ECAI 2010
H. Coelho et al. (Eds.)
IOS Press, 2010

doi:10.3233/978-1-60750-606-5-523

Extraction of Places Related to Flickr Tags

Yukino Baba[1] and **Fuyuki Ishikawa**[2] and **Shinichi Honiden**[3]

Abstract. Geographic information systems use databases to map keywords to places. These databases are currently most often created by using a top-down approach based on the geographic definitions. However, there is a problem with this approach in that these databases only contain location definitions such as addresses and place names, which does not allow for searches using keywords other than these words. Additionally, they do not give any information on the popularity, e.g., which is more popular among the places indexed by the same keyword. A bottom-up approach, based on the actual usage of words, can address these problems. We propose a method to aggregate tagging data and extract places related to a tag using the pair of a tag and a geo-tagged photo. We target the co-occurrence of a tag and the geolocation and represent the places related to a tag as a probability distribution over the longitudes and latitudes. We applied our method to data on the photo sharing service Flickr and experimentally confirmed that our method made it possible to highly-accurately extract places related to tags. Our direct bottom-up approach enables the extraction of place information that is not obtained by using traditional top-down approaches.

1 Introduction

Web mapping services (e.g., Google Maps) have recently become popular tools. With the popularization of devices equipped with global positioning systems (GPS) (e.g., mobile phones, car navigation systems), geographic information presentation services have been developed that present the information around a user's current position based on the longitude and latitude. These geographic information systems use databases to map keywords to places. For example, a web mapping service receives a user's query word that represents the position the user wants to see, and the services show the position on the map by obtaining the longitude and latitude related to the query. The geographic information presentation system derives the geolocations related to words for previously targeted resources, compares these geolocations with the user's current position, and shows the resource that is strongly associated with the user's position.

Currently, these databases are usually created using a top-down approach based on the geographic definitions. However, there are two problems with this approach.

- **These databases only have information about the addresses, place names, landmarks, and stores**
  Suppose someone was making travel plans to come to Japan. He wants to see Sumo, which is Japanese-style wrestling, but does not know where to see it. He decides to search for places using a web mapping service. Fortunately, this person knows how to write "Sumo" in Japanese characters, so he inputs this word as a query. He hopes to find a site where a competition will take place, but instead, the search engine returns a town with a name that includes the word "Sumo," but it is not related to Sumo wrestling. The reason for this is that current web mapping services seem to use only keyword matching techniques, and the names of actual Sumo competition sites do not include the word "Sumo."
- **These databases do not have information about which places are most popular if multiple candidate places for a keyword exist**
  Additionally, the tourist would like to see a large statue of Buddha. He knows that there are many statues of Buddha in Japan and wants to see a famous one. He wants to know where a famous Buddha statue is located. Current mapping services are not able to satisfy his desire because they do not have information based on the popularity of a place or an object. The person would have to look at some travel guides to get the names of the more famous Buddha statues and then check the map to locate the address of these statues. This can be bothersome.

These two troubles happen because current geographic databases are created by using top-down approaches based on the geographic definitions. We believe that a bottom-up approach, based on the actual usage of words, can address these problems. In this paper, we propose a method to estimate the places related to words from the real usage of words and to create the geographic data bottom-up to make up for the lack of information. Our target words are not only place specific words (e.g., "Tokyo") but also place related words (e.g., "Sumo" or "Buddha") in order to link the word "Sumo" to some competition sites where actual sumo exhibitions are held. We also designed our method to obtain the regions of the places mentioned and the and relevance between words and places. Thus, the results received from using our method contain the regions of the places and the relevance values between "Buddha" and them, and show the ranked places by their popularities.

We focused on the tagging data from folksonomy services as the data created from the actual usage of words. Folksonomy[14, 8] is a classification method for web resources and has been adopted in many web services such as Delicious, Flickr, and YouTube. In these services, users can freely choose a tag, which is a keyword related to a resource, and annotate it to the resource. Resources are classified by these tags, and the results of tagging are shared by the users of these services. Some resources on folksonomy services have recently included geolocation data (This is called "geotagging"). For example, some photos on Flickr include the geolocation data where the photo was taken, and these data are represented by using longitude and latitude coordinates. By tagging these resources, a tag and a geolocation are associated via a resource. Thus, we believe we can use the tagging results to obtain the places related to a word (or tag).

