

# Sightseeing Hot Spots Analysis by Using SNS's Photos and Taking Location Information Based on Image-to-Tag Method

Kohjiro Hashimoto, Tadashi Miyosawa, Mai Miyabe, Takeshi Tsuchiya, Takeshi Ozaki, and Hiroo Hirose

**Abstract**—The many sightseeing photos are posted in Social Network Service (SNS). Any interested objects for tourists are shown up in the posted photos. Therefore, it is considered that the characteristics of visit spot can be discovered by analyzing the SNS's photos. By the way, Microsoft started the cloud service of an image analysis based on Artificial Intelligence technology. The Computer Vision API in this cloud service makes image segmentation and object labeling automatically possible for SNS's photos. Therefore, it is considered that it is possible to detect the characteristics of visit spot as label information by analyzing statistically generated labels based on Computer Vision API. In this paper, an analysis method of the generated tags based on Computer Vision API is studied on. Concretely, we propose a calculation method of the degree of interest for the generated tags, and an analysis method of interest object in each region based on both the tag's interest degree and taking location information of SNS's photos.

**Index Terms**—Tourism informatics, data mining, SNS image analysis, deep learning.

## I. INTRODUCTION

The problems of declining and aging population are faced in Japan. In particular, the local area has several problems of creation of industries, employment, and so on. In order to solve these problems, it needs the activation of local area. Then, sightseeing becomes a key element thinking about the activation of rural area. In Japan, tourism is put effort into. Government set a national goal that the number of foreign tourists is 4000 ten-millions by 2020 [1]. Fig. 1 shows the number of foreign tourists to Japan [2]. It is confirmed that the number of tourists has been increasing steadily over time. On the other hands, Fig. 2 shows the number of tourists to Chino city, Nagano prefecture, Japan where our university is located [3]. There are several sightseeing spots such as Sirakaba lake and Kurumayama highland in Chino city. However, it is confirmed that the number of tourists has grown at a sluggish pace. According to this reason, our university has addressed the technical resolution of this issue in cooperation with Chino city government.

For this problem, the government tries to implement a measure to increase the tourists. However, the measure is not effective when mismatch is created between the characteristics of local area that the government recognizes

and actual tourists has interest in. Therefore, in order to avoid such mismatch and encourage improvement, it is important to clear the characteristics of local area that actual tourists have interest in.

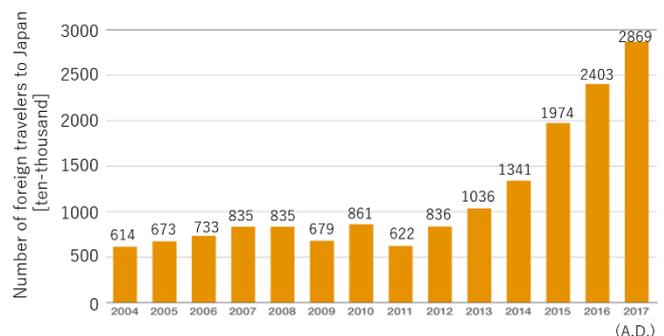


Fig. 1. The number of foreign tourists to Japan.

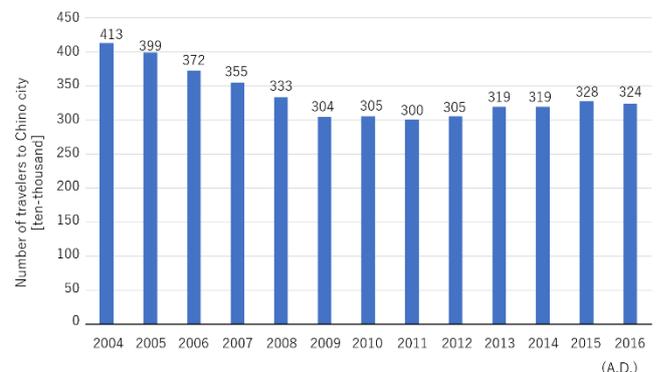


Fig. 2. The number of tourists to Chino city, Nagano prefecture, Japan.

