Grouping of Tourism Objects Using Geotagged Photo with Hierarchical Clustering Method in Bantul and Sleman

Fenny Lusiana*, Jap Tji Beng, Wasino
Information Systems Department, Faculty of Information Technology Universitas Tarumanagara

* fenny.825160026@stu.untar.ac.id

Abstract. Bantul and Sleman regency has many tourism object that can attract tourists. However, tourism objects in this area are still poorly known by many people. When tourists visit tourism objects, they like to take pictures with a smartphone. The geotagging process can be done when taking photos using a smartphone. This data consists of latitude and longitude coordinates, though they can also include altitude, distance, and place names. The purpose of this study is to visualize photos of tourism objects based on location and provide detailed information related to the geotagged photo. The method used in this study is agglomerative hierarchical clustering. Grouping photos based on a hierarchy that starts from the village, district and regency. The results of this study are an application that can help users to receive information about tourism objects and display the location of photos with google maps. Users only need to upload a geotagged photo, to get this information. This application testing will be carried out by visitors of tourism objects in the Bantul and Sleman regency.

Keywords : Agglomerative, Geotagging, Hierarchical, Tourism

1. Introduction
There is great advancement of camera phone and digital camera to share media on web services and as well as social media, on past few years[1]. Nowadays, travelers share their travel experience like photo on these web services [2]. The photos actually have information such as latitude and longitude, and it’s called metadata[3]. Therefore, the system to be developed in this research is based on a website using the php and mysql. The system can provide information about tourist objects based on geotagged photos in Bantul and Sleman Regency. It can display some information about tourist objects using google maps. In addition, this application will also use Hierarchical Clustering to be able to group geotagged photos based on the location of the photo. The goal of this paper is to study the geo-tagged photo collection of tourist objects in Bantul and Sleman Regencies.

The main contributions of this paper are: 1) We present an explanation of the process of the hierarchical clustering method when determining the location of tourist objects. 2) From the user-provided photo collection, we show how to group photos using the hierarchical clustering method. 3) Easily can find out location, where your photo is taken.

2. Method and materials
2.1 Data Collecting Method
Fact-finding method is used to collect all data that needed. Fact-finding method is a formal process using techniques such as interviews, questionnaires and observation to gather facts about the system, requirements, and preferences [4]. Survey was held on several tourism objects in Bantul and Sleman. Data collection and system requirements are carried out through direct observation of tourism objects to the Bantul and Sleman Regency. Tourist data that is entered and processed into the system are primary data and secondary data. Primary data is data from the researchers themselves, namely data from direct observation [5]. In addition there are some secondary data that is obtained from other sources. [6]
2.2 System development method

Hierarchical clustering is one of the methods from data mining for grouping data objects and analysis which seeks to build a hierarchy. Hierarchical clustering methods can be further classified into agglomerative and divisive hierarchical clustering [7]. We will use agglomerative hierarchical clustering to develop this system. The result of agglomerative clustering is shown by a diagram called a ‘constituency tree’ or ‘dendrogram’ [8].

3. Literature Review

Bantul and Sleman become the districts with the most tourism villages [9]. Tourist always take a picture when they visited a tourism object [10]. In past few years, we can take photo with our smartphone and the photo will have a metadata [11]. This metadata consist of information like latitude and longitude [12]. Geotagged photo is the process of embedding a digital photo with longitude, latitude and even altitude data [13].

4. Result and Discussion

This website application uses MySql database to store information about tourism object in Bantul and Sleman Regency. In this website, the user images are used for finding the location. The input image will has metadata consist of latitude and longitude. And system compared it with latitude and longitude of tourism objects in the database. If the latitude and longitude of tourism object is not present in database
system will alert message to user. If image latitude and longitude is found in database, our proposed system will visualize image location with map. Testing is done by uploading 15 geotagged sample photos on the system, and the location will appear on google maps. To get longitude and latitude, it is done by adding up the degree value and the minute value that has been divided by 60. Then the results will be added to the second value that has been divided by 3600. The results of these calculations will obtain latitude and longitude as in Figure 3.

This longitude and latitude from geotagged photo will be calculated by haversine formula to compare the latitude and longitude tourism objects in database as the central coordinate. Haversine formula is a method of knowing the distance between two points by taking into account that the earth is not a plane but is a plane of a degree of curvature and has a radius of 6,367.45 km [14]. This formula will used in the program coding, as can be seen in Figure 4.

