged Pictures on the Web

Carlo Torniai¹, Steve Battle, Steve Cayzer
Digital Media Systems Laboratory
HP Laboratories Bristol
HPL-2007-73
May 15, 2007*

photos, semantic
web, semantic
photos, geo,
geolocation, gps,
mobile, digital
photography,
metadata,
geographic
metadata,
photo metadata

In recent years the availability of GPS devices and the development in web technologies has produced a considerable growth in geographical applications available on the web. In particular the growing popularity of digital photography and photo sharing services has opened the way to a myriad of possible applications related to geotagged pictures. In this work we present an overview of the creation, sharing and use of geotagged pictures. We propose an approach to providing a new browsing experience of photo collections based on location and heading information metadata.

* Internal Accession Date Only

¹Multimedia Integration & Communication Centre, Universita' di Firenze, Firenze, Italy

Approved for External Publication

© Copyright 2007 Hewlett-Packard Development Company, L.P.

Sharing, Discovering and Browsing Geotagged Pictures on the Web

Carlo Torniai . Steve Battle . Steve Cayzer

***Abstract.** In recent years the availability of GPS devices and the development in web technologies has produced a considerable growth in geographical applications available on the web. In particular the growing popularity of digital photography and photo sharing services has opened the way to a myriad of possible applications related to geotagged pictures. In this work we present an overview of the creation, sharing and use of geotagged pictures. We propose an approach to providing a new browsing experience of photo collections based on location and heading information metadata.*

1 Introduction

With the growing popularity of digital photography, there is now a vast resource of publicly available photos. The availability of cheap GPS devices has made it easy to classify, organize and share *geotagged* pictures on the Web. Geotagging (or geocoding) is the process of adding geographical identification metadata to resources (websites, RSS feed, images or videos). The metadata usually consist of latitude and longitude coordinates, but they may also include altitude, camera heading direction and place names.

There has recently been a dramatic increase in the number of people using geo-location information for tagging pictures. The result of a query for pictures with *geo:lat* tag uploaded in Flickr¹ returns 16,048 results between October 2003 and October 2004, 89,514 results for the following year and 171,574 results for the period from October 2005 to October 2006. In principle, the availability of geotagged pictures allows a user to access photos relevant to his or her current location. However in practice there is a dearth of methods for discovering and linking such spatially (and perhaps socially) related photographs. In this paper we focus on geotagged pictures, describing how to add geo-location information to pictures; how geotagged pictures can be organized and shared on the web; what kind of applications can be built using pictures provided with geo-location information.

The paper is organized as follows: services and applications related to geotagged pictures are described in Sect. 2; our approach to using geotagged pictures is presented in Sect. 3; possible metadata and distributed environment enhancements, together with benefits and drawbacks of the proposed approach are discussed in Sect. 4 while in Sect. 5 we provide conclusions and some future work.

2 How to Create, Share and Use Geotagged Pictures: Services and Applications

In the web community, geotagging is becomingly increasingly prevalent in photo-sharing services that allow users to add metadata, including geo-location information, to pictures. The generated metadata are then used to classify and retrieve images. Once pictures are geotagged, different kinds of applications can be developed in order to present relations among them and explore new ways of browsing pictures. In this section we discuss services providing tools for geotagging pictures, and applications that use geotagged resources.

2.1 Tools for Geotagging Pictures

Flickr is perhaps the premier photo sharing website at the time of writing. Following the increasing number of pictures that are manually geotagged by users, Flickr has recently launched its own service for adding latitude and longitude information to a picture. The tool allows a user to select on a map the location in which a picture is taken then the corresponding latitude and longitude information is added as metadata to the picture. The process of manual geotagging is quite lengthy, especially the first time we look for a location. The service uses Yahoo Maps, and the accuracy in the location specification is not fine enough to identify the precise point in which a picture has been taken. In addition, the process of doesn't add the latitude and longitude to the picture as standard *geo:long* and *geo:lat* tags nor as EXIF information but rather in an unknown format decoupled from the picture. On the other hand, pictures already geotagged manually with the proper *geo:long* and *geo:lat* values can be automatically referenced on the map.

Zoomr² is another photo sharing service that provides a geotagging tool. If a picture with EXIF information on latitude and longitude is uploaded it is automatically placed on the map. The process of manually geo-referencing an image is similar to Flickr but here Google Maps is used, providing a more accurate and satisfying geotagging process.

Picasa³ is a desktop application for organizing digital photos. Recently, a beta version of Picasa (Picasa Web Albums) with a geotagging service integrated with Google Earth⁴ has been released. Google Earth is used to select the location in which the pictures have been taken, and the latitude and longitude information are added to their EXIF metadata. This tool is very user friendly and effective, taking advantage of the powerful Google Earth desktop application.

2.2 Applications Using Geotagged Pictures

The applications for geotagged pictures available on Flickr provide a view of nearby pictures and a browser for geotagged pictures⁵. When looking at a picture on the map, the option "Explore this map" is available and clusters of nearby pictures are displayed. Similarly, in the geotagged images browser a world map with clusters of geotagged pictures are presented. Clicking on a cluster shows thumbnails of the contained pictures (Figure 1).