By the way, the tourists post to the social networking service (SNS) the messages and photos that express the impressions and experiences when they come over to sightseeing spot. Therefore, it is effective to analyze the tourist behavior and sightseeing spot by using the information on SNS. For example, K. Murakami *et al.* [4] had analyzed the characteristics of sightseeing spot by detecting the word with high appearance frequency from blog website based on text mining library. K. Okubo *et al.* [5] had made heat map of the location that tourist taking photos based on the latitude and longitude of the taking photos posted to photo sharing site Flickr. H. Uehara *et al.* [6] had made the feature vector of sightseeing spot by using information of several website, such as Wikipedia and yahoo. The positional information of tourists has used to analyze their behavior and sightseeing route. It is effective to improve transportation condition and barrier-free environment [7], [8]. On the other hand, the posted photos are effective to analyze the interest

Manuscript received September 25, 2019; revised May 15, 2020.

Kohjiro Hashimoto, Tadashi Miyosawa, Mai Miyabe, Takeshi Tsuchiya, Takeshi Ozaki, and Hiroo Hirose are with the Department of Applied Information Engineering, Suwa University of Science, 5000-1, Chino, Nagano, Japan (e-mail: k-hashimoto@rs.sus.ac.jp).

object of tourists. Therefore, it is possible to analyze the characteristics of local area as the objects that tourists have interest in by collecting and analyzing the posted photos in any local area.

J. Wang *et al.* [9] had proposed an analysis method of the SNS's photos based on an image recognition technique. However, it is necessary to set experientially the image features of recognized target object on photo in the conventional method. It is difficult to set the image features in advance. Because it is considered that a lot of interest object of tourism exists. On the other hand, Google or Microsoft have been opened the service of image analysis on cloud in 2016. In this image analysis, deep learning technology was used. In this technique, image features can be able to set automatically, and recognize several objects. Therefore, it is considered that the above problem of the conventional image analysis method can be improved by using this technique.

In this paper, we focus on the Image-To-Tag function of Computer Vision API (CVA) that Microsoft provides. Image-To-Tag function is the function that appearance objects in an image are detected and labeled as tag. This function means that an image is converted into language as tag information. Therefore, it is considered that it is possible to detect the characteristics of local area as tag by analyzing statistically generated tags based on Image-To-Tag function. In this paper, hot spots analysis method is proposed by using Image-To-Tag function. Hot spot means the spot that tourists visit to see their interest objects. Here, their interest objects are detected by analyzing statistically generated tag information based on Image-To-Tag for SNS's photos in any area. And hot spot is analyzed by plotting the taking positions of the photos with analyzed tag information. The effective of the proposed method is evaluated through the several experiments.

## II. PHOTO SHARING WEB SITE FLICKR

Flickr is the famous photo sharing web site and is used worldwide. Posted photo of Flickr has the following characteristics. Firstly, the posted photo has several information, such as the poster's id, the photo's id, the latitude and longitude of the taking location, and title of the photo. Secondly, the posted photos with above information can be collected by making program based on Flickr API. Therefore, this Flickr's photo is used to analyze the hot spot of sightseeing in this paper. In this paper, suwa city, matsumoto city, and chino city in Nagano prefecture, Japan are set as analyzed local area. Then, these area's photos are collected from Flickr web site. The collected photos are converted into tag information by applied the Image-To-Tag function of CVA explained the next section, and the taking location information of the photo is given to the tag information.

## III. IMAGE TO TAG FUNCTION OF CVA

As Fig. 3 shows, Image-to-Tag function on CVA can convert image into tag that express object name. Fig. 3(a) shows an example of input image. On the other hand, Fig. 3(b)

shows the result of analysis of the input image. In this function, the objects shown up in the input image are detected as tag information. As Fig. 3(b) shows, "grass" and "outdoor" described on the back of "name" are tags. Note, however, that total number of tags given by using CVA is unclear. Therefore, it was confirmed that there are tags of more than 2000 by the result of executed previous experiment. The value described on the back of "confidence" express the confidence of given tag. This vale has range from zero to one.

