TOUR RECOMMENDATION SYSTEM BASED ON WEB INFORMATION AND GIS

Yukiko Kawai, Jianwei Zhang

Kyoto Sangyo University, Faculty of Computer Science

Hiroshi Kawasaki

Saitama University, Faculty of Engineering

ABSTRACT

Information recommendation and filtering techniques have been studied intensively. Traditional tour recommendation systems, which can be considered one of information recommendation systems, usually calculate the shortest path in terms of time or distance. Recently, tour recommendation systems for more general purposes have become an important research topic. In this paper we propose an efficient tourist route search system which not only recommends the path simply connecting several tourist spots, but also recommends the path with beautiful scenic sights. We focus on the visibility of scenic sights between one tourist spot and another, which is an important factor for choosing a driving route, but has not been considered in traditional tour recommendation systems. To automatically retrieve tourist spots, we propose a personalized tourist spot recommendation technique using the Web information. Although, for some regions, databases of the famous spots exist and are published, such regions are limited and usually outdated. Our method automatically extracts spots from the Web, thus our system is versatile and up-to-date for large regions. To find a route with attractive scenery, we calculate scores for paths based on the visibility of scenic sights. After generating route candidates using GIS, a 3D virtual space is constructed and the Z-Buffer method is used to decide the visibility of scenic sights for each route candidate. We implemented a prototype and tested the effectiveness of the system.

Index Terms— Navigation system, GIS, personalized adaptive web techniques, 3D virtual space

1. INTRODUCTION

Information recommendation and filtering techniques for user preferences have become an important research topic and have been studied intensively in recent years. Tourist route search systems are also one of information recommendation systems, which are actively researched and developed in the fields of Geographic Information Systems (GIS) and Intelligent Transport Systems (ITS). When a tourist decides a region to visit, he may investigate some spots where he wants to go and see and constructs a route using a car navigation system or *Google Map*. Since available databases of famous spots are only published for limited regions and not frequently updated, tourists have to manually search the spots from Web or guide books to construct a tour route. Furthermore, most of navigation systems only consider a shortest path and traffic jam avoidance as main targets, and do not consider other factors such as scenic sights along a route, weather, and traveling seasons; those are also important for deciding a tour route.

In this paper, we propose an advanced tour recommendation system, which includes the extraction of famous spots from the Web and route search based on the visibility of scenery along a route. The Web is a useful resource in which famous spots are usually introduced. We propose a method for automatically extracting spots from the Web. The visibility of scenery along a route is an important factor for deciding an actual driving route. Sometimes, a tourist would rather drive along a route with beautiful scenery even if it is devious. Our system extracts attractive spots and recommends the path with beautiful scenic sights which connects several spots. We informally define *scenic sights* mentioned in this paper as "landscapes that have great heights and can be enjoyed as a distant view", such as Mt. Fuji, Eiffel Tower, Tian'an Gate, and so on. The proposed system has two features: (i) a personalized tourist spot recommendation technique using the Web information dependent on users' preferences, and (ii) a route search technique based on the visibility of scenic sights from a path. Especially, for the latter we construct a 3D virtual space by using GIS and calculate the visibility of scenic sights by using a method called Z-Buffer. The routes with attractive scenic sights between spots are highly ranked by our system.

2. RELATED WORK

Information recommendation based on personal interests is currently an active research area. Jameson [1] proposes a prototype of a travel decision system for dealing with item recommendation to a group of two or more users. Tezuka *et al.*[2] integrate Web search with a geographic information system. Compared to these researches, our work extracts tourist information from Web pages and recommends them to users for their route search.

Pedestrian or car navigation systems can also be categorized as information recommendation in a wide sense. Maruyama *et al.*[3] propose a personal navigation system called P-Tour for tourism. This system allows users to specify multiple destinations and time restrictions on arrival and staying time of each destination, and gives the nearly best schedule using a GA algorithm. The system given by Kawabata *et al.*[4] satisfies users' special requirements, such as

providing a route without mud, a route without bumps or a route with well maintained security, rather than the shortest path. Another navigation system proposed by Akasaka and Onisawa [5] teaches pedestrians an appropriate route by using Sketch-map and fuzzy theory. Compared to these researches, our work focuses on the new aspect, the visibility of scenic sights for the route search. Hosokawa *et al.*[6] also consider landmark visibility, but their goal is to identify a user's current location, different from our route recommendation.

Several companies provide some geographic information services. Yahoo! Local Maps and Google Maps show search results on a map interface. Google Earth and Microsoft Virtual Earth allow users to explore richer geographical contents. These commercial services maintain our motivation on the research of route search considering visibility.

