A Visualization Method of Residents’ Location for Media Spots Estimation

Koya Kimura*, Yurika Shiozu**, Kosuke Ogita*, Ivan Tanev*, and Katsunori Shimohara*

Abstract: In this research, we aim to visualize activities in a local community for achieving resident-centered community design utilizing information and communication technology (ICT). For that purpose, we introduced a new core concept of “media spots” not only as places where residents could communicate with one another more frequently than in the other areas but also as a prospective platform to mediate relationality between people, “Mono” as tangible and physically perceived things, and “Koto” as intangible and cognitively conceived things. That is, “media spots” should have the potential to proactively promote resident-motivated communications and activities in local community. We have proposed, in this paper, a visualization method for media spots estimation from residents’ location information that uses DBSCAN (density-based spatial clustering of applications with noise) for cleaning enormous amount of raw location information data. The present results suggested that representative points extracted by using DBSCAN could visualize activities in a local community more effectively, and estimate media spots more accurately, than the previously used method.

Key Words: resident-centered community design, local community, media spots, location information, DBSCAN.

1. Introduction

The structure of local communities in Japan needed to be changed after the Great Tohoku Earthquake. Japan is the first country to have a super-aging society, an observation that was first noted in 2007. According to the latest Annual Health, Labor and Welfare Report, the connections have been weak among members of the communities, but the number of people who want to help in the community has increased [1]. Consequently, local communities must be vitalized for the rebuilding of relationships.

This type of social situation has prompted the development of several approaches from the computer science field. Examples of these approaches include the person-watching system of the Japan Science and Technology Agency in the Strategy Proposal of 2012 and the smart-city system [2]. However, despite the need for these approaches, most communities cannot use them casually.

The research on constructing such systems has been successful, but its usage is still in the development stage. Moreover, the introduction of these systems does not always guarantee local activation. The concept of local activation means that locals themselves notice their problems and act to solve them. In the concept of resident-centered vitalization of local communities, residents become aware of the invisible relationship that they have among each other, and understand the importance of such awareness, thereby vitalizing their community through sustaining and promoting it. An invisible relationship is one of which they are otherwise unconscious.

For that purpose, this work introduced “media spots” as not only places where communication occurs more frequently among residents as compared to other areas but also as prospective platforms to mediate relationality among people, “Mono” (i.e., tangible and physically perceived things), and “Koto” (i.e., intangible and cognitively conceived things). That is, “media spots” must have the potential to promote proactively the resident-motivated communications and activities in local communities. The resident-centered vitalization of a local community will not be realized until residents themselves find regional challenges and/or issues and tackle them spontaneously.

To visualize media spots as places for creating and strengthening relationships, and to create an opportunity to regain relationship within the community, this study explored a method of estimating media spots, which are the places where there were greater possibilities of residents communicating with each other, using residents’ location information, which was obtained through smartphones. However, enormous amounts of raw data are unsuitable for visualization because useful information is buried in them. In addition, data obtained through smartphones are often less accurate. Thus, when handling an enormous amount of location information, it is indispensable to be able to slice the data efficiently.

In this work, a visualization method of residents’ location information was proposed to estimate media spots, using density-based spatial clustering of applications with noise (DBSCAN) among the clustering algorithms.

2. Resident-Centered Vitalization of a Local Community

Ushino (1982) explained the importance of the concept of local resident-based regional development, and proposed a system called “Kande System [3].” After the industrialization and urbanization in the 1950s, the village communities in
rural areas were divided by agricultural policies and then re-integrated in the 1970s to create a new regional system (Ushino, 1982). The importance of this concept has already been a significant research topic since the 1980s.

Meanwhile, Yoshizumi (2013) analyzed the way for locals to develop regions sustainably, and suggested the “Eco Card System [4].” In this system, the locals are given a stamp card called an “Eco Card” that promotes environmental activities, thereby creating a setup for the locals to be involved in the region. This system highlights the importance of visualizing or making the locals notice the problems and thus incites them to manage local resident-based regional development efforts.

