International Journal of Engineering Inventions
e-ISSN: 2278-7461, p-ISSN: 2319-6491
Volume 2, Issue 7 (May 2013) PP: 27-33

Binocular Search Engine Using Android

[1]Santhosh Kumar S., [2]Sudarshan K.
[1] *Department of Information Science and Engineering, Srinivas Institute of Technology,Mangalore, Karnataka, India*
[2] *Department of Computer Science and Engineering, Srinivas Institute of Technology,Mangalore, Karnataka, India*

Abstract: *This project proposes and implements a Binocular Search System to search buildings and places through their photos captured by phone cameras. User need to take a picture of the building he/she wants to know with an Android phone and upload the picture to server. The server returns name and information of the buildings or places. We use Scale Invariant Feature Transform features and the bag of features method to represent and recognize building images. Euclidean distance method estimates the distance between query image and those in database. This project is implemented using android and J2EE and makes use Location based Service. We introduce a system using mobile camera, GPS information and PC server to search and recognize buildings without typing any words. With quick development of mobile techniques, a large number of people already own smart phones. Our system provides an attracting and easy way to know about the world using images captured by mobile cameras. We have achieved good performance and the method is effective.*
Keywords: *GPS, SIFT, GAP, HTTP*

I. INTRODUCTION

Many phones have cameras and GPSs, which provide useful information for users to discover and navigate their environments. The information is usually in the form of image and latitude/longitude. However, in many cases, users may want Meta information such as the names and introduction of the buildings around them. In this project, we propose a system that combines network technologies and image retrieval algorithms to address this problem. A user uploads a building photo, and then our system can return its name and other information about the image.

In our system we make use of the bag-of-words method for image retrieval due to its good performance in many image processing and computer vision tasks. The method consists of four steps: 1) extraction of SIFT features, 2) clustering the features to visual words, 3) generating the frequency vector according to the visual words, 4) image query. SIFT feature shows good behavior in efficiency and precision so that we adopt it in our system.

A key problem in our system is how to estimate the similarity between a query image and those in the database. In our approach, the image is represented by frequency vectors. Thus the problem can be reduced to calculate the distance between frequency vectors. We introduce different methods for distance calculation and compare them in experiments. Our final system makes use of the Euclidean distance method one with the best performance.

The overall objective of this project is to implement a GAP Search System (GPS Aided Photo Search System) to identify buildings through their photos captured by phone cameras. User need to take a picture of the building he/she wants to know with Android phone and upload the picture to our system. The system returns name and introduction on the buildings. We use SIFT features and the bag of features method to represent and recognize building images. Euclidean distance methods are compared to estimate the distance between query image and those in database.

The rest of the paper is organized as follows. Proposed embedding and extraction algorithms are explained in section II. Experimental results are presented in section III. Concluding remarks are given in section IV.

II. PROPOSED ALGORITHM

A. *Existing System*

Google achieves great success by moving its navigation applications, such as Google Map, to mobile phone. As tens of thousands people could afford mobile phones, how to locate in outdoor with the help of computer vision and GPS technology becomes a new hot topic. In, a hybrid image-and-keyword searching system, first, image is used to search through WebPages, and then keywords on these WebPages are identified and submitted to existing text search engine, such as Google. In, a group from Microsoft Research Asia has conducted an experiment on Photo-to-Search system, which makes use of image retrieval methods to locate around the world, gives an image retrieval system based on Content Based Image Retrieval methods.

Different from these approaches, our focus is not just on navigation but also a fresh way to provide information about the world by pictures. Let us consider the situation that a user is in an unfamiliar environment and he/she wants to get information on an unknown object (buildings, sculptures, sceneries etc). The user can take a photo of that object with their phones, and upload it to our system. After that, our system recognizes the object and returns useful information. For experiment purpose, we choose the buildings of Bapuji Campus of davanagere as our start point.

