# Combining VGI with Viewshed for Geo-tagging Suggestion

HSIANGHSU LIN February, 2011

SUPERVISORS: Dr. O. Huisman Drs. B.J. Köbben



# Combining VGI with Viewshed for Geo-tagging Suggestion

HSIANGHSU LIN Enschede, The Netherlands, February, 2011

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation. Specialization: Geoinformatics

SUPERVISORS: Dr. O. Huisman Drs. B.J. Köbben

THESIS ASSESSMENT BOARD: Prof. Dr. M.J. Kraak Ir. E. Verbree, TU Delft

#### DISCLAIMER

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

### ABSTRACT

Photo sharing websites have bloomed in recent years. Many of these provide maps to let users upload photos and tag them. To help users in tagging photos, many methods are being developed. After geotagging, the photo has the spatial reference and the annotation and then it becomes the part of Volunteered Geographic Information (VGI). This information can help researchers from different fields, especial for Earth observation. Nowadays, GPS-enabled digital cameras are becoming inexpensive and are widely available. This means the location of the camera is not the problem. Moreover, a digital compass can record the direction in which the photo is taken. These two data are useful in viewshed analysis. The EXIF file can record metadata of camera. Parts of them are useful to help researchers make advanced analysis to identify the objects in user's photo. To help users get information for the major objective in their photos, in this research, we develop a method to improve geo-tagging suggestion. This method uses two major steps: Viewshed Analysis and Clustering Analysis. Through using data in the EXIF file, the essential parameters for Viewshed Analysis are obtained so that the view sight and visible objects through camera can be determined. We use Clustering Analysis to analyze VGI data (in the form of tagged photos on public websites). After Clustering Analysis of the spatial coordinates of VGI data, Hot Spots showing popular objects can be flexibly detected presenting the likely possible objectives in a user's photo.

Key word: Geo-tagging suggestion, EXIF, Viewshed, Clustering, VGI

### ACKNOWLEDGEMENTS

Thanks for ITC course. It lets me broaden my view of GIS. At this moment, I realize the GIS is everywhere in our daily life. And it really can help human have better life. This is my first time to challenge this kind of research. Although I met a lot of problem, liked spending a lot of time to understand another fields, finally, I overcome it. During this period, I have to thank both of my supervisors, Mr. Huisman and Mr. Kobben for their help. They aren't just supervisors, but also like friends. They are full of patience and friendly. Not just giving the order, they like to guide me to think and keep thinking. It let my research look like a game, a big challenge game. Every time when we had meeting, they always let me try to present and organize my time table. It's really helpful to me, because it drives me to go forward and challenge me by myself. If I miss something, they will give me the hint, relative papers and suggestion. Finally, they also spend a lot of time to check my writing. I'm happy that I can work with them and really enjoy this study. They are great, great, great supervisors. Despite my supervisors, my senior, Mr. Deng also gave me time to discuss some ideas and offered me relative information and data. Thanks all of them. And also my good classmate, Mrs. Yang, who offered the iPhone4 to me to have my field test.

The experience of ITC life is very fresh and funny. First time, I work with so many international friends form different countries. It is a good chance to contact with them and try to understand what is the GIS in their mind. Not only for commercial purpose, but GIS can really be used to improve our daily life and also protect our world and safety of human. While I attended the UN course, I felt excited. I realized how UN peace mission undertaking and what the role of GIS in the mission is. In our leisure time, my friends like to invite me to play football!! I never played this before. And we also went to the central square to enjoy our beer with watching the World Cup Game. It was a crazy memory!! I believe I wouldn't forget this period. Thanks, all my friends.

After this study, I become more independent. If the problem exists, now, I know how to organize myself to solve it. I get stronger during this period and own more confidence. Now, there are more and more gifts full of me heart and I will bring of them back to my country. I believe this period will become the sweet memory in my life and I never forget it. Appreciate ITC and people who once had help me. Thank you. Wish you all the best.

## TABLE OF CONTENTS

1.	Introduction	1
	1.1. Motivation and problem	1
	1.2. Research identification	3
	1.2.1 Research purpose	3
	1.2.2. Research sub-objective	
	1.2.3. Research question	3
	1.3. Innovation	3
	1.4. Main Step of my Research	4
	1.5. Outline	6
2.	Literature review	7
	2.1. Geo-tagging system	7
	2.2. Visibility analysis	
	2.3. Density and Clustering Analysis	
	2.4. Summary	
3.	Digital photos and attributes	14
	3.1. Photo	14
	3.2. EXIF	
	3.2.1 Field description	
	3.2.1.1 Date	16
	3.2.1.2 Direction	
	3.2.1.3 Lens Into	
	3.2.1.4 Focal Length	19
	3.2.1.5 CCD Into and Spatial Resolution	
	3.2.2. Discussion	
4	3.3. Summary	
4.	Methodology	23
	4.1. Viewshed Analysis	
	4.1.1. Viewsned Analysis process	
	4.1.2. Implementation.	20
	4.2. Clustering analysis	
	4.2.1. Keniel Delisity method	27 28
	4.2.2. Comparison and discussion	29
	4.3 Example case	30
	4.4 Summary	32
5	Experiment and Discussion	33
5.	5.1 Study Area	33
	5.2. Data Introduction	
	5.3. Experiment and Result	
	5.4. Discussion	
	5.4.1. Benefit of our process	
	5.4.2. Elements for improvement	
	5.4.2.1. Time issue and Personal factors	40
	5.4.2.2 Location and Distance weight value	42
	5.4.2.3 New Defined Boundary	43
	5.4.3. Extended Issues	44
	5.4.3.1. Time cycle	44
	5.4.3.2. Web Service Diagram	45

6.	Conc	lusion	.50
	6.1.	Summary of the research	. 50
	6.2.	Futhur Work	. 52

## LIST OF FIGURES

Figure1.1	My Process and Element	4
Figure1.2	Workflow	5
Figure2.1	Model of STS (Marlow, Naaman et al. 2006)	8
Figure2.2	Simple Geographic Mining.	8
Figure 3.1.a	EXIF Fields	.15
Figure 3.1.b	EXIF Fields	16
Figure3.2	Azimuth	.19
Figure3.3	Formula for calculating angle	.20
Figure3.4	View Point Attribute	21
Figure 3.5	View Point Attribute Table2	22
Figure4.1	Different composition of Torre di Pis	.23
Figure4.2	Parameters of Viewshed Analysis	.25
Figure4.3	Procedure of Viewshed Analysis	25
Figure4.4	Result of Viewshed Analysis	26
Figure4.5	Output of Procedure	.26
Figure4.6	Result of KDE	.28
Figure4.7	Result of Percentage	.29
Figure4.8	Kölner Dom	.30
Figure4.9	Viewshed with H.Angel.	30
Figure4.10	Viewshed of V.Angel.	30
Figure4.11	Viewshed area of height setting	.31
Figure4.12	Final Output.	31
Figure5.1	DSM Data	.34
Figure5.2	Building Layer	.35
Figure5.3	DSM with Building Layer	.35
Figure5.4	Point Data	.35
Figure5.5	Attribute Table of iPhone4 Photo	.36
Figure5.6	EXIF File of iPhone4 Photo	.36
Figure5.7	Photo of iPhone4	.37
Figure5.8	Result of iPhone4 Photo	.37
Figure5.9	Photo of Fujifilm	.37
Figure5.10	Result of Fujifilm Photo	.37
Figure5.11	Attribute Table of Fujifilm Photo	.38
Figure5.12	Photo of City Hall-a	.39
Figure5.13	Photo of City Hall-b	.39
Figure5.14	Reduce Part	.40
Figure5.15	Saturday Evening	.41
Figure5.16	Sunday Afternoon	.41
Figure5.17	Example of Distance Weight	.42
Figure5.18	Smaller Map Scale	43
Figure5.19	Bigger Map Scale	.43
Figure5.20	Time cycles	.44
Figure5.21	3D building appearance with GoogleMap	46
Figure5.22	Web Service Diagram	.47
Figure6.1	Relationship between Chapter and workflow	51

## LIST OF TABLES

 Table1. Common Kernel Density functions(Smith, Goodchild et al. 2007)

# 1. INTRODUCTION

#### 1.1. Motivation and problem

In past several decades, technology has continued to improve. From cpu286 to dual core computer, the power of technology has also increased. Nowadays, not only computer, but also information technology and Internet have become important parts of human lives. With the computer, we can work, calculate or restore data easily; with IT and Internet, we can freely and quickly communicate with other people who are around the world or transmit our information to them.

From the result of these improvements in technology, more and more people subconsciously contribute their information on Internet. At the same time, scientists, engineers, and marketing companies try to create appropriate ways to help users to act on Internet and contribute data through Internet, while also helping themselves to collect data.

Four years ago, Goodchild(2007) gave a term "Volunteered geographic information (VGI)" to define this new phenomenon: "the widespread engagement of large numbers of private citizens, often with little in the way of formal qualifications, in the creation of geographic information, a function that for centuries has been reserved to official agencies. They were largely untrained and their actions were almost voluntary, and the results may or may not be accurate. But collectively, they represent a dramatic innovation that will certainly have profound impacts on geographic information system and more generally on the discipline of geography and its relationship to general public." (Page 212)

Flickr and Panoramio are two websites of this kind of Internet application platforms. Users can disseminate their information by uploading photos through Internet with no cost. What users have to do is connecting to the Internet and creating accounts, then uploading what they would like to share with Internet Community. This information can be used by researchers, especially, for Earth observation research. Citizens can always offer latest new local landcover changing information more quick than experts by doing official Earth observation or field work. According to Flickr historical record, there are about 4.3 million photos with Geo-tag were posted at July, 2010. This means there is a large amount of potentially useful data.

In other case, it is called "Human Sensor" which is the famous one of VGI application, too. In anterior researches, researchers set down static monitors to detect environment or capture real-time image. But, in fact, this only could cover parts of whole area and could not move so that the efficiency of these methods was limited. Nowadays, human, who can compile and interpret, what they feel and move on the surface of the planet, can be the senor. By using device, such as mobile phone, communication system, or other advanced technology, human can report what they had seen and felt. The report could include time, location and feature. It's really interesting for researchers to mine useful information from reports. In Zanzibar, such a project is underway to monitor water availability by citizens using mobile phones to report (Jürrens, Bröring et al. 2009). In this case, if the water level is too low, citizens report it immediately and government department can response positively to solve water shortage problem. The government doesn't need to ask their employees to go out and survey, but by this way, they can get the information and have action. This case shows one of the benefits of using VGI principle.

VGI seems be powerful, but how people practically act it in their reality life? There are several ways to do it. For example, websites, liked Flickr and Panoramio, can offer free space for users to upload their photos and write blogs. Moxley and Kleban(2008) created suggestion tool for geo-tags on Flickr. Mainly, GPS data was used for first step. The algorithm can reduce search range by x and y coordinates obtained by GPS device with setting geographic radius and extract the appropriate suggestion for photo annotation to user. The efficiency of this algorithm seems good.

According to NeoGeography mentioned by Goodchild (2009), this new trend is so powerful and would raise innovation. After more people upload their photos and information, the websites, as mentioned as before, become useful. For example, tourism management can make data mining to find out the popular spatial location and to offer tourist information or Local Based Service for tourists at that place. These are a huge database! Further, ranking by canonical view (Yang, Johnstone et al. 2010), the VGI data can be used to extract popular view of tourist interesting from tagged photos. And using EXIF file(JEITA 2002), it's possible to build relationship for space and time and the result could be used for movement tracking.

Although VGI seems be the positive trend, the quality of VGI should be considered as an important issue. In reality, users always rely on system functions to have their behaviour on Internet. For instance, if users want to tag the photo, they will only follow the system suggestion without concern, because they sometimes don't know what is in their photos. Here is the case. A lot of international tourists visit other countries which used different language, then, after they return home, they can't remember what is really in their photos. Although they may have some paper maps, notes, some brand with foreign language in the photo or operating digital maps like Flickr to search information, the problem is, actually, language. In general, for international tourists, the words of different languages, which are not used or learned by them, is looked like strange symbols. It means the words written in the strange language are not easy for international tourists to identify. If system can provide the relative information for the main objective in user photo, instead of mechanically clicking button and moving mouse, it could let user think of photo and interpret the photo. After this work, the VGI could probably increase the quality and get more meaningful worth.

Although several application systems on mobile device offer the function, based on the direction of face, which can probably give users information for what they probably looked for or took photo for, most of them need wireless to connect to Internet for advancing technique supports or depend on the digital map and database which store in the mobile device to implement the function. It means the ability of current function is limited. Users don't like to stand in front of the landscape to wait for connecting to Internet. In other words, if people use the device which doesn't have wireless to take photo, they can't get this function. So it's necessary to find another possible way to do this work, after they return to home and sit in front of the computer.

Our interest is to try to find possible solution to help user to identify what main objective in their photo is which may help them recall memory of travel. There are several methods to solve this problem. However, first of all, we would try to combine geo-tag photo clustering and viewshed analysis to build our process. Another key point is EXIF file will be used in this research to obtain true view sight, because people use camera to capture the scene into digital film, not eyes. The result of our suggestion process can figure out what is the possible objective in user's photo. In this research, we only focus on the core methods of geotagging suggestion. We don't build a real web service to offer this function. But the discussion of web service diagram will be done. We hope, after implementing this new process, in future, the quality of relative VGI data could get better than previous one.

#### 1.2. Research identification

#### 1.2.1. Research purpose

Our research will try to provide a way to suggest user what is the possible objective in their photos while they want to operate geo-tagging application. In other researches, the simple geographic mining method uses geo-distance to set boundary for searching relative objects. In this research, geo-distance would be changed by other possible elements for geo-tagging suggestion system. Main issue is application of viewshed analysis. Our process will try to use viewshed analysis in place of geo-distance to deal with geotagging problem. Also, the VGI data such as tagged photos on Flickr or other websites will be assistant element to improve the result of geo-tagging suggestion. We will use these two elements to build geotagging suggestion process to tell user what is the possible objective in the photo.

#### 1.2.2. Research sub-objective

- I. To realize tourist behaviors and possible elements for improving quality of suggestion.
- II. To realize what they can see at the position where they took pictures during travel.
- III. To analyze EXIF attributes and how use them into this process.
- IV. To understand how current photo database which were already tagged by users can be used in this process.
- V. To design process and get some feedback for giving other researchers suggestion to create or improve current geo-tagging suggestion system.

#### 1.2.3. Research question

- I. What is tourist interesting for taking pictures? And why?
- II. What kind of function can be used for detecting visible area? Which objects are seen?
- III. How can EXIF be used? What attributes of EXIF are useful?
- IV. What is the role of tagged photos? What kinds of information can be provided by tagged photos? How can process use it to extract famous landmarks?
- V. How does the resulting system work? Is the result good or bad? What can be the possible factors to improve it in future?