As Fig. 3 shows, the Image-To-Tag function can be regarded as the function that image data is converted into language data. In this paper, the local area's photos on Flickr are collected and converted into language data as tag information based on this function. Then, the area's characteristics which indices interest is extracted by analyzing language data based on the method explained in next section.



(a) An example of input photo (b) Analysis result(tag + confidence)  
Fig. 3. An example of execution result of Image-To-Tag function of CVA program.

## IV. ANALYSIS METHOD OF INTEREST OBJECTS

When the object that rarely is seen in other location is shown up in the photo, it is considered that the object has higher degree of interest. Therefore, the difference of object shown up in other area's photos is defined as the degree of interest. In this paper, the degree of interest is calculated for the given tag based on the Image-To-Tag function, and the tag with high degree of interest is defined as interest object. The degree of interest is calculated based on the logistic regression model.

Logistic regression model is one of the statistical regression model. This model is used to classify into two types. And this model is used to evaluate the correlative of between input and output information. The outline is described as follows.  $P$  express the probability that an event is occurred. And  $\mathbf{x}=(x_1, x_2, \dots, x_r)$  is a set of variable number to explain the occurrence of the event. In general, the  $x$  is called as explanatory variable. Then, the probability  $P(\mathbf{x})$  at the state of  $\mathbf{x}=(x_1, x_2, \dots, x_r)$  is defined as the follows.

$$P(\mathbf{x}) = \text{Pr}(\text{occurrence of event} \mid x_1, x_2, \dots, x_r) = F(x_1, x_2, \dots, x_r)$$

$Pr$  expresses a probability. Therefore,  $P$  means the conditional probability that event occurs under conditions of explanatory variable  $\mathbf{x}=(x_1, x_2, \dots, x_r)$ . In this method, this  $Pr$  is regarded as function  $F(\cdot)$  is modeled. The following logistic regression function is used to express the function  $F(\cdot)$  in this method.

$$p(\mathbf{x}) = F(Z) = \frac{\exp(Z)}{1 + \exp(Z)}$$

Here,  $Z$  is defined as the follows.

$$Z = \eta_0 + \eta_1 x_1 + \eta_2 x_2 + \dots + \eta_r x_r$$

This means that the logistic regression model is the model linked multiple regression equation  $Z$  of explanatory variable and occurrence probability  $p(x)$  that output the value with the range zero to one. This model is useful effectively for modeling of an event with two-valued output when the explanatory variable input. Therefore, this model is used as clustering method. On the other hand, this model is generated by training the coefficient of explanatory variable  $\eta=(\eta_0, \eta_1, \eta_2, \dots, \eta_r)$  as model parameters. This parameter can be regarded as the evaluation index which expresses the correlative between output value and explanatory variable. Therefore, this model also is used as data analysis. In this paper, the model parameters estimated by model training is defined as the degree of interest for tag. Concretely, the calculating method of the interest degree based on the logistic regression model is followed.

The training data of logistic regression model is the data consisted of explanatory variable  $\mathbf{x}=(x_1, x_2, \dots, x_r)$  and outcome variable  $y \in \{0,1\}$ . In this method, the explanatory variable is expressed as the summation of one-hot vector for tag given by Image-To-Tag function. For example, consider the case where the total number  $M$  of tag type given by Image-To-Tag function is five (“outdoor”, “car”, “sea”, “driver”, and “train”). And the index of these tags makes corresponding to the index of explanatory variable. When a photo input to Image-To-Tag function, the given tags were “car”, and “driver” as an example. The explanatory variable is expressed as  $\mathbf{x}=(0, 1, 0, 1, 0)$  in this method. On the other hand, for the outcome variable,  $y=1$  when the rural area of input photo is the analysis target, otherwise  $y=0$ . For example, analysis target is suwa city. The photos of each rural area are collected and converted into training data as follows by Image-To-Tag function. Then, the logistic regression model is generated by training parameters.

