A Street Name-Based Summarization Method for Voice Navigation

Tomoya Sukigara
Nagoya Institute of Technology
Aichi, Japan
e-mail: t.sukigara.939@stn.nitech.ac.jp

Yonghwan Kim
Nagoya Institute of Technology
Aichi, Japan
e-mail: kim@nitech.ac.jp

Daisuke Yamamoto
Nagoya Institute of Technology
Aichi, Japan
e-mail: daisuke@nitech.ac.jp

Naohisa Takahashi
Nagoya Institute of Technology
Aichi, Japan
e-mail: naohisa@nitech.ac.jp

Abstract— **While existing voice navigation systems provide simple and easy-to-understand directions, they have a disadvantage in that voice directions tend to be long when the number of left-right turns is large. The objective of this study is to realize a summarization method for voice directions that also provides a geographic image of the route to the destination. The system generates multiple candidate routes and simplifies the voice directions based on street names. Furthermore, we propose a method to evaluate multiple candidate routes and determine the optimal route using evaluation formulas that indicate the degree of summarization of both the route and text of directions. Experimental results showed that the proposed method yielded a higher route recall agreement rate in participants. Although the optimization method was effective, the evaluation formulas did not work well for some paths.**

Keywords-Voice navigation; Road generalization; Turn by Turn; Shortest path problem; Geographic Information System.

I. Introduction

In recent years, voice navigation systems, such as Google Maps, have become increasingly popular. Many of existing voice navigation systems employ a turn-by-turn style directions method [1], in which 1) a route is determined primarily based on a shortest path algorithm [2], and 2) the distance and direction to the next intersection are repeated by a voice; this enables guidance to a specified destination. While this has the advantage of simple and easy-to-understand guidance, it has the disadvantage of a long voice navigation text if the number of left-right turns is large. It is better to keep voice directions as short and clear as possible, especially when considering the use of voice directions on fixed digital signage, which has become increasingly popular in recent years. For example, the Mei-chan [3] and Shaberu Bus Stop [4] signage-type voice navigation systems, installed at the main gate of our university, allow users to receive voice directions while viewing a route to the required destination on a map on a screen. In other words, when a user prefers a geographical image of a route, detailed voice directions are not necessary, but summarized, short, and easy-to-understand voice directions should be generated.

The objective of this study is to realize a summary of spoken directions that also provides a geographic image of the route to the destination. In order to achieve the above objectives, we previously proposed a method [5] that 1) generate an easy-to-understand route that is not necessarily the shortest route, but is suitable for summarization, and 2) summarize the route so that the voice direction fits within the specified number of characters.

In this study, we adopt the following three approaches: 1) generate multiple candidate routes with fewer right/left turns than the shortest route using the n-Min Stroke Shortest Path (n-MSSP) algorithm; 2) simplify the routes based on street names, considering street names to be key points for generating voice directions; and 3) based on evaluation formulas that indicate the degree of summarization of the route, the best route is determined. We believe that summarizing voice directions will contribute to improved pedestrian navigation and more advanced digital signage-based systems.

In this paper, we describe related studies in Section 2, describe the proposed system in Section 3, discuss the evaluation of the proposed system in Section 4, and summarize in Section 5.

II. Related work

Although there have been many studies on summarizing natural language sentences and on road generalizations [6-8], few studies have focused on summarizing voice directions.

A road generalization is a method that draws only the major roads in a road network based on the length of road strokes. A stroke is a grouping of road networks based on cognitive psychology [6] [7], and represents a following path. Zhang et al. [8] realized road generalization by selecting distinctive roads based on their connective relationships. Within studies into the applications of road generalizations, methods based on facility search results [9] [10] and in a Fisheye-view format [11] have been proposed. However, road generalization only selects roads based on their geometry and connectivity, and does not consider the implications on voice directions.

There are also several studies on summarization in the Geographic Information System (GIS) field; Kaigun et al. [12] proposed social data and sentence summarization methods for travel route recommendations. Other existing methods include summarizing the structure of a road network in an easy-to-read manner [13], generating customized cognitive turn-by-turn directions based on a user's travel history [14], and summarizing and generating meaningful sentences from travel data [15].