[1] The University of Tokyo, Japan, ybaba@nii.ac.jp
[2] National Institute of Informatics, Japan
[3] The University of Tokyo and National Institute of Informatics, Japan

In particular, our method extracts the places related to the tags from the tagging results for geotagged Flickr photos. We refer to the adding of a tag to a photo annotated with a geolocation as the co-occurrence of a tag and a geolocation, and we use these co-occurrence data to extract the places related to the tags. We represent these places by using the probability distributions over the longitude and latitude coordinates because a probability distribution can represent the regions as well as the relevance of each region, which is suitable for our requirements.

Our approach to place representation is more versatile than other representations, such as cells [3, 12] or fractions of the Earth's surface [5]. Therefore, the extracted results provide several advantages for many different applications, other than the geographic database creation:

- Photo location prediction from tags
- Tag recommendation from photo geolocations
- Tag similarity calculations

In our work, we do not apply our method to specific applications, however, we evaluate that our method extracts information that was unavailable using top-down approaches and how well our results fit the actual Flickr usage and the necessary level of human recognition regarding places.

The contributions of this paper are summarized as follows:

- The first attempt to automatically extract the places related to the tags, which are not only place specific tags, but also place related tags.
- A proposal of a method that represents the desired places by using the probability distributions over the longitude and latitude coordinates, and a probabilistic model of users' tagging behavior to geotagged photos.
- An application of proposed method to extract places related to tags from geotagged photos on Flickr, and qualitative and quantitative evaluations of our approach.

2 Related Work

2.1 Tags and Geotagged Photos

A lot of research has focused on Flickr tags and geolocation data. Ahern et al. [1] proposed a method to derive representative tags for areas in the world by using multi-level clustering and TF-IDF based scoring of tags and to display these tags on a map. Their targeted task was different from ours because they attempted to derive tags from a specific area while we try to derive places from a tag. Rattenbury et al. [10] focused on the place and event semantics for tags and presented a method to determine whether a tag corresponds to a place and/or event based on the time and geolocation data of the resources annotated with the tag. As with our research, they were interested in tags and geolocations. However, they only tackled the classification problem and did not target the task of representing the places related to specific tags.

Some studies have been conducted to predicting the locations of photos that have been taken in order to organize the photos on a map. Hays et al. [5] proposed a method to predict the locations of photos that were taken from the tags and features of the photos. First, they collected photos annotated with tags of city or landmark names and created the feature vectors of each photo. Then they created the feature vectors of the targeted photos for prediction and used the Nearest Neighbor method to predict the places. Crandall et al. pointed out that Hays's approach limited the predicted places to cities and landmarks and proposed a new method [3]. In their approach, some primary locations are extracted at first using the photos annotated with the geolocations. Next, they classify the non-geotagged photos from each major place by using support vector machines (SVM) based on the tags, image features, and the times the photos were taken. A prediction approach using only tags was proposed by Serdyukov et al. [12]. They put grids on a map and represented each place by using an individual cell. Then they built a language model to describe the relation of a tag and a place and estimated a cell as the place where a photo was taken by using tags annotated to the photo.

These approaches target the problem of predicting the places where photos have been taken, so it can be said that they extract the places related to photos. We target the problem of extracting the places that are related to the tags, so the scope is different from these other studies. In addition, they represented places as cells [3, 12] or fractions of the Earth's surface [5]. We believe that our representation, which are the probabilistic distributions over the longitude and latitude coordinates, has a broader utility than the others.