#### 1.3. Innovation

In this research, boundary of using geo-distance would be replaced by output of viewshed analysis. In general, when geo-distance was used to be radius in algorithm, all photos within this circle would be accounted for possible objectives. In fact, the buildings have the height to affect view. Some of shorter buildings behind tall ones would be hidden. So only objects within the viewshed area are possible selection for this viewpoint. This is the major benefit by using viewshed analysis. Then this research will use VGI data to make a flexible and broad suggestion. In past time, most of geo-tagging systems like to use image content-based method to tell users what are the objects in their photos. But, actually, it's not so useful for users. In the photo, there exist several objects, if the system only use image content-based method, and

then users will get a long list of landmark or building names. Obviously, in the list of names, only one or two could be more important than others so that if the system can point out them, this will be more meaningful to user. Under this situation, VGI data can combine data from different sources and after appropriate analysis of VGU, liked clustering analysis, it also can reveal what is user's interesting. It will be better than only using image content-based method.

There are two most important issues in this research. One is using viewshed analysis to reduce search area and improve the result which displayed the visible objects. We also try to utilize EXIF file of the photo to improve the result of viewshed analysis. Another is using VGI data. Here, VGI data means tagged photos on websites. From VGI data analysis, it should be flexibly evaluated for user's interesting. Figure 1.1 shows our process.



#### 1.4. Main Step of my Research

I. Literature review

Literature review is important beginning. It can help us build principle for our process design. From other papers, we can know what kinds of relative works had done and they can offer ideas or suggestions for elements of our process.

#### II. Viewshed Analysis

To extract visible objects, this research employs viewshed analysis. The requirement of viewshed analysis is existing, especially DSM (Ashton 2010). In general, it should have more work to get DSM data. Enschede city would be selected for analysis. Because we have Enschede data, this research decides to use Enschede for study area. The viewshed of point would be created for individual layer and used in future.

III. Tagged photos clustering analysis

Obviously, tagged photos contributed by users on Flickr perform specific distribution. At different scale, the clusters of photos appear on the map. Most of clusters are for special landmark of this area. These kinds of landmark can be historical building, square, natural landscape or other interesting objectives. Through evaluating of clustering of tagged photos, this research would identify their distribution and extract representative landmarks by using Clustering Analysis within the visible area for a given viewpoint.

IV. Process building

Using ArcGIS9.3 builds the process of this research. ArcGIS9.3 is the commercial GIS software and provides a lot of geo-functions, including Viewshed analysis and Clustering analysis.

V. Experiment

At this step, I will use iPhone4 or other mobile phone or other digital camera equipped with GPS to take photos at Enschede. Two photos would be used for experiment. The result of experiment would give feedback and suggestions for improving the system.

The work flow shows as following:



#### 1.5. Outline

The thesis is structured as following patterns to achieve the research goal:

- Chapter1 introduce motivation, research objective, questions and other basic idea, including innovation
- Chapter2 will do literature review to understand how other works had done and how many useful information from other papers will be used in this research. This chapter will have three major sections: geo-tagging suggestion system, viewshed analysis and clustering analysis.
- Chapter3 will briefly explain relationship between photos' content and human thought and introduce several key attributes of camera metadata stored in EXIF file.
- Chapter4 will explain our methodology and make example case to show the benefit.
- Chapter5 will be the experiment and discussion. This part will show the result of test. It means we will have a short field work at Enschede to take several photos for testing. Also, we make the discussion and suggestion. We try to find out the possible factors to improve this research and give suggestion for future work.
- Chapter6 will be conclusion.
- Final section is Reference and Appendix.

# 2. LITERATURE REVIEW

#### 2.1. Geo-tagging system

Since the turn of the century, more and more people use digital cameras to take photos when travelling. After returning home, the photos are often uploaded to website for writing their travel blogs through Internet. This is shown in Flickr historical reports.

Various commercial interests have tried to attract Internet users to make up the Internet Community to share their experiences or information with each other. These kinds of websites, like Google Maps and Flickr, are increasingly popular. To help users to record their photos correctly, a geo-tagging suggestion system is used. Tagging means users use words to annotate photos including photo's name, and text. Geo-Tagging is when users assign a spatial point on a map to display the photo location. A Geo-tagging suggestion system is an application for user, photo and map. The general idea is that the system will suggest to users to put their photos on the specific position where they are.

There are lots of algorithms or processes for geo-tagging. Basically, there are two major categories of geotagging: tags using the *camera* location and tags of the *subject* location. The former is for camera position, and, also, photographer position. The later tries to help photographers to remind what are the subjects in their photos and where they are.

The Global Position System (GPS) which can record coordinate information plays an important role in geo-tagging system. While the system reads GPS data, the location of the camera or the user could be put on the map approximately. Although GPS can do it easily, the accuracy of GPS is still the problem. The accuracy is different depending on what kind of GPS device is used. Normally, the GPS devices have a positional accuracy of 0.5 meters. Scientists continue to try to improve GPS ability, including reducing positioning time, accuracy or efficiency.

To help the user determine both the location and objective in the photo, the system needs photo annotation to recognize the characteristics of photos. (Viana, Bringel Filho et al. 2008) argue that "Photo annotation can be set into two main categories: context-based and content-based." The characteristic of *content-based* algorithm is analysis of the image. While the system knows what the objective is and the location of it, system can predict the location of camera. This means the result can indicate location of the camera. Here EXIF should be mentioned. EXIF is a template created by Japan Industrial Agency. The major function is recording various metadata of camera automatically, when photographer used the camera to take photos. The field which recorded focal length setting could be used to obtain camera's location. Other attributes, like subject distance and exposed time also could help system to estimate orientation or other features.

Although content-based algorithm was developed several years ago and got great real application, it still has a several factors to cause bad effect the result which is automatically generated content annotation of photos (Naaman, Harada et al. 2004). Also, the accuracy of content annotation should be considered as an important issue. For example, the shape of a church would be similar. So if there are two similar buildings close to each other, the content-based algorithm cannot easily identify which one is the correct building. To improve the system, *the context-based* algorithm has been developed.

The work of context-based algorithm is trying to find out the popular or relative annotation of photos' context within the defined area. And then system could translate the annotation to relative place. For example, the word, "Louvre Museum", is relative the address "PalaisRoyal, MuséeduLouvre, 75001Paris,

France". When talking about context-based algorithm, the social tagging system should be discussed. A social tagging system relies on two major factors to work: one is shared social structure; another is structure of linguistic and thought in user community. Figure2.1 shows how Social tagging systems work. While researchers try to discuss about people annotation of photos they should consider their linguistic issue as an important part, because the linguistic issue can affect human's writing. And the social groups, like engineers, vendor or doctors, can have different habits while they use words to describe something. Even more, cultural factors can have a great effect on human thought and writing. So, a Social Network System is an important issue for context-based algorithms and for a geo-tagging system, too.

As shown in Figure 2.1, there are many relationships between different groups and within the same group. The system will try to find the regular rule for each social group and categorize each user into social group, and then use the rule to find out the appropriate annotation for each user and photo. All of them indicate an important fact: the social network of the individual can affect the result. Under Social tagging system, the benefit is improving vocabulary problem and linking in social network.



Researchers used simple geographic mining method to extract relative photo annotation to find out the popular things within the area where is defined by used both GPS information and search radius. The algorithm is shown as Figure 2.2.

$$S(a) = \sum_{i=1}^{|\mathbf{U}|} |a \cap A_i|$$
 (Moxley, Kleban et al. 2008)  
Figure 2.2 Simple Geographic Mining

This algorithm tries to find out the possible annotation for user's photo based on geo-tagged photos contributed by other internet users. "U" is the set of users who contributed at least one photo. "Ai" is a collection of annotations from each user; and "a" is geographic radius. According to the algorithm, the score is for each similar annotation value which calculated the numbers of users who used the similar annotation within the defined geographic radius. If the score is higher, it indicates this specific annotation could be the possible annotation for the user's photo. The highest score is therefore more appropriate

suggestion for user. In the above method, it shows the geo-distance, "a", is used for setting boundary for data search. First of all, location of camera from GPS is the central point, then the radius could be one hundred or two hundred meters, and the algorithm would search whole data within this radius and try to find the numbers of each similar annotations of the user's photograph.

But, there still exists a problem. Users have noted that several convenient functions on a website could cause a serious problem. For instance, user can drag an album of photo onto digital map for geo-tagging. If they do this, the resulting point has many photos. Actually, not all photos under that geo-tagged point are belonging to this place. The result would be that different photos were assigned the same tagging position.

To improve simple geographic mining method (Moxley, Kleban et al. 2008) used new idea to refine it. Based on the idea of a social tagging system, this algorithm considered other users' distribution being important material when the system searched similar annotation to give user a tag suggestion. Each user can only offer their contribution once. So if someone has multiple similar photos' annotation within the defined radius, the system only collects one of them and the others are ignored.

Although researchers try to improve the algorithm, obviously, the problem still exists. Polysemy and synonymy still are the difficult problem (Golder and Huberman 2005). Polysemy means a single word has multiple related meanings and Synonymy means different words have the same meaning. Although a Social Network System can improve this problem, the result is limited. The reason is because personal characteristics can affect their habit. To improve upon these methods, many more algorithms exist to assist in geo-tagging. SpiritTagger (Moxley, Kleban et al. 2008) is another example. After simple geographical mining, the image contents, including global colour, texture, are set as the filtering. Moreover, it uses *local frequency* in comparison to *global frequency* to increase accuracy. The idea is they compare each frequency of annotation appeared at LA which they called local frequency and represented globally which is for global frequency. For tourists, they believe the tag which local frequency is obvious higher than global frequency is more important and after their algorithm, they will add this frequency rule in weight and give suggestion for tourists. Viana, etc, (Viana, Bringel Filho et al. 2008) tried to combine mobile device and the other metadata including spatial, temporal and social character to give a possible solution. The idea uses Bluetooth device codes. Each device has its unique Bluetooth ID like"000xx00101" and when the device is working, the devices can detect each other. This method is trying to identify Bluetooth ID and grab related metadata to make the judgment for geo-tagging.

In addition, an alternative idea for geo-tagging suggestion system should be mentioned. In 2005, Japanese researchers used orientation of direction and subject distance to predict the major objective in the user's photo (Iwasaki, Yamazawa et al. 2005). Orientation of direction is sourced from compass data and the *subject distance* (which is a terminology in photograph field) is the information for the distance between camera positions to objectives. This distance is recorded in a regular unit such as meter or in specific unit defined by the designers and stored in EXIF file. The general idea is using orientation of direction and GPS data to decide the standing point of the photographer and then translating subject distance into map units to search for possible objects in a reference database which includes building attribute or other spatial attributes.

Web2.0 allows users to upload their photos through the Internet and tag them on a location map, and in doing so lets social networks extend themselves. Scientists should keep developing better solutions to deal with the geo-tagging issue. However, "new" possible elements are emerging for improving geo-tagging suggestion tools. If the accuracy of this kind data is good, they may also be very useful in many other fields.

#### 2.2. Visibility analysis

Bartie et al. (Bartie, Reitsma et al. 2010) said: "When you stand at Trafalgar Square, Nelson's column, The National Gallery and many statues and building could be seen. The space is largely defined as visual field." In general, the analysis of a visual field is called visibility analysis. Many researchers have tried to apply visibility analysis into other fields. These include scenic quality, sound reduction, urban design, civil or military observation needs, etc. A viewshed is one basic function of visibility analysis. Currently, most of commercial GIS products, like ArcGIS10.0 and Global Mapper, has this function.

Nowadays, the popular application of visibility analysis is using in navigation systems. Not only for in-car devices, but also for pedestrian devices such as mobile phones or PDAs, visibility analysis can help users to identify efficiently where they are. The GPS can collect coordinates information, and the device, at first, can show a visual field of the location. This means a user just needs to read real necessary map at this moment, not a regular square map. This result removes the redundant part of the map. On the other hand, the benefit offered by combining visibility analysis and spatial database is the specific landmark could help users more easily identify their real location, not only extensively searching on the map.

The application of visibility area is great. So how does the algorithm work? The principal functions can be set two categories: line of sight and viewshed (Smith, Goodchild et al. 2007). Line of sight is point-to-point computation; viewshed, typically, is a point or a series of point set to make a surface computation. Other functions like "Isovist analysis" are also popular and useful. Benedikt (1979) built the basic version of Isovist analysis, only using 2D to display urban appearance; Rana (2004), firstly combined with ranking rule system evaluation to improve the previous version. For further application, "Space syntax" which is related to "Isovist analysis", for instance, is another case. The major purpose of "Space syntax" is to find out the place with high connectivity so that it is useful for urban design or movement management. In 2004, Claremont and Turner had done the related study. They created specific software, "Turner's Depthmap", which included the concept of both space syntax and Isovist analysis.

During the past decade, most GIS software only used 2.5D way to store 3D data. In fact, this creates a problem. That is some small but visible objects are hidden in 2.5D way. So the height value will be incorrect. To solve the problem, currently, more projects are focus on how many spatial details could be included in viewshed analysis so that some of them try to use 3D's idea and methods to grab more spatial detail. Several scientists have done related researches (Engel and Dollner) In other case, Bartie et al., (Bartie, Reitsma et al. 2010) used spatial relationship between different objects to obtain more viewshed content. However, considering what kind of purpose for the project, researchers need to make a decision to choice appropriate one from all of current algorithm. Or time is wasted in unnecessary computation.

The sources of height value/elevation are variable. Most of them can be split into two categories: Digital Elevation Model (DEM) and Digital Surface Model (DSM). The main difference between them is DSM includes DEM and, additionally, the height of objects located on the ground. That is, at a given point, the DSM value equals DEM value plus the objective height value. The objective such as vegetation and building could be the obstacle for viewing. This implies that using DSM data would be more accurate than using DEM. However, this is not true in all cases. For example, if you walk at rural area where the major economic is agriculture, there are not many houses so that there aren't higher obstacles. Then using DEM and DSM in viewshed analysis almost get the same result. DSM is more difficult to get than DEM, and more expensive. In addition, DSM has too much detail to calculate viewshed, especially at different land using type. Time required for calculating a viewshed using DEM would commonly be less than by using DSM. Again, choosing DSM or DEM depends on what the purpose is.

In general, viewshed analysis can play an important role for setting a view boundary to search VGI data. Although viewshed analysis can get lots of advanced spatial detail, according to our purpose, only view boundary will be necessary, not the precise spatial detail within the objects. This means the deeply details of objects, such as a small sculpture on the building appearance, are not so essential that I will ignore it at this stage. But this could be possible to display after using VGI data. Basic algorithm of viewshed analysis will be implemented. It implies the time using for calculating would be less.

