Creating Visual Summaries for Geographic Regions

Adrian Popescu
Institut
Télécom/TELECOM
Bretagne
Technopôle Brest-Iroise
29238 Plouzané
+3314654
adrian.popescu@telecom-bretagne.eu

Ioannis Kanellos
Institut
Télécom/TELECOM
Bretagne
Technopôle Brest-Iroise
29238 Plouzané
+3314654
ioannis.kanellos@telecom-bretagne.eu

ABSTRACT
Queries for place names represent a hefty chunk of Web image searches. Generic image retrieval solutions do not use any geographic information to improve the results for queries with place names, while systems like *Flickr Places* [3] or *Panoramio* [11] exploit raw georeferencing information to illustrate places. In this paper, we present a method for generating informative and diversified visual summaries (photographic representations) of places by using a geographical gazetteer. The gazetteer is built from user-contributed data and its content reflects a community-based perception of places. We evaluate our image retrieval method against *Flickr* and *Flickr Places* and show that the obtained results are significantly improved compared to these two baseline applications.

Categories and Subject Descriptors
H.3.3 **Information Search and Retrieval**

General Terms
Algorithms, Experimentation, Human Factors.

Keywords
Place names, image retrieval, *Flickr*, *Flickr Places*, gazetteer, georeferenced images.

1. INTRODUCTION
Uploading photos of places is a popular activity on sharing platforms, such as *Flickr* or *Panoramio* [11]. As shown in [15], [7] or [12], these images and their associated metadata (tags, geo-tags, temporal information) can constitute the basis for extracting a shared view of place semantics. The authors of [19] discuss the utility of Web content in general for the acquisition of a vernacular gazetteer. However, user-contributed data are often noisy and efficient filtering methods are required for extracting knowledge. We review some information filtering and structuring methods that lead to the automatic mining of a large-scale geographical gazetteer [13] from user-contributed sources (*Flickr*, *Panoramio*, Wikipedia). We also present a reliable way to add a spatial inclusion relation to the gazetteer elements using information downloaded via the *Flickr* API.

Geographic image search is an application that received considerable attention both from commercial search engines (Google Maps, Flickr Places) and from the research community ([1], [13], [8]). Existing dedicated applications (*Flickr Places* [3], *Panoramio*) exploit raw georeferencing metadata associated to photos in order to propose visual summaries of places. An interesting multimodal image clustering technique aimed at obtaining a diversified representation of landmarks is presented in [8]. While effective for landmarks, a visual clustering is unlikely to work equally well if applied to larger geographic regions (cities, countries) which have very complex and diversified visual representations. We argue that the use of semantic relations encoded in a geographical gazetteer improves the quality of visual summaries for complex geographic objects because it enables the automatic selection of a diversified range of representative encompassed place names. For instance, a visual summary of *Paris* will be composed of images of landmarks situated in this city (*Eiffel Tower, Notre Dame de Paris, Louvre, Arc de Triomphe* etc.). Our approach is applicable both to textual queries with place names and to geographic queries expressed in map based interfaces like Google Maps (which consist in the selection of a bounding box representing a region of the world).

The remainder of this paper is structured as follows: Section 2 reviews related work; Section 3 presents a methodology for an automatic construction of a geographical gazetteer; Section 4 describes the role of a geographical gazetteer in georeferenced image retrieval; before concluding, Section 5 presents an user-centered evaluation of our approach.

2. RELATED WORK
The constitution of large-scale geographical databases received sustained attention but remains a hot research topic. Initial efforts [5] addressed the proposition of domain models for geography and the transformation of preexisting gazetteers into digital libraries. These long and tedious efforts led to the creation of gazetteers like Alexandria [5] or Geonames [4], which contain several millions of place names and their characteristics. The minimal description of a location in a geographical gazetteer includes: its name (*Louvre*), its spatial position (*48°51'39.79N, 2°20'08.99''E*) and its parent category (*museum*) [5]. All these dimensions and other information are described in databases like Geonames. Despite the fact that spatial inclusion is a fundamental relation in the geographical domain [9], it is rather poorly defined in existing gazetteers making it impossible, for instance, to infer that the *Louvre* is a museum in *Paris*. Another problem with manually constituted geographical databases built from preexisting sources is the variability in coverage of different regions of the world [4]. Finally, the lack of a ranking score associated to place names constitutes an important drawback from an information retrieval perspective. For instance, it is impossible

to answer a question like "Which are the most representative/salient geographical objects in *France*?"

