Analysis of the Difference of Movement Trajectory by Residents and Tourists using Geotagged Tweet

Shintaro Fujii
Faculty of System Design
Tokyo Metropolitan University
Hino-shi, Tokyo
Email: fujii-shintarou@ed.tmu.ac.jp

Masaharu Hirota
Faculty of Informatics
Okayama University of Science
Okayama-shi, Okayama
Email: hirota@mis.ous.ac.jp

Daiju Kato
WingArc1st Inc.
Shibuya-ku, Tokyo
Email: kato.d@wingarc.com

Tetsuya Araki
Faculty of System Design
Tokyo Metropolitan University
Hino-shi, Tokyo
Email: araki@tmu.ac.jp

Masaki Endo
Division of Core Manufacturing
Polytechnic University
Kodaira-shi, Tokyo
Email: endou@uitec.ac.jp

Hiroshi Ishikawa
Faculty of System Design
Tokyo Metropolitan University
Hino-shi, Tokyo
Email: ishikawa-hiroshi@tmu.ac.jp

***Abstract*—In recent years, tourism industry has been drawing attention in various countries. Tourists are on an increasing trend all over the world, and it is estimated that the value exceed 1.8 billion in 2030. Consumer behavior by tourists brings high economic effects to many industries such as transportation, lodging, manufacturing. Therefore the increase in tourists is an important issue for governments and tourism agency. According to a survey by tourism agency, 60% of foreign tourists visiting Japan are repeaters. In other words, it is considered important to increase repeaters to increase tourists. Compared with the first tourist, there is a need for repeaters to visit sightseeing spot that many residents visit and tourists do not know. One must analyze data of resident to discover these sightseeing spots. Nevertheless, most studies conducted to extract hotspots (areas where many photographs are taken) and recommend sightseeing routes using movement trajectories do not consider user attributes. Therefore, by considering user attributes, this study was conducted to extract hotspots that many residents visit but are not know to tourists. Additionally, we extract movement trajectories from residents and tourists to ascertain differences in sightseeing areas and to analyze them by visualizing those results on a map.**

***Keywords*—Tourism; Geospatial analysis; Cities and towns.**

I. INTRODUCTION

The number of tourists worldwide is increasing every year. It is predicted that they will be 1.8 billion in 2030 [1]. Tourism occupies an important position as a key industry in many countries. Consumption activities related to tourism positively affect industries such as transportation, lodging, and manufacturing. Therefore, increasing the number of tourists represents an important issue for governments and companies. In Japan, where the Tokyo Olympics Games are to be held in 2020, the Japanese Government and companies are actively conducting activities to increase foreign visitors to Japan, such as the Visit Japan Campaign [2] and promotion [3]. As a result, foreign visitors Japan have increased year by year, reaching a record high of 28.69 million in 2017 [4]. It is necessary to analyze tourists data to increase tourism. According to a survey by the Tourism Agency, 60% of foreign visitors to Japan are repeaters [5]. We consider that increasing repeaters is important to increase tourists. The repeater described here refers to a person who visits a specific sightseeing area more than once. The need exists for repeaters to experience more local culture and visit local spots more than first-time tourists do [6]. The local spot described here refers to a place that many residents know: not famous sightseeing spots that many tourists visit. For this study, we define a local spot as a hotspot that many residents visit but few tourists visit. Discovering local spots is important to increase tourists. Therefore, we extract hotspots that many residents visit but few tourists visit.

Additionally, several needs exist for tourism agencies as tourists increase. It is necessary to ascertain the movements and interests of tourists in sightseeing areas. Tourism agencies perform more effective PR methods for sightseeing areas and recommend sightseeing plans to satisfy tourist needs and attract more tourists in sightseeing areas by knowing the movements and interests of tourists in sightseeing areas. We extract hotspots and movement trajectories of tourists to analyze the movements and interesting spots of tourists in sightseeing areas. Therefore, one must extract the respective hotspots and movement trajectories of residents and tourists to satisfy the needs of tourism agencies and tourists.