The records coordinates of tourism objects in database that already stored is displayed using Table 1.
5. Conclusion and Future Works
The advantages of this system are; 1) geotagging allows users to visualize tourist object photo in new and interesting ways. 2) Easily can find out location, where your photo is taken. 3) This system is very useful to recognize the real location of tourism object in Bantul and Sleman Regency. From trials conducted by uploading 15 photos, the accuracy level is 85%. This can be seen from the results table that can be seen in Figure 5.

| No. | Coordinate Central | Result     |
|-----|--------------------|------------|
|     | Latitude           | Longitude  |             |
| 1   | -7754.4646576      | 11049.571990967 | F008,F011  |
| 2   | -7742.1388261      | 11050.452423096 | F002      |
| 3   | -7761.49988174     | 11047.456648682 | F003,F007 |
| 4   | -7784.11102295     | 11051.127624512 | F009      |
| 5   | -7769.444446564    | 11048.532104492 | F001      |
| 6   | -7766.72202230     | 11047.25870361  | F10       |
| 7   | -7.9225556         | 11044.721497222 | F010,F012 |
| 8   | -7.9093333         | 11042.4444444   | F015      |
| 9   | -7930.83333        | 11043.1111111   | F014      |
| 10  | -7947.02778        | 11.046.805.556  | -          |

Figure 5. The Result of Hierarchical Clustering

6. References
[1] Kádár, Bálint & Gede, Mátéyás. (2013). Where Do Tourists Go? Visualizing and Analysing the Spatial Distribution of Geotagged Photography. Cartographica The International Journal for Geographic Information and Geovisualization. 48. 78-88. 10.3138/carto.48.2.1839.

[2] Ráthonyi, G., & Várallyai, L. (2011, September). Web 2.0 and Tourism. Proceedings of the International Conference on Information and Communication Technologies (pp. 27-39).

[3] Chaitali, Sapna & Anuskha. (2017). Mining Geo-tagged Images for Location Inference. International Research Journal of Engineering and Technology (IRJET). Volume 04 Issue: 04. e-ISSN: 2395-0056.

[4] Connolly, Thomas and Begg, Carolyn. 2010. Database System: A Practical Approach to Design Implementation and Management. 5th Ed. Pearson Education.

[5] Mills, J. A., Teplitsky, C., Arroyo, B., Charmantier, A., Becker, P. H., Birkhead, T. R., ... & Bushuev, (2015). Archiving primary data: solutions for long-term studies, 30(10), 581-589.

[6] Mueller, C. E., & Hart, C. O. (2010). Effective use of secondary data analysis in gifted education research: Opportunities and challenges. Gifted Children, 4(2), 3.

[7] Aljumily. (2016). Agglomerative Hierarchical Clustering: An Introduction to Essentials. (1) Proximity Coefficients and Creation of a Vector-Distance Matrix and (2) Construction of the Hierarchical Tree and a Selection of Methods. Global Journals Inc. Online ISSN: 2249-460x

[8] Rani and Rohil. (2013). A Study of Hierarchical Clustering Algorithm. International Journal of Information and Computation Technology. Volume 3, Number 11, pp. 1225-1232

[9] Vitasurya, Hardirman & Sari. (2018). Geographical Conditions and Cultural Tradition as Determinants in sustaining tourist village program, Case Study Tourism Villages in Yogyakarta. International Journal on The Academic Research Community Publication. ISSN online: 2537-0162
[10] Terzidou, M., Stylidis, D., & Terzidis, K. (2018). The role of visual media in religious tourists' destination image, choice, and on-site experience: The case of Tinos, Greece. *Journal of Travel & Tourism Marketing*, 35(3), 306-319.

[11] Shen, R., Yan, H., Sun, Q., & Li, X. (2019). Study on the Behaviors of Inbound Tourists based on Geotagged Photo Metadata-A Case of the City Group in the Upper Reach of the Yellow River in China. *Abstracts of the ICA, 1*.

[12] Gavrilis, D., Makri, D. N., Papachristopoulos, L., Angelis, S., Kravvaritis, K., Papatheodorou, C., & Constantopoulos, P. (2015, September). Measuring quality in metadata repositories. In *International Conference on Theory and Practice of Digital Libraries* (pp. 56-67). Springer, Cham.

[13] Lee, C. M., & Thomas, J. J. (2017, November). Travel Route Recommendation Based on Geotagged Photo Metadata. In *International Visual Informatics Conference* (pp. 297-308). Springer, Cham.

[14] Hartanto, Siahaan, & Fitriani. (2017). Haversine Method in Looking for the Nearest Masjid. *International Journal of Recent Trends Engineering And Research. ISSN : 2455-1457*