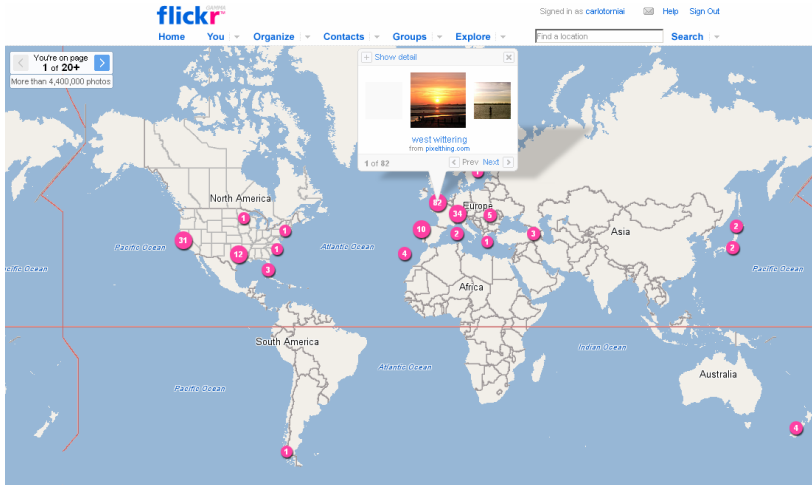


Figure 1. Flickr geotagged images browser

Zoomr provides a similar application for visualizing pictures on a map. The “browse nearby pictures” feature presents both a map and a textual navigation based on pictures clustered according to the distance from the current picture (Figure 2).

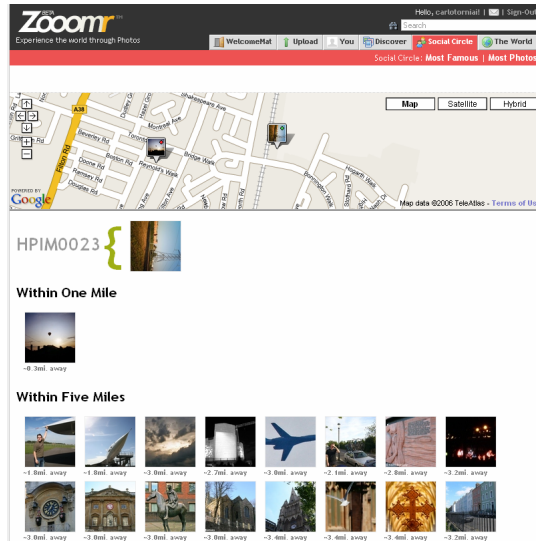


Figure 2. Zoomr nearby pictures view

Picasa, as mentioned, uses Google Earth to visualise geotagged images. It can also be combined with Flickr or Zoomr to upload already geotagged pictures.

Other web based services for geotagging pictures are available. Zoto⁶ provides services similar to Flickr and Zoomr but with less features, while jpgEarth⁷ allows user to upload pictures related to a location picked up from a google map, but no search or clustering features are available.

2.3 Interaction with Geotagged Pictures

The services and applications described so far provide tools for geotagging pictures and applications that use geotagged data to obtain a cluster view of images on a map, or to find nearby pictures. Other interesting applications take advantage of geotagged resources, building new paradigms of interaction.

Lo.ca.lise.us⁸, a service built on top on Flickr and Google Maps, displays geotagged pictures and provides tool for geotagging pictures and uploading them directly into Flickr. The interesting feature is the possibility to interact with tags and users in order to create and share custom ‘views’ of maps, users and related pictures. Other interaction possibilities are provided by flickr-based greasemonkey⁹ scripts which enable browsing of pictures based on location information. GeoRadar is a script to search closest photos. A radar screen is displayed in the picture page and green points on the radar indicate the locations of nearby photos. Thumbnails of nearby pictures are displayed in order of distance from current photo; clicking on thumbnail causes the corresponding green point on the radar to turn red and a small compass to appear showing the direction from the current picture to the one selected (Figure 3).

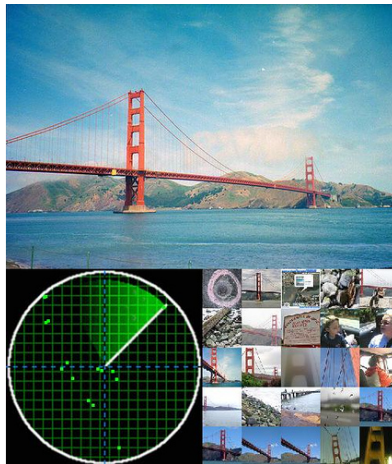


Figure 3. GeoRadar screenshot



Figure 4. Photo Compass screenshot

Flickr Photo Compass is another script that displays the 8 closest photos to the actual one in the cardinal and intercardinal directions: N, NE, E, SE, S, SW, W, NW. By clicking on the direction icons the user can move around and find other photos (Figure 4).