Several studies have been conducted to extract hotspots and recommend sightseeing routes using movement trajectories [7]–[9]. Nevertheless, these studies do not consider user attributes. Many tourists visit famous sightseeing spots in sightseeing areas and post many contents from those locations. Tourist contents continue to increase in sightseeing areas year by year. By contrast, contents of residents for sightseeing areas have not changed much numerically. These indicate that more contents uploaded in sightseeing areas to social media site are posted by tourists than by residents. Therefore, when we do not consider user attributes and extract hotspots from their contents, it is difficult to extract local spots in sightseeing areas because tourists post the most contents in sightseeing

 ISBN: 978-1-61208-687-3

areas. Therefore, for this study, we extract hotspots of residents and tourists respectively using geotagged tweets to discover local spots. Furthermore, when classifying users, we adapt the method proposed in [10] because our research goal is discovering local spots and not user classification.

Actually, [11] and [12] are studies applying [7] and [8]. Also, [11] and [12] extract sightseeing spot by considering user attributes. For this study, in addition to discovering hotspots considering user attributes, we extract movement trajectories of residents and tourists. Considering user attributes, we combine hotspot and movement trajectories and thereby discover sightseeing routes that many residents use, but which tourists do not use. Discovering these routes contributed to recommendation of new sightseeing routes that many residents know, thereby relieving congestion in sightseeing areas. Therefore, for this study, we extract sightseeing routes that many residents use and tourists do not use by clustering and visualizing resident and tourist movement trajectories.

The structure of this paper is the following. In Section II, this report presents some related research efforts. In Section III, we describe our proposed method to discover local spots and to assess differences of residents' and tourists' movement trajectories. Section IV presents experiments and results obtained using the proposed method. Section V, presents discussion of the results of visualizing hotspots and movement trajectories of residents and tourists obtained using our proposed method. Section VI, we conclude this paper and describe avenues for future work.

II. RELATED WORK

In this section, we describe research related to this research.

A. *Hotspot*

Research on hotspot extraction is actively conducted using geotagged tweets and photographs posted on social media. Crandall et al. [7] proposed a method to discover popular spots using spatial clustering with large amounts of geotagged photographs and image features. Kisilevich et al. [8] proposed a method to discover hotspots using PDBSCAN: an improved DBSCAN algorithm. Yang et al. [13] proposed an algorithm to extract hotspots of various sizes: Self-Tuning Spectral Clustering. Lacerda et al. [14] extracted hotspots using geotag information and intersections of photograph orientations. Zhijun et al. [15] divided areas into grids and extracted and visualized geographical features in the grid using geotag information and text tags attached to photographs. Li et al. [16] proposed a method to classify Flickr [17] users as residents or tourists, calculate their relative proportions in each of the five cities, and compare them to ascertain and analyze differences in cities. Zhuang et al. [11] proposed a method to discover Anaba (sightseeing spots that are less well-known, but still worth visiting) using geotagged photos. They evaluate the scenery quality by considering both social appreciation and the contents of images shot around there. Van et al. [18] first extracted hotspots by clustering Flickr photographs. Then they analyzed Twitter [19] text and extracted areas of interest. Furthermore, they confirmed and investigated the places using data from Foursquare [20]. Zhuang et al. [12] proposed a method to discover obscure sightseeing spots that are less well-known, but which are still worth visiting. They aimed to overcome challenges that classical authority analysis based methods do not encounter: how to discover and rank spots based on popularity (obscurity level) and on scenery quality. For the present research, we extract hotspots of residents and tourists and discover local spots in sightseeing areas.

