Sharing, Discovering and Browsing Photo Collections through RDF geo-metadata

Carlo Torniai
Multimedia Integration and Communication Center
University of Florence
Firenze, Italy
Email: torniai@micc.unifi.it

Steve Battle
and Steve Cayzer
Hewlett Packard Labs
Bristol, United Kingdom
Email: steve.battle@hp.com, steve.cayzer@hp.com

***Abstract*— In recent years the growth in popularity of digital photography, together with the development of services and technologies to annotate and organize data on the Web, have extended the possibilities for managing and sharing large numbers of pictures. Our work explores the kinds of metadata that can be captured at the time a photo is taken, and ways to share these metadata in order to create a browsing experience of distributed photo collections based on their spatial information and relations. We present a prototype system in which an RDF description of pictures, including location and compass heading information, is used to discover geo-related pictures from other users. A browsing interface that allows users to explore pictures according to the spatial relationships discovered is proposed.**

I. Introduction

With the growing popularity in digital photography, there is now a vast resource of publicly available photos on the web. 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 Flickr[1] 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. 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.

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. Our work explores the kinds of metadata that can be captured at the time a photo is taken, and ways to link photos together according to these metadata. The objective of our work is to create an experience where someone can view a photo on the web, then jump, for instance, 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 build RDF descriptions of 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. A web application then uses these descriptions to provide a browsable interface. This interface allows users to explore shared photo collections through their spatial relationships with each other.

The paper is organized as follows: the process of choosing metadata and building RDF description of a photo collection is discussed in Sect. II. Algorithm for building relations between pictures is described in Sect. III. In Sect. IV the architecture of the distributed environment and the process of image discovery is presented. Sect. V discusses possible metadata and architecture enhancements while in Sect. VI previous works based on geotagged images are presented. Finally, in Sect. VII we provide conclusions and some future works.

II. Photo collections and metadata

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

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

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

We have used both RDF translation of the EXIF [1] standard and Basic Geo (WGS84 lat/long) vocabulary [2] for latitude and longitude. 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 [3] 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 the *dc:coverage* tags, 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. 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/jpg" ;
  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 ontology[2]. 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>
```

III. Discovering picture relations

The RDF picture descriptions are used to determine the spatial relations between pictures. We have chosen to define a light-weight computation algorithm that provides 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 Fig. 1. 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 $image_b$.

The algorithm for field of view evaluation is shown in Alg. 1.

The $FOV\_THRESHOLD$ 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.

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 we consider the pictures to be taken in the same location so that their relative spatial position is given by the difference between their heading information. Referring to Fig. 2 we say that one can turn right from A to B.

[2]http://www.mindswap.org/∼glapizco/technical.owl

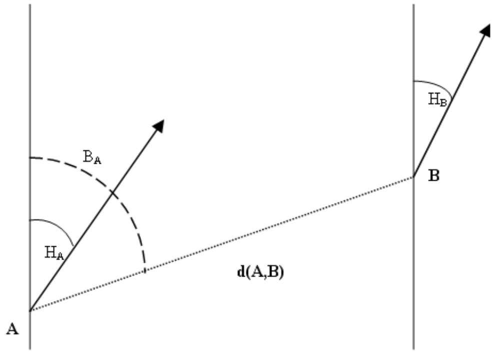

Fig. 1. Field of view evaluation. If $|H_A - B_A|$ is less than a given threshold point B is in the field of view of point A. If $|H_A - H_B|$ is less than a given threshold then the pictures taken in A and B have similar heading direction. If both these conditions are met then $image_b$, taken at B is in field of view of $image_a$ taken at A.

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 $image_a$
      AND $(|H_A - H_B| < T_{head})$
      // ie $image_b$ *and* $image_a$ *have similar headings*
      then set $fov\_relation(image_a, image_b)$

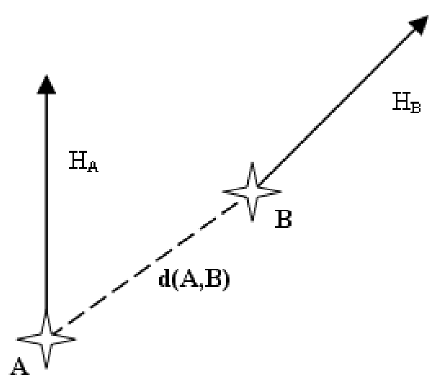

Fig. 2. Spatial relations evaluation. If $d(A, B)$ is less than a given threshold then the spatial relation is given by $(H_A - H_B)$.

The algorithm for spatial relations discovering is shown in Alg. 2. $DISTANCE\_THRESHOLD$ is typically set to 15 meters taking into account the GPS accuracy.

The output of the algorithm is an RDF model describing the relations discovered between the pictures.

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 $spatial\_relation(position, image_a, image_b)$

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). An example of an RDF relations model is shown in listing 3.

Listing 3. Example of RDF relations file