Recent works on geographical databases try to automate the extraction of spatial semantics from user-contributed data. Rattenbury et al. [15] made one of the first attempts to extract place names from large scale unstructured multimedia data. They implemented a multiscale burst analysis to separate geographical locations from others tags found in georeferenced *Flickr* photos' descriptions. Their analysis is fully automatic, purely statistical and the authors reported an extraction precision of 85% (with 50% recall). The obtained database contains names, ranking scores and coordinates for the discovered entities, but it includes neither categorization nor spatial inclusion information. In [12], we proposed a method for exploiting heterogeneous data sources (*Wikipedia*, *Panoramio* and a Web search engine) in order to automatically constitute a geographical gazetteer (*Gazetiki*). We used a domain vocabulary and linguistic rules to extract, categorize and rank geographical names and we averaged photo coordinates in order to situate discovered entities. The evaluation, performed on areas around 15 cities showed that we increase both the precision of the extraction (by 5%) and the coverage (3 times) of the results compared to [15]. The structure of *Gazetiki* is richer compared to the geographical database in [15] but it still does not contain any spatial inclusion information.

Spatial inclusion relations can be derived from spatial footprints of place names. Although broadly studied (see [17] for a review), the automatic extraction of spatial footprints for places remains a difficult research topic. Schockaert and De Cock [17] discuss the vagueness of boundaries, concluding that fuzzy representations of spatial footprints seem more adapted than geometric shapes in most cases. The authors address the description of the spatial extent of city neighborhoods and build their model relying on Yahoo! Local API (a service that provides lists of businesses for given areas), and on NLP techniques applied to the text describing a business. This method exploits spatial inclusion relations served by the Yahoo! Local API. Following a similar approach, we mine spatial inclusion relations from metadata associated to *Flickr* georeferenced images. Given the geodesic coordinates of gazetteer elements, we look at nearby photos and extract the encompassing entities from Flickr metadata.

Recent works exploit geographical information in image retrieval. *Flickr Places* [3] is described as a "page on *Flickr* for every place in the world". The application takes advantage of georeferencing information associated to *Flickr* photos to propose representative images. The images are selected on the basis of *Flickr*'s "interestingness" metric of a photo, which includes the number of tags, of comments, of views and the number of times the image is marked as a favorite. Except for spatial proximity between the place name coordinates and the photo coordinates, no other geographical information seems to be exploited in *Flickr Places*.

Ahern et al. [1] implemented *World Explorer*, a map-based visualization tool for georeferenced *Flickr* images. They exploit the results in [15] and associate a relevance value to each discovered entity. Tags are overlayed on a scalable map using their coordinates and the relevance ranking. By selecting a displayed tag, we get associated georeferenced images for a visual exploration. With the introduction of clickable tags, World Explorer exploits both geographical coordinates and textual descriptions of images to propose results whereas in *Flickr Places* only coordinates were used. Nevertheless, World Explorer ignores the type of place name a tag represents and the pictures that represent tags are simply selected from the *Flickr* georeferenced corpus.

The tag displaying system is also open to linkage with other external geographical databases. We combined the map interface and a content based image retrieval (CBIR) software in ThemExplorer [13], a georeferenced image search engine exploiting *Gazetiki* [12]. ThemExplorer is more interactive than World Explorer because it allows the user to express preferences with the selection of interesting geographical concepts; it also offers the possibility to explore visually similar images of an object. However, in ThemExplorer we focused on the presentation of precise geographical concepts (i.e. *Notre Dame de Paris, Tower Bridge*) and thus it was not possible to query for more general concepts like *city* or *country* names. Log file studies showed that these generic entities are queried more often than specific objects in textual search [16] or in image retrieval [14].

Kennedy and Naaman [8] discuss the generation of diversified results for landmarks using image clustering techniques. While their approach is highly effective for differentiating representations of precise objects like the *Golden Gate Bridge* or the *Eiffel Tower*, it is doubtful whether it can scale to more general place names (*cities, regions* or *countries*). These place names types are probably too complex from a visual point of view to discriminate their sub-concepts using only visual processing.