2.2 Tag Relationships

Our method is designed to extract the relationships between the tags and geographical locations. To our knowledge, there is no existing work that tries to connect tags to non-textual information. In the existing research, the similarity relations or hierarchical relations between tags have been investigated. For a similarity relation, Hotho et al. [7] proposed a method to calculate the tag similarity by using PageRank. Cattuto et al. [2] analyzed five measures of tag similarity. They mapped the tags that were considered to be similar based on each measure on WordNet and calculated the distance between each tag. Noh et al. [9] tried to translate a tag into multiple languages. They built a tag graph for each language based on the co-occurrence of the tags and estimated which tags had the same meanings in the various languages by using the graph similarity.

For hierarchical relations, Schmitz et al. [11] proposed a method to introduce an ontology of tags based on the co-occurrence of tags. Tang et al. [13] focused on tags and documents, built a tag topic model by using the words in the documents, and estimated the hierarchical relations based on the distribution in the topics. These studies represented the tag semantics as the relationships between tags. However, we treat the semantics as the distributions over the longitude and latitude, so on this point, our method is different from the existing research.

3 Problem Definition

In this section, we formalize the co-occurrence data used for extraction and describe the representation of the places related to tags.

3.1 Formalization of Co-occurrence data

Hotho et al. [7] defined the data structure of folksonomy as follows.

Definition 1 *Folksonomy is a tuple* $\mathbb{F} := (U, T, R, Y)$ [4]. $U, T, and R$ *are the sets whose elements are called users, tags, and resources.* Y *is a ternary relation between them, i.e.,* $Y \subseteq U \times T \times R$, *called tag assignments.*

[4] Hotho et al. introduces a user-specific subtag/supertag-relation as one element of Folksonomy. However, we omit it because we have not included it in our method.

when $y = (user_l, tag_i, resource_j)$ for $y \in Y$, y represents "$user_l$ annotates tag_i to $resource_j$."

With this data structure, we represent the "$resource_j$ is annotated with $location_j \in L$" as $(resource_j, location_j) \in X$. Here, $L = \{(lat_j, lon_j) | -90 \leq lat_j \leq 90, -180 \leq lon_j \leq 180\}$

Using X, we describe D, a set of co-occurrence data, as follows.

$$D = \{(tag_i, location_j) \mid (resource_j, location_j) \in X, \\ (user_l, tag_i, resource_j) \in Y\}$$

3.2 Probabilistic Representation of Place Related to Tag

We decided to represent the desired places as the probability distributions over the longitude and latitude coordinates to describe the regions and the relevance of each region. Thus, the next question is how to represent the "places related to a tag" as a probability distribution.

To determine the representation by using the probability, we model the users' tagging behavior. Here, we assume that when a user annotates a tag to a geotagged resource, the user associates the tag and the geolocation that are annotated to the resource. Thus, the tagging model is built using the following process.

1. A user encounters $resource_j$, which is annotated with $location_j$ with probability $p(location_j)$
2. The user annotates a tag to the resource. At the same time, the user associates the tag and the geolocation with the probability $p(tag_i|location_j)$.
3. Then, the tag and the geolocation co-occur with the probability $p(tag_i, location_j)$.

In this model, "the probability of a resource annotated with tag_i having $location_j \in L$" is represented as a probability $p(location_j|tag_i)$. We are able to interpret this probability as "the probability of a tag being associated with a geolocation." Thus, this probability distribution represents the place related to a tag, and our goal is to estimate $p(location_j|tag_i)$ from the co-occurrence data, D.

4 Extraction of Places

In this section, we model the probability of a co-occurrence in order to estimate the desired probability, $p(tag_i|location_j)$, and describe the method to obtain the parameter values of the probability model.

In our method, we use the co-occurrence of two variables: tags and geolocations. Some probability models with latent variables have been proposed to highly-accurately estimate the probabilities of two co-occurring variables. Zhang et al. showed that one such model, called Aspect Model [6], was effective at representing the relationship between tags and resources [15]. Therefore, we have adopted the Aspect Model to represent the co-occurrence of tags and geolocations. With this model, we are able to describe the probability of the co-occurrence of a tag and a geolocation, $p(tag_i, location_j)$, as follows.

$$p(tag_i, location_j) \\ = \sum_k p(location_j)p(z_k|location_j)p(tag_i|z_k) \quad (1)$$

This model is based on the following tagging model.