<Training data>

*Photo1(suwa city) :  $y = 1, \mathbf{x} = (0, 1, 0, 1, 0)$*   
*Photo2(matsumoto city) :  $y = 0, \mathbf{x} = (1, 1, 0, 0, 0)$*   
*Photo3(chino city) :  $y = 0, \mathbf{x} = (0, 0, 0, 0, 1)$*   
*Photo4(suwa city) :  $y = 1, \mathbf{x} = (1, 1, 0, 1, 0)$*   
*Photo5(chino city) :  $y = 0, \mathbf{x} = (0, 0, 0, 1, 0)$*   
 $\vdots$

In the trained model, the model parameter  $\eta=(\eta_0, \eta_1, \eta_2, \dots, \eta_r)$  is high in order to increase outcome variable close to  $y=1$  when the image of analysis target is input. On the other hand, the parameter is low in order to reduce outcome variable close to  $y=0$  when the image of extra-analysis target is input. Therefore, if there are tags that rarely is appeared in other rural area’s photos, the coefficient for the tags is higher than other one. In this paper, the logistic regression model of suwa, matsumoto, and chino city are generated by the above method, and trained parameter  $\eta=(\eta_0, \eta_1, \eta_2, \dots, \eta_r)$  is regarded as the degree of interest for each tag. Note that the maximum likelihood estimation method is used as training method, and this method is executed by using machine

learning library scikit-learn. Moreover, the total number  $M$  of the tag type of Image-To-Tag function was unclear. Therefore,  $M$  was set to the total number of appeared tags when all images of rural area are applied to Image-To-Tag.

## V. POSITION PLOTTING OF TAKING PHOTOS WITH TAGS HAVE HIGH DEGREE OF INTEREST

In this paper, a hot spot is defined as the spot that tourists visit to see their interest objects. Here, their interest objects are regarded as the tag with high degree of interest calculated based on the previous analysis method. Therefore, by plotting the taking position information of the photos with tag has a high degree of interest on a map, the distribution of hot spots can be visualized on the map.

## VI. EXPERIMENT

### A. Outline

In this experiment, suwa, matsumoto, and chino city in Nagano Prefecture of Japan are set as the analyzed local area. The number of collected photos of suwa, matsumoto, and chino city are 3,215, 10283, and 3,638 photos. Each photo was taken during 2010 to 2017. The collected photos are applied to Image-To-Tag function and converted into tag information. Then, the interest degrees were calculated for each tag based on analysis method described in 4 section. These tags were arranged according to the interest degree. Then, the validity of the calculated degree of interest was evaluated by comparing on the known sightseeing characteristics of each local area. Note that the total number of tags is 1,821 when all collected photos is applied to Image-To-Tag function.

### B. Analysis Result Based on the Interest Degree

In this section, the degree of interest calculated based on the logistic regression model was evaluated. Table I, II, and III show the calculated degree of interest for each tag. In these table, the top 10 tags of the calculated degree of interest are shown. As Table I shows, “Fireworks” tag is top. Suwa lake fireworks festival held in suwa city is famous festival and attracts about 500 thousand spectators every year. Fig. 4(a) shows an example of firework photo. The collected photos for suwa city contain the “Fireworks” photo. On the other hand, the collected photos in other city did not contain the fireworks. Therefore, it is considered that the high degree of interest was calculated. Moreover, “shrine” tag is ranked at top 10. This is because that the collected photos in suwa city contain the photos of Suwa Taisha Shrine as Fig. 4(b). In this city, strange festival called “Onbashirasai” is held at Suwa Taisha Shrine once every six years. The used photos in this experiment were taken during 2010 to 2017 and the “Onbashirasai” was held over this time period. According to this reason, it is considered that the “shrine” tag which maybe express the Suwa Taisha Shrine is ranked at top 10. Fig. 4(c) shows Suwa lake photo. This lake is famous sightseeing spot in suwa city. The “sea” tag is corresponded to this lake photos. Fig. 4(d) shows the photo that corresponds to “can” tag. This photo shows the beer released locally. Therefore, this tag

calculates the high degree of interest.