B. *Movement trajectory*

Actively conducted research efforts contribute to each industry by analyzing movement trajectories from geotagged data. Yuan et al. [21] proposed a method to discover areas of different functions in the city by combining taxi trajectory data and data of a person's area of interest obtained from social media. Nanni et al. [22] proposed a method to adapt density-based clustering algorithms to trajectory data based on the simple concept of distance between trajectories. Additionally, to improve trajectory clustering, they proposed an algorithm incorporating time information. Kori et al. [23] proposed a method to recommend sightseeing routes using user blogs to extract movement trajectories that are produced during sightseeing. Sun et al. [24] proposed a system that recommends the best sightseeing route for users using geotagged photographs that had been posted on Flickr. They defined the best sightseeing route recommendations as one for which many users visit and for which each landmark distance is close. Memon et al. [25] proposed a method to recommend sightseeing routes particularly addressing the posting times of geotagged photographs posted on Flickr. Garcia et al. [26] proposed a method to examine route generation and route customization and to analyze them to solve the tourist planning problem. They present an heuristic that is able to solve a tourist planning problem in real-time using public transportation information and the Time Dependent Team Orienteering Problem with Time Windows (TDTOPTW). Zhang et al. [27] proposed an efficient tourist route search system that not only recommends a route simply connecting several tourist spots, but which also recommends a route with beautiful scenic sights. Xin et al. [28] propose to leverage existing travel clues recovered from 20 million geo-tagged photographs to suggest customized travel route plans according to user preferences. For the present study, we extract movement trajectories of residents and tourists and discover sightseeing routes that many residents use and which many tourists do not use.

III. PROPOSED METHOD

In this section, we describe our proposed method to extract hotspots of residents and tourists and their respective movement trajectories in the sightseeing area. The procedure that is followed to accomplish the proposed method is the following.

1) We apply preprocessing.
2) We classify users as residents and tourists.

3) We cluster movement trajectories.
4) We visualize the hotspots and movement trajectories of residents and tourists.

For this study, we define users who post many tweets in specific sightseeing areas as residents, and users who post many tweets outside specific sightseeing areas as tourists.

A. Preprocessing

This section specifically explains how to obtain data and how to preprocess the data. From Twitter, we obtained tweets with annotated geo-tag information. At that time, we eliminated tweets posted from countries other than Japan. Next, we applied preprocessing to the tweets we obtained. We deleted tweets including auto-generated texts from other social media sites, replies, retweets and tweets by bots.

B. Classification of users

This section presents a description of methods used to characterize users as residents and tourists and methods of extracting a series of tweets within a specific sightseeing area. First, we sort the user tweets to arrange them in chronological order. Additionally, we calculate the proportion of tweets by latitude and longitude within a specific sightseeing area. Subsequently, we define specific sightseeing areas by latitude and longitude. Nozawa et al. [10] classified Twitter users as residents or tourists. Users who posted over 30% of tweets within a specific sightseeing area were inferred as residents, and were otherwise inferred as tourists. We apply this classification method to classify users as residents or tourists because our research is not aimed at user classification. Next, we extract tweets posted during a specified sightseeing period. We extract a series of tweets posted from the time tourists start tweeting within this area until they are out of range. For this research, a series of tweets within the area is called as a tourism tweet. Furthermore, for residents, we extract tourism tweets by classifying everyday tweets within a range. Through this process, we extract many movement trajectories suggested by tourism tweets posted by residents and tourists.

C. Clustering of movement trajectories

This section presents an explanation of a method to cluster tourism tweets extracted in Section III-B. The purpose of clustering is to clarify differences in movement trajectories between residents and tourists. First, we ascertain tourism tweets as those of residents or tourists. Subsequently, we classify them accordingly. Next, for each tourism tweet of residents and tourists, we extract the distance of each tourism tweet using Dynamic Time Warping (DTW) for every round. Then, we calculate the distance of all tweets included in tourism tweets. We adopted DTW in this study because the length of tourism tweets is different depending on the user. DTW allows duplication of correspondence between two time series and is applied to time series data of different lengths. We use this extracted distance to cluster tourism tweets using kmeans++. Kmeans that is non-hierarchical clustering depends on the initial value because the initial centroid is allocated as a random number. Therefore, we adopted kmeans++ in this study to avoid the problem of assigning the cluster to the one in which the kmeans method should not be frequently used as a cluster. The clustered movement trajectories show where the residents and tourists frequently move.