```
<rdf:Description rdf:about="http://bigpicture/pictures/
     HPIM1375.JPG">
 <bigpicture:has_in_fov rdf:resource="http://bigpicture/
      pictures/HPIM1351.JPG"/>
 <exif:gpsImgDirection>223.00</j.0:gpsImgDirection>
 <dc:title>Watershed from Peto Bridge</dc:title>
 <exif:gpsLongitudeRef>W</exif:gpsLongitudeRef>
 <exif:gpsLatitudeRef>N</exif:gpsLatitudeRef>
 <geo:long>-2.5976999</geo:long>
 <geo:lat>51.4496125</geo:lat>
</rdf:Description>
  ...
<rdf:Description rdf:about="http://bigpicture/pictures/
     HPIM1351.JPG">
 <bigpicture:Back_Right rdf:resource="http://bigpicture/
      pictures/HPIM1350.JPG"/>
 <exif:gpsImgDirection>210.00</j.0:gpsImgDirection>
      <dc:title>A red Boat</dc:title>
      <exif:gpsLongitudeRef>W</exif:gpsLongitudeRef>
      <exif:gpsLatitudeRef>N</exif:gpsLatitudeRef>
      <geo:long>-2.5976999</geo:long>
      <geo:lat>51.4496125</geo:lat>
</rdf:Description>
  ...
```

IV. Distributed environment

A distributed test environment has been implemented in order to evaluate the picture discovery process and the algorithm for relations evaluation across different photo collections.

The distributed environment is composed of a set of “clients” (Fig. 3). Each client exposes its photo collection(s) (i.e. RDF metadata) to its peers by means of SPARQL [4] 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 Alg. 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 Picture discovery algorithm

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

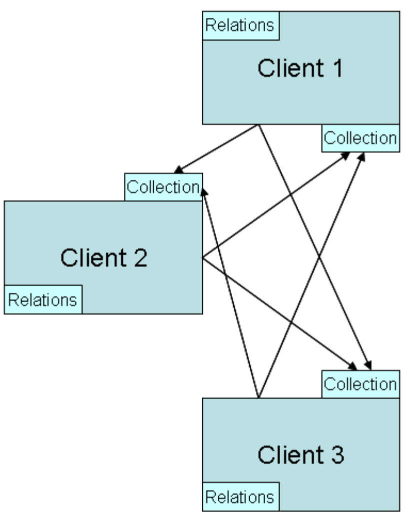

Fig. 3. Distributed Environment

The first step is the expansion of hierarchical *dc:coverage* tags (Sect. II) 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 as in its own collection. This is performed by means of SPARQL queries against (similarly expanded) *dc:coverage* tags. 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 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 their 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 Fig. 4.

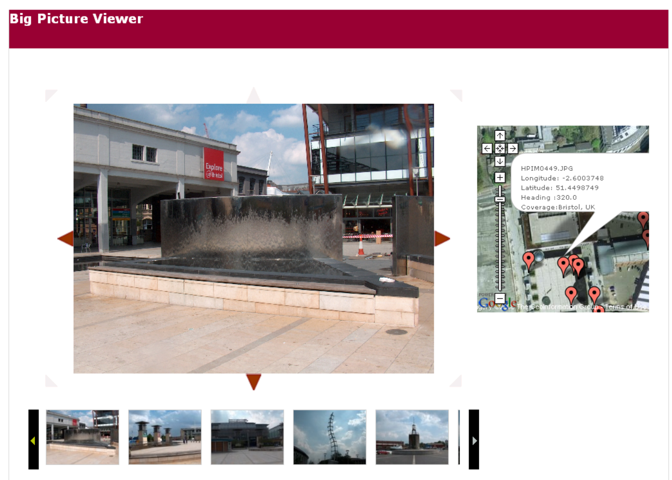

Fig. 4. Picture browsers interface

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. The pictures related by means of the field of view relations 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 G9[3] watch at the time the pictures were taken and then later injected in the EXIF data for

[3]http://www.suunto.com

each picture. The RDF collection files were created by a batch program reading the EXIF information directly from the pictures. The test distributed environment was composed by 4 clients. Each client was implemented using Joseki[4] SPARQL server running as a web application under Apache Tomcat. The browsing interface was developed as a web application using Jena[5] and Velocity[6].

V. Discussion: alternative representation, additional metadata, scalable architecture

In our approach we used RDF as format to describe photo collections and metadata related to the pictures they contain. This 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 the pictures can be shared and reused more easily.
- RDF is “syntax neutral”; different RDF vocabularies share the same syntax. This allows us to mix different vocabularies, and load any vocabulary into any tool.
- Many ontologies related to pictures 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 [5] 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 [6].
- Space can be described using precise geographical descriptors, like latitude and longitude[7]. To represent hierarchical relations such as “England contains London” we could use formal approaches like the ‘space’ ontology[8]. A more ambitious, though incomplete, schema based on ISA standards has also been proposed[9]. Differing degrees of accuracy can be catered for by taking a ‘layered’approach[10](‘within 10m’, ‘within 100m’, ‘within 10km’). An alternative approach is to consult a controlled vocabulary with concrete place names.
- Device metadata is provided within a photo in EXIF format, for which an RDF version exists. Other terms relevant to cameras such as focal length are represented in Morten Frederickson’s Photography Vocabulary[11] and in Roger Costello’s Camera ontology[12].
- 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[13].

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

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 may 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

[4]http://www.joseki.org/
[5]http://jena.sourceforge.net/
[6]http://jakarta.apache.org/velocity/