1. A user encounters $resource_j$, which is annotated with $location_j$ with probability $p(location_j)$.
2. The resource makes the user think of a concept z_k. At the same time, the user associates the concept and the geolocation with probability $p(z_k|location_j)$.
3. The concept z_k triggers the user to use tag_i with probability $p(tag_i|z_k)$.
4. Then, the tag and the geolocation co-occur with probability $p(tag_i, location_j)$.

Using the observable probability of the co-occurrence, (1), we estimate the desired probability distributions.

First, the desired probability distribution, $p(location_j|tag_i)$, is represented as follows.

$$p(location_j|tag_i) \\ = \sum_k p(z_k|tag_i)p(location_j|z_k) \\ = \frac{\sum_k p(tag_i|z_k)p(z_k)p(location_j|z_k)}{\sum_k p(tag_i|z_k)p(z_k)} \quad (2)$$

Thus, three probabilities, $p(tag_i|z_k), p(z_k)$, and $p(location_j|z_k)$ are needed to estimate $p(location_j|tag_i)$. Using the Bayse' theorem,

$$p(tag_i, location_j) \\ = \sum_k p(tag_i|z_k)p(z_k)p(location_j|z_k) \quad (3)$$

The observable probability of co-occurrence is composed of the three required probabilities, $p(tag_i|z_k), p(z_k)$ and $p(location_j|z_k)$, and we are able to estimate the parameters of these three probabilities.

Additionally, we assume $p(location_j|z_k)$ is a two-dimensional Gaussian distribution to represent the regions by using the probability distributions using a continuous value, $location_j$. In our method, a Gaussian distribution corresponds to a place. Finally, the desired distributions are obtained as a Gaussian mixture with the mixing rate π_k.

$$\pi_k = \frac{p(tag_i|z_k)p(z_k)}{\sum_k p(tag_i|z_k)p(z_k)} \quad (4)$$

The parameters we need to estimate are $p(tag_i|z_k), p(z_k), \mu_k$ and Σ_k. Here, μ_k and Σ_k are the mean and the variance of the Gaussian mixture given by the following formula.

$$p(location_j|z_k) = \mathcal{N}(location_j|\mu_k, \Sigma_k) \quad (5)$$

We decided to estimate these parameters using EM algorithm [4].

5 Evaluation

We applied our method to real Flickr data and conducted an experiment. First, we confirmed that with our method it was possible to extract information that was not available using traditional top-down approaches. Second, we evaluated the accuracy of the results by checking how the output fit the Flickr test data and compared the accuracy of our method against two baselines. Finally, we evaluated the results using human judgment.

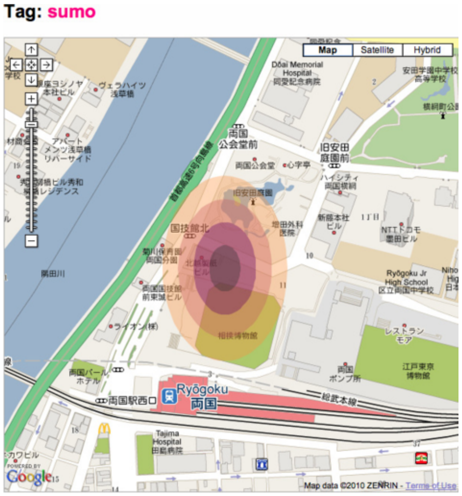

Figure 1. Part of results for sumo tag. Ellipses with dark colors indicate extracted regions, where an actual sumo competition site is located.

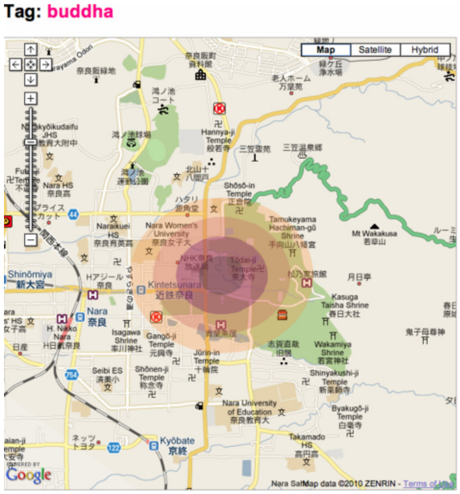

Figure 2. Part of results for buddha tag. This map shows the region where a famous statue of Buddha exists.