B. *Proposed System*

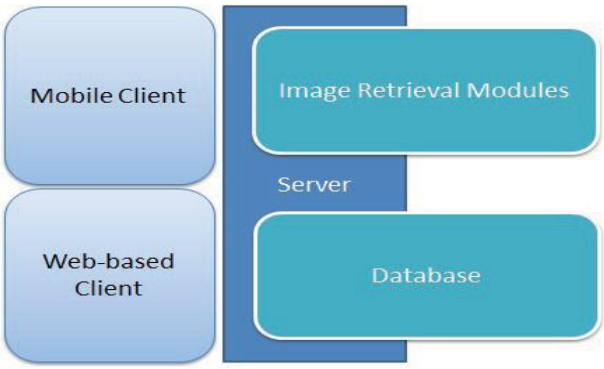

Figure 1. Structure of the system

Our system is composed of three layers:

1. The client
2. The server
3. The image retrieval component.

There are two types of the client: the web-based client and the mobile client which is based on Android mobile system. As the web-based client, users have to manually enter the latitude and the longitude where the picture is taken. System administrators are able to sign into the system and manage all the building information, picture information, and user uploaded photos. As the mobile client, users only need to upload building photos to perform a search. Latitude and longitude are directly retrieved from GPS instruments.

The server is a conjunction of the client and the image retrieval component and has a database of the information and images of the buildings to be retrieved. Given an input image, the image retrieval component finds its nearest image in the database, which indicates the building this image belongs to.

The image retrieval component contains the key algorithm of the whole system. We made use of the bag-of-words method for image retrieval due to its good performance in many image processing and computer vision tasks. The method consists of four steps: 1) extraction of SIFT features, 2) clustering the features to visual words, 3) generating the frequency vector according to the visual words, 4) image query. SIFT feature shows good behavior in efficiency and precision so that we adopt it in our system.

III. SYSTEM DESIGN

A. *Architecture of Binocular Search System-*

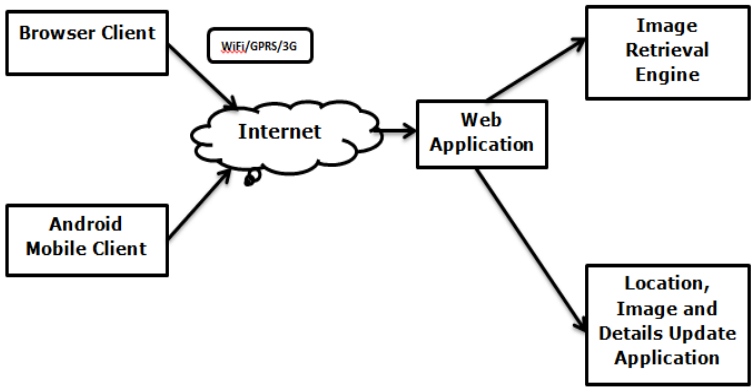

Figure 2. Architecture of Binocular Search System

The whole system is constructed to be flexible and scalable. The architecture of our system is shown in Fig 2. Here Browser client and Android mobile client can communicate to server through Wi-Fi/GPRS/3G.The System consists of two parts

1. The Client.
2. The Server.

The server is implemented using j2ee and the client is implemented using Android.

Client

The client is mainly composed of three modules: Interface, Data handling and Network connection.

Interface

This part provides users with simple and convenient ways to search for buildings and review searching history. To make user experience to the best, User Inter phase design of the client follows the design philosophy and principles of android application, including a clear dashboard, consistent theme and title bar, etc.

Data handling

This part is responsible for data handling and transferring between Views module and Network connection module. Its function includes generating search query, dealing with the response of server (both when searching for a building and posting a new building), and managing the local records of buildings that user searched before.

Network connection

This part serves to make client connect to server, post search query or new buildings and receive results. By using Http Client, the network part sends query as formatted entities to the server, receives the response texted as JSON and passes it to the Data handling module.

The client provides functions as below:

a. Forming a query by taking a photo with camera or selecting a picture file.
b. Uploading the query, including picture from the user and GPS information generated by the client, and show detailed result both as text contents and markers on the Google Map.
c. Generating formatted date of a new building with information provided by user and posting it to the server.
d. Providing history of buildings that user has searched for.

Server

The main function of the server is to manage the information of the available buildings, including the latitude, longitude and the related photos.