#### 2.3. Density and Clustering Analysis

Before discussing methodologies of Clustering Analysis, it's necessary to define what Clustering is and what a Hot Spot is. Generally, Clustering means there are similar things occurring within a given area. According to Lawson (Lawson 2010): "clustering of a spatially-referenced feature is broadly defined by the term 'unusual aggregation' of events" (page 232). It is obvious that the high density of events at spatial location represents a cluster. Suppose we obtain traffic accident data as points, and we want to know where the most dangerous area is. Areas with higher density in the result mean higher possibility of area to have traffic accidents. This is clustering. Key elements of clustering are spatial location and size. These will be discussed later.

As the basic definition of clustering is given above, what is a Hot Spot? The general idea is that if the density is higher than threshold value, this area could be a Hot Spot. The word "threshold value" is set individually in each different project. For example, in researching traffic accidents, we want to know which road segment is more dangerous than other ones. The threshold value might be set to 10 accidents each month per road segment. That is, if the value of density for each road segment is higher than 10, they point to a dangerous area which could be described as a Hotspot.

The data type which was referred before is also an important thing in Clustering Analysis. To display data, it can be divided into two different types: continuous data type and discrete data type. Continuous data types include surfaces and contour lines; and the point data type can be used to represent the discrete data such as events. Actually, there exist several algorithms for continue data type for variable research purpose. Although continuous data type can also be used for clustering analysis, from general definition of clustering analysis, point data which displays the events on the map will be better (Lawson 2010). However, before selecting methods of Clustering Analysis, we should consider data type and research purpose as important critical issues.

Different data types and different analytical purposes have given rise to many different Clustering Analysis algorithms. These range from simple to complicate. At first, for example, using point data of traffic accident, the spatial location is only considered in point density method. But, we can also include the element of time, not only spatial relationships. The terminology of this kind of method is called spatio-temporal clustering. This can be applied for many purposes. Disease transmission is a well-known case. Scientists not only want to know how many people were sick and where they were, but also when they got sick. From "number", "location" and "time", scientists can imitate the ways of disease transmission and furthermore predict the next danger area to give residents warning.

In 1964, the first researcher, Knox, referred the idea of spatio-temporal clustering in his research for child leukemia cases. It should be aware that Knox set the specific critical value for Euclidean distance and time interval by himself. Based on his work, Mantel, at 1967, developed the widely adopted methodology:

 $Z = \sum_{i,j} XijYij$ ,  $i \neq j$  (Smith, Goodchild et al. 2007)

The algorithm shows it not only measure Euclidean distance, but also time. After implementing, the result is the total numbers of close pairs. But the problem is how to decide the value for distance and time interval? And do they consider the situation changes? For example, in traffic accident management, population change means that numbers of cars also change. So is the result accurate?

After Mantel's work, experts suggested a Monte Carlo simulation approach can help to decide the critical value for distance and time interval. On the other hand, scientists like Jaquez (1996), developed K nearest neighbours method which doesn't use a fixed value for distance to improve Mantel's work. In 1999, Kulldorf and Hjalmars tried to think about the population change problem and added this variable into the model to let the result more closely resemble reality.(Smith, Goodchild et al. 2007)

The other ways for Clustering Analysis include Point Density and Kernel Density Estimation. Density is the basic idea of the original algorithm: event numbers per zone area. The idea of 'Point density within a polygon' method mainly focuses on the point data type, but also can be used for analyzing line data type. The work of Point Density within a polygon is to count number of events and use a well defined grid. This well defined grid means the area that would be affected by the point. The cell size of well defined grid could, for example, be fixed 5 x 5 in map units. Nowadays, researchers have developed a method which uses political districts or uniquely defined polygons to replace fixed quadrants. However, this meets the same problem, because the situations of small parts within the zone are not same as each other. For example, the population in a city will change a lot from edge to downtown. The general assumption of Point Density is all situations within the cell/zone are the same. Obviously, if we want to know the Clustering for specific goals, this density idea is not good enough.

The other researchers proposed the Kernel Density Estimation (KDE) to solve the Clustering Analysis problem. The key difference between KDE and Point Density of Polygon is KDE focuses on point relationship, not on fixed well-defined grid or specific polygons. At the first step of KDE, the algorithm should consider each point as the starting point and decide on the affected area. Within the affected area, each cell gets value like 1; out of the area, it gets 0. Finally, counting total value of each cell, the result with a higher value means this cell is the unusual aggregation place. This is also known as a Hot Spot. To let readers more easily identify, researchers input whole cell value into GIS software and make thematic map, setting classes(Brimicombe 2007). The terminology of radius of affected area could be called the bandwidth of matrix. This is an important issue. The other issues are how many classes should be set and what are the values of each class.

According to different condition of different study area, several researchers suggest KDE should include the effect of (Euclidian) distance, so they have proposed several different distance weight matrixes to improve KDE. Table 1 shows parts of these kinds of algorithm for different distance weight matrix. It's impossible to say which way is correct or incorrect. The appropriate one depends on the situation t.

Nowadays, as a result of great development of technology, the computational ability of computers have resulted in more and more software and complicated algorithms built for Clustering Analysis. For example, most Geographic Information Systems, such as ArcGIS, can do Clustering Analysis easily. In ArcGIS, the tool package is pre-programmed. The button named "Hot Spot Analysis with Rendering" is used for Hot Spot Analysis. In other software or websites, including SatScan, Flexscan and R, the application tools could be easily found.

Kernel	Formula	Note t=distance/kernel bandwidth
Normal(or Gaussian)	$1/2k e^{(-t^2)/2)}$	Unbounded, hence defined for all t. The standard kernel in Crimestat; bandwidth h is the standard deviations.
Quartic(spherical)	$3/k(1-t^2)^2, t \le 1$	Bounded. Approximates the Normal. K is a constant
(Negative)Exponential	$Ae^{-kltl}$ , $ltl \leq 1$	Optionally bounded. A is a constant and k is a parameter. Weights more heavily to the central point than other kernel.
Triangular(conic)	$1 - \text{ltl}, \text{ltl} \le 1$	Bounded. Very simple linear decay with distance.
Uniform(flat)	k, ltl $\leq 1$	Bounded. K=a constant. No central weighting so function is like a uniform disk placed over each event point.
Epanechnikov (paraboloid/quadratic)	$3/4 (1 - t^2),  t  \le 1$	Bounded; optimal smoothing function for some statistical applications; used as the smoothing function in the Geographical Analysis Machine and in ArcGIS
	Table1 Common Kernel Density	utions/Smith Goodchild et al. 2007)

#### 2.4. Summary

Today, more and more methods for geo-tagging suggestion system are developing. This research aims to find better ways to solve the geo-tagging problem by using geographical components to help identify what is the objective in the photo. Although Image content is very important, the geography issue is thought to be equally if not more so.

Viewshed analysis offers an efficient method to narrow down searching geographic area. And the result of viewshed analysis is reasonable, because only what can physically be seen could be the objectives in your photos. The result of using viewshed analysis will be better than using method of fixed geo-distance search radius.

An important issue is that each person has different idea of how to organize the objectives' composition in a photo, which means it is very difficult to indicate absolutely what really major objective in user's photo is. To solve this problem, existing VGI data (in the form of geo-tagged photos) can do more. In general, collecting VGI point data of geo-tagged photos, and then using Clustering Analysis to find out the order of the possibility for each objective. After that, the output could list the name of landmarks or buildings in order. Through this list, may the first one or second or third is key point to answer user's problem.

This research will not use any image content-based method. As a result, the calculating time should be less. And I hope from flexible and dynamic VGI data, the suggestion can respond to user the real situation which a tourist is really interested in.

## 3. DIGITAL PHOTOS AND ATTRIBUTES

#### 3.1. Photo

What is a photograph? Why do people like to take photos? For citizens or tourists, not expert, it could be generally said that a photo is the image taken using a camera for the purpose of capturing an interesting or important moment. Using a camera can quickly do this job well. Today's cameras are simple to use and range in price from very affordable to very expensive. Simply by pressing a button, the scene is automatically focused and the image stored in the digital file. It's a very simple process, hence more and more people like to use photos to record their daily life and travel memory.

What is the possible objective in the photo? People always like to take photos for the objectives which are of interest to them. This could be a landscape, an old building, a music performance on the street, or a shop selling something special. If the locations of people are close to each other, generally, they would take the photos for the same objective which is more attractive than others.

While taking a photo, people believe they can remember the information of the objective. But, in fact, most of them tend to forget it. This is especially so for tourists, who often make longer journeys. The other problem is that tourist usually gets their information through travel books or magazines. These kinds of books tend to focus on restaurants or special shops, like desert shops, bakeries or other specific shop. Tourists interest in them, but not in deep motivation. Sometimes, they just want to tell their friends they had visited the shop or the place. So they may get there and take the photo to show the 'evidence' that they visited there. After they return home and organize the photos, perhaps they don't remember what is in this photo, or even where exactly it was taken.

In other cases, people decided their travel route 'on-the-fly'. As a result, they couldn't keep every place they visit in their mind for the entire journey, and almost lost the memory of the objectives in their photos. For example, tourists arrived at the train station of famous travel city and used the tourist guide to navigate around. Then they made a plan and walked across each street to visit special landmarks of this city. During that time period, they got a lot of information and took many photos. Finally, after returning to home, there was possibly more than 1 thousand photos on their memory card. It would be a big work to identify each photo and matched the objectives with the correct information, including landscape name. If they had GPS information with their photos, it could only tell them where the photo was taken. For further information, including name of possible objective or others, the system couldn't represent them effectively. Only knowing position, but no idea of what is in this photo means it will not be good. Under this situation, people might make a wrong annotation of photo and it seems cause researchers the bad effect while they try to analyze photo' annotation.

Generally, the very specific objects in real environment like church, the splendid building can attract people's interesting and let them take photo. For example, if the building is assigned by "United Nations Educational, Scientific and Cultural Organization", the uniqueness of this building is more attractive than general landmark so that most of people would like to take photo for it within the visible area. But, there is another interesting phenomenon. That is if we make a distinction between people with different personal background, although people stand at almost same place, they may not take the photo for the same object. It could be related to their personal difference, for example, the culture and the language.

Here is a case. "L'OCCITANE" is a shop selling body which can be easily found in European cities. Also, it has the shops in Taiwan. But for Taiwanese, the difference is the price and the new products. Although they can find the shops in Taiwan, the price is higher and all products are not the same as in European

shops. Most of the time, Taiwan's shop has to wait for several months to get the latest products. So while Taiwanese travel in Europe, they may like to find the shop and buy some goods for their friend. When they arrive at the shop, a photo is taken of it. This type of phenomenon can be found in personal blogs.

Also, the church's photo is a good case, too. Asian people see the grand church building with beautiful decoration, the first response is "Wow~ a beautiful church~" and then take a photo of it. But for a European resident, it's not as interesting, just a part of their daily life. So they may not take the photo for the church. The above two cases show the phenomenon in more detail and introduce why different people have different interesting at the same place.

Ignored the "smaller" difference between people, while browsing the Flickr Photo Map, the photos always gather at some special place, like the Kölner Dom. It could be said that the photo cloud, the "Hot Spot", which humans take the photos for at the given place is existence.

#### 3.2. EXIF

Exchangeable image file format (EXIF) used in digital cameras is a special regular file built by Japan Electronic Industries Development Association. It keeps updating. The version of EXIF used to discuss in this chapter is 2.2, released in April, 2002. EXIF is used in lots of different image types, included JPEG, Tiff and others. EXIF is a metadata standard for digital photographs and stores information such as focal length, GPS information, etc. To read EXIF files, it needs the specific software such as"EXIF 3.0", and the attributes will be displayed in the table. The field table is looked as Figure3.1.a and 3.1.b (JEITA 2002)show.

Tag Namo	Tag Namo Eiold Namo		D	Tuno	Count
Tay Name	Field Name	Dec	Hex	Type	Count
A. Tags Relating to Version		-		0001100000	
Exif version	ExifVersion	36864	9000	UNDEFINED	4
Supported Flashpix version	FlashpixVersion	40960	A000	UNDEFINED	4
B. Tag Relating to Image Data Chai	acteristics				
Color space information	ColorSpace	40961	A001	SHORT	1
C. Tags Relating to Image Configur	ation	8			
Meaning of each component	ComponentsConfiguration	37121	9101	UNDEFINED	4
Image compression mode	CompressedBitsPerPixel	37122	9102	RATIONAL	1
Valid image width	PixelXDimension	40962	A002	SHORT or LONG	1
Valid image height	PixelYDimension	40963	A003	SHORT or LONG	1
D. Tags Relating to User Informatio					
Manufacturer notes	MakerNote	37500	927C	UNDEFINED	Anv
User comments	UserComment	37510	9286	UNDEFINED	Any
E. Tag Relating to Related File Info	Tag Relating to Related File Information				
Related audio file	RelatedSoundFile	40964	A004	ASCII	13
F. Tags Relating to Date and Time		-			
Date and time of original data generation	DateTimeOriginal	36867	9 <mark>00</mark> 3	ASCII	20
Date and time of digital data generation	DateTimeDigitized	36868	9004	ASCII	20
DateTime subseconds	SubSecTime	37520	9290	ASCII	Any
DateTimeOriginal subseconds	SubSecTimeOriginal	37521	9291	ASCII	Any
DateTimeDigitized subseconds	SubSecTimeDigitized	37522	9292	ASCII	Any
G. Tags Relating to Picture-Taking	Conditions				
See Table 5		-			
H. Other Tags					
Unique image ID	ImageUniqueID	42016	A420	ASCII	33

Figure3.1.a EXIF Fields

Exposure time	ExposureTime	33434	829A	RATIONAL	1
Fnumber	FNumber	33437	829D	RATIONAL	1
Exposure program	ExposureProgram	34850	8822	SHORT	1
Spectral sensitivity	SpectralSensitivity	34852	8824	ASCII	Any
ISO speed rating	ISOSpeedRatings	34855	8827	SHORT	Any
Optoelectric conversion factor	OECF	34856	8828	UNDEFINED	Any
Shutter speed	ShutterSpeedValue	37377	9201	SRATIONAL	1
Aperture	ApertureValue	37378	9202	RATIONAL	1
Brightness	BrightnessValue	37379	9203	SRATIONAL	1
Exposure bias	ExposureBiasValue	37380	9204	SRATIONAL	1
Maximum lens aperture	MaxApertureValue	37381	9205	RATIONAL	1
Subject distance	SubjectDistance	37382	9206	RATIONAL	1
Metering mode	MeteringMode	37383	9207	SHORT	1
Light source	LightSource	37384	9208	SHORT	1
Flash	Flash	37385	9209	SHORT	1
Lens focal length	FocalLength	37386	920A	RATIONAL	1
Subject area	SubjectArea	37396	9214	SHORT	2 or 3 or 4
Flash energy	FlashEnergy	41483	A20B	RATIONAL	1
Spatial frequency response	SpatialFrequencyResponse	41484	A20C	UNDEFINED	Any
Focal plane X resolution	FocalPlaneXResolution	41486	A20E	RATIONAL	1
Focal plane Y resolution	FocalPlaneYResolution	41487	A20F	RATIONAL	1
Focal plane resolution unit	FocalPlaneResolutionUnit	41488	A210	SHORT	1
Subject location	SubjectLocation	41492	A214	SHORT	2
Exposure index	ExposureIndex	41493	A215	RATIONAL	1
Sensing method	SensingMethod	41495	A217	SHORT	1
File source	FileSource	41728	A300	UNDEFINED	1
Scene type	SceneType	41729	A301	UNDEFINED	1
CFA pattern	CFAPattern	41730	A302	UNDEFINED	Any
Custom image processing	CustomRendered	41985	A401	SHORT	1
Exposure mode	ExposureMode	41986	A402	SHORT	1
White balance	WhiteBalance	41987	A403	SHORT	1
Digital zoom ratio	DigitalZoomRatio	41988	A404	RATIONAL	1
Focal length in 35 mm film	FocalLengthIn35mmFilm	41989	A405	SHORT	1
Scene capture type	SceneCaptureType	41990	A406	SHORT	1
Gain control	GainControl	41991	A407	RATIONAL	1
Contrast	Contrast	41992	A408	SHORT	1
Saturation	Saturation	41993	A409	SHORT	1
Sharpness	Sharpness	41994	A40A	SHORT	1
Device settings description	DeviceSettingDescription	41995	A40B	UNDEFINED	Any
Subject distance range	SubjectDistanceRange	41996	A40C	SHORT	1

Figure3.1.b EXIF Fields

The specific software 'EXIF 3.0' will be used for reading EXIF file data in this research. There are several useful information fields in the EXIF data for geo-tagging application. For instance, GPS data from EXIF of photo could be extracted and combined with reference data so that the geo-tagging application system will tell user where the photo is. For example, the iPhone4 equips the function which can read photo's spatial reference and then put the photo on the map to show where you took photo. The next section will introduce several important fields of EXIF.