TABLE I. DESCRIPTION OF GRID COLOR-CODED INTO 7 COLORS.

Grid color	Difference in proportion of users between residents and tourists
Green	low order 2%
Blue	low order 2 ˜4%
Purple	low order 4 ˜6%
No color	other
Yellow	superior 6 ˜4%
Orange	superior 4 ˜2%
Red	superior 2%

D. Visualize hotspots and moving trajectories on the map

This section describes a method to discover local spot and sightseeing routes that many residents use and which many tourists do not use. We visualize hotspots based on the posting position of tweets to analyze areas where a user is interested in the sightseeing area. To analyze details of the visited places, we map areas into sixth-order meshes (125-meter square grids), which is the smallest grid size provided by the Geographical Survey Institute in Japan. We count the users in each cell. We define a threshold in the cell and a hotspot cell according to the proportion of the number of users.

To assess movement trajectories, we visualize the resident and tourist tourism tweets as clustered in Section III-C on the map. First, for all the clusters classified in Section III-C, we calculate the movement proportion of the user between the grids. Next, as a result of clustering, in clusters classified in the same sightseeing area, we calculate the difference of the movement proportion between the resident and the tourist grids. The one that exceeds the threshold is visualized.

IV. EXPERIMENT

In this section, we describe experiment conducted based on the proposed method.

A. Data set

We compiled and used a data set that was especially intended for this experiment. We obtained geotagged tweets for Twitter using Twitter API [29]. The data collection period was January 1, 2017 to December 31, 2017. The total number of data was 2,793,207. We used these data to classify users as residents or tourists. Data used for the visualization of hotspots and movement trajectory were tweets posted in Tokyo during April 1, 2017 – May 31, 2017. For that time, we deleted replies, retweets, tweets posted by bots and tweets that included auto-generated texts from other social media sites such as FourSquare. We consider these tweets are noise because we analyze hotspots and consider users' text. In addition, we deleted tourism tweets that were only single

 ISBN: 978-1-61208-687-3

tweets or tweets sent from the same place because we need to analyze the movement trajectory and extract consecutive tweets. Results show that resident tweets were 190,091, users as resident were 27,231, tourist tweets were 100,444, and users as tourist were 13,712. Tourism tweets classified using the proposed method were 14,582 for residents and 17,119 for tourists.

B. Clustering

This section presents a description of procedures used for clustering tourism tweets posted by residents and tourists after extraction using this proposed method. We used the elbow method to ascertain the optimal number of clusters because kmeans++ must be determined beforehand. The elbow method is widely used as a method for determining the optimum number of clusters. We adopted widely used methods because our goal is analysis of sightseeing areas. Results show that the number of clusters of resident movement trajectories was 34; that of tourist movement trajectories was 29. The movement proportion between the grids is calculated by dividing the number of movements between the grids by the total movement number for each cluster. We explain related details with Figure 1 as an example. We extract movements between grids when users tweet on different grids. In Figure 1, the respective grid movement numbers of residents and tourists are 100 and 200. This grid movement number is the result of clustering and is classified in the same area. The top two figures in Figure 1 show the number of residents and tourist movements in the grid. The numbers in parentheses represent the movement proportion. The difference between residents and tourists is calculated as shown in the figure below. We calculated the difference between residents and tourists and visualized the movement trajectory of the top 0.5%. Hotspots and movement trajectories of residents and tourists extracted using the proposed method are portrayed in Figure 2, Figure 3 and Figure 4. Figure 2 presents the difference between the proportion of residents and tourists in each grid around the Tokyo Skytree. In Figure 2, the relation between the grid color and the difference in proportion of users between residents and tourists is presented in Table I. As a result of the difference in proportion of users between residents and tourists, more than 90% of the grids existed at 25% or less. Therefore, we visualized result as shown in Table I. Additionally, if no user exists in the grid, then the grid itself is not displayed in Figures 2–4. The area around the Tokyo Skytree is a popular sightseeing area in Tokyo that many tourists visit.