TABLE I: THE CALCULATION RESULTS OF THE DEGREE OF INTEREST BASED ON VARIANCE STANDARD METHOD (SUWA CITY)

Ranking	Tag	Degree of Interest
1	Fireworks	5.54
2	Automotive	2.38
3	Shrine	1.98
4	Can	1.95
5	Box	1.85
6	Sea	1.79
7	Feet	1.78
8	Election	1.76
9	Steam	1.74
10	Festival	1.73

TABLE II: THE CALCULATION RESULTS OF THE DEGREE OF INTEREST BASED ON VARIANCE STANDARD METHOD (MATSUMOTO CITY)

Ranking	Tag	Degree of Interest
1	Pagoda	2.73
2	Camping	2.55
3	Hiking	2.29
4	Herb	2.28
5	Trek	2.25
6	Cycling	2.13
7	Preschool	2.12
8	Mailbox	2.11
9	Clothing	2.09
10	Temple	2.08

TABLE III: THE CALCULATION RESULTS OF THE DEGREE OF INTEREST BASED ON VARIANCE STANDARD METHOD (CHINO CITY)

Ranking	Tag	Degree of Interest
1	Snow	1.67
2	Little	1.64
3	Pulling	1.00
4	Table	0.96
5	Platform	0.95
6	Bunch	0.94
7	Ramp	0.87
8	Highway	0.85
9	Standing	0.84
10	Jacket	0.83



(a) "Fireworks" photo in a set of suwa city photos



(b) "Festival" photo in a set of suwa city photos



(c) "Sea" photo in a set of suwa city photos



(d) "Can" photo in a set of suwa city photos

Fig. 4. An example of photos contained in a set of Suwa city photos.

As Table II shows, "pagoda" tag is top. It is considered that this tag expresses the matsumoto castle. Fig. 5 (a) shows the matsumoto castle photos. Matsumoto castle is famous

sightseeing spot in matsumoto city. On the other hand, "camping" and "hiking" tag which express the activity are ranked at top 10. In fact, it was confirmed that the collected photos in matsumoto city contain the photos of the camping (Fig. 5(b)), hiking, and cycling (Fig. 5(c)) situation. Therefore, this result can be regarded as adequate result.



(a) "Pagoda" photo in a set of matsumoto city photos



(b) "Camping" photo in a set of matsumoto city photos



(c) "Cycling" photo in a set of matsumoto city photos

Fig. 5. An example of photos contained in a set of matsumoto city photos.

As Table III shows, "snow" tag is top. This is because that the collected photos in Chino city contain the many photos (Fig. 6(a)) of hillwalking at winter season. Chino city is famous for hillwalking and ski. On the other hand, the collected photos in Chino city also contained the photos with children (Fig. 6(b)). Therefore, it is considered that the "little" tag in Table III expresses the children. According to this result, it can be assumed that Chino city has a location to play with family.



(a) "Ski" photo in a set of chino city photos



(b) "little" photo in a set of chino city photos

Fig. 6. An example of photos contained in a set of chino city photos.

According to these results, it was confirmed that the proposed analysis method has a tendency that the tags which does not appear in other local area are detected. Moreover, it is also interesting that the activity tag is ranked higher by using this method. However, it is confirmed that the abstract tag, such as "Pulling" and "Box", were also ranked higher. This is caused by the redundancy and abstractness of given tag by Image-To-Tag function.

### C. Analysis Result of Hot Spots

Fig. 7 (suwa city), Fig. 8 (matsumoto city), and Fig. 9 (chino city) show the distributions of taking location of photos with tag that calculates high degree of interest. As Fig. 7 shows, it can be confirmed that the taking locations of photos are concentrated in around the suwa lake. In addition, it is possible to confirm what the tourists are photographing. On the other hands, it can be confirmed that the taking locations vary in chino city, as Fig. 8 shows. It is considered that this is big difference from other cities. This suggests that

attractive features may exist in various spots compared to other cities.