#### 3.2.1. Field description

#### 3.2.1.1 Date

In an EXIF file, there are several date fields, such as "Date Time", "Date Time Original", and "Date Time digitized". The first one and second one are date data for taking this photo. And the final one is recorded automatically while user takes a photo. The date information for taking photos is important, because everything changes; Our Earth is dynamic. The field of 'Date Time digitized' will not be considered as an essential one so it will not be discussed here. The issue of time scale is discussed below.

The Earth is dynamic! Apparently, it is relative to time scale. For large time scale, our world changes slowly. For example, a new city was built, replacing the grass plain. The forest may disappear and factories may be built. Human always try to change the Earth. This kind of processes may need tens of years or more so it's not easy to find something having apparent change within one or two years, even several months. Here is a case about how people change the appearance of Earth. At 2006, the one of greatest human building, 'Three Gorges Dam' located at Yichang City, Hubei province, China, was built completely. After completing the build project, the water flooded 19 counties along the Changjiang River. One of the places is a famous old city, Kaixian County. Kaixian County is a historical city; it appeared in the Chinese history 1800 years ago. It was full of historical buildings and cultural heritage. Before flooding, most tourists wanted to visit there. But now, everything is different. No old buildings are left, and no cultural heritage. All that is left is only water body, a big water body. Checking photos of this location at popular websites or blogs, two different groups of results will be retrieved: these with pictures of historical buildings and those with pictures of a big water body. What does this indicate? Without time issues, grabbing photos from large sources will meet with problems. This case shows the effect of humans and demonstrates the time issue in large time scale.

For small time scale, the four seasons could be a good illustration. Different areas of the Earth are in different climate categories. And at different climate categories, the seasonal change can have different effects, especial for vegetation. The "Tropic Zone", "Temperate Zone" and "Frigid Zone" are the major ones. In the "Tropic Zone", the effect of vegetation from seasons change is not obvious, but in the "Temperate Zone" and "Frigid Zone", it makes a great difference. The leaves will fall from the trees and the grass will fade during winter and autumn. Generally, the vegetation could be the obstacle for tourists while they want to take photos and the big trees or high grasses are in front of them.

For 'smaller' time than previous scale, the unit could be an hour, day or week. While environmental change of this smaller time scale is not so obvious in wilderness or rural areas, in the city, it becomes possible. For example, some of musicians perform on the streets or at squares where city tours and tourists, from other countries like to take photos of them and share with their friends what they saw during travel. These kinds of events could be special, because in their home country, it isn't a common performance style. Taking such photos is a common behaviour for international tourists. In this way a city square has a different appearance at different time of the day. This kind of difference could be called a temporal phenomenon as it will appear and disappear again and again.

In the previous section, the time scale is roughly described in three different groups; in fact, it is possible to make a deep discussion and precisely categorize it. Absolutely, time scale is an important issue. It can be discussed from different aspects, for instance, land use type would be a simple but efficient critical factor. In the future, making a good filter selection setting of time issue can help geo-tagging suggestion system to choose suitable geo-tagged photos to make analysis.

#### 3.2.1.2 Direction

Many researchers try to build algorithm to point out the orientation of face direction. These algorithms work without image content-based methods – meaning they do not look at what might be the subject of the photos. "GPS Trajectory", "Subject Distance" and "Electronic Compass" are three common ways discussed by the experts. Each of them has advantages and disadvantages as discussed in the following paragraphs.

The Global Positioning System (GPS) is an important technology development, because it can let holder know where he is. Today, simple and inexpensive receiver is built into devices and offer the function such as cameras can record their location automatically. When a user returns from their holiday, s/he can

operate the specific software to read GPS record and compare with his EXIF date information to get a reasonable position for each photo. Then, the software can generate the tourist's movement path direction. Since most of the time, tourists face the same direction as they walked on their visit route, the assumption is that the direction they face is the same as their movement path.

Although GPS Trajectory seem be the simple method, the accuracy is not good enough. In fact, when people stop, they often do other things, not just facing toward. They can turn, they can look back, and they can do lots of different things, especially, while they are close to visible famous landmarks. The most important behaviour is stopping and taking photos. This small time-period is quite important, so the assumption of direction should not be accepted in viewshed analysis.

The second method is about "Subject Distance". In the EXIF file, the field "Subject distance" is designed, but not all cameras have this record. The definition of "Subject Distance" can be simply said the distance between objects to the camera. EXIF file can store the data for "Subject Distance" with unit: Meters, but in some special cases, manufacturer of camera use their unit for this field. Another important element is building footprint layer which could be stored in geo-database. This indicates the first step needed is to translate the used unit in EXIF file to the used unit in geo-database, like meters or feet. After that, the system tries to use building data to calculate Euclidean distance between objects and position of the camera. In general, if there is a similar distance as Subject Distance, the system would say this building is the objective in the photo and decide the direction.

The above method is interesting, but seems not to deliver a good result. For tracking objective when taking photos, it is based on CCD and other optical technology. The light, the objective's material, there are still a lot of things that can affect a sensor for deciding major objective and then calculate Sub-distance. Getting precise data is not an easy job. On the other hand, people often like to let their companions be the major objective and the other things, including buildings, just be the background. So the Sub-distance in this situation is for distance between companion and camera's position. The system would make obvious mistake in this instance.

Despite "GPS Trajectory" and "Subject Distance", the rest one is "Electronic Compass". Although the compass is not a common additional requirement for digital camera, the situation is changing. Compass will likely be more used in future, because it is quite cheap and knowing direction is very important.

Nowadays, GPS is widely used, and already included in many digital cameras and mobile devices. While devices are equipped with GPS, the compass is not the problem. Not using traditional compass, the electronic compass is more convenient and smaller. The price of electronic compass is also cheaper than other advanced additional devices. Several years ago, the members of Japan Electronic Industries Development Association discussed about the compass field and finally, they decided to add the new field for recording compass data. In EXIF 2.2 version, the field named 'GPS Direction' was added. And the content of 'GPSDirection' is for recording camera direction.

Also, according to a Sony commercial report (Sony, January, 2010), the new series of digital cameras, DSC-HX5V and DSC-TX7, produced by Sony will be equipped with GPS and electronic compass. It shows this will be the trend in future for digital cameras. On the other hand, electronic compasses for digital cameras can be easily and cheaply found on the Internet. Although the compass is not common at this moment, the news indicates that it will be used more widely in the future.



The compass uses the term "azimuth" to record orientation of its direction. The angle between facing direction and North is called "azimuth". The direction of North is 0 degrees, and the degree of azimuth follows clockwise. The Figure 3.2 shows azimuth. While azimuth is used in viewshed analysis, the degree of azimuth will be set as start direction in ArcGIS.

#### 3.2.1.3 Lens Info

If you want to take photos, the first thing needed to do is using your eye to search beautiful or specific objects which you really interested in. But, in fact, while you press the button for taking photos, it's another case. Through camera lens, the scene is different from that captured by only your eye. This phenomenon is caused by optics and the material used for the lens. All of these can change the "true" angle of view for an image. In general, the wide-angle lens and normal-angle lens are the two main groups. The wide-angle lens means it can have wider angle of view for getting view into files. Some of these can see more wide than a human eye. Moreover, parts of them can obtain an angle of view more than 180 degrees. A normal-angle lens can grab a view using smaller or similar angles as the human eye.

Basically, each different lens has different ability of angle of view without changing focal length. That is, the original angle of view is different and their change level is also different. They can use more glasses to change the structure of lens. This should be considered as an essential issue for calculating angle of view in viewshed analysis.

The "Model" records the type of camera and the "Make" has the information foe manufacture. So combining "Model" and "Make", it could predict the type of used Lens. Actually, the designer can change their lens design for having different ability of view angle. So the view angle of a normal-angle lens or wild-angle lens could be different between different lenses. Considering what lens the user used, the system could automatically change viewshed analysis parameter, or angle of view, to get more precise output.

#### 3.2.1.4 Focal Length

In the above section, the lens information is referred. However, not only the lens, but also focal length can affect the angle of view. If a photographer zoomed in our out, the view is changed; also, angle of view has changed. Zooming in, the angle of view decreases; zooming out, it increases. The EXIF file can record the focal length information so that the data could be used to revise parameters of viewshed analysis: i.e. the

view boundary. There are two fields for focal length: "FocalLengthIn35mmFilm" and "Focal Length". The "Focal Length" is the actual focal length of this camera; "FocalLengthIn35mmFilm" is a comparison value of this camera by 35mm film camera. Figure3.3 shows the basic formula for calculating angle of view by using EXIF data.



This diagram shows the simple formula for calculating view angle of lens. All necessary parameters for this formula can be sourced from the EXIF file.

This formula is not the absolute correct one. Revision is necessary so that the new index: " $\kappa$ " is set. While each manufacturer uses different technology, components or materials, the value of  $\kappa$  would be very different so that it implies value of  $\kappa$  could be offered by their Maker. Now, new formula is as Formual3.1 below shows.

$$\theta = K \star 2 \operatorname{Arctan}(|d/2|_{f})$$

Formula3.1 Revised formula

Actually, value of index ' $\varkappa$ ' is related to 'Lens Info'. In this thesis, the value of  $\varkappa$  will not be further discussed, and the value 1 will be used for  $\varkappa$ .

To calculate the view angle, the Image Dimension should be imported into the formula. The value of Image Dimension is different, and the reason is because the 'digitization' that takes place in the camera. In the past, film was used in traditional cameras, but in digital cameras, instead of standard film, a CCD is used to digitize the signal. The CCD size is different between cameras. By checking CCD size, we can get the diagonal of CCD which could be the Image Dimension. Using diagonal of CCD, the system can calculate the maximum view angle of the camera, and combine this with focal length information, so that the new view angle can be calculated.

#### 3.2.1.5 CCD Info and Spatial Resolution

In EXIF file, there are several groups of data fields used for distinct purposes. Under Main information group, it can predict CCD Info through "Model" and "Make". Here, the maximum spatial resolution of this camera is shown. Although a human eye can see remote objects under better weather, and the camera

also can do that, in reality the detail of the image stored in the digital file is limited. The way of digital camera for storing image is different than a traditional camera and human's sight. This is governed by how the "Charge Coupled Device" (CCD) works.

Before discussing CCD, the terms "Effective Pixel" and "Maximum Pixel" should be clarified. Effective Pixel means the pixels of CCD which truly contribute to the imaging process. The Maximum Pixel includes Effective Pixel and other parts of camera, like other auxiliary imaging technology, so that the value of Maximum Pixel is greater than Effective Pixel. The total area of Effective Pixel decides how many details can be recorded. This shows an important issue: the level of detail that can be recorded with a digital camera relies on the CCD ability.

Based on the pixel of CCD, it can be said the ability of CCD decides how far your camera still can get objects into clear image in final digital file. If the resolution of image is bad, details of objects are blurred in images, then the photo isn't meaningful for a tourist. It means if people want to get the clear image for identifiable objects in their photo, the maximum distance between objects and camera is limited. If the object is very far from camera, and user doesn't have other assistant devices, he has to walk close to the object for getting the clear image in photo. So this limited distance can also be used for viewshed analysis for finite calculating. If the search radius is not limited, the calculation will be infinite when the global digital surface model is used for viewshed analysis.

On the other hand, the focal length changing also can affect spatial resolution so it is an interesting topic to discuss how CCD and focal length data can translate into searching radius for viewshed analysis. It would be feasible to grab this kind of information from their design team.

#### 3.2.2. Discussion

Several specific EXIF fields are discussed and here will show how they can work in viewshed analysis. We use ArcMap9.3 for the case to show how software translates EXIF relative information to necessary parameters for viewshed analysis and Figure3.4 and Figure3.5 show all attribute table of point data used for viewshed analysis and the parameters calculated from EXIF.



The "AZIMUTH1" and "AZIMUTH2" are for horizontal view angle. First of all, the attribute of direction extracted from EXIF compass's data can be translated into field "AZIMUTH1" and "AZIMUTH2". The analysis will start at "AZIMUTH1" and stop at "AZIMUTH2". The way of compass's data to AZIMUTH1 is: value of compass's data minus half horizontal view angle. Then, the value of "AZIMUTH2" is "direction of compass's data plus half horizontal view angle. The horizontal view angle can be calculated from Formula1.