3. SYSTEM OVERVIEW

Figure 1 gives an overview of our proposed system. First, a user provides a region name to visit. Then, our system retrieves tourist spot and scenic sight information related to the region from the Web. Next, using the Web based interface of our system, the user can select several tourist spots and a scenic sight from the recommended information. The system receives the user's selection and recommends routes based on the selected spots and scenic sight by checking the visibility of the scenic sight for each route. Our system recommends the route satisfying the user's two main preferences: visiting good tourist spots and viewing good scenic sights during a driving. The uniqueness of the proposed system is the recommendation of spots and scenic sights based on Web retrieval and analysis, and the recommendation of routes considering the visibility of scenic sights.

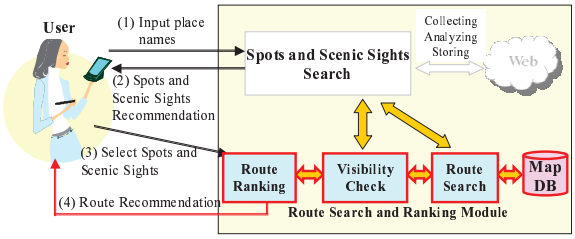


Fig. 1. System overview

4. SPOTS AND SCENIC SIGHTS RECOMMENDATION

For recommendation, tourist spots and scenic sights near the location entered by the user are retrieved by searching the Web. A standard search engine is queried using the entered location and supplementary keywords. The examples of supplementary keywords are those which often appear in tourist Web pages, such as "sightseeing", "tourist spot", "landscape". We make the further morpheme analysis on the search result pages for discovering tourist spots and scenic sights. The proper nouns including "general", "region", "organization", are extracted. We sort these nouns by their term frequency (tf) values in the descending order and take the high ranked ones as recommendation candidates. Our experimental results show this method is feasible though it is relatively simple.

After spots and scenic sights are extracted automatically, highly ranked ones are exhibited to the user. The user picks up the spots which he wants to visit and the scenic sights which he wants to view along the route. In our current system, spots and scenic sights are separately selected by the user. Automatically distinguishing scenic sights from general spots is greatly helpful to users, and thus, developing such system is our future research.

After receiving the selected spots and sights, our system computes and recommends the best scenic route.

5. ROUTE RECOMMENDATION

5.1. Generation of route candidates

Route candidates for calculating visibility are generated by connecting several tourist spots provided by the method described in the previous section. In our current system, we generate route candidates via only one tourist spot; this means we connect a start spot, a tourist spot to pass through, a goal spot, to make a route candidate. Multiple route candidates are generated corresponding to different tourist spots to pass through. We use a GIS system called STIMS (Spatio-Temporal Information Management System) [7] to retrieve the shortest path passing through these three spots. The input of this system is the latitudes and longitudes of selected spots. The system first obtains the points on the road nearest to the three spots, and then finds the shortest path taking into account the type and width of the roads. As a result, the system outputs a route graph which consists of nodes set on the road and links connecting these nodes.

5.2. Visibility check of scenic sights

Z-Buffer (Figure 2) is one usual solution of the visibility problem, which decides whether a rendered scene is visible. We construct a virtual 3D space and use the idea of Z-Buffer to decide the visibility of the scenic sight from a route. We obtain the three-dimensional data using a Digital Elevation Map (DEM), which is currently published by governments in many countries. Using this 3D data, the visibility can be calculated by rendering the target object (scenic sight) and checking whether it can be seen from each node on the route. In our current research, the visibility value for each node is defined as a binary level; "1" means the target object can be seen and "0" means it is invisible.

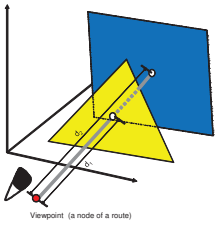


Fig. 2. Z-Buffer

Given a route candidate, the flow of the algorithm of visibility check is shown in Figure 3. First, the DEM data is read and locations of the route and scenic sight are determined. Then, a 3D space is constructed in the computer. Next, the visibility check is conducted by flying the eyes from the current viewpoint (a node of the route) toward the scenic sight. The depth d_1 from the viewpoint to the object which can be seen is computed by the Z-Buffer method (step a), and then the distance d_2 from the viewpoint to the scenic sight is calculated (step b). After that, the visibility value is set to 0 or 1 by comparing d_1 and d_2 (step c); the visibility value becomes 1 if $d_1 = d_2$, whereas 0. The visibility data is stored into a file for later ranking of routes (step d). After moving the viewpoint to the next node and obtaining its altitude, steps from (a) to (d) are repeated until no nodes exist on the route candidate. The image of the visibility check from the whole route is shown in Figure 4.