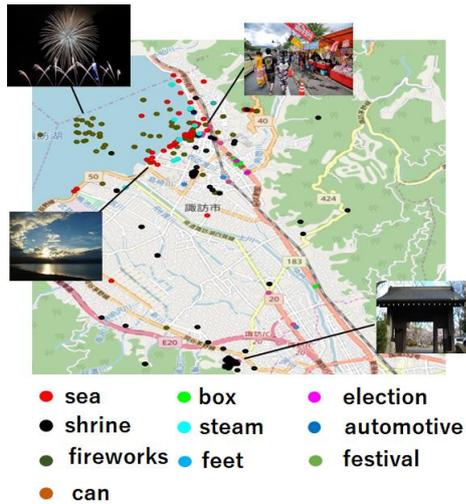


Fig. 7. Analysis result of taking location of photos (suwa city).

According to these results, analyzing tourists' location of visit and taking object of photos is effective in discovering the characteristics of the area. From the results of the experiment, it was confirmed that the proposed methods was effectiveness because the differences from other areas was able to be discovered.

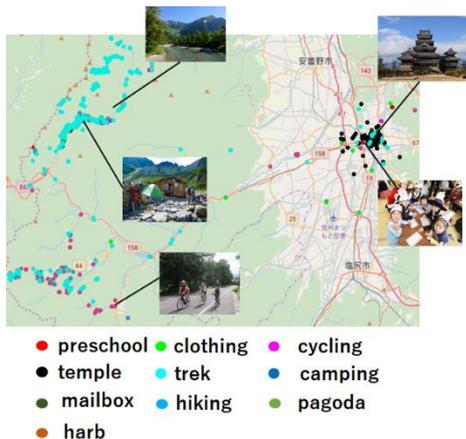


Fig. 8. Analysis result of taking location of photos (Matsumoto city).

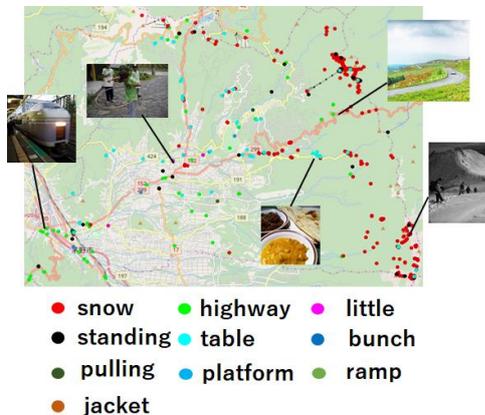


Fig. 9. Analysis result of taking location of photos (Chino city).

## VII. CONCLUSION

In this paper, we focused on the cloud image analysis service based on the deep learning provided by Microsoft as

tools to analyze sightseeing characteristics of rural area. We tried this service to SNS's photos posted by the tourist in the rural area. Concretely, the Image-To-Tag function that input image data is converted into language data as tag was programmed. The collected SNS's photos were converted into tag by using the made program. And a set of the converted tag was analyzed based on the logistic regression model to discover the interest object of sightseeing for rural area in the given tags. In this method, the given tags in analysis target rural area are evaluated by comparing on the tags of other rural area. Therefore, this method has a characteristic that the high degree of interest is calculated for the tags that rarely is appeared in other rural area. In the experiment, the SNS's photos in suwa, matsumoto, chino city was analyzed based on the proposed method. According to the experimental results, it was confirmed that the tags which does not appear in other rural area has high degree of interest. Moreover, it was confirmed that the activity tag is ranked higher by using this method. As the analysis result of hot spots, it was confirmed that distribution of the hot spots in each city was different and the attractive features may exist in various spots compared to other cities was suggested.