VERTI	VERT2	RADIUS1	RADIUS2
90	-22	0	2
0.11	[		

Figure3.5 shows other attributes of viewpoint. "VERT1" and "VERT2" are for vertical angle of view. "VERT1" is for upper; "VERT2" is for down. The way of Vertical view angle can be obtained as horizontal view angle. In ArcMap9.3, the original value of horizontal eye's direction isn't setting. Normally, if tourist wants to take photos for landscape, especially at flat area, she/he should look for upper or horizontally, only few cases are need to look down. At the flat area, for example, Enschede, The Netherland, if the eye direction is down, it means the objective is smaller and beside their foot while they took photos on the road. It's not the common case. If we can't get data for angle of elevation, the value for "VERT1" will be 90. The first reason for setting 90 to be upper angle is this is the maximum setting value in ArcGIS9.3. Secondly, based on physical principle, in normal standing posture, the human can't have their eyes to stare straight for upper 90 degree. Although eye's sight can over 90 degree, the over part is very small and if people keep in this posture, they will feel uncomfortable. Although some cases with over 90 degree might be happened, they are even exceptional cases. On the other hands, if the object's location is over than 90 degree, it means this object is behind human. To take photos, the common behaviour of tourist is he would like to turn back and do it in easy posture. So if human change their direction, it shows viewshed analysis should recalculate. If we don't know angle of elevation, the calculated value of vertical angle of view is only used for "VERT2". Assuming the  $\theta$  is the value of vertical view angle, and then the value of "VERT2" is " $-\theta/2$ ".

The last attributes, "RADIUS1" and "RADIUS2" are for setting search radius. Undoubtedly, the searching area is starting at camera position so that the value of "RADIUS1" should be 0. Sometimes, zooming in can make users feel confuse that some of front objectives disappear; it seems the search radius distance is decreasing. In fact, it is the result of the vertical angle of view decreasing, not searching radius problem. Value of "RADIUS2" can be obtained from relationship between CCD and Spatial Resolution.

#### 3.3. Summary

This section tries to give an example to show how EXIF is used in viewshed analysis. In this chapter, description of EXIF data and examination of its fields was represented, and the reasoning of why people take photos discussed. Briefly, the people's motivations affect the content of the photo, and different camera settings give different results. Through EXIF data, we can further realize the actual parameters while the photographer took a given photo, including time, location, distance to target object, etcetera. We will use relative EXIF information to develop our methodology for our geo-tagging suggestion tool to achieve research purpose.

# 4. METHODOLOGY

Iwasaki et al. (2005), they used orientation of direction and subject distance to predict the major objective in the user's photo. Their research is really interesting and very different from other major researches. The benefit of this idea is that, at this stage, the requirement of calculation is less than image content-based methods and the result shows it is feasible.

However, there are several critical issues. Firstly, how can they know the major objective while they use EXIF file to read related fields for calculation? As photographers know, if you take photos, there are several possibly main objectives in your sight. For example, if photographer's friend asked photographer to take photo for him with the landmark, which was the major objective, the human or the landmark? At the same time, the camera could automatically detect the objectives and the screen of camera might have several flash squares to label the possible objectives. According to optical theory, camera could change camera setting to get better quality of photo. So if someone is standing in front of the landmark, what is the result? The man will possibly be the major objective for camera setting so that the subject distance will be between man and camera, not landmark and camera. And then if the system uses this subject distance in its calculation, it could be seen that the result would be in inaccurate. Secondly, another problem results from accuracy of GPS. Actually, the error of common GPS could be within half a meter. For calculating real distance and matching, the system error could be more affected. Moreover, the error from GPS accuracy is not easy to precisely be overcome, which reduces the use of this idea.



Figure 4.1 Different composition of Torre di Pisa

To solve the problem, this research will employ *viewshed analysis* to replace the calculation of subject distance. The reason is that after viewshed analysis, the output could show whole possible objects which can be seen from a given viewpoint. And the system does not need to imitate the photographer idea for composition. For example, the Figure4.1 shows the possible place for the major objective in each photo. The photographer may like to put the major objective, the **Torre di Pisa**, at the corner of the photo or put it at the top part of the photo, not only in the middle of the photo. And, the effect of accuracy of GPS for viewshed analysis is not as sensitive as subject distance calculation. Furthermore, it is possible to use advanced algorithm of viewshed analysis to remove the error. For instance, it is also possible to use GSI software to make a buffer around the location point representing the GPS inaccuracy. For example, the inaccuracy of GPS is 0.5 meter; the radius of buffer should be 0.5 meter. Then translating the boundary of buffer into polyline type, and using it to replace the viewpoint for executing viewshed analysis, the GPS error could get possible improvement.

Human behaviour is very difficult to imitate and predict. So the VGI which come from broad users may offer an immediately usable solution. The idea is, after VGI analysis, it is possible to indicate what people like to do at this place. For example, through analysis of annotation of photos, researchers can find the popular things here. Clustering analysis could be used by VGI data and find out the "unusual" aggregation place such as *photo cloud*. In other words, from the result of Clustering analysis, the popular landmark which tourists take photo of could be presented. In this research, we use the Frequency idea to be the Clustering analysis. Although Frequency is a simple function, it can reflect the popular level in the quick and effective way.

This chapter will first introduce viewshed analysis as a tool for spatially referenced layers. Next, it reviews methods for detecting "Hot Spot". Finally, a brief summary of the chapter is shown.

#### 4.1. Viewshed Analysis

The major purpose of viewshed analysis in this research is trying to find out the true visible buildings/objectives at that given viewpoint. Although there are several useful algorithms for geo-tagging suggestion systems, most of them use geo-distance to be the searching radius. As discussed above, this is not sufficient. So to improve this part, the viewshed analysis is proposed.

#### 4.1.1. Viewshed Analysis process

In this research, the major purpose of viewshed analysis is trying to know which objects could be seen at a given viewpoint, not only to know view area. To achieve the main purpose, Viewshed Analysis process proposed here includes several parts: data preparation, computation by GIS software and post-processing. ESRI, ArcGIS9.3, is the software which used in this research.

Data preparation is the first step for viewshed analysis. Most of GIS software is had limited drawing ability so that lots of people use other drawing software to make a 3D surface data for building. For example, AutoCAD is s popular one, which produces '.dxf' file type. These kinds of data type are different from '.shp' or 'raster' data type. They use various ways to record surface data. For viewshed analysis, the one of essential elements is height. Actually, each GIS software uses different data type for computing viewshed area. In this case, it is necessary to translate other data type into appropriate data type for getting each unit's height value. For instance, Global Mapper can use '.dxf' directly; ArcGIS needs 'raster'.

For altitude information, despite of drawing software production, there are other sources. For example, the remote sensing image also can produce digital surface model, including satellite image or laser scanning technology. From them, the 'TIFF' file type or point data type are the major data type. Each grid records the value for height so that it could be used directly in several GIS software.

In ArcGIS, there already exist tool package for viewshed analysis. In the Arc Tool Box, the function for viewshed analysis could be found as following path: Spatial Analyst Tools/ Surface/ Viewshed. For the viewshed analysis function, both of the raster file with height value and viewpoint layer are necessary inputs. On the dialogue window, several parameters can be changed, like scale, grid size or z value. But if one wants to set other parameters for computation process, the only way is by changing the attribute table of the viewpoint layer. The several fields with specific name should be added and compiled. They are "OFFSETA", "OFFSETB", "AZIMUTH1", "AZIMUTH2", "VERT1", "VERT2", "RADIUS1", and "RADIUS2", shown as Figure4.2.



"OFFSETA" is setting for height of observer above the surface. Here the real height is the sum of this value and the surface height. This parameter can be set for different observer heights. Because the observer's height decides the eye's height and this can affect the computation result. In other words, that setting "OFFSETA" can obtain better accuracy of viewshed analysis. "OFFSETB" is using for objective's height above the surface to be considered.

"AZIMUTH1" and "AZIMUTH2" are setting for horizontal angle of computation. The rule is as same as Azimuth. It increases from 0 to 360 degree, following clockwise and starts at North. It starts at "AZIMUTH1" and stop at "AZIMUTH2". The value of "AZIMUTH2" minus 'AZIMUTH1' " is the horizontal angle of view. And "VERT1" and "VERT2" is similar as "AZIMUTH1" and "AZIMUTH2", but for vertical angle of view.

The last common parameters are "RADIUS1" and "RADIUS2". The major purpose of them is for setting radius of calculating boundary. It starts at "RADIUS1" and ends at "RADIUS2". If "RADIUS1" and "RADIUS2"are not set well, the calculation will be probably infinite, especially, when whole world DSM data is used! So it is necessary to set these at a reasonable value. The reasonable value could be derived from EXIF file for each different camera, but this is outside the scope of the current research. If the value of parameter is 'default', it means the analysis will not use that parameter.

After setting each parameter well, the ArcGIS can calculate the result automatically. Within several seconds, the result is displayed on the screen. At this moment, the map is raster data: the value '0' is not visible; value '1' is visible.



For next step, the raster data type needs to be converted into polygon to allow topological operation. To grab building/landmark polygon data, 'Clip' function, 'Spatial Joint' and 'Union' function can produce the output data layer for next analysis. The procedure for visible area is as Figure4.3. The result of the viewshed analysis process is the polygon layer which contains topology and attributes.

#### 4.1.2. Implementation

At this moment, there are no specific parameters for viewshed analysis, so all of them are set initially to 'default'. Using ArcGIS 9.3 to implement viewshed analysis and get the result as Figure 4.4 show. Firstly, the output is the raster data which only has two values: 1 and 0, "1" means visible; "0" isn't visible.

For the next step, it needs to translate raster data to polygon data. The "Conversion" tool can do the work, and then we get the polygon data. Now, the polygon data of viewshed doesn't have any attributes. To obtain building information, we use "building layer" and "clip" function to clip down building polygon following viewshed boundary from complete "building layer".

On the screen, the layer is a polygon layer with lots of small polygons of building footprints. For final product, we joint complete building layer with this "cutting" polygon layer and select "keep only matching records". So the screen only shows the related parts of whole building polygon. To avoid duplicate parts, the polygon layer of viewshed should erase the building parts. The last step is let "cutting" polygon layer unite with "erasing" viewshed polygon layer. The output is as Figure 4.5 show.



#### 4.2. Clustering analysis

According to the major purpose of this research, the output of the research is trying to discover what the mainly possible objective is in tourist's photo. Even if using image content-based method, it is only possible to discover whole buildings/objects in tourist's photos, but it's not always meaningful to do that, and it also has high requirement of device to implement the algorithm and enough image data. To be more effective and meaningful, the previous section discussed the output of viewshed analysis could extract visible buildings/objects in photos. Now, Clustering analysis is the way to solve the remaining part of question: "What is the possible objective which tourist took photo for at that view point." The idea is

to use clustering on existing volume of geo tagged photos. Geo-tagged photos can come from popular Internet community, private blogs and famous photo sharing websites, like Flickr and Picasa. After Clustering analysis of this data, it is possible to state what the most likely objective in the photo might be.

The previous section, it referred that the spatial detail which was ignored while the viewshed analysis was implemented. But now it is possible to recover. This is not done with an image content-based algorithm, but the spatial detail will be obtained from VGI idea. The reasons for why VGI data can possibly recover the spatial detail which was ignored before are:

• VGI is flexible, is dynamic.

VGI data is flexible and dynamic. Because people came from around world contribute their data daily, the database keeps updating and increasing. Also, the contributors are from different area, race, and country and culture so that their information can show different important aspect for the same phenomenon. This means their information can cover different aspects of the same phenomenon. For example, at Koln, tourists from Asia may be interested in the Dom church building, so they take beautiful photos for whole building and share with their friends; tourists from other Europe countries may think the sculpture on the building wall is more interesting than whole building appearance, so they only take photos for the specific sculpture. It can be said that according to distinct user group, the process can possibly produce more reasonable description of the specific spatial detail for users. It will be better than the 'fixed' result which only used image content-based methods.

• VGI is a huge database, including different sources.

It is feasible to get enormous database for VGI. Until now, the websites, like Flickr and Panoramio, keep receiving uploads of geo-tagged photos on an hourly basis. This VGI data is not only geo-tagged photos on the special websites, but also can collect from other ways. And spatial coordinates are not the only useful information that can be extracted; it is possible to grab information in the description of the photo. According to Flickr report, about 4.3 million geo-tagged photos are currently on the website. And this number will increase in future, especially as more and more users access and use the internet. It should be noted that VGI data is potentially very huge, so how to select relevant sets and analyze it will be an important issue.

As previously discussed, Clustering analysis will be the used to find out Hot Spot in the VGI data. Before go forward, it's necessary to explain what Clustering analysis is and what the purpose if Clustering analysis is. Generally speaking, Clustering analysis is using statistical method to calculate the density/intensity of specific events which have spatial location and meaningful attribute. So the major goal of Clustering Analysis can be said that it will find where the unusually gathering places which may cause researchers interesting in are. For example, for a minister of Traffic department, he/she feels interesting to know where the traffic jam is. To avoid this problem happened day after day, they can focus on the intersection with unusually intensity of traffic jam to have some ideas. Other experts give the terminology "Hot Spot" for this kind of places.

#### 4.2.1. Kernel Density method

Based on Point Density method, Kernel Density Estimation adds the distance weight matrix to make appropriate estimation of Hot Spot and tries to smooth the output layer. This idea believes the distance can have different effect of the point for the affected area so that it should use distance weight to improve the result of analysis of Clustering. There are several different distance weight matrix listed at literature review, the section 2.3, Table2.1. Considering purpose of research and point data spread situation, the researchers should choose the appropriate one. And then researcher decides the bandwidth of matrix.

The value of affected area of each point will be recorded as 1; no affected area is 0. In this research, the function "Kernel Density" in ArcGIS 9.3 is used. The output is a raster layer where each cell records the total value. For example, the cell is affected by 4 points, and the total value would be the 4, if the algorithm only uses the normal KDE distance weight. After the analysis, the raster layer could be visualized as Figure 4.6 shows. Through the layer, it's clear to see the place with deep colour is affected by lots of points so this could be the Hot Spot..



#### 4.2.2. Frequency and Percentage method

The previous section talked about Kernel Density Estimation, but actually, there are other ways to deal with this kind of problem. The Frequency is also a good way to find the Hot Spot. The Frequency is a terminology of statistics and the major idea is to calculate how many times the objective occurred within a given area or zone. In the context of finding likely objectives in photographs, Frequency can present the popular level of this objective. That is, if the value of Frequency is higher than others, it means this objective is more popular than others at this location. In this case, the point of a photo can represent the building/landscape polygon. So, if the system calculates the number of points within the building/landmark polygon, this value can be considered as the frequency of occurrences of this polygon.

Can principle of polygon of buildings be ignored? Or can we use point data to represent the polygon data? This is an interesting issue in this research. While discussing this issue, it has to go back the moment when users upload the photos and put the point which represents the photo on the map. What is user's mind of point, polygon and map? For experts, they know point and polygon are different data type to represent different spatial meaning; the common users, they don't deeply have this kind of conscious. Surely, while they upload a photo and put the point within the polygon of the building on the map, they believe this point represents the building, no idea of polygon is necessary. Under this situation, while we deal with this VGI issue, the point data should be considered as representation of the specific object. It means if the object is the building, the point can really represent the polygon. So it's reasonable to use Frequency of point to do this work.