Some of the photos are marked as CRITICAL only if the photos are greatly taken and are typical to represent the looking of the building. The GPS information is also very useful in our search system. To speed up the search and increase the accuracy, GPS information is used to filter out the impossible buildings and the related photos even if the GPS information is not so accurate. Within the area that is filtered by the GPS information, the server is able to perform a search in a very delightfully fast speed. The server uses all the photos marked as critical to build an index.

The server also includes the management system with the user interface of the web-based client, system administrators are allowed to add buildings and the related photos and mark the critical photos. But the system administrators do not need to build the clusters and index, it's automatically done if there is some change on the critical photo set.

The image retrieval component is running on the same server, but it's not eventually deployed. As long as there's a chance, the server is able to run the image retrieval component on the different servers or even the supercomputing server.

B. Data flow diagram

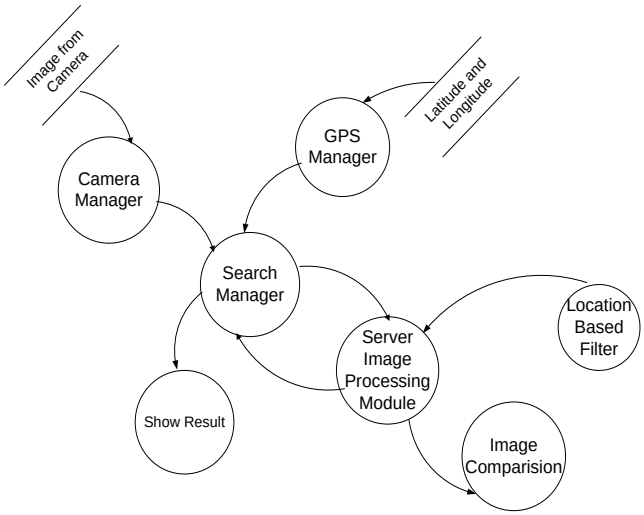

Figure 3: Data flow diagram

Figure 3: Indicates data flow diagram of our system. Client uploads building photos, latitude and longitude to perform a search. The server is a conjunction of the client and the image retrieval component and has a database of the information and images of the buildings to be retrieved. The GPS information is also very useful in our search system. To speed up the search and increase the accuracy, GPS information is used to filter out the impossible buildings and the related photos even if the GPS information is not so accurate. Given an input image, the image retrieval component finds its nearest image in the database, which indicates the building this image belongs to. Using image comparison algorithm fetch matched image in database and send details to client. If data is not available Generate formatted date of a new building with information provided by user by taking photo and its description and posting it to the server

C. Sequence diagram

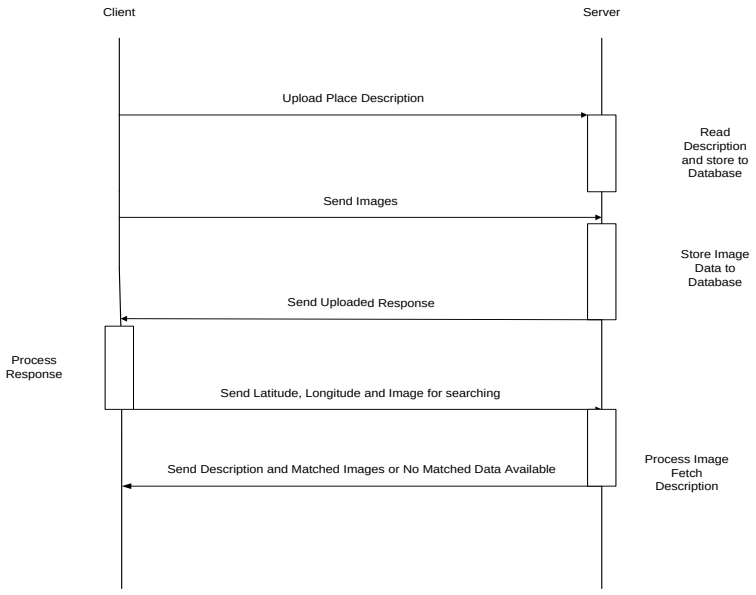

Figure 4. Sequence diagram

Figure 4. represents sequence diagram of the project. The above diagram shows the sequence of message flow between the client and the server. The Server Provides two kinds of features upload details of a monument / building and providing information about the building / movement using image based searching of description. The User takes the image and uploads the image with the latitude and longitude of image location to the server, the server process the image data compares the image for matching regions and filters image with latitude and

longitude and sends the description of matched location to the client if images are matched else send no description available to the client.