Research in environmental psychology is relevant for the creation of visual summaries of places. Milgram [10] conducted a study with Parisians in order to study the mental representation of their city. One of the findings he reports is that landmarks are key components of the representation inhabitants have of their environment. Tversky [18] analyzes spatial knowledge and observes that "the information about the spatial world that we encode and remember should be general and varied enough to serve purposes, both known and not yet known". A second interesting result of the same study concerns the need to summarize information about spatial entities and the central role of landmarks in these summaries. These observations are pertinent for geographical image retrieval systems and should be taken into account in order to build applications that mirror the way users imagine places. Hierarchical and spatial relations in gazetteers, combined with efficient landmarks ranking, can constitute the basis for creating such applications.

3. GAZETTEER CONSTRUCTION

We dealt with the automatic construction of a geographical gazetteer using heterogeneous data sources in [12]. Each discovered entity was characterized by: a name, a set of geographical coordinates, a parent category and a popularity rank. No spatial inclusions were extracted in this work. We briefly describe how place names are extracted from unstructured or semi-structured documents, how these names are categorized, spatially situated and ranked. We then explain how we enrich our gazetteer with spatial inclusion information extracted from *Flickr*.

3.1 Initial Version of Gazetiki

In [12], we built a vocabulary of geographical concepts (*museum*, *church*, *bridge*, *tower* etc.) and used it to extract geographical names from *Panoramio* photo descriptions and *Wikipedia*. We briefly review the methods we developed in the cited paper in order to familiarize the reader with our approach. A detailed description of the gazetteer construction process and of the implementation choices is provided in [12].

We first filter a list of all georeferenced articles from *Wikipedia* and consider their titles to be geographic names. Place names often contain explicit geographic concepts and, based on this fact, we mined *Panoramio* for such multiword place names. Instance names are usually written in capital letters and we wanted to retain only capitalized multiwords that contain geographic concepts. In order to obtain complete place names, low-case stop words were not taken into account when appearing between two other capitalized terms. We extract place names such as *Museum of Modern Art, Saint-Eustache Church* or *Union Square*.

To associate coordinates to geographic names extracted from *Panoramio*, we search for the first photo described by a candidate name and then average the latitudes and longitudes of all photos described by the same name which are found within a predefined radius. In order to deal with place names polysemy, whenever photos tagged with the same place name are found outside the radius, we include a new element in the gazetteer. The coordinates of Wikipedia articles are extracted using the structure of the article and patterns that introduce coordinates.

The first sentences of georeferenced Wikipedia articles are often definitions and we use them to extract a geographical category for the entity described in the article. We categorize candidate geographic names extracted from *Panoramio* by employing the text of the snippets obtained from a search engine, a procedure similar to approaches used in question answering [2]. We launch a query with the candidate name, extract 50 snippets and compare the geographical vocabulary to the snippets in order to extract elements that are most often associated with the candidate. We then launch definitional queries ("*candidate* IS A *concept*") and retain as final category of the candidate the concept appearing the most often in definitional queries.

Gazetteer elements are ranked using a popularity measure which is computed over the Panoramio corpus. We first limit the search space to a radius around the element's coordinates and then combine the total number of images (a classical term frequency measure) with the number of different users (a community-based relevance assessment) having uploaded those images. We assume that, if a geographical object was photographed by several users, it is more representative than an object photographed by only one person; it should thus be ranked higher.

3.2 Mining Inclusion Relations from Flickr

In *Flickr*, each georeferenced image has associated spatial inclusion relations which enable an automatic enrichment of metadata with localization information. For instance, a georeferenced image taken at the *Louvre* will be tagged *Paris 1er arrondissement, Paris, Île de France, France* (encompassing entities of the *Louvre*). The spatial inclusion data are also available via the *Flickr* API and we associate them to *Gazetiki* element in the following way: given the geodesic coordinates of an place name, we launch a query for neighboring images and infer that the encompassing entities of the closest photo are equally those of the analyzed place name if the photo is closer than a threshold. We set the threshold at one km in order to avoid inclusion errors. Given the high volume of georeferenced images in *Flickr* (over 100 million items as of March 2009), it is possible to mine spatial inclusion information for a large panel of regions of the world. For the 56 cities in the evaluation set the *Flickr* API spatial inclusion information is correct in all cases.