5.1 Dataset

We collected photo data from Flickr. The photos in the collection were taken in Japan between January 2004 and December 2007. In our dataset were the number of photos $|R| = 512,356$, the number of tags $|T| = 71,223$, the number of users $|U| = 7,457$, and the number of tag assignments $|Y| = 3,826,253$.

5.2 Qualitative Evaluation

We displayed the estimated probability distribution on a map and checked to see what type of information was extracted by using our method. We chose two tags, sumo and buddha, to see the results. We applied our method to a subset from our dataset including sumo and four other tags that occurred with almost equal frequency throughout the whole dataset. We also created another subset that included buddha and four additional tags. When using our method, it is necessary to indicate the number of latent variables in advance. We set 25 as this number for the five tags.

Figures 1 and 2 show the results for the sumo and buddha tags. Based on the estimated probability distributions, we draw ellipses on the figures.

Figure 1 shows part of the results for the sumo tag. We extracted the area around Ryogoku Kokugikan, the most popular sumo competition hall, as the places related to the sumo tag. Actually, our method extracted multiple regions with this tag; in fact, three other venues for sumo competitions were extracted.

Figure 2 shows part of the results for the buddha tag. We extracted the region around Todai-ji, a temple where a large and very famous statue of Buddha is located. In addition, we extracted the area around Kamakura, where another large and very popular Buddha statue exists.

In this experiment, we confirmed that we can extract information other than the addresses, place names, landmarks, and stores, and that we can extract information about which place is most popular.

5.3 Quantitative Evaluation

Next, we conducted a quantitative evaluation that confirmed the accuracy of the results by checking how the extracted probabilistic distributions of each tag fit the distributions on the Flickr test data. In addition, we compared this accuracy with two baseline methods.

Tags
2005, 2006, 2007, **asakusa**, asia, autumn, canon, **chiba**, city, festival, flower, food, **fukuoka**, geotagged, **hakone, harajuku, hiroshima, hokkaido**, japan, japanese, japon, **kamakura, kanagawa, kansai, kyoto, miyajima, nagoya, nara**, night, **nikko**, nikon, **okinawa, osaka**, park, people, sakura, **sea, shibuya, shinjuku**, shrine, sky, summer, temple, **tokyo**, travel, trip, **yokohama, KYOTO**, JAPAN, **TOKYO**

Table 1. Top 50 tags used for quantitative evaluation. Words in boldface indicate the tags used for human judgment. Words in capital letters are written in Japanese characters in the actual Flickr data.

For this experiment, we chose 50 tags in descending order of frequency. Then we randomly selected the co-occurrence data including these tags and created two subsets of our dataset. One was for training (5,000 data items) and the other was for the test (50,000 data items). In Table 1, we list the 50 tags used for this experiment. The Tags in bold type are described in 5.4 We set the number of latent variables at 50.

The baselines were a Gaussian estimation and K-means. The Gaussian estimation uses only the co-occurrence data that includes a targeted tag (i.e., when extracting the places for Tokyo, it only uses co-occurrence data that includes a "Tokyo" tag). Additionally, the proposed method uses K-means to set the initial values for the EM algorithm. We use these initial values as a base-line method, which is called K-means.