III. PROPOSED SYSTEM

The configuration of the proposed system is outlined in Figure 1. First, given a departure point and an arrival point, a set of the n (n=1,2,...) MSSP routes is computed based on the n-min stroke shortest path (n-MSSP) function; these are considered as candidate routes. Then, the candidate routes are simplified based on a street-based path simplifier function. If the simplification of the candidate paths is not sufficient, a Vertex based path simplifier function, based on the number of vertices, is used to further simplify them. Next, the text for the simplified candidate routes' navigation is summarized using on a word count-based text summarization function. Finally, the optimization of voice directions is achieved using the optimal path determination function based on evaluation formulas. This generates a summarized route and the corresponding guidance text. The MMDAgent [16] and MMDAgent EDIT [17] mechanisms were used to generate the voice directions.

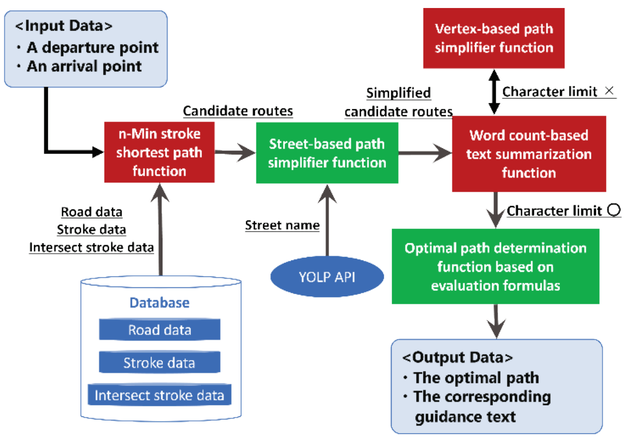

Figure 1. Configuration of the proposed system.

A. n-Min Stroke Shortest Path Function

The n-MSSP search method [18] searches for multiple candidate routes between two points based on given rules, which will be outlined later. The n-MSSP function can obtain the shortest route from the routes that require fewer right or left turns than the shortest route.

Specifically, a road set, stroke set, etc., are obtained from the road database, and the algorithm proposed by Hiura et al. [18] is used to search for and obtain a set of the n-MSSP routes. The definition of the n-MSSP is as follows.

- 1-MSSP

The shortest path among the paths with the fewest number of left or right turns between two points.

- n-MSSP

The path with the smallest amount of left or right turns among the paths shorter than the [n-1]-MSSP route.

B. Street-Based Path Simplifier Function

The street-based path simplifier function simplifies routes by using major roads with street names as key points. In other words, even if there are multiple routes to a major road, the system greatly simplifies route guidance based on the hypothesis that any route can reach the major road if it has a general direction.

The simplification algorithm is shown in Figure 2. First, for every route obtained by the n-MSSP search function, data on the behavior of the current node is stored when progressing to the next node. Next, for each path, the existence of a street name for the next link connected to the current node is checked sequentially in the direction from the end point to the beginning point. If a street name exists, the loop terminates. If there is no street name, the search continues until the next node of the starting point is reached.

n_i : The i-th node
a_i : Action at the i-th node

for (i ← 0 to n-1) {
  a_i= next action at the i-th node (go straight : -1, turn left : 0, turn right : 1)
  Store ∞ at the start and end points (a_0,a_(n-1))
}
for(i ← n-2 to 1) {
  if($a_i \neq -1$) {
   Check if there is a street name on the link in the direction to the end point connected to that node.
   if (Street name ≠null) {
    exit
   }
  }
}

Figure 2. Street name-based route simplification algorithm.

C. Word-Based Text Summarization Function

A word-based text summarization function generates an optimal guide text within the specified number of characters. The five key components of a guide text are "minimum information", "time required", "total distance", "distance", and "turning point". "Minimum information" refers to the information expressed only in terms of the direction of the turn, and the other four components are considered to be supplementary information to make the directions easier to understand.

Taking the initial value of each component of the guide consisting of only the "minimum information," we seek a combination that maximizes the sum of values under the condition that the word limit is satisfied. In other words, the problem of generating a guidebook is reduced to the 0-1 knapsack problem, where "number of characters = weight". The value of each component of the guide text is determined based on a hierarchical analysis method [19], as shown in

 ISBN: 978-1-61208-983-6

Figure 3. This allows us to determine the optimal combination for the audio guide text using dynamic programming.

