USING SR-TREE IN A CONTENT-BASED AND LOCATION-BASED IMAGE RETRIEVAL SYSTEM

Hien Phuong Lai[1,2,3], Nhu Van Nguyen[1,2,3], Alain Boucher[2,3] and Jean-Marc Ogier[1]
[1] L3I, Université de La Rochelle, 17042 La Rochelle cedex 1, France
[2] Institut de la Francophonie pour l'Informatique, MSI, UMI 209, Hanoi, Vietnam
[3] IRD, UMI 209, UMMISCO, IRD France Nord, Bondy, F-93143, France

Keywords: Content-based image retrieval, Location-based information, Multidimensional indexing, SR-tree.

Abstract: This paper presents an approach for combining content-based and location-based information in an image retrieval system. With the performance for nearest neighbour queries in the area of multidimensional data and for spatial data structuring, the SR-tree (Katayama and Satoh, 1997) structure is chosen for structuring the images simultaneously in location space and visual content space. The proposed approach also uses the SR-tree structure to organize various geographic objects of a Geographic Information System (GIS). We apply then this approach to a decision-aid system in a situation of post-natural disaster in which images describe different disasters and geographic objects are monuments registered in GIS data in the form of polygons. The proposed system aims at finding emergencies in the city after a natural disaster and giving them an emergency level. Some scenarios showing the interest of using content-based and location-based search in different ways are also presented and tested in the developed system.

1 INTRODUCTION

This work aims to develop an information retrieval model based simultaneously on visual content and geographic location information of images. To our knowledge, there are only few works combining these two information types. The SnapToTell system (Chevallet et al., 2007) tries to first use geographic information for reducing the number of images that have to be examined. Secondly, it uses content information for finding the most similar image in the reduced image database. The MobiLog system (Cemerlang et al., 2006) sends the image with location information to the SnapToTell server to get the information about the scene visited by user; this information will be then suggested to add in the user's blog. In our system, each image is represented by two descriptors: a visual content descriptor and a geographic location descriptor of image. Another type of information useful for some application is geographic objects. Geographic objects can be a house, a building, a street, etc. They are often registered in GIS system in the form of point, multipoint, polyline or polygon.

The proposed approach uses the SR-tree structure (Katayama and Satoh, 1997) which is based simultaneously on the structure of the R-tree (Guttman, 1984) and the SS-tree (White and Jain, 1996) for structuring simultaneously the visual content and the geographic location of images and for organizing the geographic objects. By incorporating bounding spheres and bounding rectangles, we can reduce the overlapping regions comparing with the case of R-tree and SS-tree. This enhances the performance for nearest neighbours queries in the area of multidimensional data and for indexing spatial data objects that have non-zero size (Guttman, 1984) of SS-tree and R-tree.

For implementing our approach, we develop a decision-aid system in a situation of post-natural disasters which is in the context of the IDEA project[1]. The main idea is to exploit the images collected after a natural disaster by using a camera network located throughout a city (surveillance cameras, cameras mounted on patrolling robots, cameras installed in buildings, aircrafts, inhabitant mobile phones, etc.) to help local decision makers in organizing rescue.

[1] Images of natural Disasters from robot Exploration in Urban Area (IDEA), http://www.ifi.auf.org/IDEA/

Phuong Lai H., Van Nguyen N., Boucher A. and Ogier J. (2010).
USING SR-TREE IN A CONTENT-BASED AND LOCATION-BASED IMAGE RETRIEVAL SYSTEM.
In *Proceedings of the International Conference on Computer Vision Theory and Applications*, pages 491-494
DOI: 10.5220/0002831504910494

2 CONTENT-BASED AND LOCATION-BASED IMAGE RETRIEVAL USING SR-TREES

We state the hypothesis that our image database is partly annotated by image class. A small amount of images have been manually annotated, while the remaining is unknown. This is consistent with our application (IDEA project), where initial images describe the knowledge of interest for the application, while more images will be coming in real-time afterward. From this, we choose to organize the visual content of images into several SR-trees corresponding to different image class. The advantage of this approach is to use the visual content nearest neighbours search with annotated images for each new incoming non-annotated image for determining its class.