The goal of this evaluation is to confirm how the extracted probabilistic distributions of each tag fit the distributions on Flickr test data. Due to the difficulty in comparing two continuous distributions directly, we quantize the distributions and compare them. The evaluation for accuracy is conducted as follows. We apply our method to the training data and each tag. Then, we place grid regions on a map and calculate the integral values of the extracted probability distribution for each grid. Next, we calculate the ratio of the test data located on each grid for each tag. If the estimation is done at a high level

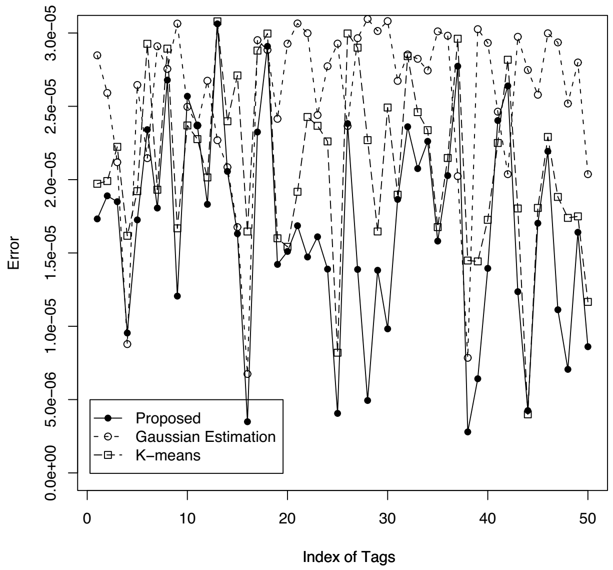

Figure 3. Error across tags

of accuracy, there will not be a lot of difference between the integral values from the extracted data and the ratio from the test data. The details of this evaluation process are as follows:

1. We prepare a set of grid cells S_v, covering all the geolocations that appeared in the training data and test data; the length of a cell is v_{lat} and v_{lng}. In particular, we set both the v_{lat} and v_{lng} at 0.05.
2. For $tag_i \in T$ and for a region $s_{v_c} \in S_v$, the integral values of extracted probability distributions are calculated, $p_{d_{i,c}} = \int_{s_{v_c}} p(location_j|tag_i)$.
3. For $tag_i \in T$ and for $s_{v_c} \in S_v$, the ratio ($p_{e_{i,c}}$) of the test data E_{flickr} with the geolocations that was located in s_{v_c} is calculated.

$$p_{e_{i,c}} = \frac{\left|\{(tag, location) \in E_{flickr} | tag = tag_i, location \in s_{v_c}\}\right|}{\left|\{(tag, location) \in E_{flickr} | tag = tag_i\}\right|}$$

4. For $tag_i \in T$ and $s_{v_c} \in S_v$, the errors between the extracted probability and the ratio are calculated, $|p_{d_{i,c}} - p_{e_{i,c}}|$, and the average value for each region is obtained. If the error is small, the accuracy of the result will be high.

Figure 3 shows the error values for each tag, and Table 2 lists the average error value for each method. The proposed method obtains a lower value than the Gaussian estimation for 42 out of 50 tags and a lower value than the K-means for 46 out of 50 tags. Thus, we confirmed that our proposed method achieves a higher accuracy than the two baseline methods. The average value of error is about 65% of the Gaussian mixture and 79% of K-means.

In addition, we manually classify the 50 tags into three categories and calculate the average value of error for each category. The three categories are: (a) Place Specific Tags (e.g., Kyoto, Nikko), (b) Place Related Tags (e.g., city, festival, flower), and (c) Place Unrelated Tags (e.g., autumn, canon, night). The results are given in Table 3. We observed that the place specific tags obtained a higher accuracy than the tags in the other categories.

	Proposed	Gaussian Estimation	K-means
Average of errors	1.67e-05	2.56e-05	2.11e-05