IV. EXPERIMENT AND RESULT

There are five modules used in the proposed system of the project. They are

a. Data Upload Module
b. Http Communicator
c. Image Comparison Module
d. Location Manager Module
e. Gallery Viewer

Data Upload Module

The module runs at server which parses the request and fetches the image data from request and stores the image on the server for comparison.

Http Communicator

Http Communicator runs at the client side and allows user to send image data for Comparison, fetches the result and displays on the mobile.

Image Comparison Module

Image comparison module match the images uploaded by the user with the image present in the database by using region based comparison technique and filters the matched image based on GPS Co-Ordinates.

Image Comparison Pseudo Code:

```
Image_Data=Read the uploaded image data
ULatitude=Read Latitude
ULongitude=Read Longitude
Latlng[]=Fetch Latitude and Longitude of all the images in the database
For each (latlng in Latlng[])
Dlatitude=latlng["latitude"]
Dlongitude=latln["longitude"]
Distance=calculateDistance(Ulatitude,Ulongitude,Dlatitude,Dlongitude)
If (Distance<=1)
Fetches images that matches the latitude/longitude
Usignature=Rescale uploaded image and calculate the signature
Foreach (image from matchedimage)
Tsignature=rescale uploaded image and calculate the signature
Difference=calculateDifference (Usignature,Tsignature)
If (Difference<Pre_Defined_Score)
Add image to matched list
Fetch Description about the image
Send the description and matched images to the client
Else
Send "No Data available Response to Client"
```

Here the algorithm fetches the image data and latitude, longitude from the request and stores the uploaded image for comparison, first the algorithm fetches the latitude, longitude of all the images in the database and calculates the distance between the 2 GPS points, if the distance is less than 1km, it fetches the images of that latitude and longitude and passes for image comparison algorithm, The image comparison algorithm first rescales the uploaded image and calculates the signature by getting sum of all the values of RGB values in the signature box and also calculates the signature of all the matched images and calculates the difference between the uploaded image and matched image, if the difference is less than the pre-defined range, application send the matched images and description of the matched images to the client.

Location Manager Module

This Module allows user to interact with the GPS system to fetch latitude and longitude from the GPS System.

Gallery Viewer

This module allows user to take a picture of a building or monument with location information of the captured image and store in the phone database for later upload.

Result Analysis

Test Cases for Server Module

Test case	Description	Outcome	Status of execution Pass/fail
Image Comparison Module	Matches the image using region based matched technique and send the information about the uploaded image to the client	Performs Image comparison using the RGB values of the partitioned region of the image and filter the matched image using location based filter system	Pass
		May do wrong comparison if the matched building image location are within the location of 1 KMS and color matrix of the images in the database are similar	Fail
Data Upload Module	Upload the details of the image and image data to the database for comparison	Parses the request for latitude, longitude and description and fetches the image data from request and upload to the database	Pass

Test Cases of Client Module

Test Cases	Description	Outcome	Status of execution Pass/fail
Http Communicator Function	This Module Makes a HTTP communication with the web server to send and receive data between server and the client.	Makes HTTP Connection to the server and passes the data (image, latitude, longitude) to the server	Pass
Gallery Viewer	Module allows user to take a picture from the camera and fetches latitude, longitude of the captured location and stores in the phone database for later use	Module allows user to take a picture from the camera and fetches latitude, longitude of the captured location and stores in the phone database for later use and allows user browse stored image and allows for image upload	Pass
Camera Manager	Allow user to open the phone camera, take the picture, store the image on the sdcard	Module links the application with the camera of the phone and transfer the data between the client application and camera application	Pass

Snapshots

Search Camera Application

Fig (a)

The above fig (a) shows the snapshot of Search Camera Application. Here there are two options, one is camera and another is gallery of images that have captured through camera.