A is more important than B.				A is as important as B.				B is more important than A.
9	7	5	3	1	1/3	1/5	1/7	1/9

A \ B	total distance	time required	distance	turning point	Geometric mean	Weight
total distance	1	1/3	1/9	1/9	$\sqrt[4]{1 \cdot \frac{1}{3} \cdot \frac{1}{9} \cdot \frac{1}{9}} \approx 0.25$	$\frac{0.25}{6.46} \approx 0.04$
time required	3	1	1/5	1/7	$\sqrt[4]{3 \cdot 1 \cdot \frac{1}{5} \cdot \frac{1}{7}} \approx 0.54$	$\frac{0.54}{6.46} \approx 0.08$
distance	9	5	1	1/3	$\sqrt[4]{9 \cdot 5 \cdot 1 \cdot \frac{1}{3}} \approx 1.97$	$\frac{1.97}{6.46} \approx 0.30$
turning point	9	7	3	1	$\sqrt[4]{9 \cdot 7 \cdot 3 \cdot 1} \approx 3.70$	$\frac{3.70}{6.46} \approx 0.57$
				Sum	6.46	1

Figure 3. Value setting for components.

D. Vertex-Based Path Simplifier Function

The vertex based path simplifier function uses the Visvalingam-Whyatt algorithm [20] to simplify a path by reducing the number of vertices on the guided path one by one. The Visvalingam-Whatt algorithm is a method that geometrically simplifies Polyline. Although geometrical simplification omits some guidance, such as small turns, we considered it to be sufficient for the purpose of this study, i.e., to grasp the geographical image of the guided route.

Specifically, the Visvalingam-Whyatt algorithm is applied iteratively until the number of vertices within the character limit of the audio guidance text generated by the "minimum information" is reached. For each decrease in the number of vertices, a guide sentence is generated, which is then judged regarding whether it satisfies the termination condition (i.e., the character count limit) or not.

E. Optimal Path Determination Function

In the optimal path determination function, all candidate routes are evaluated using the evaluation formulas to determine the optimal route and guide sentences. First, the function obtains the value A_i from the result of the street based path simplifier function, which indicates how well the route is expressed, i.e., how well it concisely captures the main points of the route. The definition of A_i is as follows:

$$A_i = \frac{X}{x_i},$$

where x_i is the number of components in the simplified guide for the i-th candidate route, and X is the number of components in the pre-simplified candidate route directions with the greatest number of right or left turns.

Next, we obtain the value B_i from the vertex-based path simplifier function, which is a measure of how much of the guide text is summarized in the parts that are not related to the main parts of the route. The definition of B_i is as follows:

$$B_i = \frac{n_i}{N_i},$$

where N_i and n_i are the number of vertices of the i-th path before and after applying the vertex based path simplifier function, respectively.

When determining the optimal route for a user, it is necessary to consider the balance between the number of left or right turns and the distance. Therefore, we define C_i and D_i, which are standardized numbers for left/right turns and distances, respectively, as follows:

$$C_i = \frac{c_i - \bar{c}}{\sigma_c},$$

where c_i is the inverse of the number of right/left turns, $\bar{c}$ is the mean of c_i, and σ_c is the standard deviation of c_i, and

$$D_i = \frac{d_i - \bar{d}}{\sigma_d},$$

where d_i is the inverse of the distance, $\bar{d}$ is the mean of d_i, and σ_d is the standard deviation of d_i.

Using these values, we evaluate each path using the following equation, called the evaluation formula (1):

$$E_i = A_i B_i + \omega(C_i + D_i), \qquad (1)$$

where ω is a small value, $A_i B_i$ represents how precisely and clearly the guidance text is expressed, and $C_i + D_i$ represents how well balanced the number of right and left turns and distances are.

IV. Evaluation

The purpose of our experiment was to verify whether the proposed method of summarizing directions is more effective for users to remember routes and for expressing directions more clearly than the conventional method.

A. Experimental Condition

An evaluation experiment was conducted on 12 subjects. Generated directions and map images were presented to the subjects, and 1) the time taken to memorize the route (average measurement time) was recorded; 2) the accuracy of subject' route recall (route matching rate) was analyzed; and 3) a 5-step questionnaire survey was conducted. Table 1 shows the combinations based on the comparison experiments.

Two types of paths (shortest paths and the ⌈n/2⌉-th n-MSSP) and a combination of three path simplification methods (no summary, vertex based path simplifier, street based path simplifier) were evaluated. The ⌈n/2⌉th n-MSSP is the ⌈n/2⌉th path obtained by the n-MSSP algorithm. Subjects performed evaluation experiments on different paths for each method. Methods 1-2, 1-3, 2-2, and 2-3 are the proposed methods.