With the introduction of information communication technology (ICT), it is temporarily possible to solve the challenges of a local community. However, to vitalize a local community continuously, residents must solve them positively. For residents to solve the challenges of a local community by themselves, they must be conscious of these challenges. Thus, “resident-centered” development means that “residents themselves solve the challenges faced by their local community.” In this research, we aimed to establish a methodology that enables them to do so by visualization.

3. Experiment Method

3.1 Overview of Field Experiment

The residents’ activity data were acquired through the smartphones lent to them[5],[6]. This study referred to “activity data” as “location information,” “sent and received emails,” “telephone reception and transmission,” and “data passed among each other via Bluetooth.” In this work, “location information” was analyzed to improve the media spot estimation method.

3.2 Experimental Area and Cooperators’ Attribute

This field experiment was conducted in the Makishima area in Uji City, Kyoto, Japan. Uji City is located south of Kyoto, on the south side of Kyoto City. As of April 1, 2016, Uji City has a population 190,000; 15,000 (7.9%) of them live in the Makishima area. Uji City has attracted much attention as a residential area near Kyoto, Osaka, and Kobe since the early 1960s. As a result, parcels of residential land were developed in Uji City, and the population has increased remarkably. The Makishima area is among the development areas in Uji City. It is a densely populated area, with many housing complexes. Although many parts of this area, which were developed in the early 1960s, are currently aging, the population of the whole Makishima area remains slightly increasing.

The experimental cooperators lived in the Makishima area and were members of the nonprofit corporation Makishima Kizuna-no-Kai. Table 1 shows the attributes of the experimental cooperators. Many of the experimental cooperators were over 65 years old for the reason that people who retired at mandatory age mainly join the regional development.

Table 1 Attributes of field experiment.

| Area               | Makishima, Uji City, Kyoto, Japan |
|--------------------|-----------------------------------|
| Cooperators        | 20 to 50                          |
| Age                | 30 to 70 years old                |

3.3 Experimental Installation

Table 2 shows the periods of the field experiment. The experimental cooperators were instructed to use the lent smartphones at all times for the duration of the experiment. The location information was acquired every minute, except in the following situations:

- From the perspective of informed consent, the experimental cooperators were instructed to switch off the smartphone when they did not want to inform others of their location information.
- The smartphone has run out of battery, or the experiment cooperator forgets to carry the smartphone.
- The smartphone cannot transmit location information as it is out of range.
- The timing when Android finishes to acquire location information is not determined exactly.

Table 2 Periods of field experiment.

| Period      | Dates                |
|-------------|----------------------|
| 1st period  | Nov. 11, 2013 to Dec. 10, 2013 |
| 2nd period  | Feb. 11, 2015 to Mar. 27, 2015 |
| 3rd period  | Jul. 11, 2015 to Jan. 11, 2016 |

3.4 Method of Data Cleansing

For data cleansing, DBSCAN [6], which is a density-based clustering algorithm, was used. Location information gathered by Android smartphones has meter accuracy. The accuracy is defined as the 68% confidence interval, or true with a probability of 68% in the circle with accuracy as the radius, according to Android specification. The location information gathered by smartphones cannot have pinpoint accuracy, and can be noisy.

DBSCAN is robust with noise, and matches excellently for processing location information. Data cleansing with DBSCAN was executed using the following steps:

1. Cluster location information into experimental cooperator daily.
2. Extract the location information with the highest accuracy as a representative point.
3. Execute clustering of location information for all representative points on each experimental cooperator for the
Table 3 Specification of smartphones.