Our approach represents independently data in both visual content and geographic spaces as follows: (1) A SR-tree is built to represent feature vectors of all annotated images. (2) All images are represented into two spaces (total of 1+n SR-trees, where n is the number of image classes) by a SR-tree organizing geographic descriptor (a point (x,y) representing longitude and latitude) and many SR-trees organizing feature vectors (each SR-tree corresponding to one image class). (3) All geographic objects are represented into a single SR-tree; the region of each leaf is the intersection of the bounding rectangle and the bounding sphere covering all the points of an object.

We can realize some simple manipulations using these SR-trees: (1) Nearest neighbours search can be done for each non-annotated image using the SR-tree of annotated images to determine the input image type. (2) We can find in the SR-tree of geographic objects all objects that are geographically close within a given radius from any image. (3) Using the SR-tree of geographic location of images, we can find all images that are geographically close to a geographic object or to another image. (4) For finding images which are most similar to an input image, the nearest neighbours search is used in the SR-tree corresponding to the type of the input image. More complex manipulations can be realized based on these simple ones according to different scenarios.

3 DECISION-AID SYSTEM IN A SITUATION OF POST-NATURAL DISASTERS

This section describes a decision-aid system in a situation of post-natural disasters for the IDEA project implementing the approach proposed in section 2. Each image in this system represents a disaster occurring in the city. It contains geographic information of GPS type (latitude and longitude). We characterize the visual content information using only the RGB color histobins (vectors of 48 dimensions) of images (more visual descriptors are planned to be added). Geographic objects of interest are various monuments of the four following types: houses, buildings, hospitals, schools. This system aims to identify emergencies in the city and to assign each of them an emergency level according to the proximity between similar situations and to different monuments around each disaster. For example, a fire occurring inside a hospital is more urgent than three consecutive residential houses on fire.

Experimental Data Sources. We build our own database using the "Earthquake Image Archives"[2] and several images representing emergency situations retrieved from the website Flickr[3]. All images are of five different types of disasters (fire, wounded people, damaged building, damaged road and flood), each type contains between 300 and 350 image. Two thirds of these images are annotated by its type. Regarding GIS data, we simulate them using the Colorados GIS database[4] and we add to this database information identifying four different monument types (house, hospital, building and school).

In this system, we have six visual content SR-trees (one for annotated-images and five for non-annotated images corresponding to the above five disaster types), one SR-tree for geographic location of non-annotated images and one SR-tree for geographic objects. Note that in the GIS database, each monument is represented by a polygon identifying its position and its form. Each leaf of the geographic objects SR-tree represents a monument. Therefore, if the geographic position of a disaster falls into the region of a leaf identifying a hospital for example, we can say that this disaster occurs inside the hospital.

Determination of Emergency Levels. Disaster type identification of an image is based on the number of different disaster types resulting of the k nearest neighbours search applied into the visual content SR-tree containing the annotated images. The problem is to assign an emergency level for each disaster. It is difficult to use visual content for this task, as we

[2] http://geot.civil.metro-u.ac.jp/archives/eq/index.html, Tokyo Metropolitan University.

[3] http://www.flickr.com/, All images used in this work are under the Creative Commons licence.

[4] 2007 TIGER/Line Shapefiles, http://www.census.gov/geo/www/tiger/tgrshp2007/tgrshp2007.html

have no information on the zooming factor or camera distance for each image acquisition. We assume that the emergency level is assigned only basing on geographic information, which are the proximity between the similar disasters and the type of monuments that are around each disaster. Other criteria like disaster nature, influence of different disasters, etc, are not considered in this paper. We determine $\mathbf{r}_1$, the radius denoting the geographic proximity between situations, and $\mathbf{r}_2$, the radius determining the proximity between a situation and a monument. After a natural disaster, there may be a lot of images that are sent to the server. It is useful to group similar situations according to their proximity in geographic space and show to the user general information about these groups. Two similar situations A and B will be grouped together if there is a series of similar situations between A and B that each is close to another.