After this step, the elements of *Gazetiki* are characterized by the vector (name; geodesic coordinates; parent concept; encompassing entities); this information can be exploited for image retrieval. For instance, the entry for Louvre will be (Louvre; (*48°51'39.79N, 2°20'08.99''E*); *museum*; (*Paris 1er arrondissement, Paris, Île de France, France*)).

4. PLACE NAMES IMAGE RETRIEVAL

Gazetiki can be used to propose a structured image search for queries with place names. Here we focus on place names having a significant surface, such as *neighborhoods, cities, regions* or *countries* which are often queried in Web image search [16]. Note however that the method is equally applicable to map based interfaces, where the query is a bounding box representing a region of the world. The only difference is that, instead of being expressed using place names, the spatial inclusion is expressed using geodesic coordinates. Given a query with a place name, we select relevant entities in the gazetteer using the following rules:

- Specific entities must be included in the spatial footprint of the place name. This spatial inclusion relation in the gazetteer is used in order to select entities lying inside a place name.
- Specific entities must belong to punctual classes like *museums, churches, palaces, monuments* etc. We exclude non-punctual entities such as *parks* or *rivers* because they are hard to represent using a single image. This step of the algorithm exploits the hierarchical organization of the gazetteer to select only specific entities that belong to punctual classes.
- Specific entities must be among the most popular associated to the queried place name. The popularity ranking of Gazetiki elements is used to return a ranked list of elements.
- Specific entities must range in different categories. A page of results includes a limited number of pictures and we aim at proposing a conceptually diversified representation of place names. We limited the number of elements belonging to the same category at 3.

The above conditions ensure that a place name is represented by a series of specific entities that are spatially associated to the entity, belong to different geographical categories and are popular. We illustrate the results of our approach with the first images representing *Paris* (figure 1a) and compare the gazetteer based technique to the set of images obtained with *Flickr* (figure 1b). The gazetteer based description of a place name is a kind of community leveraged view of the respective place and it ensures their diversified visual representation.

The content of the gazetteer can also be used to enrich the interactivity of the application. In [14], we showed that a categorical restriction over the content of the gazetteer is generally well perceived by the users. By default, an image retrieval system is unaware of user preferences and cannot adapt the results; after seeing the first results, the user can select categories of interest (for instance *museums, bridges* and *arches*). Another interesting option concerns the capacity to focus the results on interesting specific elements. Once a user selects a specific element, like the *Eiffel Tower*, the system can present only images of this element. Another option for rapidly browsing interesting images is to introduce a CBIR among the photos of the same specific object similarly to [13].

(a)

(b)

Figure 1. Visual summaries of Paris using Gazetiki (a) and *Flickr* (b).

The use of the gazetteer also solves ambiguous queries. In *Flickr Places*, when querying for *Paris* the system asks the user to choose between the different localities having this name in the database. In *Gazetiki*, the elements are ranked; the system proposes the most popular denotation of an ambiguous name first, indicating that there are other places with a identical name.

5. EVALUATION

We evaluate our visual summaries generation technique against two existing applications that support geographical queries (a generic search engine and a dedicated application). *Flickr* is an image search engine that answers a query by leveraging textual image descriptions while *Flickr Places* uses geo-tagging and implicit user feedback (via the "interestingness" metric introduced in section 2). The images presented using our technique are also downloaded from *Flickr*, ensuring thus that global quality of the images is similar for the three methods. Places are complex visual objects and we considered that an evaluation of sets of images is more appropriate than an evaluation of individual images.

5.1 Experimental setting

We designed a user satisfaction test[1] in which participants were asked to compare the three approaches driven on sets of pictures representing 56 well known cities around the world (examples include *Athens*, *New York*, *Paris*, *Sydney* or *Tokyo*). For each city and each method we only considered the first 10 images returned on the same day; we assumed that these answers correspond to a visual representation of the cities. Each page of answers contained three columns of image results corresponding to the three assessed methods. Since a priori knowledge is needed to evaluate the quality of the results, we asked the participants to select only those cities they knew well enough. We also instructed participants to skip selected cities if they thought they were unable to assess them.