The computational cost of the above process is high because the 3D model must be redrawn from all the nodes of the route. Therefore, we also propose a method to efficiently check the visibility. The 3D model is rendered only once from the scenic sight surrounding 360 degree (Figure 5). By this way, there is no need to redraw the 3D model, and thus, computational cost is significantly reduced, especially when a route is long.

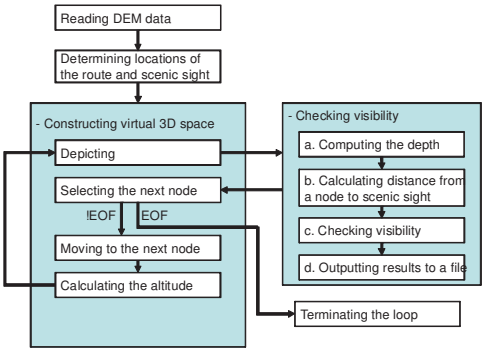


Fig. 3. Flow of visibility check

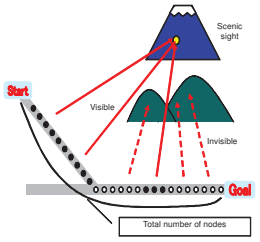


Fig. 4. Visibility check from nodes

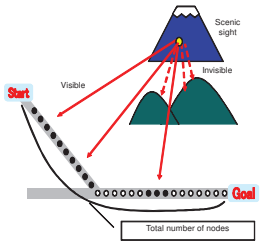


Fig. 5. Visibility check from scenic sights

5.3. Ranking of route candidates

Route candidates are ranked based on the average of the visibility values of all nodes. The calculation formula of the visibility rate s for a route candidate is $s = \frac{1}{n}\sum_{i=1}^{n} v_i$, where v_i is the visibility value of the scenic sight from a node i on the route, and n is the total number of nodes on the route.

We have considered a scenic sight as a single point on the map. However, in the case that a scenic sight has a large region (*e.g.*, a mountain), we sample multiple points from the region and calculate s for each sampled point. The total visibility rate S for a route candidate is calculated by taking the logical summation of m points of the scenic sight as follows: $S = \frac{1}{n}\sum_{i=1}^{n}\{\bigvee_{j=1}^{m} v_{ij}\}$, where v_{ij} is the visibility value of the scenic sight's sample point j from a node i on the route. That is to say, the visibility value for one node on the route is set to 1, if one or more sample points of the scenic sight can be seen from it.

6. EXPERIMENTS

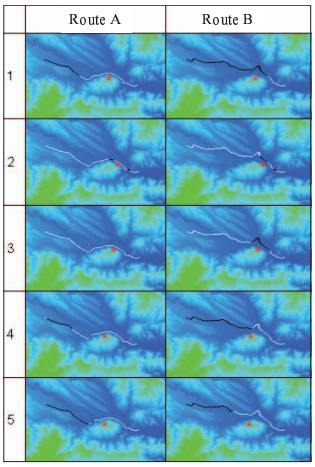
6.1. Extraction of spots and scenic sights

We carried out experiments to extract tourist spots and scenic sights in Japan. The accuracy was calculated as the ratio of appropriate spots to all the 20 recommendations (the top 20 spots with the highest tf values). The appropriateness was subjectively judged by 3 individuals. For example, when we input the place name "Mt. Fuji" as the search keyword, the recommendation results were as follows: 1st Lake Kawaguchi (tf value = 0.087), 2nd Shiroito waterfall (0.063), 3rd Country club (0.051); the 1st and 2nd were judged as appropriate ones by the individuals, however, the 3rd was regarded as inappropriate. The average accuracy for a large area such as a prefecture was 63%, whereas that for a small geographical region such as municipalities was 86%. This was mainly because search results for the former case were collected from some portal Web sites which included many sightseeing information pages across all the country.

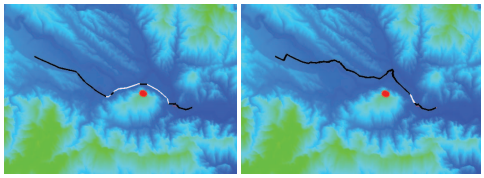
6.2. Calculation of visibility of scenic sights and ranking of route candidates

We conducted an experiment to show the effectiveness of the calculation of visibility of scenic sights and the ranking of route candidates. We selected "Mt. Buko in Chichibu-gun of Saitama prefecture, Japan" as a scenic sight, calculated the visibility, and ranked some route candidates.