V. DISCUSSION

In this section, we discuss the results presented in Section IV-B. First, we explain Figure 2. Figure 2 portray the area surrounding Tokyo Skytree. We show Figure 2, Figure 3 and Figure 4 in five areas to support several points of the discussion. We shall specifically discuss areas numbered as Area1, Area2, ..., Area5 and describe each area in Table II. From Figure 2, it is proven that many tourists visit Ueno

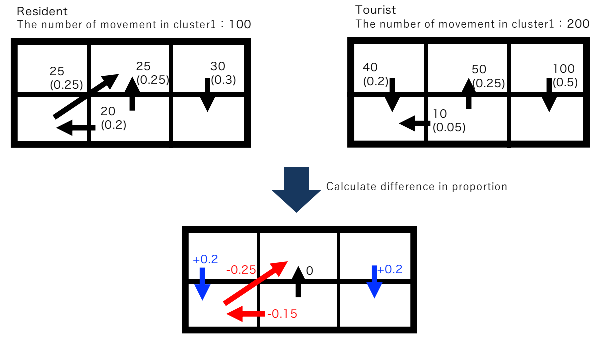

Figure 1. Movement trajectories of residents(Blue) and tourists(Red).

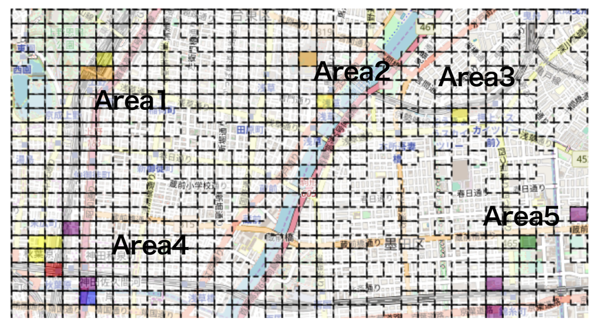

Figure 2. Difference between hotspots of resident and tourist around Tokyo Skytree.

in Area1, but fewer residents there. Many sightseeing spots exist around Ueno, such as Shinobazu-no-ike Pond and the Ueno Zoological Gardens. Especially at the Ueno Zoological Gardens, attendance has increased recently [30] because of the birth of a panda. Results demonstrate that it has become a popular sightseeing spot for tourists.

Therefore, we regard the area around Ueno as a sightseeing area of interest for tourists rather than residents. Conversely, many residents visit Kameido Temple and Kinshicho Park in Area5, but there appear to be few tourists among the users. Kameido Temple, located near the Tokyo Skytree in Area3, is a sightseeing spot where the main shrine and the Tokyo Skytree can be photographed together. In addition, because many wisteria flowers grow within its precincts, it is possible in the spring to take photographs of the Tokyo Skytree as well as wisteria flowers in the main shrine. A festival, called the Fuji Festival, is held there and is visited by many people. As Kinshicho Park is famous for cherry blossoms, many people visit in spring. Therefore, the possibility exists that these sightseeing spots are the local spot that is an object of this research.

Next, we discuss Figure 3. Many more tourists than residents move to Ueno in Area1, Asakusa in Area2, and Tokyo Skytree in Area3. The reason for this result is that many pamphlets and web sites have presented this area as a series of sightseeing areas. However Kameido Temple and Kinshicho

 ISBN: 978-1-61208-687-3

TABLE II. DESCRIPTION OF THE AREA AROUND SHIBUYA AND ASAKUSA.

Area	Area description
Area1	Near Ueno, with the Ueno Zoological Gardens and their well-known panda attraction
Area2	Near Asakusa, with its many temples such as Sensoji Temple
Area3	Near Tokyo Skytree
Area4	Near Akihabara, with its many famous electronics mass merchandisers and animation goods retailers
Area5	Near Kinshichou that has many taverns and restaurants
Area6	Near Shimokitazawa that is famous as a fashion like old clothes
Area7	Near Shibuya where many young people visit
Area8	Near Harajyuku where there are many stylish cafes and shops