|                  | First period        | Second and third periods                |
|------------------|---------------------|----------------------------------------|
| Manufacture      | Fujitsu             | ASUS                                   |
| Model number     | ARROWS Kiss F-03E   | ZenFone 5 A500KL                      |
| OS               | Android 4.0.4       | Android 4.4.2                          |
| Network career   | NTT docomo          | III Mobile (MVNO of NTT docomo)        |
| CPU              | Qualcomm Snapdragon S4 | Qualcomm Snapdragon 400              |
| Clock frequency  | 1.5 GHz             | 1.2 GHz                                |
| Core             | Dual Core           | Quad Core                              |
| RAM              | 1 GB                | 2 GB                                   |
| Location information | GPS               | GPS and GLONASS                      |
| Bluetooth        | 4                   | 4                                      |
| Sensor           | G-Sensor            | G-Sensor (E-Compass/Proximity Light/Hall Sensor) |

where $n$ is the number of dimensions, $SD$ is standard deviation, and $q(0.75) - q(0.25)$ is the interquartile range acquired by subtracting the first quartile from the third quartile.

3.6 Estimation Method of Media Spot

The 2D location information (the point showing a position) consisted of latitude and longitude that recorded a person’s action at a specific time. The point that showed many positions at the spot where the person stayed would be plotted, and then density would become higher. The part that is high in density appears as the maximum value. Subsequently, the probability density function of this location information was measured using KDE, and then the place where each person was usually found was estimated by counting the number of maximum value locations. The inspection of the hypothesis followed. A media spot was defined as the place where communication was active in an area. “Active communication” referred to the place where many residents gathered in an area. In this work, media spots were estimated using KDE from the data cleansed by DBSCAN.

Numpy, scipy, and scikit-learn, which are libraries for high-level scientific calculations of Python, were used for the inspection.

4. Results and Discussions

This section describes the results of the analysis on the proposed method, using the data of the third period.

4.1 DBSCAN

Figure 2 shows the scatter diagram of raw location information of one of the experimental cooperators. Figure 3 shows the scatter diagram of a sliced location information of the same one using DBSCAN. Figure 4 shows the scatter diagram of raw location information of all experimental cooperators. Figure 5 shows the scatter diagram of a sliced location information of all experimental cooperators using DBSCAN.

As a result of aggregating the representative points of each cooperator, 101 location information could be acquired. The location information was then converted into real addresses by reverse-geocoding with Google Maps Geocoding application programming interface (API). Table 4 shows a result of checking whether the location information and address match. “Un-
Fig. 2 Scatter diagram of raw location information data (one of experimental cooperators) (2015-07-11 - 2016-01-11).

Fig. 3 Scatter diagram of location information data after DBSCAN (one of experimental cooperators) (2015-07-11 - 2016-01-11).

determinable” means that there were multiple meeting places of, among others, towns and commercial facilities around the location information. “Exclusion” means that there must be a residential quarter around the location information. This result shows that location information could be compressed while ensuring measurable accuracy, by extracting the representative points by DBSCAN.

This result shows whether the representative points of the location information derived by DBSCAN point to public facilities or commercial facilities on the map. Considering the accuracy of each location information, this result shows we could derive the representative points more accurately by the proposed method.

Table 4 Results of checking whether the location information and the address match (samples: 101).

|                | Accord | Not accord | Undeterminable | Exclusion |
|----------------|--------|------------|----------------|-----------|
|                | 13     | 32         | 26             | 30        |

4.2 Media Spots Estimation

Figures 6 and 7 show the results of applying KDE to location information (period: 2015-08-01 to 2015-08-31) using the previous method [7]. Figure 6 shows 2D figures. The vertical and horizontal axes represent latitude and longitude, respectively. The dark part is the place of the point’s high density. Meanwhile, Fig. 7 shows 3D figures. Figures 8 and 9 show the results of applying KDE to a sliced location information (period: 2015-08-01 to 2015-08-31) by DBSCAN. The left and right axes represent latitude and longitude, respectively. The upper axis represents density. The dark part means the place of the point’s high density.