If the output is a table to list each value of frequency of each objective, it wouldn't be a good idea, because the reader has to identify the intensity of each value by themselves. To avoid this problem, this research translates the frequency value into percentage type. That is the system grabs all relative points within the viewshed area: N; and each polygons sum the number of point within it: n. Then calculating percentage for each polygon using the formula: n/N \* 100, the result can easily let users compare the intensity between each polygon. Finally, using symbology to make a thematic map can give more appropriate viewing for users, for example, seen Figure 4.7.



#### 4.2.3. Comparison and discussion

Kernel Density Estimation, Frequency and percentage measures show the potential Hot Spots. All of them show these effectively and the position of the hot spots should be correct. However, they are apparently different. The problem is the result of Kernel Density Estimation is describing a "special" zone to present the hot place; the result of Frequent and Percentage are shown at the 'building' level. For the result of KDE, the difficult thing is to point this special zone to present a building or landmark precisely, but this is not a problem for the output of Frequency and Percentage. Because the spatial distribution of points can affect the Kernel Density Estimation result, this will not be easily controllable. As the Figure 4.6 shows, t the zone may go outside of the polygon and make a connection to another polygon. The other problem is it might have two Hot Spots within the same polygon. These two problems are not easy to solve.

For Frequency and Percentage, it also can find the problems. The most important one is how to define the boundary of each building. If you browse the Flickr or Google Map, it can easily find that some of photos belonged to the specific landmark are put outside the polygon of the landmark. Can the building have new defined boundary? It is common phenomenon that the building has a square or roads around it. Figure4.8 shows a good example. That building is Kölner Dom and the square. Many tourists take photos at that square and put the point which represents the photo on the square. So the new definition of the building should include this part. Focusing on major purpose of this research, the major work is trying to find ways to answer user's question "what is the objective in the photo" and cheek if this way is feasible. The issue of boundary redefinition should make advanced study for solve this problem in future work. And related work will be discussed in the final discussion in this thesis.



The conclusion for clustering analysis is that although Kernel Density Estimation can use all points for analysis, the Frequency and Percentage can do this work more easily. For the needs of this research, the Frequency and Percentage method will be used

#### 4.3. Example case

At the start of this chapter, viewshed analysis is discussed. A simple example case will show the result after importing parameters of viewshed analysis from EXIF.



The following parts are tests for each different parameter setting. First of all, the most essential parameter for viewshed analysis is horizontal angle of view. According human body structure, the maximum horizontal angle of view of eyes is 200 degrees. This means if the direction is set (here "AZIMUTH1"), the viewshed field can remove unnecessary part from the analysis: 360-200=160 degrees. This indicates the efficiency is increasing 180%. Furthermore, if the angle of view of digital camera is set, the efficiency will increase also. For example, the angle of view of digital camera calculated from "Focal Length" and "Image Dimension" is 90 degrees, the efficiency is 4 times the original result and about 2 times of result of human sight. Obviously, this can save calculating time for viewshed analysis. As Figure4.9 shown, after setting horizontal angle of view, the total area (the green part) is smaller than the yellow part which is without setting horizontal angle of view.

Then, the vertical angle of view can also be improved with EXIF data. The formula of vertical angle of view is same as horizontal angle. Without specific setting, the original value is +90 to -90. Actually, the vertical angle of eyes is 120 degrees and, again, from result of calculation, the value can be set so that the efficiency can be improved. Figure 4.10 shows the example of setting vertical angle of view. In the picture, the blue part is visible area and green part is non-visible area. Apparently, the total area of result reduces.

But, it is essential to keep in mind that the setting for vertical angle of view needs other data, while the system can't know the angle of elevation. Especially, if the area is not flat, tourist might keep their camera tilted down. For instance, while tourists stand at the top of mountain, they always like to take photos for overview of the closed town so that their angle of elevation is not horizontal. It can be said if the angle data is not available, considering the situation of the different area, flat or mountain area, it's necessary to change start angle of elevation setting.





Finally, the height of the observer should be set. In reality, the common GPS device in a camera or other consumer device is unable to detect elevation. There several reasons for this and some of them are not easy to be overcome, like budget. Ultimately, it's very difficult to obtain precise height of the camera while a tourist took photos. To conquer this problem, the rough information for height may be sourced through Internet and technology of web 2.0. Each website for sharing photos asks users to create their account. The general personal information for registration includes Gender and Country. These two data can be used to extract information for (average) height. That is the web system can use Country data to connect their official office to get average height information. Some of countries have made different statistical database for female and male. So it is possible to automatically derive the average height for a given user. Additionally, average height of each age group is possibly recorded. So it implies it's possible to grab still more precise data for each user. The result of height setting is as Figure4.11 shown. And Figure4.12 shows the final output. In this layer, we own each polygon attribute and can keep forward for next analysis.

#### 4.4. Summary

This chapter has described the methodology developed for this research, and given an example case. The models built in ArcMap for Viewshed analysis process and Clustering are presented in Appendix One

The next chapter will develop these methods further using real-world experiments. Two photos will take at Enschede City and processed. The results are be evaluated and discussed in detail.

# 5. EXPERIMENT AND DISCUSSION

This chapter will test and evaluate the process developed in Chapter 4. Photos were taken in down town Enschede City close to the open Market for this purpose. Also, the data used will be introduced and described. The equipment used for taking photo are camera: Fujifilm F200 EXR, GPS: eTrex and Digital Compass and iPhone4.

#### 5.1. Study Area

The study area of the research is Enschede. Enschede is located in the East of the Netherlands, closed to Germany. According to the GeoNames geographic database, the population of Enschede is 153,655. The latitude and longitude of Enschede is 52.2183 and 6.89583. The average elevation above sea level is 45 meters.

There is one major university, Twente University, including ITC at Enschede. At 2010, ITC joint into Twente University and become the sixth department. Twente University has a lot of international students so that it means when new international students arrive, they explore and take photos around the city which will be part of their daily life. And they perhaps will upload their photos at several Internet Community websites. While Enschede is one of the bigger cities in the East of the Netherlands and agriculture is part of the major economical activities, it is a diverse environment to test for each relative character of VGI data and viewshed analysis.

The other reason for using Enschede for our study area is that the precise (base) data, like DSM and building footprint layers, already exists. Most of these data are created by ITC in cooperation with other organizations so that the accuracy of data is high. Although Enschede is not really a famous tourist destinations, the reasons described above more than make up for this. Enschede City is selected for the study area in this research.

#### 5.2. Data Introduction

#### • DSM

The Digital surface Model (DSM) image is an important data for viewshed analysis. Several decades ago, the first data type for saving height information is Digital Elevation Model (DEM). The DEM only records the z value of ground elevation level, excluded the objective's height on the ground. But the DSM includes the height of the objective. There are many papers discussing the effect of viewshed analysis by using DEM and DSM. The general idea is, at urban area, the result of viewshed analysis by using DSM will be better than using DEM. The DSM image can be obtained from remote sensing image, laser scanning or using DEM combined with 3D building data might produced by expert drawing software, like AutoCAD. In this research, the DSM data shown as Figure5.1 is using the same method as the paper (Vosselman 2008).



#### Building Footprint Layer

The major purpose of this research is trying to transmit the information of the possible objective in the photo to user so that it needs the attribute for each area. To get the spatial attribute, the building layer "TOP10vector" is used. The building layer "TOP10vector" was created by "Topografische Dienst Kadaster" in the scale 1:10,000, and published in 1998. "TOP10vector" has multiple data types, including polygon and polyline, but for this case, only building polygons and attributes were used. The view is shown in Figure 5.2.

#### Point data of geo-tagged photos

For Clustering analysis, the x, y location of the photo is necessary. These can be collected through several websites: Flickr, Panoramio, Webshots, Visit Enschede, Photobucket and Virtual Tourist. Flickr is one of the most popular websites for sharing photos, and it offers a map to let user make geo-tagged photos and assign a location to the photo. This means you can view all geo-tagged photos on a map. Panoramio is another example. The Main purpose of Panoramio is showing geo-tagged photos which were permitted by GoogleMap on the map. But it also can show other photos not collected by GoogleMap, too. Other examples are travel websites which offer information and city photos for tourists.

The points are stored into Excel file containing a tag name. If they have x, y coordinate information, this can be used for the geo-reference directly. If they don't have explicit coordinate information, their position on the website map and tag name are considered, and the geo reference is manually assigned. According to our DSM image boundary, only the points of photos within the boundary of DSM and neighborhood are collected. The total number of points is 953. Figure 5.3 shows the area of DSM and Figure 5.4 shows the study area with point data.



#### 5.3. Experiment and Result

Photos were taken at the chosen location in Enschede shown as Figure 5.7 and Figure 5.9. The first test used the "iPhone4". This is one of the most popular mobile phones and it is equipped with a digital camera, GPS and digital compass. This means the EXIF file of iPhone4's photo can have all data needed to calculate view angle, including focal length, direction and coordinates of location of camera. iPhone4's CCD is 1/3.2. So the "Image Dimension" is 5.68mm. The Figure 5.6 shows the EXIF file of the test photo. And then we calculate the angle of view of this photo as 73 degrees. Next, we translate all relative elements into necessary parameters for viewshed analysis and get the viewpoint attribute table as Figure 5.5 shows.

F	FID   Sh	ape *	Id	AZIMUTH1	AZIMUTH2	VERT1	VERT2	OFFSETA	OFFSETB
	0 Point	200	0	37.58	110.58	90	-36.5	1.7	C

🖞 ExifReader - iphone_1.jpg					
ile <u>I</u> nformation <u>H</u> elp					
Open D:\Work\Data\	est photo\iPhone\iphone_1.jpg		<	5/8	>
	ItemName	Information			1
humbnail Image	Iteminanie       GPSInfoOffset       SubjectLocation       ExposureMode       WhiteBalance       SceneCaptureType       Sharpness       Sub Information       ExposureTime       FNumber       ExposureProgram       ISOSpeedRatings       ExitVersion       DateTimeDigitized       ComponentConfiguration       ShutterSpeedValue       ApertureValue       MeteringMode       Flash       FocalLength       FlashPixVersion       ColleSpace	Internation         562         1295,967,699,696         Auto         Auto         Standard         Hard         1/60Sec         F2.8         Program Normal         80         0221         2011:01:30 16:08:49         2011:01:30 16:08:49         YCbCr         1/60Sec         F2.8         Average         Not fried[Compulsory]         3.85[mm]         0100         -BCB			
	ExilimageWidth ExilimageWeithd ExilimageHeight SensingMethod GPSLatitudeRef GPSLatitudeRef GPSLongitude GPSAltitude GPSAltitud	1336         2592         OneChipColorArea sensor         N         52 13.14 [DM]       52 13' 8.40" [DMS]         E         6 53.76 [DM]       6 53' 45.60" [DMS]         Sea level       17355/362 meters         17355/362 meters       15:08:48         Magnetic direction       75:33?         OLDJPEG       72/1         72/1       1nch         958       5760			

Figure 5.6 EXIF File of iPhone4 Photo





After running our process described in the previous chapter, Figure 5.8 shows the result. The "Red area" is the most likely objective in the photo. In reality, the building is Enschede City Hall. That is the big objective in our photo on the left side. The Enschede City Hall is one of the more famous buildings in the Enschede area and several websites have written an introduction for it. Lots of new international students or city visitors who come to Enschede will likely take photos of it, especially at night. The result also shows a big building with light yellow in the bottom. This should normally not occur. But, we don't have a more detailed layer. If we did have a more detailed building layer, this big building will disappear and only small one will be displayed. And it is worth to mention that the second higher percentage in the result is the yellow area. In reality, this area is a spring and several sculptures so that many people like to take photos here, too. In our research, the major used layer is only focus on building so that in our layer, there don't have any information for other landscapes. So if we using the tourism map layer, this kind of landscapes will be analyzed by our process. Although there are small defect, generally, this result can really reflect the information of the objective in their photo to users.



We also used other devices to take photos in this experiment. The EXIF attributes: Focal Length: 6.4 mm, Direction: 85 degree and CCD diagonal: 10mm are used to calculate the view angle: 76 degree. We put all relevant information into viewpoint data shown as Figure5.11. Finally, the model created by this research runs and the result is shown in Figure5.10. It can be seen that like previous result, this one also successfully evaluates the possible objective in the photo.

#### 5.4. Discussion

While the results are as Figure 5.8 and Figure 5.10 show, there are several problems. First of all, the building layer is not accurate enough. That is, the building layer does not have detailed enough information of each shop or house, and does not include other environmental objects such as statues/sculptures on the street. If a more complete layer could be used the result will be better. Secondly, the view angle of a digital camera may not completely match the real photo sight. In this research, we use relative information to calculate the maximum view angle of the camera under the used focal length. This problem could be overcome if the manufactures can offer more information on their lens and CCD. There is also the slight problem of direction. The probable reason is accuracy of our digital compass is for 5 degree intervals, meaning some generalization happens here. The other problem is the accuracy of GPS. It is almost certain that most of these issues will be improved in future.

#### 5.4.1. Benefit of our process

Regardless of the problems of the experiment discussed above, the effect of the method developed is clear. The benefit of this process for the geo-tagging suggestion system can be divided into two major parts: viewshed analysis and VGI data. The viewshed analysis is the one of the basic functions in GIS software, but it can offer the effective work for searching "true" visible objects for the viewpoint. The VGI is a new issue and is becoming more popular in recent years. It is resulting in data made available voluntarily,

resulting in huge volumes. This can provide a flexible data source for researchers to understand the common behaviour of people. These two elements provide the basis for geo-tagging suggestion system.

As mentioned above, viewshed analysis can find the visible buildings/landmarks for a given viewpoint. Using this method to replace fixed geo-distance, the search will be more effective and the result will be more correct. For example, using fixed geo-distance, the search is for 360 degrees with a probable radius of 500 meters. The result should be incorrect, especially in the city area. The reason is buildings which are located behind obstacles will not be seen. But the viewshed analysis method doesn't have this problem. Here, the viewshed area should be re-defined in this research. The viewshed area in this research is the area that could be seen through the camera, not human eyes.

According to human physiology, 200 degrees is the maximum number of view angle for two eyes' sight(Wang 2008). In reality, the common digital camera relies on the ability of the lens to decide how wide a view angle can it have and it also can change view angle by zooming in and out. Generally, the angle of view of digital camera would be less than eyes' sight. This means if the parameters of the view angle from a digital camera could be used, viewshed analysis could limit the search boundary. It could be argued that the result will be better and the calculating time of the process will reduce.