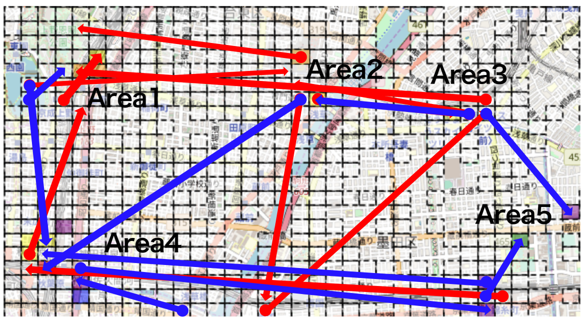

Figure 3. Movement trajectories of residents(Blue) and tourists(Red) around Tokyo Skytree.

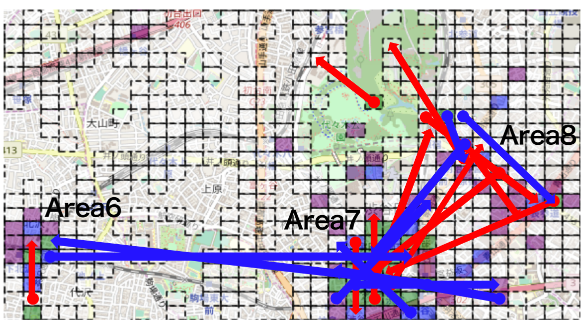

Figure 4. Movement trajectories of residents(Blue) and tourists(Red) around Shibuya.

Park in Area5 is near the Tokyo Skytree in Area3 by movement trajectory of residents and sightseeing spots that many tourist visit. We infer the possibility that the route of Kameido Temple and Kinshicho Park is a sightseeing route that many residents know, but tourists do not know. These results are regarded as useful information for tourism agencies when recommending sightseeing plans and sightseeing spots for tourists.

Next, we discuss Figure 4. Figure 4 portrays the area surrounding Shibuya and Harajyuku. We show Figure 4 in three areas to support several points of the discussion. We shall specifically discuss areas numbered as Area6, Area7, and Area8 and describe each area in Table II. Many more residents than tourists move to Shimokitazawa in Area6 and Shibuya in Area7. Conversely, many more tourists than residents move to Harajyuku in Area8 and Shibuya in Area7, although both Shimokitazawa in Area6 and Harajyuku in Area8 are famous for fashion. The reason for this result is that Harajyuku is known to many more people than Shimokitazawa because Harajyuku is reported frequently in media such as television and dramas. However, Shimokitazawa is a fashion town that anyone living in Tokyo knows. Many magazines publish the area and many residents are visit there. Possibly, the route of Shimokitazawa and Shibuya is a sightseeing route that many residents know, but which tourists do not know. In addition, many reviews [31] state dissatisfaction with sightseeing because Shibuya and Harajyuku are extremely crowded by many tourists on holidays. The result of our experiment points to resolution of this difficulty if tourism agencies have performed PR for Shimokitazawa, Shibuya, and Harajyuku as a series of sightseeing areas and if tourists who visit Harajyuku visit Shimokitazawa.

As for the implementation of discussion, we discover hotspots and sightseeing routes that many residents use but many tourists do not use. These result have the possibility of local spots and new sightseeing routes. As reported herein, we have discovered differences in the movement trajectories of residents and tourists in sightseeing areas.

VI. CONCLUSIONS

This study used latitude and longitude information given along with huge volumes of data obtained from social media sites. By classifying contents into those of residents and those of tourists, and by performing DTW and kmeans++ analyses, we clustered the movement trajectories, visualized hotspots and movement trajectories, and analyzed them further. Based on those results, we were able to discover sightseeing spots that many residents and tourists visit respectively around the Tokyo Skytree. Especially, sightseeing spots that many residents visit, other than tourists, can become new sightseeing spots for increased tourists. We also discovered sightseeing route that many residents use and few tourists use around Shibuya by movement trajectories.