The participants assessed the global quality of the answers, for each city and for each method, using a four-level scale (0 bad representativeness and 3 for very good representativeness). In order to ensure unbiased comparisons, the users were not informed about the tested retrieval techniques. Also, to avoid the formation of preference patterns, the columns were presented in different orders on the results pages and the users were warned about this random presentation. The participants were contacted via e-mail and the participation was voluntary. In order to preserve the participants' interest during the test, we informed them that they can quit at any time. We received 33 sets of answers, totalizing 218 votes. 28 sessions out of 33 contain at least 3 assessed items and this distribution reflects the users' knowledge of the cities in the proposed panel, as well as their interest during the test. There are 3 sessions including over 15 votes which reflect both extensive knowledge of the domain and a high interest in our test. The participants found the test engaging and interesting as proven by the fact that some of them voluntarily contacted us to discuss results and methods and to express their willingness to participate in future evaluations.

5.2Results – users view

We compared the gazetteer based image retrieval method to *Flickr* and *Flickr Places*. We report the results in figure 2. For readability reasons, we plotted only sessions including at least 5 votes (20 sessions out of 33) as well as the average scores for the three methods over the entire panel of users. The averaged results in figure 2 indicate that our method outperforms *Flickr* and *Flickr Places* (2.16 vs. 1.79 vs. 1.36). The differences between the three methods are significant (0.37 between our method and *Flickr* and 0.8 between our method and *Flickr Places*, respectively). On average, the results obtained with our technique are situated between a good and a very good representation of place name. *Flickr Places* and *Flickr* are situated between a poor representation and a good representation (closer to a poor representation for *Flickr Places* and closer to a good representation for *Flickr*).

At an individual level, the gazetteer based method is judged best by 21 participants out of 33, *Flickr* is considered as the best system by 5 participants and *Flickr Places* by only 1 participant. For the entire panel of users the gazetteer based image retrieval is considered at least as good as the two baseline systems by 26 participants out of 33. When analyzing only sessions that include

[1] Online at http://comupedia.org/city_explorer/ (in French)

at least 5 votes, our method outperforms the other two in 13 cases out of 20 and ties with *Flickr* as the best system in two cases.

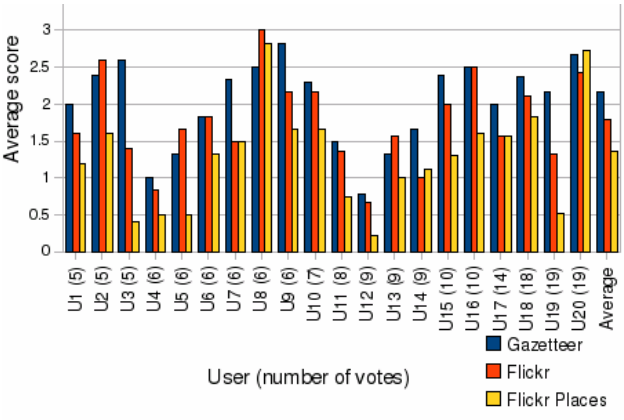

Figure 2. Comparison results between the 3 image retrieval techniques for place names – User-based view. The maximum score (3) corresponds to high quality visual summaries of cities.

The results in figure 2 show that there are important differences between the average scores attributed by different participants, accounting for the inherent subjectivity of the task. The most strict and demanding participant, U12, ranks the three techniques between 0.22 (*Flickr Places*) and 0.78 (gazetteer) while the less exigent, U8, ranks between 2.5 (gazetteer) and 3 (*Flickr*). The current panel of participants is not large enough to establish statistically significant correlations between participants' characteristics (gender, age, occupation) and preferable image retrieval method. It would be certainly interesting to analyze such correlations in larger scale studies.

Our first conclusion is that, even if our technique is not unanimously considered better than *Flickr* and *Flickr Places*, a large majority of the participants prefer it to these baseline systems.

5.3 Results – cities view

It is also interesting to analyze results for individual cities. 46 of the 56 proposed cities received at least one vote. In figure 3, we present only the 20 cities that were voted at least 4 times; this limitation seems necessary because a small number of votes does not ensure a good separation of the three tested techniques. The gazetteer based image retrieval method outperforms *Flickr* and Flick *Places* for 14 cities out of 20, *Flickr* is best in 5 cases and *Flickr Places* in 1 (figure 3). Major differences in favor of our method are obtained for cities like *Paris*, *Brussels*, *New York*, *Milan* or *Strasbourg*. They indicate that these cities are well represented using salient artifacts like *churches*, *towers* or *museums*. *Amsterdam* is the only city for which *Flickr* significantly outperforms the gazetteer based method. This situation is explained by the fact that *Amsterdam* canals are highly representative of the city and they do not appear among the types of landmarks we considered in our selection algorithm. The results for individual cities reflect well those obtained for each user. They furthermore confirm that our method outperforms the two baseline systems.