Table 1. Geotagged images applications and services overview

Applications / Services	Goal	Geo related services	Standard format	Other services and technologies
Flickr	Photo Sharing	-Geotagging tool -Geotagged picture browser	None	Yahoo Maps
Zoomr	Photo Sharing	-Geotagging tool -Geotagged picture browser	None	Google Maps
Picasa	Photo Organizer	-Geotagging tool	EXIF	Google Earth
Lo.ca.lise.us	Geotagged pictures browsing service	-Geotagging tool -Social network related to picture	None	- Flickr - Google Maps
GeoRadar	Enhanced Flickr interaction	Location based image browsing	None	- Flickr - Greasemonkey
Photo Compass	Enhanced Flickr interaction	Location based image browsing	None	- Flickr - Greasemonkey

Table 1 presents an overview of these services and applications. Notice that most of the services and applications are related to one community and to one service (Flickr). The main mode of interaction is to locate pictures on, and browsed using, a map. In our view this is only one of the potential benefits of georeferenced data, and we discuss in the next section some recent research projects which use geo-location information to create novel photo browsing experiences.

2.4 Research on Geotagged Pictures

In Sharing places¹⁰, multimedia annotation (photo, video and audio) is associated with physical locations to create a '*mediascape*'. These trails, based on GPS information and enriched with annotations, can be accessed over the web or downloaded to a suitable device (e.g. PDA) and experienced in the real world. The trails can be tagged, published for others to find, remixed and shared.

Images are arranged according to their location in the World-Wide Media Exchange (Toyama et al. 2003) while time and location are used to cluster images in PhotoCompass (Naaman et al. 2004). Realityflythrough (McCurdy and Grishwold 2005) presents a very friendly user interface for browsing video from camcorders equipped with GPS and tilt sensors, and a method for retrieving images using proximity to a virtual camera is presented in (Kadobayashi and Tanaka 2005).

In Photo Tourism (Snavely et al. 2006) a system for interactively browsing and exploring large unstructured collections of photographs is presented. Using a computer vision-based modelling system, photographers' location and orientation are computed along with a sparse 3D geometric representation of the scene. Full 3D navigation and exploration of the set of images and world geometry, along with auxiliary information such as overhead maps and geo locations is provided by the photo explorer interface.

These approaches provide a user experience enhanced by geo-information but don't rely on standard format for metadata nor provide a distributed environment for exchanging metadata. As already pointed out (Cayzer and Butler 2004) we believe that metadata related to pictures and their locations should be expressed in a common and sharable standard so that they may be used by other applications. Sharing picture metadata across a distributed environment using an open standard such as RDF (W3C-RDF 2002) can lead to interesting evolutions in the way in which pictures and other multimedia geotagged content are shared, discovered and browsed.

3 Building Applications with Geotagged Pictures

Our contribution in applications related to geotagged pictures explores the kinds of metadata that can be captured at the time a photo is taken, and ways to link photos together according to this metadata. The objective of our work is to create an experience where someone can view a photo on the web, then jump to other photos in the field of view or taken nearby. It draws on the network effect of the web by including not only the user's own photos but any photo that can be discovered with

suitable metadata. This includes location (GPS or other mobile location) and heading information to identify the position and direction of the camera. The photos discovered may have been taken by different people and are shared on the web. The key to this linking is location and heading metadata attached to the photo. There are no explicit hyper-links between photos, making it easy for people to contribute. Automatic linking is achieved by the discovery of photos on the *semantic web*.

The main idea is to capture RDF metadata related to pictures and photo collections and share these descriptions in a distributed environment. Spatial relations between nearby pictures are discovered by means of inference over their RDF descriptions. We have implemented a proof of concept system comprising the algorithm for inferring spatial relations between different pictures (see Sect. 3.2), a distributed system for sharing metadata and picture discovery, and a web client that uses these RDF descriptions to provide a browsable interface, allowing users to explore shared photo collections through their spatial relationships with each other (see Sect. 3.3).

This interface can be seen also as a new paradigm of interaction with maps and urban landscape providing different ways of touring and exploring places. Once different users can share metadata about their photo collections, and when mechanism to relate pictures, locations, events and users according to these metadata are provided, the urban landscape and maps in general can be transformed in a multi-layered virtual space with different levels of interaction. Pictures can be the entry points to move towards other pictures and locations, look for other places, discover events and meet people: the map, the users and the georeferenced pictures together with their metadata, become part of one big “mashup interface” (Figure 5). Suppose that, using this enhanced interface, I’m looking at pictures of Time Square in New York moving around the place through pictures taken from different users. At a certain point I want to see the same place at night, during the beginning of the year eve. Switching seamlessly from location-based to time-based browsing I will be prompted from pictures of the place at that particular time and date. An interesting shot of the square has been taken by Leo and I want to see other picture