The character count limit was set to 40 characters, which was considered to be an appropriate length for the voice guidance text based on a preliminary survey. The questionnaire questions, presented on a 5-point scale, were as follows:

1. Is the route easy to remember?
2. Is the guidance easy to understand?
3. Did the voice guidance help ease your anxiety about the road?

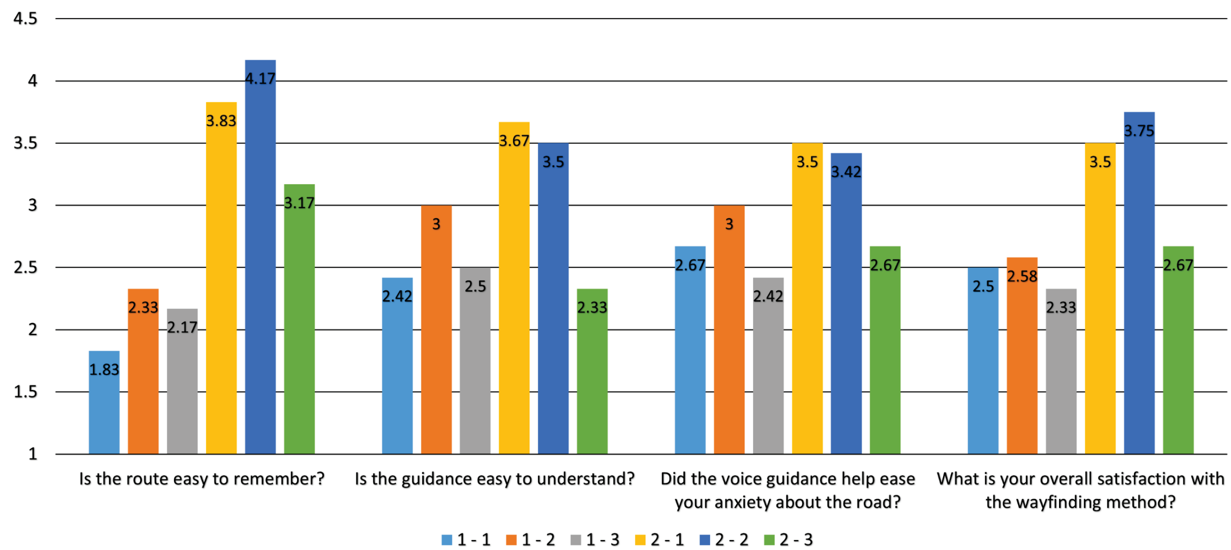

Figure 5. Questionnaire results.

4. What is your overall satisfaction with the wayfinding method?

TABLE I. PATH SEARCH ALGORITHM AND SIMPLIFICATION METHOD.

Method	Path search	Simplification method	Word count
1-1	Shortest path	normal	-
1-2	Shortest path	Vertex based simplifier	40
1-3	Shortest path	Street based simplifier	40
2-1	⌈n/2⌉-th n-MSSP	Normal	-
2-2	⌈n/2⌉-th n-MSSP	Vertex based simplifier	40
2-3	⌈n/2⌉-th n-MSSP	Street based Simplifier	40

B. *Experimental Result*

Figure 4 shows the results of the route matching rate. The route matching rate is the percentage of the actual route that matched the participant's memorized route after they listened to the voice guidance text. The methods based on the $\lceil n/2 \rceil^{th}$ n-MSSP (2-1, 2-2, and 2-3) yielded higher route matching rates than the methods based on the shortest paths (1-1, 1-2, and 1-3). Since the method based on the $\lceil n/2 \rceil^{th}$ n-MSSP requires fewer left or right turns than the shortest path, this result suggests that the path with fewer left or right turns is easier for a user to remember.

Table 2 compares the average time taken for a participant to memorize the route by looking at the voice guidance text and route map for each method. Users required less time to memorize the route for Methods 1-2 and 2-2 than for the other methods. This result indicates that the vertex-based path simplifier function is effective in reducing the time required to memorize a route.

Finally, the results of the questionnaire are shown in Figure 5. For the question "Is the route easy to remember?", the summary generated from the vertex-based path simplifier function (Methods 1-2 and 2-2) obtained the highest evaluation for each route. The summary based on the street-based path simplifier function (Method 2-3) did not achieve such a high value. This indicates that there is room for improvement of the street name-based route summarization algorithm. In response to the question "Is the guidance easy to understand?", Method 1-2 received a higher rating when the number of right or left turns was higher.