Open Camera & Gallery viewer

Fig (b)

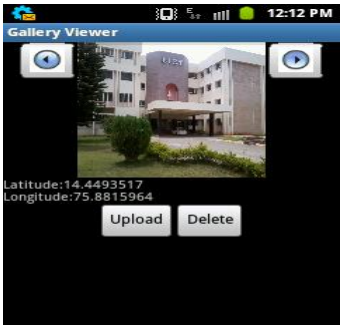

Fig (c)

The above figures fig (b) shows the snapshot of Open Camera. Here we capture images of building, and we can save it for later upload purpose or upload directly and fig (c) shows the snapshot of Gallery Viewer. Here images captured through camera are stored. We can upload image to server.

Final Outcome

Fig (d)

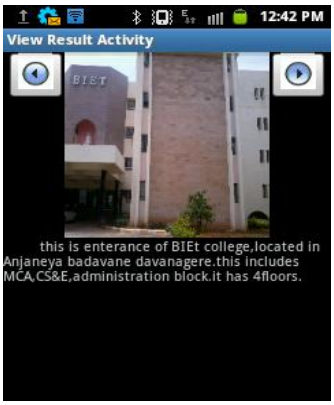

Fig (e)

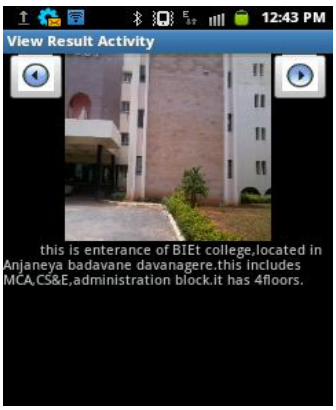

Fig (f)

The above figures shows the snapshot of image and image description after uploading image to server, it returns name and introduction of building and matched images of that particular building.

V. CONCLUSION

In this paper, we introduce a system using mobile camera, GPS information and PC server to search and recognize buildings without typing any words. With quick development of mobile techniques, a large number of people already own smart phones. Our system provides an attracting and easy way to know about the world using images captured by mobile cameras. We have achieved good performance. The method is simple but effective. We resize images 300x300 pixels, extract SIFT feature descriptor to describe each image in database. For a query image, the system calculates the frequency vector just as the images in database, selects candidate images by GPS, estimates the scores for each candidate image, ranks and lists the results.

REFERENCES

[1] Yeh, T., Tollmar, K., and Darrell, T., "Search the Web with Mobile Images for Location Recognition," Computer Vision and Pattern Recognition, vol. 2, pp. 1063-6919, July 2004.
[2] Jianguo Zhang, Marszalek, M., Lazebnik, S., Schmid, C., "Local features and kernels for classification of texture and object categories: A comprehensive study," CVPR, June 2006.
[3] Jégou, H.; Douze, M.; Schmid, C.; Pérez, P., "Aggregating local descriptors into a compact image representation," CVPR, June 2010.
[4] Menglei Jia, Xin Fan, Xing Xie, Mingjing Li, Wei-Ying Ma, "Photo-to-Search: Using Camera Phones to Inquire of the Surrounding World," Mobile Data Management, pp. 46 - 46, May 2006 Ulrich Steinhoff, Dušan Omer_evi_, Roland Perko, Bernt Schiele,
[5] Aleš Leonardis, "How Computer Vision Can Help in Outdoor Positioning," International Joint Conference on Ambient Intelligence, pp.124-141, AmI, 2007
[6] Xu Wangming, Wu Jin, Liu Xinhai, Zhu Lei, Shi Gang, "Application of Image SIFT Features to The Context of CBIR," vol. 4,pp. 552 - 555, Dec. 2008.
[7] Sivic, J., Zisserman, A., "Video Google: A Text Retrieval Approach to Object Matching in Videos," page(s) 1470, International Conference on Computer Vision, pp. 1470 - 1477, Oct. 2003.
[8] G. Csurka, C. Bray, C. Dance, and L. Fan, "Visual Categorization with Bags of Keypoints," European Conference on Computer Vision, pp. 1-22, 2004.