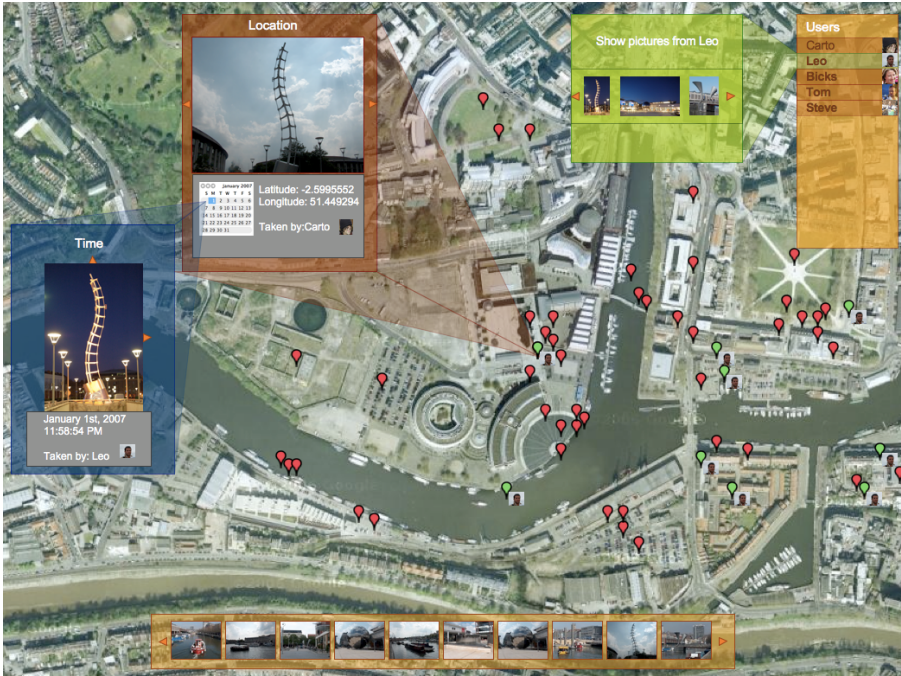


Figure 5. An example of multi-layered map interface

that Leo took. I discover that Leo has lot of pictures taken at the Madison Square Garden during a Bruce Springsteen gig I attended too. I switch to event/social-based navigation looking for who took pictures of that event that night and I am prompted with users avatars displayed on the New York City map according to the place they live. My attention is captured by the high concentration of users in the Long Island area and I begin to explore the place getting back to the picture/location-based navigation. The above example shows that, providing appropriate sharing mechanisms of metadata expressed in a standard format, location/time/event/user-based navigation of maps through geotagged pictures, can be enabled.

3.1 Metadata Definition

To define the structure and the content of metadata for picture description we consider the existing RDF schemata that capture the following information:

- Latitude
- Longitude

- Heading information
- Author
- Date and time
- Title
- Annotation about location
- EXIF metadata

We used both an RDF translation of the EXIF standard (W3C-Exif 2003) and Basic Geo vocabulary (W3C-Geo 2003) for latitude and longitude. Heading information and camera related data (focal length, focal plane resolution and so on) are expressed using the RDF format of the EXIF standard. Dublin Core (DCMI 2006) was selected for defining author, title, date, time and annotation about location.

To describe the location context we used the Dublin Core *dc:coverage* tag. The purpose of *dc:coverage* is to define the extent or scope of the content of a resource and typically includes spatial location (a place name or geographic coordinates), temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity). Additionally, we introduced a hierarchical order into the values of this tag, namely: Place or area, City, Country. For instance values representing a picture taken at the Watershed in Bristol would be, “Watershed, Bristol, UK”. Furthermore, this hierarchical tag could be used to generate a less specific tag, “Bristol, UK”, providing more flexibility in the discovery process.

A collection of pictures is expressed in RDF as a list of images with a title and a creator expressed through the *dc:creator* and the *dc:title* tags. An example of an RDF description of a picture is shown in Listing 1.

Listing 1. Example of an RDF picture description in N3 notation

```
@prefix mindswap: <http://www.mindswap.org/~glapizco/technical.owl#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix exif: <http://www.w3.org/2003/12/exif/ns#> .
@prefix geo: <http://www.w3.org/2003/01/geo/wgs84_pos#> .

<http://bigpicture/pictures/HPIM0459.JPG> a mindswap:Image;

  # Coverage data
  dc:coverage "Bristol, UK" ;

  # Geo Information :
  # Latitude in decimal degree notation (WGS84)
  geo:lat "51.4496826" ;

  # Longitude in decimal degree notation (WGS84)
  geo:long "-2.5976958" ;

  # Latitude in degree-minutes-seconds notation
  exif:gpsLatitude "51 26 58.0" ;

  # Latitude reference
  exif:gpsLatitudeRef "N" ;
```

```
# Longitude in degree-minutes-seconds notation
exif:gpsLongitude "2 35 51.0" ;

# Longitude reference
exif:gpsLongitudeRef "W" ;

#Image Direction
exif:gpsImgDirection "320.00" ;

dc:creator "Carlo Torniai" ;
dc:date "2007:04:18T15:48:59" ;
dc:format "image/jpeg" ;
dc:title "Cabot Tower from waterfront" ;
dc:type "image" ;
exif:brightnessValue "2389/256" ;
exif:componentsConfiguration "48 51 50 49" ;
exif:contrast "0" ;
exif:customRendered "0" ;
exif:dateTimeDigitized "2007:04:18 15:48:59" ;
exif:dateTimeOriginal "2007:04:18 15:48:59" ;
exif:focalLength "44.63" ;
exif:focalPlaneResolutionUnit "3" ;
exif:focalPlaneXResolution "20000000/555" ;
exif:focalPlaneYResolution "20000000/555" ;
exif:gpsVersionID "2 0 0 0" ;
exif:imageLength "1952" ;
exif:imageWidth "2608" .
```