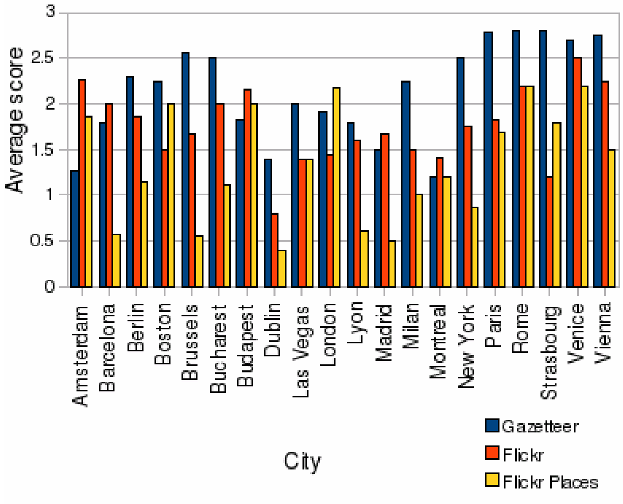

Figure 3. Comparison results between the 3 image retrieval techniques for place names – City-based view. The maximum score (3) corresponds to high quality visual summaries of cities.

5.4 Discussion

The use of a geographical gazetteer in geographical image retrieval improves the perceived quality of results when compared to existing retrieval techniques. When compared both to a generic retrieval framework (*Flickr*) and to one that exploits georeferencing information (*Flickr Places*), a large majority of participants preferred our method. The presented experimental results confirm that landmarks constitute an important part of the visual representation of place names. They confirm our hypothesis concerning the utility of a conceptually structured representation of place names. The participants were asked to evaluate only results for city names but the landmark selection technique presented in this paper is generic and applies also to more generic place names (*countries*, *regions* etc.) or to more specific ones (*neighborhoods parks,*, *museums*). Further tests are necessary in order to assess its applicability at different scales. We centered the discussion on administrative divisions because the majority of geographical queries correspond to such names; since gazetteers also contain information concerning other geographical entities (*islands*, *natural parks* etc.), it is easy to adapt the method to one offering a structured representation of these entities.

When comparing only *Flickr* and *Flickr Places*, the preference of the participants for the first system is somewhat surprising because *Flickr Places* exploits georeferencing information, ensuring that the presented photos were taken at the desired location. The poor perception of *Flickr Places* may be explained by the fact that this system uses only raw georeferencing information (geodesic coordinates), neglecting textual annotations of the images. The annotations are only indirectly considered via the interestingness measure used by *Flickr Places* while they play a central role in *Flickr*. It seems that, combining georeferencing and textual annotations associated to images increases the perceived quality of results.

Ensuring results diversity in order to respond to different user needs is a crucial topic in image retrieval evaluation [6]. Image clustering [8] is one way to obtain diversity but, as we already observed, it only applies to visually coherent concepts, like landmarks. Obtaining diversified results is implicit when hierarchical and spatial relations between geographical concepts are exploited. *Flickr* and *Flickr Places* results often contain several images depicting a single aspect of a place name; consequently, these applications do not ensure sufficient diversity. Our experiment supports the hypothesis that presenting a diversified view of a concept has a positive influence on the results assessment in image search.

The participants liked the experimental setting. Although we did not solicit free text feedback, some of them sent us their impressions concerning the test protocol. Their reaction is important for two reasons. First, the experimental setting where people are asked to perform a global evaluation of methods (on sets of answers) is more engaging for participants compared to the assessment of individual photos. Second, some user comments are helpful for improving the proposed image retrieval technique. Our method gives preferential treatment to the presentation of landmark images and neglects “atmosphere” pictures (images of streets or panoramic views). Several users indicated that they consider “atmosphere” images to be highly representative and they sometimes preferred results containing such pictures. In order not to overload the answers pages, we limited the number of presented images to 10 per method but, in real retrieval situations, we can easily present up to 20 images on an answer page. To ensure an even higher diversity of results, we could complement the use of the gazetteer and add some “atmosphere” images among the retrieved images.