The media spots estimation method, together with the previ-
ous method [7], entailed the following steps:

- Find the local maxima of places where location information is in high density using KDE from each person’s location information.

- The local maxima show places where a person was usually found, as they show the density of location information.

- Delete location information around the global maxima, which are expected to be main activity places, such as the home and workplace.

- Collect the dataset with deleted information on the home and workplace.

- Apply KDE to the collected dataset and then find the local maxima as media spots.

Location information was acquired at regular intervals. Therefore, the density of the location information became high in fixed and long-staying places (main activity places) such as their home or workplace. When we estimated the media spot using raw data, places where there was a high possibility that communication was active, with the exception of their home and workplace, were evaluated as relatively worthless.

To prevent such phenomena, we deleted location information around the global maxima, which were expected to be main activity places such as their home and workplace.

As a result of the comparison with the previous method, it appears that the proposed method for media spots estimation was more accurate than the previous one, because the deviation of the data was reduced. The previous method had a problem that the local maxima were significantly influenced by the data volume of each person, since places where a person who acquired large amounts of location information stayed long were visualized as high density.

As a result of interviews presenting these results to the organizer of the experimental cooperators, we found that the area where the experimental cooperators lived and had meetings in the community had high density.

Figures 10 and 11 show the results of applying KDE to a sliced location information (period: 2015-07-11 to 2016-01-11) by DBSCAN. As a result of the comparison of it with the result of the whole period, we found that the figure was almost the same as the figure described by the whole period. We also found that there was hardly any change in the place where people usually stayed in the whole period and during a certain period.

5. Conclusion

In attempt to achieve the resident-centered community design by utilizing ICT, “media spots” were defined as not only the places where communication occurred more frequently among residents as compared to other areas but also prospective platforms to mediate relationality among people, “Mono” (i.e., tangible and physically perceived things), and “Koto” (i.e. intangible and cognitively conceived things).

To visualize media spots as places for creating and strengthening relationships, and to create an opportunity to regain relationships within the community, a visualization method of residents’ location information was proposed to estimate media spots, using DBSCAN as among the clustering algorithms.

The present results suggested that representative points ex-
extracted by using DBSCAN could visualize activities in a local community more effectively, and estimate media spots more accurately, than the previously used method.

Further studies are needed to improve the accuracy of representative points and evaluate the validity of media spots. Moreover, we will work on how to develop the potential of media spots to promote proactively the resident-motivated communications and activities in a local community.

Acknowledgments
This work was supported by JSPS KAKENHI Grant Numbers JP16J03602, JP16K03718, and JP17K0086.

References
[1] Ministry of Health, Labour and Welfare: Annual Health, Labour and Welfare Report 2013-2014, 2014.
[2] Japan Science and Technology Agency: Research and development on fundamental technologies of cyber physical systems and their social implementation: A case study on promoting aged people to social activities, CDS-FY2012-SP-05, 2013.
[3] T. Ushino: Comprehensive district plan by inhabitants and “Kande” system, Journal of Rural Planning Association, Vol. 1, No. 3, pp. 19–29, 1982.
[4] M. Yoshizumi: A study on actively community-based environmental town planning toward sustainable communities: A case study on the Eco-community program in Nishinomiya, Hyogo, Japan, Journal of the City Planning Institute of Japan, Vol. 48, No. 3, pp. 831–836, 2013.
[5] K. Kimura, Y. Shiozui, I. Tanev, and K. Shimohara: Visualization of relationship between residents using passing-each-other data toward resident-centered vitalization of local community, Proceedings of the Second International Conference on Electronics and Software Science (ICES2016), pp. 122–127, 2016.
[6] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu: A density-based algorithm for discovering clusters in large spatial databases with noise, Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD’96), pp. 226–231, 1996.
[7] K. Kimura, Y. Shiozui, I. Tanev, and K. Shimohara: A leader and media spot estimation method using location information, HIMI 2016, Part II, LNCS 9735, pp. 550–559, 2016.