Each image is defined according to the Image class described in the mindswap ontology¹¹. The annotation about location is included in the *dc:coverage* value. Latitude and longitude information in degree-minute-second (d-m-s) notation are represented by *exif:gpsLongitude* and *exif:gpsLatitude* while *geo:lat* and *geo:long* contain the decimal degree (WGS84) notation. North or south latitudes are indicated by *exif:gpsLatitudeRef*; while *exif:gpsLongitudeRef* specifies whether a longitude is east or west. The *exif:gpsImgDirection* indicates the direction of the image when it was captured. The range of values is from 0.00 (north) to 359.99. A collection of pictures is defined as an RDF list of images with a title and a creator as shown in Listing 2:

Listing 2. Example of RDF pictures collection

```
<rdf:Description>
  <dc:creator>Carlo Torniai</dc:creator>
  <dc:title>collection_3</dc:title>
  <rdf:type>http://hp.co.uk/semPhoto/photo#Collection</rdf:type>
  <rdf:first>
    <mindswap:Image
rdf:about="http://bigpicture/pictures/HPIM0428.JPG"/>
  </rdf:first>
  <rdf:rest rdf:parseType="Collection">
    <mindswap:Image
rdf:about="http://bigpicture/pictures/HPIM0429.JPG"/>
    <mindswap:Image
rdf:about="http://bigpicture/pictures/HPIM0432.JPG"/>
    ....
```

```
</rdf:rest>  
</rdf:Description>
```

3.2 Discovering Pictures Relations

RDF descriptions capture the spatial relationships between pictures. We define a simple algorithm that extracts the following information:

- Field of view evaluation (moving forward - zoom)
- Spatial relations (turning - pan)

The field of view relation describes the fact that from a picture taken at A ($image_A$) one can *move towards* the picture taken at B ($image_B$). The way in which the field of view is evaluated is shown in Figure 6. This states that for $image_B$ to be in the field of view of $image_A$, one must be able to see point B in

$image_A$, and $image_B$ must have a similar heading direction to $image_A$.

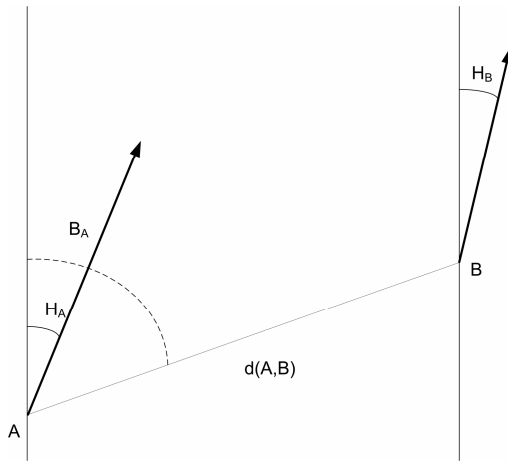


Figure 6. Field of view evaluation. If $|H_A - B_A|$ is less than a given threshold point B is in the field of view of point A. If $|H_A - H_B|$ is less than a given threshold then the pictures have a similar heading. If these conditions are met then $image_B$, taken at B is in field of view of $image_A$ taken at A.

The method for field of view evaluation is shown in Algorithm 1. $FOV_THRESHOLD$ has been set to 150 meters, while the bearing angle threshold T_{bear} and the heading direction threshold T_{head} have been heuristically set to 20 degrees.

Algorithm 1. Field of view evaluation algorithm

```
for each image pair ( $image_A$ ,  $image_B$ ) in the collection  
  evaluate distance  $d(A, B)$  // distance between A and B  
  if  $d(A, B) < FOV\_THRESHOLD$  then
```

```

evaluate  $B_A$  // bearing angle between A and B
if ( $|H_A - B_A| < T_{bear}$ ) // ie point B can be seen in imageA
    AND ( $|H_A - H_B| < T_{head}$ ) then // ie imageB and imageA have similar
    // headings
    set  $fov\_relation(image_A, image_B)$ 

```

Spatial relations refer to the direction in which you have to turn, standing in A, in order to see the picture taken at B. If the pictures $image_A$ and $image_B$ have been taken within a given range of each other we consider the pictures to be taken at the same location so that their relative spatial position is given by the difference between their heading information. Referring to Figure 7 we can say that you can turn right from A to B.

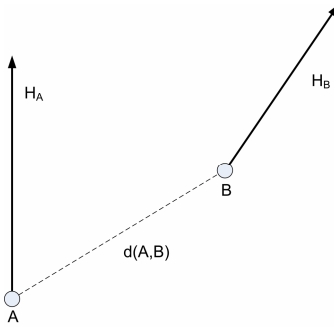


Figure 7. Spatial relation evaluation. If $d(A, B)$ is less than a given threshold then the spatial relation is given by $(H_A - H_B)$

The algorithm for spatial relation discovering is shown in Algorithm 2. $DISTANCE_THRESHOLD$ has been set to 15 meters taking into account the GPS accuracy.