Table 2. Average of errors for all tags

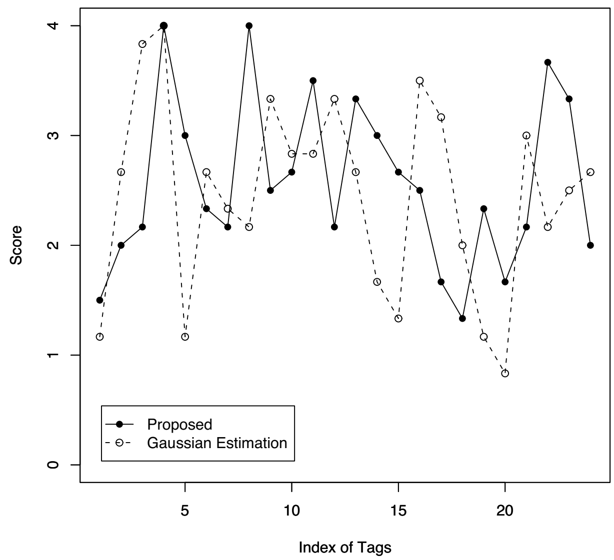

Figure 4. Average score of human judgment

5.4 Human Judgment

In the previous section, we evaluated the accuracy of the extracted results to the test data on Flickr. However, it is difficult to judge whether the results are positive or not for human. Therefore, we conducted another experiment using human judgment to evaluate whether the extracted places fit the necessary level of human recognition.

We displayed the places extracted by the proposed method and by the Gaussian estimation on a map and asked six volunteers to score the results for each tag from 1 (low) to 4 (high). We chose 24 out of 50 tags used in the accuracy evaluation because the results of 26 of the tags did not display specific areas and could have been confusing to the volunteers. The 24 tags are listed in Table 1 in boldface type. The scoring criteria are described in Table 4.

We show the average scores of all the volunteers in Figure 4. Our method received a score over 2 in 83% of the tags and over 3 in 33% of them. If we consider that scores above 2 are correct, our method achieves 83% for precision. In tags with a score of 2 or 3, only the sizes of the regions were mistaken, so it will be possible to extract a suitable size for a region with a larger number of data.

	(a)Place Specific Tags	(b)Place Related Tags	(c)Place Unrelated Tags
Average of errors	1.44e-05	1.93e-05	1.95e-05

Table 3. Average of errors for each categories

Score	Criteria
4	Extracted places are related to tag and coverage and centers of regions are correct
3	Extracted places are related to tag but coverage of regions is slightly narrower or wider and/or center slightly misses point
2	Extracted places are related to tag but coverage of regions is narrower or wider and/or center misses point
1	Some extracted places are related to tag but others are not related; Or regions are too wide and include unrelated places.
0	Extracted places are not related to tag

Table 4. Criteria for human scoring

6 Discussion

In this section, we discuss some points that need to be considered.

6.1 Place less related tags

In 5.3, we saw that place-related tags and place-unrelated tags were less accurate than place specific tags. We believe the reason for this is that the estimated probability $p(location_j|tag_i)$ is composed of three parameters: $p(location_j|z_k), p(z)$ and $p(tag_i|z_k)$. Of these three, only $p(tag_i|z_k)$ affects the difference in the accuracy across the tags. We assume that the strength of the relevance between the tag and a place influences the accuracy. Here, $p(tag_i|z_k)$ is the probability of tag_i occurring based on the concept z_k. We only focus on the places amongst the elements of a concept, so we assume that z_k represents a place.

Tags having a strong relevance to places are related to one or a few places. Thus, the $p(tag_i|z_k)$ of these tags has a specific pattern and may have a high tolerance for noise, and they are therefore able to maintain the high accuracy. However, tags with a weak relevance are expected to have a uniform distribution of $p(tag_i|z_k)$ for any place z_k. Thus, they easily suffer from the effects of noise data, and the accuracy might be reduced.

Preliminarily removing the place-unrelated tags is one approach to tackle this problem. For example, a method proposed by Rattenbury et al. [10] to identify whether a tag is related to a place or not could be applied. In our method, tags estimated at the same time directly affect the results. Thus, we believe that removing place-unrelated tags is one effective way to improve the accuracy of other tags.

6.2 Human recognition

The results in 5.4 indicate that the proposed method achieves a higher accuracy for most tags than the Gaussian estimation, but the fitness to human recognition varied according to the tags. We presume the reason for this is that the results will not always fit directly match the human recognition even if the result fits well to the Flickr data. For example, the results for Nikko had an error value of $9.82e-06$, which is smaller than the average. However, this tag did not obtain a high score for human judgment. A volunteer said, "The results for Nikko only displays a very narrow area related to Nikko." The reason that this poor result for human judgment achieved a high accuracy in the Flickr data is that there is a feature in the geolocation data on Flickr that centers on the popular places for taking photographs.