Figure 5.12 Photo of City Hall-a



Figure 5.13 Photo of City Hall-b

As discussed in earlier chapters, EXIF automatically records the digital camera setting parameters while the camera is used for taking photos. The one of most important attributes is the "Focal Lens" which can be used to calculate the view angle of the camera. And the formula for calculating view angle is as shown in Formula 3.1. After calculation, the more precise information for the view angle could be found and used to improve the viewshed analysis. Figure 5.12 and Figure5.13, which were taken at the same position, show the effect. The true view angle of the left one is 76 degrees; the other is 35 degrees. For the right photo, if the true view angle of camera isn't used, the result of viewshed analysis will include the two side houses and shops, so the analysis result will be the same as the left one. But this is not true. Based on wrong boundary of viewshed area, the process would retrieve unnecessary tagged photos and make the bad effect for clustering analysis. As Figure5.14 show, the blue line displays the original area of the camera, and it will look like as Figure5.12. The green lines show the true view angle after zooming in and it roughly represents Figure5.13. Finally, the red area is removed after using true view angle of the right photo. So it could be said using true view angle of the camera can remove this unnecessary problem and the result will be more appropriate for the analysis.



Why do photographers take a photo at a given place? The reason could be they found something interesting at that place and wanted to record it. The "something" can be explained as the spatial detail or unusual events at that place or the object which own more attraction for human. This object may attract a person, or group of people, even more, people of the country. And then, after people take a photo, upload and tag it on a public website, this kind of information becomes part of VGI data. This indicates, with appropriate analysis for this kind of VGI data, the result the result could reveal the information for human thought the human mind. This approach may be an effective solution for this kind of application which is trying to imitate human's thought. After clustering analysis of point data which represents the position of photos that have been uploaded, the result could indicate a Hot Spot which attracts people to that area.

Basically, this research is the initial study of this process of geo-tagging suggestion system. Therefore we don't develop an advanced way to grab VGI data and do the clustering analysis. But it is still possible to see the benefits. We define the spatial relationship between point (photo) data and (building) polygons, and then using Frequency analysis to represent popular objectives. The most popular polygon could be considered as the Hot Spot. Although Frequency analysis is a simple function, it still can reflect VGI's character. That is, the result can flexibly predict user's objective and give the geo-tagging suggestion.

#### 5.4.2. Elements for improvement

This research doesn't put more attention on advanced Clustering Analysis algorithm; this is outside the scope and it could is an important area for future work. Another possible Clustering Analysis way to improve this work is to calculate the "Ranking Value" for each objective with variable elements of photo's file and social network of users. The key is to compare all of them to find the highest total value which can be defined as the Hot Spot. This idea should be divided to two parts: one that uses the EXIF file and the other that looks at users personal details.

#### 5.4.2.1. Time issue and Personal factors

Another interesting phenomenon is the time issue. The Figures 5.15 and 5.16 are the two photos taken at almost the same position, in the Enschede Open Market, but at different times. The time of the left one is at Saturday evening; the other is taken on Sunday afternoon. The content of the left image shows the scene of the open market. The major purpose of this photo seems to be to record the phenomenon of open market while vendors finished their work. But on the right is the steel sculpture which is part of the parking space building. Why are they different? What is the reason to for the difference, and why were they taken at the same location? It is a special temporal phenomenon. In the Enschede area, the Open Market is only held on Tuesday and Saturday. So if people stay at the square where it is held, and take a photo between 9AM and 5PM on Tuesday and Saturday, the image of photo is most likely for recording the scene of the Open Market, not for the steel sculpture. But at other times, the market doesn't exist in the square, so the image should be for other buildings or the steel sculpture. This case points out an important character of VGI data. That is "timing" issue.





Figure5.16 Sunday Afternoon

Our world is built in 5 Dimensions, the x, y, z, time and human. The general database can store attributes for all x, y, z and time. But the problem is it's not an easy job do deal with human Dimension, because human can have their own behaviour. Our world appearance is affected by human behaviour. This means scientists should think a better way to draw our world appearance in 5 Dimensions. The VGI data could be used for this purpose. Apparently, from this kind of data mining skill, the researcher could control variable world more easily. For other applications, such as tourism management department, they can try to find out where is the popular location for provisional music performance at different time. Through the VGI data, now, it's feasible. So it could be concluded that VGI data offers a flexible way to represent interaction between people, space and time, and also can possibly give some insight into human thought while they took photos.

While VGI relates to a user's personal information, social network systems should be considered to play an important role in this. If users want to upload and share their photos, they need to create a personal account on the website. Their personal details are necessary for registration, including gender, age, and country. They could let users be categorized into different groups. Generally, the country might be the most important one of these. The country will show the user's background, especially, their culture. It could be said that different culture can make people be different in their behaviours and interests. For example, Asian countries, like Japan, their residents consider a temple building to be a common building in the city, but a church is a very different building. So when they have an international trip, if they see a church with splendid decorations, they will feel excited and decide to take photos for it. But for Europeans, the church may not be interesting at all. On the contrary, they will be interested in Asia's traditional temple building. The culture can affect human's thought and the country, roughly, can approximate user's culture. So country might play an important role in this condition.

In other cases, the gender and age can also obviously affect human thought. Youths may feel interest for special shops, for example, the small shops which offer personal designer goods. The adults may pay more attention to shops of quality goods, like LOUSI VUITTON. The gender is another key element. Females and males have their own separate interests. For instance, women like shopping for clothes, but men may like stereos or knives. All of these variables can affect the human's thought and motivation for taking a photo. Basically, VGI is contributed from people of the whole world and is related to people's thoughts, so that it's possible to categorize it into several groups. When the system tries to collect VGI data, it can use these categories to better reflect the human thought and motivation.

#### 5.4.2.2. Location and Distance weight value

In some cases, the locations of people could have different meaning, while they took photos. In Figure5.7 and Figure5.9, the spring and several sculptures are displayed in the photos. Actually, the location of Figure5.7, the spring, is far from the photographer's location. The sculptures and spring are small objects, and some of the sculptures are hidden. If people want to have clear view of them, they will probably stand close to the spring and sculptures and take photos. This shows an important issue. Within the visible area for the specific object, each different position has a different role. It is possible to find the 'favourite' area within visible area where people take photos for this specific object by looking at clustering of VGI data. We can consider this factor as slight regional difference and it's possible to use distance weights to estimate this effect.

In the EXIF file, the location is for the GPS position of the camera, not the objective. But, on these kinds of photo sharing websites, the spatial reference is for the photo, not for camera. So these are obviously different. Then we can calculate the distance between two users' positions in order to derive the distance weight. As the Figure 5.17 shown, the brown line "d", the definition of distance is the Euclidean distance between the position of user's camera, the point A, and the GPS position of geo-tagged photo, the point B, which can obtain GPS information by checking photo's EXIF file.



After getting Euclidean distance between two points, the rule for distance weight can be set. Several rings with different radii could be draw and each circle would have a different weight value. For example, for the first ring, the red one, with radius 10 meters, the value of distance weight could be 100%, so that this VGI point's value would be 2 calculated from the formula: value= 1\*(100+distance weight) %. In the second ring, the blue one, the value of weight could be 75%, using the same formula as above, this means the point's value would be 1.75; the rest could be 50%; others without distance weight would be 0. According to this rule, the value of C would be 1.5 and value of D is 1. Finally, calculating total value for each polygon, the result can show the different popular level for Hot Spot detection. In Figure 5.17, the ObjectM gets new value: 2.5 and it is higher than ObjectN: 2. It could be said the ObjectM could be the mainly objective at this view place.

The benefit of this approach is that is it adds the small regional differences into Clustering Analysis algorithm. That is, if people stood close to this viewpoint and took photos, most of them should have done so for the same objective. This could be considered as spatial relationship between each user. In other words, they could discuss the effect of Direction and try to add this parameter into algorithm. It could let algorithm be more flexible and reasonable.

#### 5.4.2.3. New Defined Boundary

While dealing with the VGI data, some problems arise. The general problem is VGI data is very difficult to control for quality, because the contributors are not in the same level of skill or background. Some of them may not have any spatial training so that they don't have any "specific feeling" for why they put the photo on the map and then they tag photos on the map just following the system suggestion. Other contributors may be experts so they may tag photos carefully. They might also think this photo is for the specific part of the building so it should be put on the specific place within the building polygon. Due to differences of users, the quality of VGI is also a problem.

While collecting data, we find a common problem is the error caused by map scale and point size. Under the smaller map scale, for example, 1:100,000, the users believe they put the points into the polygon, as Figure 5.18 show. In fact, zooming in, under bigger map scale, looked like Figure 5.19; the points are outside the polygon.



To solve this kind of problem, there are some options. Firstly, we could use map scale and point size to estimate the buffer radius for each polygon. But this seems not a good idea, because while different users operate the map to check, the map scale for the same building might not be the same value. Finally, there

is another interesting option to be discussed here. In most cases, the photos which are shared on websites have a tag and a title. Users use words to describe the photos. If we mine the context of the description or the tag, the title and then match to the spatial location, we could draw a new boundary for each objective. For example, if checking the photo map at Flickr for Kölner Dom, you will find that there are a lot of photos on the square around the Dom. And most of them are for the Kölner Dom. After collecting them, and getting the location of each point, we can use spatial geo-processing to draw a new polygon from the point set. This new boundary of each polygon is created from information supplied by the wide users around the world. So it's reasonable.

This approach is interesting, but more study and discussion is needed. For example, should all of points with tag content "Kölner Dom" be used? Actually, the answer should be no, because some of points are far from the building. If we still insist on using all of them, the new polygon will cover other irrelevant polygons. So how to choose the points for analysis is a big issue.

#### 5.4.3. Extended Issues

#### 5.4.3.1. Time cycle

It's necessary to mention the "time" issue again. Although we do not account for the time issue in this research, it's still an important point for collecting VGI data. In different time intervals, the VGI data possibly can represent the true view of the world. The time interval could be four seasons, the months, weekday and weekend, or the days. For example, at Enschede, Tuesday and Saturday have the Open Market at the central square, but in other times, the central square has nothing.

It is a common phenomenon in our daily life. In different times, the appearance of the world is different. But, in general, this kind of change is possibly temporal clustering. For small time intervals such as daytime and night, the appearance of the city is different. In the daytime, there are markets, but no pubs or clubs are open until 4 or 5pm; at night, all pubs and clubs are opening and attract people. So people will visit pub and club at night, not in the daytime. This means a city changes its appearance and people have different behaviours over a day. As a result, the photos for different time periods will not be alike. This phenomenon repeats day after day, just like a cycle. So we call it a "Time Cycle". A daily cycle could be the blue line in Figure 5.20.



For bigger time cycles, the unit of time interval could be day. The special events are held on regular days and at the same place. So this place changes its appearance regularly. The Enschede Open Market is a good case which we mentioned before. It's Time cycle is as Figure 5.20 shows, the green line. Furthermore, the time cycle could be months or seasons. For example, the National Park has beautiful natural scenery in spring and summer and then people come here to enjoy the natural environment and may visit the museum. In winter, there is little to see in terms of natural environment, no animals, and few green trees. So if people visit the park, they might prefer to spend time on museum and take more photos for paintings, sculptures or buildings. The different environments can make people take different photos.

Figure 5.20 shows the different time cycles and the time intervals from half days to weeks. From these different cycles, there are some interesting things which could respond to human possible behaviours. For example, the Open Market event of Enschede city isn't easy to find in a half day time cycle, but it appears in a weekly time cycle. Through appropriate analysis and estimation, time cycles could be used for flexibly estimating the content within the photo taken by people. If the appropriate time cycle could be found, the improved method could use spatial-temporal clustering analysis to do a more precise analysis. And the result will be possibly better. On the other hand, this kind of phenomenon could also be used in other researches, for example, tourism management or clustering of Human behaviour.

#### 5.4.3.2. Web Service Diagram

In this section, we will discuss how translate our core method of geo-tagging suggestion into web service. Before creating diagram of web service, the issue: "visualization of spatial data" is essential. The final production of this process is the spatial data. To let users easily realize the spatial data, we need to create a good map to pass the explicit information from the relative spatial data.

There are several common mapping methods to build the good statistic map, such as "Chorochromatic map", Choropleth map", or "Isoline map". In our case, the output statistic data for each polygon is the percentage. Percentage should be categorized to "Relative Ration" data. So the "Choropleth map" is a good way to deal with this kind of data. The major character of "Choropleth map" is using Value to identify order between classes. The darker grey value is usually used for more intensive area. Following this principle, we can use this kind of mapping method to create a good map and display on user's browser.

In recent years, "Virtual Reality" is a specific skill in visualization and broadly discussed. The "Virtual Reality" builds the view in 3D mind. Although 2D map is useful to transmit the information to map reader, the 3D map seems attract people's more attention. The Website, CEBRA(CEBRA), offers the projects to show how it works in urban design. It will be the interesting trend and also possibly be used in map to pass information to map reader. If we can combine VR skill and principle of statistic map, this kind of new application map will be interesting. For example, while building is belonging to higher ration class, the designer can give the building a flashing and colourful border to emphases its importance. At the same time, it also provides the 3D environment. While users put their contribution on the 3D building appearance, it could probably be used to store user's contribution in 3D format.

While using the traditional mapping ideas, the technology development could offer a great help to make the map. The map designers can combine new technology with traditional mapping principle to create a more "technological" map. For example, comparison animation map with Space-time Cube, the technology provides a good ability on cube environment, for example, rotating cube to look at the data, and it can overcome some problem of mapping. Designers can build other tools and show on the map to help user read the information passed from map easily such as spatial query table which can display the relative data on the screen. It would be a good idea to think can we use advanced technology in mapping design. In our case, for example, it's possible that we can put on the 3D building surface, which represented the Hot Spot and could be the major objective in user's photo, to attract users' attention and help them operate the geo-tagging suggestion system.

In conclusion, visualization of spatial data is an important issue for passing the information from map to users. Considering user's background as an important point, especially in VGI issue, it's necessary to find a balanced way between map and technology to build a good map to let wide users use it.

The final production which is a map with relative ration data or buildings' attribute should be displayed on the user's screen. How we do this work? The Web service which could offer this geo-tagging suggestion function should be built. Not all of users have good knowledge in technology and might not own good device to run GIS software. Especially, if they only use mobile device to do run geo-tagging suggestion function, the mobile device almost doesn't have a good ability to implement the process. It means while we design the web service diagram, clients should be considered as the "Thin Client" which is lack of capacity for running calculation on their own devices. The Figure 5.22 shows our major web service diagram.

According our geo-tagging suggestion process, the necessary data are: DSM, Building/Landmark spatial layer and VGI data. First of all, the DSM is the most difficult one. Where can we find the free DMS to use is a big challenge. Fortunately, there are several websites and organizations had done relative work. The GoogleMap cooperates with other Internet community and specific organizations such as govern departments and education institutes to make a 3D spatial data for building and puts on the GoogleMap, as Figure5.21 shown. Experts can use other drawing software, like AutoCAD, to draw building structure in 3D and then put it on the GoogleMap. Until now, there are many famous place own this kind of data liked Paris and Amsterdam. Combining 3D building structure with DEM, the DSM data could be extracted. In other case, the OpenStreetMap also has the relative work for 3D data and the project is called "osm-3D" (University of Heidelberg 2010). The major contribution of "osm-3D" is trying to build 3D spatial data infrastructure for Europe. Also, NASA and other international organizations had released the free DEM data. Generally, the DSM data could be generated from the relative Internet resources.