Algorithm 2. Spatial relations discovering algorithm

```

for each image pair ( $image_A, image_B$ ) in the collection
    evaluate distance  $d(A, B)$  // distance between A and B
    if  $d(A, B) < DISTANCE\_THRESHOLD$  then
         $diff\_angle = H_A - H_B$ 
        case  $diff\_angle$ 
            0 to +22.5 OR -337.6 to -360 : position = Front
            +22.6 to +67.5 OR -292.6 to -337.5 : position = Front_Right
            +67.6 to +112.5 OR -247.6 to -292.5 : position = Right
            +112.6 to +157.5 OR -202.6 to -247.5 : position = Back_Right
            +157.6 to +202.5 OR -157.6 to -202.5 : position = Back
            +202.6 to +247.5 OR -112.6 to -157.5 : position = Back_Left
            +247.6 to +292.5 OR -67.6 to -112.5 : position = Left
            +292.6 to +337.5 OR -22.6 to -67.5 : position = Front_Left
            +337.6 to +360 OR -0.1 to -22.5 : position = Front
        set  $spatial\_relation(position, image_A, image_B)$ 

```

The output of the algorithm is an RDF model describing the relations discovered between the pictures. We have defined simple properties describing the field of view (*has_in_fov*) and spatial relations (*Front*, *Left*, *Right*, *Back_Left*, *Front_Right*, and so on).

3.3 Distributed Environment

A distributed test environment has been implemented in order to evaluate the pictures discovering process and the algorithm for relations evaluation across different photo collections. This environment is composed of a set of “clients”. Each client exposes its photo collection(s) (i.e. the RDF collection descriptions files) to its peers by means of SPARQL (W3C 2006) endpoint(s). The clients hold, but do not need to share, the inferred spatial relations between pictures.

The process of discovering related pictures is described in Algorithm 3. Discovery is performed through queries against remote clients, and does not require the relatively expensive computation of spatial relations. Instead, photos are selected by their coverage, expressed as relatively simple location hierarchies.

Algorithm 3. Pictures discovering algorithm

```
expand the coverage tags in the collection
for each distinct coverage
  for each client
    query client for coverage entries
    evaluate relations(client_collection, virtual_collection)
```

The first step is the expansion of hierarchical *dc:coverage* tags in a client’s own collection. This allows a SPARQL query to retrieve photos at varying degrees of granularity. For example, given a picture with the coverage “Peto Bridge, City Center, Bristol, UK ” the expanded coverage tags will be the following:

```
<dc:coverage> Peto Bridge, City Center, Bristol, UK </dc:coverage>
<dc:coverage> City Center, Bristol, UK </dc:coverage>
<dc:coverage>Bristol, UK </dc:coverage>
```

The client asks other known clients for pictures that have the same coverage entries than the ones related to its own collection. This is performed by means of SPARQL queries against (similarly expanded) *dc:coverage* tags. It would also be possible to use GPS latitude and longitude information in the SPARQL queries but this would be relatively expensive. As a result of this query process a list of images is returned to the client. Only when potentially relevant photos have been discovered and their metadata retrieved from a remote client do we begin to evaluate the specific spatial relationships between them. These images can be considered as a virtual collection of images; candidates that may have some relation with the pictures in the client’s own photo collection. The client executes the algorithm for relations evaluation between its collection images and the candidate images. Every relationship discovered is added to the RDF model. At the end of this process the client

will hold all the relations between its own pictures and pictures of the remote clients.

The distributed environment and the algorithm for relations evaluation permit the growth of the RDF relations model. This holds the information required for building the browser interface for picture collections. The interface is shown in Figure 8.

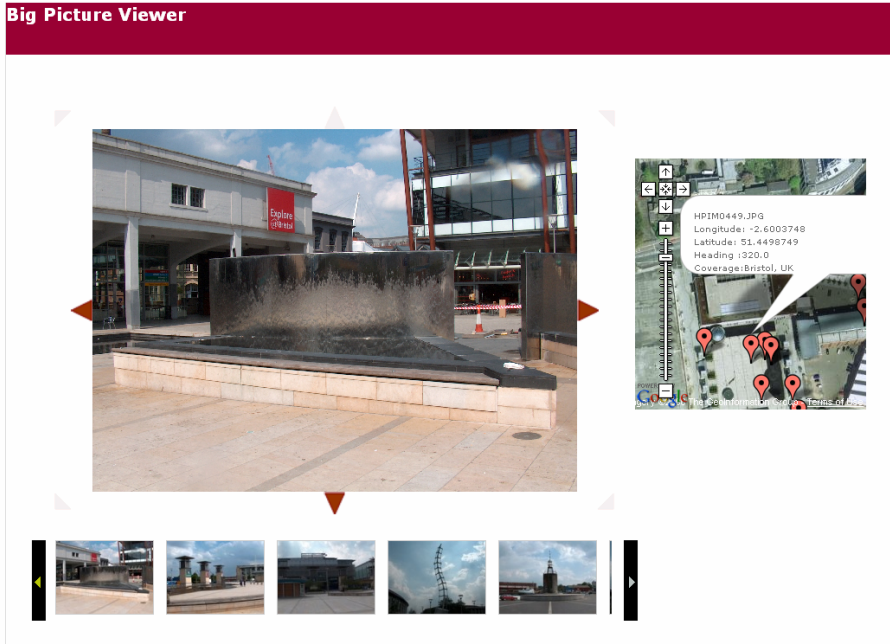


Figure 8. Browsing interface

The pictures described in RDF can be accessed by a thumbnail menu or a Google Maps panel. Moving the mouse over the markers on the map causes the latitude, longitude, heading and coverage information for the corresponding picture to be displayed. The user can browse the pictures by means of the navigation arrows surrounding the pictures that show the direction in which a user can move from the perspective of the current picture. Pictures in the 'field of view' can be reached by clicking on the current picture.