As future work, we expect to conduct quantitative evaluation experiments and improve the proposed method. As described in this paper, we consider different hotspots and movement trajectories of residents and tourists based on visualization

 ISBN: 978-1-61208-687-3

results. However, in future work, we plan to evaluate them more quantitatively. For improvement of the method, we focus on a certain cell, calculate the movement proportion of the next cell, and extract the ranking of the cell movement proportion. Then we must adapt this method to all resident and tourist cells. By adapting the Spearman's rank correlation coefficient to the calculated data, the difference between the movement trajectories of residents and tourists is quantified. As a different method, map matching are regarded as revealing details of differences between residents and tourists when assessing the roads that they used. Additionally, for this study, users were classified as residents or tourists, but user attributes of many types exist. Studies assessing them and their characteristics are being conducted actively. The main targets of estimation are gender [32], age [33] and residence [34]. As future work, we expect to consider these user classifications and to analyze their movement trajectories in sightseeing areas. Additionally, we do not consider the user's preference in this study, however we also experiment with them in future work.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 16K00157, 16K16158, and Tokyo Metropolitan University Grant-in-Aid for Research on Priority AreasResearch on social big data.

REFERENCES

[1] "By 2030 overseas travelers expand to 1.8 billion people annually," 2019, URL: https://www.travelvoice.jp/20171005-97777 [accessed: 2019-01-09].
[2] "Tourism promotion business to Japan," 2019, URL: http://www.mlit.go.jp/kankocho/shisaku/kokusai/vjc.html [accessed: 2019-01-09].
[3] "Some tourism promortions," 2019, URL: http://www.sangyo-rodo.metro.tokyo.jp/plan/tourism/chapter6_3_2018.pdf [accessed: 2019-01-09].
[4] "The number of foreign visitors Japan records 28.69 million," 2019, URL: https://www.travelvoice.jp/20180116-104077 [accessed: 2019-01-09].
[5] "60% of foreign visitors to Japan are repeaters," 2019, URL: https://www.travelvoice.jp/20180403-108446 [accessed: 2019-01-09].
[6] "The need of repeaters," 2019, URL: http://www.mlit.go.jp/kankocho/news02_000346.html [accessed: 2019-01-09].
[7] D. J. Crandall, L. Backstrom, D. Huttenlocher, and J. Kleinberg, "Mapping the world's photos," in *Proceedings of the 18th international conference on World wide web*. ACM, 2009, pp. 761–770.
[8] S. Kisilevich, F. Mansmann, and D. Keim, "P-dbscan: a density based clustering algorithm for exploration and analysis of attractive areas using collections of geo-tagged photos," in *Proceedings of the 1st international conference*, 2010, pp. 38:1–38:4.
[9] H. Yin, X. Lu, C. Wang, N. Yu, and L. Zhang, "Photo2trip: an interactive trip planning system based on geo-tagged photos," in *Proceedings of the 18th ACM international conference on Multimedia*. ACM, 2010, pp. 1579–1582.
[10] N. Yuya, E. Masaki, E. Yo, H. Masaharu, Y. Syohei, and I. Hiroshi, "Inferring tourist behavior and purposes of a twitter user," *AIAI*, 2016.
[11] C. Zhuang, Q. Ma, X. Liang, and M. Yoshikawa, "Anaba: An obscure sightseeing spots discovering system," in *2014 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2014, pp. 1–6.
[12] ——, "Discovering obscure sightseeing spots by analysis of geo-tagged social images," in *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*. IEEE, 2015, pp. 590–595.
[13] Y. Yang, Z. Gong, and H. U, Leong, "Identifying points of interest by self-tuning clustering," in *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 2011, pp. 883–892.