For determining the emergency level, we propose the following approach. For each found disaster, we assign an emergency level according to the disaster type and to the monuments that are around this disaster. Suppose that a disaster occurs at monument M_j and we found a set M_1,...,M_k of monuments which are near in a radius r_2 from the disaster, emergency level L of the disaster is then computed as follows:

$$L = \mu + \alpha_j + \sum_{i=1}^{k} \beta_i \qquad (1)$$

where: (1) μ is the emergency level corresponding to each disaster type. In this paper, we use $\mu = 1$ for all different disaster types. For giving more realistic values, we are planning to consult rescue experts in the future. (2) α_j is an emergency factor added knowing that a disaster occurs at the monument M_j. (3) β_i is an emergency factor added knowing that a disaster occurs at a position near to the monument M_i.The value of α and β depend on the type of monument. Using the SR-tree of external descriptor of all current images, we can group the disasters in different proximity groups as defined above by finding in the geographic SR-tree all images which are close together in a radius r_1 and then assign to each group an emergency level that is equal to the sum of the emergency levels of all disasters belonging to this group.

4 RESULTS

We describe in this section some scenarios which combine content-based search and location-based search in different orders and test them in our decision-aid system in a situation of post-natural disaster. The quantitative evaluation is planned to be done in a further collaboration with rescue experts.

Scenario 1: CBIR → Location-based Search. Figure 1 presents an example of this scenario type which aims to find all monuments which can be affected (geographically close) by one of the disasters which are similar to an input disaster (query image, eg. a fire). After determining the type of disaster (image class) using the nearest neighbour search in the SR-tree containing the annotated images, the system performs a CBIR search in the SR-tree corresponding to the type of the query image to find similar images and then performs a location-based search for each retrieved image from the previous step in the SR-tree of monuments.

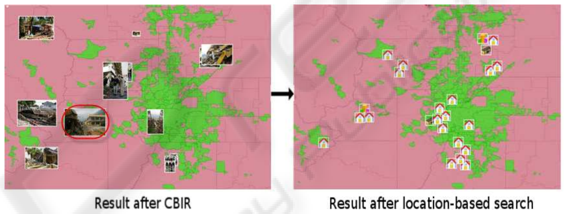

Figure 1: Scenario 1: CBIR → location-based search.

Scenario 2: location-based search → CBIR. This scenario is opposite to the previous scenario in using first location-based search in geographic location SR-tree for finding all images which are geographically close to a query image or an input geographic object and secondly using CBIR for finding all images which are similar to each retrieved image from the location-based search. Figure 2 shows an example of results using this scenario.

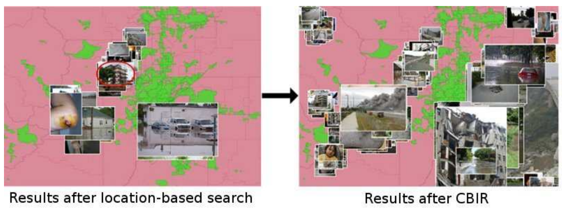

Figure 2: Scenario 2: location-based search → CBIR.

Scenario 3: CBIR and Location-based Search Simultaneously. We can perform CBIR and location-based search together and select within both set of retrieved results. A scenario using together CBIR and location-based search is to determine the emergency level of an image. CBIR is used in the SR-tree of annotated images for determining the disaster type of the input image and the location-based search is used both in the geographic location SR-tree

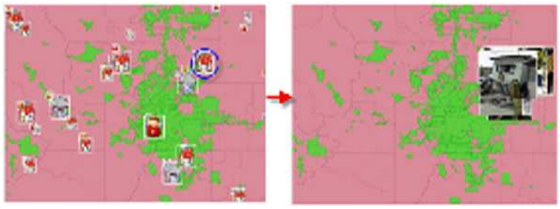

Figure 3: Scenario 3 - The system gives an overview of the distribution of disasters and also their emergency level corresponding to the symbol size. User can choose to view all disaster images within a proximity group.