For our experiments we used a set of 100 pictures related to 3 different cities. Latitude, longitude and heading information were collected on a Suunto G9¹² watch at the time the pictures were taken and then later injected in the EXIF data for each picture. The RDF collection files were created by a batch program reading the EXIF information directly from the pictures. The test environment was composed of 4 clients. Each client was implemented using a Joseki¹³ SPARQL server running as a web application under Apache Tomcat. The browsing interface was developed as a web application using Jena¹⁴ and Velocity¹⁵.

4 Discussion: Alternative Representations, Additional Metadata, Scalable Architecture

In our approach we used the semantic web recommendation Resource Description Framework (RDF) to describe photo collections and metadata related to the pictures they contain.

Among other metadata formats (EXIF or XML for instance) RDF was chosen because we want to deal with metadata decoupled from the actual resources in order to be able to store, process and expose the information about pictures (among them the location as the URI of the resource) independently of storing the actual photograph. Moreover, we want to be able to define and extend relations between metadata and have the possibility to take advantage of RDF inference capabilities that are not available in XML.

In addition RDF offers the following advantages:

- RDF is expressly designed to provide a standard, extensible format for machine readable metadata. RDF is an open standard, allowing widespread deployment and consumption. Using RDF means that metadata can be shared and reused more easily.
- RDF is 'syntax neutral'; different RDF vocabularies all share the same syntax. This allows us to easily mix different vocabularies, and load any vocabulary into any tool.
- Ontologies for image metadata are already available in RDF format.

The following ontologies are examples of those that can be used in order to define pictures metadata:

- W3C (W3C 2002) suggests three simple schemata - Dublin Core (for title and description), a technical schema (for camera type, lens) and a content schema (oft-used tags like Baby, Architecture and so on).
- Time can be dealt with as a Dublin Core tag or by treating events as first class entities (W3C – Cal 2002)
- Space can be described using precise geographical descriptors, like latitude and longitude and for which there are already¹⁶ (and see Sect. 3) ontologies available. To represent hierarchical relations such as "England contains London" we could use formal approaches like the space namespace ontology¹⁷. A more ambitious, though incomplete, schema based on ISA standards has also been proposed¹⁸. Differing degrees of accuracy can be catered for by taking a 'layered' approach¹⁹ ('within 10m', 'within 100m', 'within 10km'...). An alternative approach is to consult a controlled vocabulary with concrete place names.
- Device metadata is often provided within a photo in EXIF format, for which the RDF version exists. Other terms such as focal length relevant to cameras are represented in Morten Frederickson's Photography Vocabulary²⁰ and in Roger Costello's Camera ontology²¹.

- Topic tags can be mapped to Flickr tags as the URI for a Flickr tag is simply its URL. The RDF property used to connect a photograph to a Flickr tag would, however, need to be a custom property. The tag hierarchy can be represented within RDF using `rdfs:subClassOf` or `skos:broader`²².

Our ontology reuses some of these existing ontologies for EXIF and Basic Geo (WGS84 lat/long) metadata. Heading information and camera related data (focal length, focal plane resolution and so on) are expressed using an RDF version of the EXIF standard. Dublin Core describes author, title, date, time and annotation about location. We have introduced our own vocabulary for defining field of view and spatial relations as described in Sect. 3.

Our approach for hierarchically structured locations uses the *dc:coverage* property and the values it may contain. This approach is very lightweight compared to relations defined more formally but has the following advantages:

- simple expression of the 'Place or area, City, Country' order
- tag-like format that users can easily create
- more accessible than a series of properties values

The advantages of letting users define their own vocabulary for classifying information has already been demonstrated by the growth of tagging community, while the effectiveness of folksonomies in information classification and retrieval is becoming more and more relevant. One could extend our approach using constraints on tag-like format of property values, or indeed link photographs using controlled vocabularies. Other metadata can be added to the proposed picture description. In particular, it would be interested to add social metadata related to pictures so that social relations, other than spatial, can be discovered and presented to the users providing a *social exploration* of shared picture collections.

Our prototype has been a useful proof of concept but is not yet suitable for real deployment. A P2P architecture would provide an optimization of query caching and routing between the different clients at the expense of complexity in the client implementation. However, a centralized server, which would act as the repository of the pictures' metadata and evaluate the spatial relationships between users' pictures with batch processes, allows the development of a simple web based service without the need of a client-side application. This is a lighter-weight solution for users who wouldn't have to download and install a full software application.

Compared to other approaches and applications, our system has the benefit of standard metadata descriptions that can easily be shared and reused in many different applications and services. The browser application built on top of these descriptions is an example of what can be done using our approach. RDF provides flexibility in how spatial information is encoded, processed and computed. One can imagine for example a browser based on social networks or an algorithm combining latitude, longitude, coverage and geographic thesauri for more accurate spatial labeling. The lightweight approach proposed for computing picture relations, and indeed the choice to rely purely on metadata rather than on information gathered from

heavyweight image processing, makes our solution suitable for real time and web based applications.