[14] Y. A. Lacerda, R. G. F. Feitosa, G. Á. R. M. Esmeraldo, C. d. S. Baptista, and L. B. Marinho, "Compass clustering: A new clustering method for detection of points of interest using personal collections of georeferenced and oriented photographs," in *Proceedings of the 18th Brazilian symposium on Multimedia and the web*. ACM, 2012, pp. 281–288.
[15] Z. Yin, L. Cao, J. Han, C. Zhai, and T. Huang, "Geographical topic discovery and comparison," in *Proceedings of the 20th international conference on World wide web*. ACM, 2011, pp. 247–256.
[16] D. Li, X. Zhou, and M. Wang, "Analyzing and visualizing the spatial interactions between tourists and locals: A flickr study in ten us cities," *Cities*, vol. 74, pp. 249–258, 2018.
[17] "Flickr," 2019, URL: https://www.flickr.com [accessed: 2019-01-09].
[18] S. Van Canneyt, S. Schockaert, and B. Dhoedt, "Discovering and characterizing places of interest using flickr and twitter," *International Journal on Semantic Web and Information Systems (IJSWIS)*, vol. 9, no. 3, pp. 77–104, 2013.
[19] "Twitter," 2019, URL: https://www.twitter.com [accessed: 2019-01-09].
[20] "Foursquare," 2019, URL: https://ja.foursquare.com [accessed: 2019-01-09].
[21] J. Yuan, Y. Zheng, and X. Xie, "Discovering regions of different functions in a city using human mobility and pois," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2012, pp. 186–194.
[22] M. Nanni and D. Pedreschi, "Time-focused clustering of trajectories of moving objects," *Journal of Intelligent Information Systems*, vol. 27, no. 3, pp. 267–289, 2006.
[23] H. Kori, S. Hattori, T. Tezuka, and K. Tanaka, "Automatic generation of multimedia tour guide from local blogs," in *International conference on multimedia modeling*. Springer, 2007, pp. 690–699.
[24] Y. Sun, H. Fan, M. Bakillah, and A. Zipf, "Road-based travel recommendation using geo-tagged images," *Computers, Environment and Urban Systems*, vol. 53, pp. 110–122, 2015.
[25] I. Memon, L. Chen, A. Majid, M. Lv, I. Hussain, and G. Chen, "Travel recommendation using geo-tagged photos in social media for tourist," *Wireless Personal Communications*, vol. 80, no. 4, pp. 1347–1362, 2015.
[26] A. Garcia, O. Arbelaitz, M. T. Linaza, P. Vansteenwegen, and W. Souffriau, "Personalized tourist route generation," in *International Conference on Web Engineering*. Springer, 2010, pp. 486–497.
[27] J. Zhang, H. Kawasaki, and Y. Kawai, "A tourist route search system based on web information and the visibility of scenic sights," in *Universal Communication, 2008. ISUC'08. Second International Symposium on*. IEEE, 2008, pp. 154–161.
[28] X. Lu, C. Wang, J.-M. Yang, Y. Pang, and L. Zhang, "Photo2trip: generating travel routes from geo-tagged photos for trip planning," in *Proceedings of the 18th ACM international conference on Multimedia*. ACM, 2010, pp. 143–152.
[29] "Twitter API," 2019, URL: https://help.twitter.com/ja/rules-and-policies/twitter-api [accessed: 2019-01-09].
[30] "Tourists visiting the Ueno Zoological Gardens has increased," 2019, URL: https://www.nikkei.com/article/DGXMZO28548380U8A320C1000000/ [accessed: 2019-01-09].
[31] "Reviews of Shibuya and Harajyuku," 2019, URL: http://xayataka.hatenablog.com/entry/2017/10/15/114420 [accessed: 2019-01-09].
[32] J. D. Burger, J. Henderson, G. Kim, and G. Zarrella, "Discriminating gender on twitter," in *Proceedings of the conference on empirical methods in natural language processing*. Association for Computational Linguistics, 2011, pp. 1301–1309.
[33] D. Rao, D. Yarowsky, A. Shreevats, and M. Gupta, "Classifying latent user attributes in twitter," in *Proceedings of the 2nd international workshop on Search and mining user-generated contents*. ACM, 2010, pp. 37–44.
[34] Z. Cheng, J. Caverlee, and K. Lee, "You are where you tweet: a content-based approach to geo-locating twitter users," in *Proceedings of the 19th ACM international conference on Information and knowledge management*. ACM, 2010, pp. 759–768.