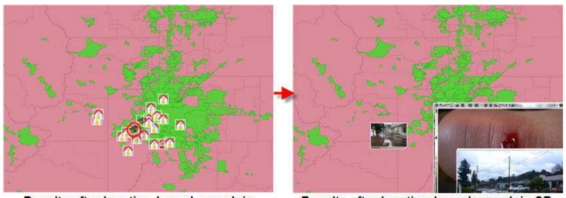

Figure 4: Scenario 4 - This example uses two successive location-based searches first in the SR-tree of geographic monuments and secondly in geographic location SR-tree of images.

of images and in the geographic object SR-tree for finding disasters and monuments of interest near the input query image. All this information is used for determining the emergency level corresponding with the input query image using the method described in section 3. The system gives an overview of all disasters in the city so that user can observe the distribution of disasters and also their emergency level corresponding to the symbol size in order to make rescue decisions quickly (see Figure 3).

Scenario 4: Many Location-based Searches. We can perform location-based search both in the geographic location SR-tree of images and in the SR-tree of geographic objects in different orders for retrieving different information. Figure 4 presents results for finding all disasters that can affect one of monuments which are around another monument.

5 CONCLUSIONS

The presented approach, using the SR-tree structure for representing images into two different spaces (visual content and location information) and for representing geographic objects, allows merging content-based image retrieval and location-based search in different ways according to requirements of different applications. Different scenarios are presented here for an application of decision-aid for rescue management, but it can be applied to different applications. More specifically, by applying our approach to the decision-aid system in a situation of post-natural disaster, we provided for the user an overview of the disasters in an urban zone (position, emergency level of each disaster). Thus, the user can coordinate appropriately rescue teams. Moreover, using multiple SR-trees for representing information in different ways, it allows the system to manipulate very different types of information (visual content and location-based) and to provide the appropriate result by searching only in the right SR-trees.

But there are some limitations that will need to be improved as further work. Concerning the genericity of the system, location-based information could be found within the image and not necessarily given externally. Text detection and recognition could provide addresses or location names within the images. Recognition of known buildings or monuments from the images could also give clues about the location of the images. Concerning the specific application for natural disasters, we are planning to integrate experts and interactive experiments for determining all parameters in our system.

ACKNOWLEDGEMENTS

This project is supported in part by the ITC-Asia IDEA project from the French Ministry of Foreign Affairs (MAE), the DRI INRIA and the DRI CNRS. The authors also thank the Region Poitou-Charentes (France) for its support in this research.

REFERENCES

Cemerlang, P., Lim, J. H., You, Y., Zhang, J., and Chevallet, J. P. (2006). Towards automatic mobile blogging. In *Proceeding of IEEE ICME 2006*, pages 2033–3036.

Chevallet, J. P., Lim, J. H., and Leong, M. K. (2007). Object identification and retrieval from efficient image matching. snap2tell with the stoic dataset. *Inf. Process. Manage.*, 43(2):515–530.

Guttman, A. (1984). R-trees: A dynamic index structure for spatial searching. In *Proceeding of the 1984 ACM SIGMOD, ICMD*, pages 47–57.

Katayama, N. and Satoh, S. (1997). The sr-tree: An index structure for high-dimensional nearest neighbor queries. In *Proceedings of the 1997 ACM SIGMOD, ICMD*, pages 369–380.

White, D. A. and Jain, R. (1996). Similarity indexing with the ss-tree. In *Proceeding of the 12th ICDE*, pages 516–532.