5 Conclusions

In this paper we have explored ways to create, share and use geotagged pictures available on the web. As an example of application using geotagged pictures we have implemented a prototype system providing ways to:

- share geotagged pictures
- discover pictures through geotag metadata
- present geotagged pictures and their spatial relationships

An algorithm for inferring spatial relations between different pictures using location and compass heading information embedded in the RDF description of the pictures has been presented. A testing environment for metadata sharing and picture discovery has been implemented so that users' photo collections are enhanced by relations with other users' pictures. We have shown how, based on geographical metadata expressed in RDF, it is possible to build a service for discovering, linking and browsing geographical related photos in a new way. Our future work will deal with experiments on large test beds in order to obtain meaningful performance evaluation, improve scalability, and improve the user interface.

References

- Cayzer, S. and Butrler, M. (2004). "Semantic Photos", *Hewlett Packard Labs Tech. Rep.* . <http://www.hpl.hp.com/techreports/2004/HPL-2004-234.html> {04-10-2006}.
- Campbell, N., Muller H. and Randell, C. (1999) "Combining Positional Information with Visual Media", *The Third International Symposium on Wearable Computers*. Ed. The IEEE Computer Society. 203-205.
- DCMI Usage Board. (2006). *Dublin Core Metadata Initiative*, DCMI. <http://dublincore.org/documents/dcmi-terms/> {01-10-2006}.
- Kadobayashi, R. and Tanaka, K. (2005). "3d viewpoint-based photo search and information browsing", *SIGIR '05: Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*. New York, NY, US: ACM Press. 621-622.
- Naaman, M., Paepcke, A. and Garcia-Molina, H. (2003). "From Where to What: Metadata Sharing for Digital Photographs with Geographic Coordinates.", *Proceedings of the 10th International Conference on Cooperative Information Systems*.
- McCurdy, N. J. and Griswold, W. G. (2005). "A Systems Architecture for Ubiquitous Video.", *MobiSys '05: Proceedings of the 3rd international conference on Mobile systems, applications, and services*. New York, NY, US: ACM Press. 1-14.
- Rodden, K. and Wood, K. (2003). "How do People Manage Their Digital Photographs?", *Proceedings of the SIGCHI 2003 conference on Human factors in computing systems*. New York, NY, US: ACM Press. 24-26.

- Snaveley, N., Seitz, S. M., Szeliski, R. (2006). "Photo tourism: Exploring photo collections in 3D", *ACM Transactions on Graphics (SIGGRAPH Proceedings)*. New York, NY, US: ACM Press 835-846.
- Toyama, K., Logan, R. and Roseway, A. (2003). "Geographic location tags on digital images.", *Proceedings of the eleventh ACM international conference on Multimedia*. New York, NY, US: ACM Press. 156-166.
- World Wide Web Consortium. (2003). *Exif vocabulary workspace - rdf schema*, W3C. <http://www.w3.org/2003/12/exif/> {01-10-2006}.
- W3C Semantic Web Interest Group. (2003). *Basic Geo (WGS84 lat/long) Vocabulary*, W3C. <http://www.w3.org/2003/01/geo/> {01-10-2006}.
- World Wide Web Consortium. (2002). *Describing and retrieving photos using rdf and http*, W3C. <http://www.w3.org/TR/photo-rdf/> {01-10-2006}.
- World Wide Web Consortium. (2002). *RDF Calendar Workspace*, W3C. <http://www.w3.org/2002/12/cal/> {01-10-2006}.
- World Wide Web Consortium. (2006). *SPARQL Protocol And RDF Query Language*, W3C. <http://www.w3.org/TR/rdf-sparql-query/> {04-10-2006}.
- World Wide Web Consortium. (2002). *Resource Description Framework*, W3C. <http://www.w3.org/RDF/> {04-10-2006}.

¹ *Flickr*. <http://www.flickr.com/>.

² *Zoomr*. <http://zoomr.com/>.

³ *Picasa*. <http://picasaweb.google.com/>.

⁴ *Google Earth*. <http://earth.google.com/>.

⁵ *Flickr Map*. <http://www.flickr.com/map/>.

⁶ *Zoto*. <http://www.zoto.com/>.

⁷ *Jpgearth*. <http://www.jpgearth.com/>.

⁸ *Loc.alize.us*. <http://loc.alize.us/>.

⁹ *Greasemonkey*. <http://greasemonkey.mozdev.org/>.

¹⁰ *Sharing Places*. <http://www.sharing-places.com/>.

¹¹ *Mindswap Ontology*. http://www.mindswap.org/_glapizco/technical.owl

¹² *Suunto*. <http://www.suunto.com>

¹³ *Joseki*. <http://www.joseki.org/>.

¹⁴ *Jena*. <http://jena.sourceforge.net/>.

¹⁵ *Jakarta Velocity*. <http://jakarta.apache.org/velocity/>.

¹⁶ *GeoOntologies*. <http://www.mindswap.org/2004/geo/geoOntologies.shtml>.

¹⁷ *Spatial Ontologies*. <http://space.frot.org/ontology.html>.

¹⁸ *Geographic Ontologies*. <http://loki.cae.drexel.edu/~wbs/ontology/iso-19115.htm>.

¹⁹ *GeoOnion*. <http://esw.w3.org/topic/GeoOnion/>.

²⁰ *Photography Vocabulary*. <http://www.wasab.dk/morten/2003/11/photo>.

²¹ *Camera OWL Ontology*. <http://www.xfront.com/camera/camera.owl>.

²² *SKOS Core Vocabulary Specification*. <http://www.w3.org/TR/swbp-skos-core-spec/>.