[7]http://www.mindswap.org/2004/geo/geoOntologies.shtml (accessed October 2006)
[8]http://space.frot.org/ontology.html (accessed October 2006)
[9]http://loki.cae.drexel.edu/ wbs/ontology/iso-19115.htm (accessed October 2006)
[10]http://esw.w3.org/topic/GeoOnion (accessed October 2006)
[11]http://www.wasab.dk/morten/2003/11/photo (accessed October 2006)
[12]http://www.xfront.com/camera/camera.owl (accessed October 2006)
[13]http://www.w3.org/TR/swbp-skos-core-spec/ (accessed October 2006)

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.

VI. Related Work

There has been much interest recently in using geo-location information to relate different picture and to create an enhanced photo browsing experience.

In *Sharing Places*[14], multimedia annotation (photo, video and audio) can be associated with physical locations to create a 'mediascape'. These trails, based on GPS information and enriched with annotation, 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. This approach relies on a central server and doesn't provide annotation in a standard sharable metadatada format.

Images are arranged according to their location in the *World-Wide Media Exchange* [7] while time and location are used to cluster images in *PhotoCompas* [8]. *Realityflythrough* [9] 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 [10].

In *Photo Tourism* [11] 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 overhead maps is provided by the photo explorer interface. In contrast to our system (based on a distributed environment in which metadata related to photo collections are exchanged in real time between users in order to discover relationships between pictures) a complex computer-vision based algorithm is used to provide spatial relationship between images.

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 [12] 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 can lead to interesting evolutions in the way in which pictures and other multimedia geotagged content are shared, discovered and browsed.

[14] http://www.sharing-places.com

VII. Conclusion

In our work we have presented a prototype system providing ways to:

- represent geographical metadata related to pictures
- discover pictures relations according to the metadata
- present the geotagged pictures and their relations

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 relations discovering 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 photo in a novel 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

[1] W. W. W. Consortium, "Exif vocabulary workspace - rdf schema," W3C, http://www.w3.org/2003/12/exif/, Tech. Rep., accessed 1 October 2006.

[2] ——, "Basic geo (wgs84 lat/long) vocabulary," W3C, http://www.w3.org/2003/01/geo/, Tech. Rep., accessed 4 October 2006.

[3] D. U. Board, "Dcmi metadata terms," Dublin Core Metadata Initiative, http://dublincore.org/documents/dcmi-terms/, Tech. Rep., accessed 1 October 2006.

[4] W. W. W. Consortium, "Sparql protocol and rdf query language," W3C, http://www.w3.org/TR/rdf-sparql-query/, Tech. Rep., accessed 4 October 2006.

[5] ——, "Describing and retrieving photos using rdf and http," W3C, http://www.w3.org/TR/photo-rdf/, Tech. Rep., accessed 1 October 2006.

[6] ——, "Rdf calendar workspace," W3C, http://www.w3.org/2002/12/cal/, Tech. Rep., accessed 1 October 2006.

[7] R. Toyama, K. Logan and A. Roseway, "Geographic location tags on digital images," in *MULTIMEDIA '03: Proceedings of the eleventh ACM international conference on Multimedia*. New York, NY, USA: ACM Press, 2003, pp. 156–166.

[8] A. Naaman, M. Paepcke and H. Garcia-Molina, "Metadata sharing for digital photographs with geographic coordinates," in *Proccedings of 10th International Conference on Cooperative Information Systems*, 2003.

[9] N. J. McCurdy and W. G. Griswold, "A systems architecture for ubiquitous video," in *MobiSys '05: Proceedings of the 3rd international conference on Mobile systems, applications, and services*. New York, NY, USA: ACM Press, 2005, pp. 1–14.

[10] R. Kadobayashi and K. Tanaka, "3d viewpoint-based photo search and information browsing," in *SIGIR '05: Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*. New York, NY, USA: ACM Press, 2005, pp. 621–622.

[11] N. Snavely, S. M. Seitz, and R. Szeliski, "Photo tourism: exploring photo collections in 3d," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 835–846, 2006.

[12] S. Cayzer and M. H. Butler, "Semantic photos," Hewlett Packard Labs, http://www.hpl.hp.com/techreports/2004/HPL-2004-234.html, Tech. Rep., accessed 4 October 2006.