The visibility images for five of sample points are shown in Figure 6(a), where the black parts of lines represent the parts of the routes from which the scenic sight can be seen, and the white parts represent invisible. The images of total visibility rates from two routes are depicted in Figure 6(b). The total visibility rate for Route B (96.7%) was higher than Route A (68.8%). Although Route A and Route B were actually almost parallel, the difference of visibility rates was 28%. In this experiment, we confirmed that the visibility rate is subjectively correct by checking constructed movies. Figure 7 shows a sample view from a virtual camera toward the scenic sight. The left image shows the scenic sight can not be seen from route A due to the obstacle, and the right one shows it is possible to view the scenic sight at the front.



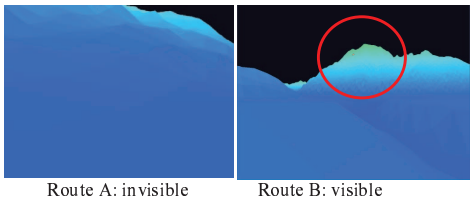
(a) The images of visibility rates for sample points



Route A: S=68.8% Route B: S=96.7%

(b) The images of total visibility rates

Fig. 6. Results (Black lines are visible and white are not)



Route A: invisible Route B: visible

Fig. 7. A sample view from a virtual camera toward the scenic sight

6.3. Prototype

We have implemented a prototype of the proposed system. Figure 8 shows the Web interface of our proposed system. The upper right part provides recommended spots which can be selected by the user, and the bottom right part shows recommended routes with their visibility rates. The upper left map gives a route search result on the Google Maps, and the bottom left frame presents a movie of a virtual 3D view.

7. CONCLUSIONS AND FUTURE WORK

In this paper we proposed a system which can recommend popular tourist spots and scenic sights based on Web information analysis, and calculate routes with the best view of a scenic sight. As shown in the experimental results, the tourist spot information was extracted with high accuracy. Moreover, by using three-dimensional data and geographical information from a digital map, the visibility rates of routes were calculated effectively. We also implemented a prototype of a navigation system which can successfully recommend tourist routes with good scenery.

There are still a number of interesting work which need to be carried out. Using a numeric value for the visibility of a scenic sight for a node is expected to improve the results. The generation of route candidates which allows multiple passing spots is also an important extension. We plan to infuse more factors in addition to temporal information, such as weather, seasons, and traveling time.

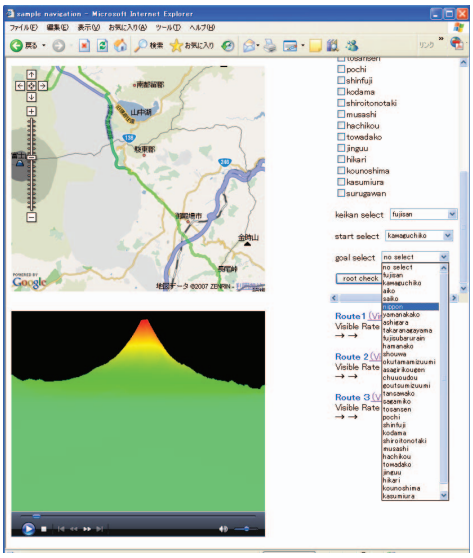


Fig. 8. Web browser interface

8. ACKNOWLEDGMENT

This work was supported in part by SCOPE (Ministry of Internal Affairs and Communications), and by MEXT Grant-in-Aid for Scientific Research in Priority Areas, Cyber Infrastructure for the Information-explosion Era (#19024073).

9. REFERENCES

[1] A. Jameson, "More than the Sum of Its Members: Challenges for Group Recommender Systems," In *AVI 2004)*, pp. 48–54, 2004.

[2] T. Tezuka, T. Kurashima, and K. Tanaka, "Toward Tighter Integration of Web Search with a Geographic Information System," In *WWW 2006*, pp. 227–286, 2006.

[3] A. Maruyama, N. Shibata, Y. Murata, K. Yasumoto and M. Ito, "A Personal Tourism Navigation System to Support Traveling Multiple Destinations with Time Restrictions," In *AINA 2004*, pp. 18–22, 2004.

[4] M. Kawabata, R. Nishide, M. Ueda and S. Ueshima, "Graph-based Approach to Context-adaptable PNS and Its Application Scenarios," In *ICDE Workshops 2005*, pp. 1250, 2005.

[5] Y. Akasaka and T. Onisawa, "Individualized Pedestrian Navigation Using Fuzzy Measures and Integrals," In *ICSMC 2005*, pp. 1461–1466, 2005.

[6] Y. Hosokawa, N. Takahashi and H. Taga, "A System Architecture for Seamless Navigation," In *ICDCS Workshops 2004*, pp. 498–504, 2004.

[7] STIMS. `http://www.mm.ics.saitama-u.ac.jp/stims/`