Figure 5.21 3D building appearance with GoogleMap



For building layer, several websites, like GoogleMap, YahooMap and OpenStreetMap, can offer the possible using layer. And the good news is they keep uploading and revising by a big number of users. As Goodchild mentioned at 2007, the VGI can offer a lot of useful function, especially for earth observation, people try to compile our world map and will give a map with more detail. So it's credible that the detailed layer is also available and free. Moreover, lots of internet communities and organizations try to develop open source for GIS software, such as QGIS. Although the ArcGIS is used in this research, it's possible to change it by using other free GIS software. Also, it could be combined with other free programming and statistical software which contains spatial principle, like Python and R to build complete system.

Now, we know all necessary elements to build the web service. The next step is to contact with camera manufactures and ask them to release the relative detail of camera lens and CCD. If the link can be built, after users upload their photo, the system can read the EXIF file and search the relative indexes to improve formula for calculating view angle and other essential parameters for viewshed analysis. Figure 5.21 shows the major geo-functions, resources and links. It's necessary to make an introduction for several terminologies for web technology. They are "Web Browser", "Thin Client", "API", "User Interface", "GML", "WFS" and "WMS".

The "Web Browser" is the browser software which provides user ability to link service and web page and view its content. XML is the standard format to let user pass their data files on Web. The "GML", Geography Markup Language, used XML grammar is a standard format which developed by OGC members, but contains geospatial information. The "WFS", Web Feature Service, allows users to update and obtain geo-data from multiple services. Web Map Service, abbreviated as WMS, provides the service to display map and data feature.

While talking about "Thin Client", the opposite terminology is "Thick Client". "Thin Client" means the client owns limited ability to implement software program and the calculation. So it's only asking client to have a platform to display the result which might be passed by server. "Thick Client" is different. The major difference between them is the device of "Thick Client" should have the ability to run the calculation or the specific application software such as ArcGIS. It means the requirement of device for "Thick Client" is higher than "Thin Client".

The major work of "User Interface" can let users operate functions easily. In the "User Interface", engineers can put on the essential tools such as "login" and "upload photo" to help users work easily. After users input something on the "User Interface", the server can pass this data to other software or system to ask for response. Between different systems and software, the "Application Programming Interface" (API) offers the communication way to both of them. The "API" is the interface between different software programs so that through setting rule of API, the software program can access this data and requirement and make response to the asker.

The first step for building web service is we need to provide a custom service to implement the geofunctions included viewshed analysis and clustering analysis and create our own web page with the appropriate "User Interface". Users will use their web browser to connect our web page and use the "User Interface" to upload their photos. After this operation, our server can get the photo and personal information and call specific software to read EXIF file. Finally, connecting to relative manufacture's resource, the server can calculate the necessary parameters for viewshed analysis and store in "GML" format. At this step, the GML file will contain the "User ID", "Coordinate", "View angle", "Spatial resolution" and "Personal detail". Using "Coordinate" to grab necessary "building layer" and "DSM" data through "WFS", our custom server can run viewshed analysis and produce the view area with spatial information liked the name of building. According to the spatial relationship and topology, the server will extract the VGI data from other servers and make the clustering analysis. The final step is visualization of our spatial data came from clustering analysis in good way and use "WMS" principle to show the map.

The map will display on users' browser and let users keep operating their geo-tagging work. The key points of this web service are:

- 1). Users are considered as the "Thin Client".
- 2). Our used data is spatial data.
- 3). It should consider using OGC standard being an important issue.

To build the web service to provide geo-tagging suggestion system, the designer should consider these three points as essential issues.

# 6. CONCLUSION

#### 6.1. Summary of the research

According to benefit of technological development, the digital camera equipped with GPS will be a major product in future. It means it's not hard work to get the location of user. In that case, what is the next possible information which could be mined from user? We believe the VGI could be used to come with more meaningful information than current approaches. To achieve this purpose, the good idea is trying to let user do more work for realizing and compiling their contribution. In geo- tagging application, this means letting a user to locate the objective, as well as where they took the picture. After that, the researcher can use this information to better study behaviour. While we talk about VGI, the important issue is that the user doesn't like to spend much time on this, because they can't get any benefit from it. The researcher should create a good environment to help users, while they operate the application. This is the main purpose of our research. We want to provide good assistance for users to help them quickly identify the major objective in their photos. To achieve our purpose, several questions should be answered.

1. What is tourist interesting for taking pictures? And why?

Generally, the object which is tourists took photos of should be an interesting thing such as a special building. It could be said these kinds of objects are unusual for users. This is discussed in chapter3.

2. What kind of functions can be used for detecting visible area? Which objects are seen?

GIS have many functions for spatial analysis. To solve this particular question,"viewshed analysis" can do this work well. Viewshed analysis is one of the common functions in GIS software. The necessary data for viewshed analysis is a viewpoint location and a DSM. Especially, in urban areas, DSM data is important. In this research, the major work of viewshed analysis is to extract the area visible by the camera; and also we make the process to connect to a GIS database to attach the spatial data such as Building Name or ID for visible objects in the area. In Chapter 4, we discuss the method developed to do this and the final output of this viewshed analysis process.

3. How can EXIF be used? What attributes of EXIF are useful?

EXIF is camera metadata. In our case, the 'true' view is that of the camera, not of human eyes. This means we should calculate the viewshed for the camera location. The two attributes of camera, "Focal Length" and "Image Dimension", can decide the field of view. Using EXIF, the information of both of these is available. We can use these for viewpoint attributes to get more accurate viewshed analysis. The discussion of EXIF fields is made in Chapter 3.

4. What is the role of tagged photos? What kinds of information can be provided by tagged photos? How can process use it to extract famous landmarks?

The tagged photo is the source of VGI data. The VGI data is an interesting topic, because the information from VGI data is useful to many fields. In our research, we use tagged photo to analyze the popular objects for different users that took photos from the same place. In the future, the volume of VGI data will only increase. This means, through appropriate analysis such as clustering analysis, the VGI data may eventually reflect our real world. The method of clustering analysis is shown in Chapter 4.

5. How does the resulting system work? Is the result good or bad? What can be the possible factors to improve it in future?

The results of our experiment show this process for geo-tagging suggestion is effective. But there are still several elements to improve this process. The advanced algorithm for clustering analysis is essential. In Chapter 5, we discuss the result and discuss the further extensions. The "Ranking Value" included "Distance Weight" and "Social Network System" is important issue for future study.

According to the research questions above, we developed the workflow as Figure 6.1 shows. It also shows the relation between chapters and components of the research.



This research provides new idea to give suggestion for geo-tagging application. We use geo-processing --Viewshed Analysis, combined with EXIF data -- to deal with visibility problem and this result would be better than using geo-distance, because only visible objects would display on the screen. And essential parameters for viewshed analysis can be calculated from the EXIF file which is automatically attached to photos. This means the system doesn't need to do extra work for searching the additional source.

Clustering Analysis of VGI data offers an important improvement. That is the result can be flexible and case by case. Here we employ Frequency as a Clustering measure. Frequency is a simple function, under well defined spatial relationship between each polygon and point, it can reflect clustering in easily way. At a given viewpoint, the system can extract the more popular objective which can attract people's attention and let them take photos.

Previous researches, has collected VGI data and analyzed it to predict the possible objective in user photo. These methods are based on analysis of annotation of photo, geo-distance and image content-based algorithms. Despite geo-distance and GPS, it could be said that the spatial principle is useless in this method. In the method proposed in this research, it's different. We make a complete definition for spatial relationship. For example, we define the boundary of visible buildings and then combine their polygons with spatial coordinates extracted from geo-tagged photos using topology principles "Contain" and "Intersect", to calculate frequencies. This method is more effective than previous ones.

Because we identify the visible building in photos so that the attribute of the building, such as Name and address, it is possible to build a semantic algorithm to extract other photos which don't have spatial information, but with title of description shared on the websites. For example, "Enschede City Hall" could make two rules to search photos: "Enschede" and "City Hall". If the title or description of photo includes both of "Enschede" and "City Hall", this photo should be collected for building" Enschede City Hall". This shows ways to extend our process using growing collection of VGI data..

#### 6.2. Futhur Work

For future work, we list several things which are related to our study, but not done in this research, including advanced algorithms, web service building and other interesting applications.

#### • Advanced algorithm for Ranking Value to analyze VGI data

This extended study for *Ranking Value* includes two major parts: Distance Weight Value and Personal Factors. Through Distance Weight Value, considering small spatial differences as an important issue, the system can try to add the different spatial relationships between each different small region and objects. Further, based on Social Network System, the Personal Factors can be used to interpret VGI data in more meaningful ways. All of these can help system improve the method developed in this work.

#### Web service building

We provide a core idea of geo-tagging suggestion system. This could be implemented by building a web service, as discussed in Chapter 5. This web service can provide free functions because all necessary data used in the process are possibly free. Future work should build a prototype for testing and development.

#### • Other related applications

When discussing 3D technology and visualization, we also find an interesting idea. At this moment, researchers are trying to build a 3D way to effectively store user's contributions. They hope 3D ways can replace of 2D maps. But this seems not be a success yet. There are many reasons, but the basic one is users do not like to spend so much time to compile this kind of information, especially, while this kind of information is not related to them. If the system has a 3D environment without any help, users have to rotate and try to identify the objective. Additionally, if the photo contains only a small part of a building, the work will become more difficult. Most users don't have enough patience. If they lose their patience, they may locate the photo on the map without concern for accuracy. Obviously, the quality of this result will be bad.

Our process can detect the most likely objective for users and it can save them work. While the system shows the 3D building appearance, users only need to focus on the detail of the specific building which is the Hot Spot of the result and put the photo on the 3D building appearance. On the other hand, the

system can show photos on the 3D building appearance. Even more, it's possible to display "small scale" Hot Spot within specific building, called "second Clustering Analysis" in three dimensions, if we have enough VGI data. In this way, users are the compilers and make a story for the building. This kind of feeling will probably be a motivation to let users keep contributing. Finally, the photos with x, y and z information can make a *3D photo wall* which could make this application more attractive and interesting.

Under limitation of time, it has not been possible to build and implement a complete geo-tagging suggestion application. Our research is the initial research and builds the core methods for geo-tagging suggestion. We combine logical and fixed algorithm with a flexible methods of data collection and analysis to enable further development of the geo-tagging suggestion application.

## LIST OF REFERENCES

(2010) Sony introduces Digital Camera with GPS and Compass.

."CEBRA."from

http://www.cebra.eu/index.php?option=com\_content&view=frontpage&Itemid=1&lang=en.

University of Heidelberg, D. o. G. (2010). "OpenStreetMap." from <u>http://www.osm-3d.org/home.en.htm</u>.

"Flickr." from <u>http://www.flickr.com/</u>.

"Photobucket." from <u>http://photobucket.com/</u>.

"Virtual Tourist." from http://www.virtualtourist.com/.

"Visit Enschede." from http://www.visitenschede.nl/.

"webshots." from http://www.webshots.com/.

Abela, J. C. and E. M. Aguilar (2007). "Panoramio." from http://www.panoramio.com/.

Wiki, E. (2001). "EXIF." from http://en.wikipedia.org/wiki/EXIF.

Python.org. "Python." from <u>http://www.python.org/</u>.

Bartie, P., F. Reitsma, et al. (2010). "Advancing visibility modelling algorithms for urban environments." <u>Computers, Environment and Urban Systems</u> **34**(6): 518-531.

Brimicombe, A. J. (2007). "A dual approach to cluster discovery in point event data sets." <u>Computers,</u> <u>Environment and Urban Systems</u> **31**(1): 4-18.

Engel, J. and J. Dollner "Approaches Towards Visual 3D Analysis for Digital Landscapes and Its Applications."

Golder, S. A. and B. A. Huberman (2005). "The Structure of Collaborative Tagging Systems." <u>CoRR</u> **abs/cs/0508082**.

Goodchild, M. (2007). "Citizens as sensors: the world of volunteered geography." <u>GeoJournal</u> **69**(4): 211-221.

Goodchild, M. (2009). "NeoGeography and the nature of geographic expertise." <u>Journal of Location</u> <u>Based Services</u> **3**(2): 82-96. Iwasaki, K., K. Yamazawa, et al. (2005). <u>An Indexing System for Photos Based on Shooting Position</u> <u>and Orientation with Geographic Database</u>. Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on.

Jürrens, E. H., A. Bröring, et al. (2009). A Human Sensor Web for Water Availability Monitoring. <u>In</u> <u>Proceedings of: OneSpace 2009 - 2nd International Workshop on Blending Physical and Digital Spaces</u> <u>on the Internet</u>.

JEITA (2002). Exchangeable image file format for digital still cameras: Exif Version 2.2.

Lawson, A. (2010). "Hotspot detection and clustering: ways and means." <u>Environmental and</u> <u>Ecological Statistics</u> **17**(2): 231-245.

Marlow, C., M. Naaman, et al. (2006). Position paper, tagging, taxonomy, flickr, article, toread. In <u>Collaborative Web Tagging Workshop at WWW'06</u>: 31-40.

Moxley, E., J. Kleban, et al. (2008). Spirittagger: a geo-aware tag suggestion tool mined from flickr. <u>Proceeding of the 1st ACM international conference on Multimedia information retrieval</u>. Vancouver, British Columbia, Canada, ACM: 24-30.

Naaman, M., S. Harada, et al. (2004). <u>Context data in geo-referenced digital photo collections</u>. MULTIMEDIA '04: Proceedings of the 12th annual ACM international conference on Multimedia, New York, NY, USA, ACM.

Smith, M. J. d., M. F. Goodchild, et al. (2007). <u>Geospatial Analysis: A comprehensive guide</u>, Troudbador Publishing Ltd.

Viana, W., J. Bringel Filho, et al. (2008). "PhotoMap: from location and time to context-aware photo annotations." Journal of Location Based Services **2**(3): 211-235.

Vosselman, G. (2008). ANALYSIS OF PLANIMETRIC ACCURACY OF AIRBORNE LASER SCANNING SURVEYS. <u>XXI ISPRS Congress: Silk Road for Information From Imagery, Beijing, China</u>. **37(part B3a):** 99-104.

Yang, L., J. Johnstone, et al. (2010). "Ranking canonical views for tourist attractions." <u>Multimedia</u> <u>Tools and Applications</u> **46**(2): 573-589.

Wang, Y.-h. (2008). Augment GPS Route Guidance with LED and Auditory Display into Motorcycle Helmet Design. <u>Industrial Design</u>. Tainan City, Taiwan, National Cheng Kung University **Master**.

#### Appendix

Model built in ArcMap9.3