6. CONCLUSION AND FUTURE WORK

We discussed a way of improving the generation of visual summaries for geographic region by exploiting the content of a geographical gazetteer. The results reported in this paper are promising and stimulating since they improve the quality of visual summaries compared to a generic image retrieval system (*Flickr*) and to a dedicated application (*Flickr Places*).

We currently assess the introduction of image processing techniques in the retrieval framework. Following [8], we can perform an image clustering for each specific element representing a place name in order to present only the most relevant images to the user. We will also explore ways to select “atmosphere” images in order to further diversify the image sets representing place names. A third line of work regards the extension of our model to non-punctual entities which, as we have seen for *canals* in *Amsterdam*, can be sometimes highly relevant. Finally, we will test the applicability of the approach to more specific entities (*neighborhoods, museums, churches*) and more generic entities (*regions, countries*). In the first case, image clustering can be a suitable solution to propose good quality visual summaries. In the last case, it is probably necessary to complement the conceptual diversity with a spatial diversity of the landmarks.

7. ACKNOWLEDGMENTS

This research is part of Georama, a French research project funded by the ANR.

8. REFERENCES

[1] Ahern, S., Naaman, M., Nair, R. and Yang, J. 2007. World Explorer: Visualizing Aggregate Data from Unstructured Text in Georeferenced Collections. In *Proc. of* JCDL 2007 (Vancouver, Canada, June 2007).

[2] Brill, E., Lin, J., Banko, M., Dumais, S. and Ng, A. Data-intensive question answering. 2001. In *Proc. of the TREC-10 Conference* (Gaithersburg, Maryland, USA, Nov. 2001).

[3] *Flickr Places* - http://flickr.com/places

[4] Geonames - http://geonames.org

[5] Hill, L. L., Frew, J. and Zheng, Q. 1999. Geographical names – the implementation of a gazetteer in a georeferenced digital library. *CNRI D-Lib Magazine* (January, 1999).

[6] ImageCLEF Photo Retrieval Task http://www.imageclef.org/2008/photo

[7] Kennedy, L., Naaman, M., Ahern, S., Rahul, N., Rattenbury, T. How flickr helps us make sense of the world: context and content in community-contributed media collections. In Proc. of ACM Multimedia 2007 (Augsburg, Germany, Sep. 2007).

[8] Kennedy, L. and Naaman M. Generating diverse and representative image search results for landmarks. In Proc. of WWW 2008 (Beijing, China, April 2008).

[9] Larsen, R., Frontiera, P. Spatial Ranking Methods for Geographical Information Retrieval (GIR) in Digital Libraries. In Proc. of *EDCL 2004* (Bath, UK, Sep. 2004).

[10] Milgram, S. Environmental Psychology: People and Their Physical Settings, 2nd ed. Holt, Rinehart and Winston, New York, USA, 1976, ch. Psychological Maps of Paris,104–124.

[11] *Panoramio* – http://panoramio.com.

[12] Popescu, A., Grefenstette, G., Moëllic, P.-A. *Gazetiki*: automatic construction of a geographical gazetteer. In Proc. of JCDL 2008 (Pittsburgh, PA, June 2008).

[13] Popescu, A., Moëllic P.-A., Kanellos, I. ThemExplorer: Finding and Browsing Georeferenced Images. In Proc. of CBMI 2008 (London, UK, June 2008).

[14] Popescu, A. Structures linguistiques pour la recherche d'images sur Internet. PhD, Télécom Bretagne, 2008.

[15] Rattenbury, T., Good, N., Naaman, M. 2007. Towards Automatic Extraction of Event and Place Semantics from *Flickr* Tags. *In Proc. of SIGIR 2007* (Amsterdam, The Netherlands, July 2007).

[16] Sanderson, M., Han, Y. Search Words and Geography. In Proc. of the GIR 2007 Workshop. (Lisbon, Portugal, 2007)

[17] Schockaert, S., De Cock, M. Neighborhood Restriction in Geographical IR. In Proc. of *SIGIR 2007* (Amsterdam, The Netherlands, July 2007).

[18] Tversky, B. Levels and structure of spatial knowledge. In Kitching, R., Freundschuh, S. (eds.) Cognitive Mapping. Past, present and future. Routledge (2000).

[19] Twarock, F., Jones, C., Abdelmoty, A. Acquisition of a vernacular gazetteer from Web sources. In Proc. of LocWeb Workshop – WWW 2008. (Beijing, China, 2008).