Therefore, the results from our method achieve a better fitness to the Flickr data compared with the other baseline methods. However, it is difficult to say that our method fits the human recognition better than the other methods. To improve the fitness of the human recognition, one approach might be to combine the data on other types of services and then do an estimation.

7 Conclusion

We proposed a method for extracting places related to Flickr tags using the co-occurrence of a tag and a geolocation. To represent the places using probability distributions over the longitude and latitude coordinates, we considered "the probability of a resource annotated with a tag having a geolocation" as "the probability of a tag being associated with a geolocation" and estimated this probability based on the co-occurrence information. We applied our method to actual data on Flickr and confirmed that our approach enables the extraction of the place information that is unavailable using top-down approaches. Additionally, we conducted a quantitative evaluation to determine whether the proposed method could be used to achieve a higher accuracy than the baseline methods. We also evaluated how well our results fit the human recognition regarding places.

To our knowledge, this paper represents the first effort to represent tag semantics using a format other than the relationships between tags. This time, we used the co-occurrence data of tags and geolocations; however, we are planning to apply other types of data. For example, if we apply our method to time information, we can extract the "time related to tags." Moreover, if we apply it to other information, it will be possible to represent the tag semantics in a more detailed way.

REFERENCES

[1] Shane Ahern, Mor Naaman, Rahul Nair, and Jeannie Hui-I Yang, 'World explorer: visualizing aggregate data from unstructured text in geo-referenced collections', in *Proc. JCDL '07*, (2007).
[2] Ciro Cattuto, Dominik Benz, Andreas Hotho, and Gerd Stumme, 'Semantic grounding of tag relatedness in social bookmarking systems', in *Proc. ISWC '08*, (2008).
[3] David J. Crandall, Lars Backstrom, Daniel Huttenlocher, and Jon Kleinberg, 'Mapping the world's photos', in *Proc. WWW '09*, (2009).
[4] A. P. Dempster, N. M. Laird, and D. B. Rubin, 'Maximum likelihood from incomplete data via the EM algorithm', *Journal of the Royal Statistical Society*, (1977).
[5] James Hays and Alexei A. Efros, 'im2gps: estimating geographic information from a single image', in *Proc. CVPR '08*, (2008).
[6] Thomas Hofmann and Jan Puzicha, 'Statistical models for co-occurrence data', (1998).
[7] Andreas Hotho, Robert Jäschke, Christoph Schmitz, and Gerd Stumme, 'Information retrieval in folksonomies: Search and ranking', *The Semantic Web: Research and Applications*, (2006).
[8] Adam Mathes, 'Folksonomies-Cooperative Classification and Communication Through Shared Metadata', (2004).
[9] Tae-Gil Noh, Seong-Bae Park, Hee-Geun Yoon, Sang-Jo Lee, and Se-Young Park, 'An automatic translation of tags for multimedia contents using folksonomy networks', in *Proc. SIGIR '09*, (2009).
[10] Tye Rattenbury, Nathaniel Good, and Mor Naaman, 'Towards automatic extraction of event and place semantics from flickr tags', in *Proc. SIGIR '07*, (2007).
[11] Patrick Schmitz, 'Inducing ontology from flickr tags', in *Workshop on the Collaborative Web Tagging Workshop*, (2006).
[12] Pavel Serdyukov, Vanessa Murdock, and Roelof van Zwol, 'Placing flickr photos on a map', in *Proc. SIGIR '09*, (2009).
[13] Jie Tang, Ho-fung Leung, Qiong Luo, Dewei Chen, and Jibin Gong, 'Towards ontology learning from folksonomies', in *Proc. IJCAI '09*, (2009).
[14] Thomas Vander Wal. Folksonomy coinage and definition. `http://www.vanderwal.net/folksonomy.html`, 2007.
[15] Lei Zhang, Xian Wu, and Yong Yu, 'Emergent semantics from folksonomies: A quantitative study', *Journal on Data Semantics VI*, (2006).