The follow contents will be considered as the future works. We will propose a conversion method of SNS's photos into tags with concrete information. For example, for a photo of a person, if the tag with object, behavior, and situation, such as "Children do fishing at suwa lake.", can be given, it is easy to image the characteristics of rural area. Fortunately, a Image-To-Text method which converts image into caption was proposed. Therefore, the analysis method based on the Image-To-Text method will be studied.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## AUTHOR CONTRIBUTIONS

Kohjiro Hashimoto wrote this paper. Kohjiro Hashimoto and Tadashi Miyosawa conducted and analyzed this research based on Image Processing. Mai Miyabe analyzed this research based on the Nature Language Processing. Takeshi Tsuchiya, Takeshi Ozaki, and Hiroo Hirose constructed the network system of collecting SNS's photos; all authors had approved the final version.

## REFERENCES

- [1] Ministry of Land, Infrastructure, Transport and Tourism, Tourism Agency in Japan. Program of sightseeing version 2018. [Online]. Available: <https://www.mlit.go.jp/kankocho/news02000354.html>
- [2] JTB Tourism Research and Consulting Co. in Japan. Japan-bound Statistics. [Online]. Available: <https://www.tourism.jp/tourism-data-base/stats/inbound/>
- [3] Nagano Prefecture Office in Japan. 2017 sightseeing spot user statistics findings results. [Online]. Available: <https://www.pref.nagano.lg.jp/>
- [4] K. Murakami and H. Kawamura, "Travel in Japan from foreigners' perspectives: Analysis of destination image through English blogs," *Journal of Japanese Society for Artificial Intelligence*, vol. 26, no. 3, pp. 286-293, 2011.
- [5] K. Okubo, T. Mochizuki, and T. Watanabe, "A comparison of POIs obtained from multiple social with location information," in *Proc. Forum on Information Technology*, 2019, pp. 481-482.
- [6] H. Uehara, K. Shimada, and T. Endo, "Sightseeing location recommendation using tourism information," *Technical Report of IEICE*, vol. 112, no. 367, pp. 13-18, 2012.
- [7] K. Suzuki, "Investigation and data mining of tourist activities: Cases of Hakodate MachiNavi and circular tour with rent-a-car," *Journal of*

Japanese Society for Artificial Intelligence, vol. 26, no. 3, pp. 248-255, 2011.

- [8] T. Miyosawa, H. Hirose, T. Tsuchiya, T. Ozaki, and N. Kurata, "Induction experiments to sightseeing spots using location based AR in large-scale events and user analysis by composite behavior analysis method," *Journal of Infoscionomics Society*, vol. 12, no. 1, 2017.
- [9] J. Wang, H. Hirose, T. Takahashi, D. Deguchi, I. Ide, and H. Murase, "Creation of a sight-seeing map with visual classification of photos on the web," *Journal of Information Processing*, vol. 52, no. 12, pp. 3588-3592, 2011.

Copyright © 2020 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).



**Kohjiro Hashimoto** received the Ph.D. degree from Nagoya University, Japan in 2014. He is an assistant professor at the Department of Applied Information Engineering, Suwa University of Science, Japan. His current research interests include image recognition and deep neural network to model vehicle driver's behavior and construct work support system.



**Tadashi Miyosawa** received the Ph.D. degree from Waseda University, Japan in 2009. He is a professor at the Department of Applied Information Engineering, Suwa University of Science, Japan. His current research interests include augmented reality and virtual reality technique and tourism informatics.



**Mai Miyabe** received the Ph.D. degree in Wakayama University, Japan in 2011. She is an assistant professor at the Department of Applied Information Engineering, Suwa University of Science, Japan. Her current research interests include natural language processing and human interface, and human communication support system.



**Takeshi Tsuchiya** received the Ph.D. degree from Waseda University, Japan in 2009. He is an associate professor at the Department of Applied Information Engineering, Suwa University of Science, Japan. His current research interests include information and communication network, and distributed cooperative systems.



**Takeshi Ozaki** received the Ph.D. degree in Tokyo University of Science, Japan in 2013. He is an assistant professor from the Department of Applied Information Engineering, Suwa University of Science, Japan. His current research interests include project based learning of educational technology and learning analysis.



**Hiroo Hirose** received the Ph.D. degree from Daito Bunka University, Japan. He is a professor at Applied Information Engineering, Suwa University of Science, Japan. His current research interests include e-learning of educational technology and user support system.