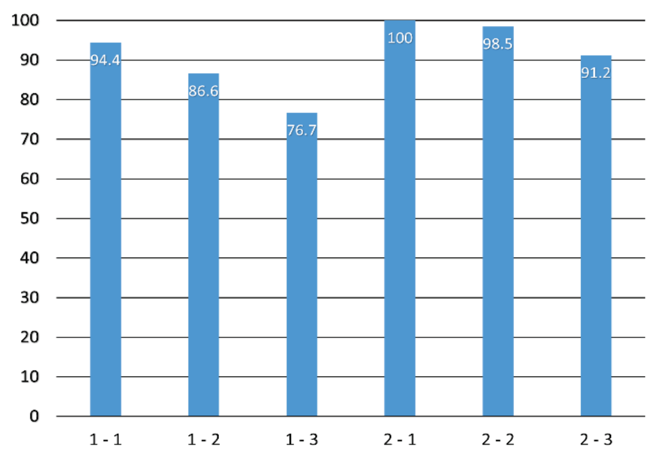

Figure 4. Route matching rate: the degree of agreement between the memorized route and the actual route after listening to the voice guidance text.

TABLE II. TIME TAKEN TO MEMORIZE THE ROUTE.

Method	1-1	1-2	1-3	2-1	2-2	2-3
Mean time	69.6	48.4	65.4	47.3	36.5	51.2

C. *Evaluation of Optimization Methods Based on Evaluation Formulas*

Next, we evaluated the optimization function based on the evaluation formula (Eq.1). 12 subjects were interviewed about which of the routes between Destinations A and F generated by the following three methods they found easier to understand:

Shortest path The route determined by the shortest path algorithm.

n-MSSP path（proposed method） The route with the fewest number of right/left turns, determined by the n-MSSP algorithm.

Optimal path (proposed method） Overall optimal path generated based on the evaluation equations.

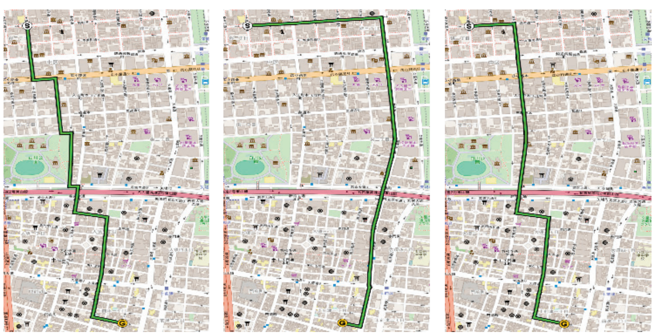

Figure 6. Example of simplified route (destination B). Left: shortest path method. Middle: n-MSSP method. Right: optimal path method.

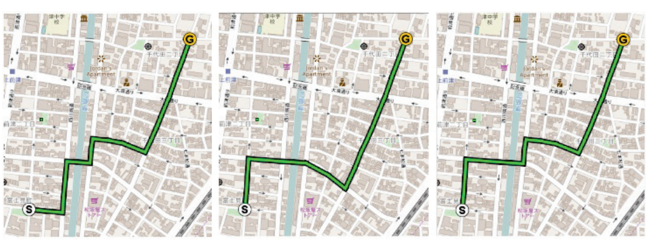

Figure 7. Example of simplified route (destination C). Left: shortest path method. Middle: n-MSSP method. Right: optimal path method.

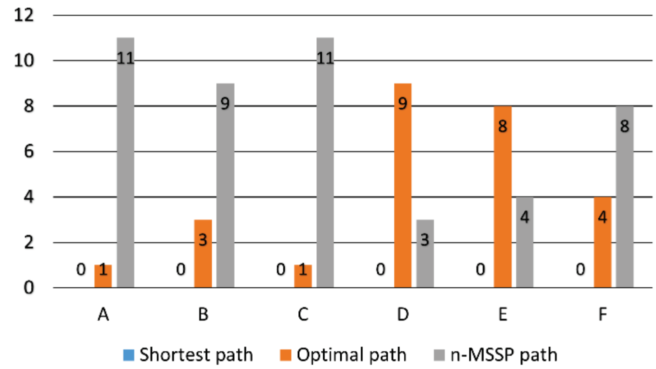

Figure 8. Interviews results regarding optimal routes.

Examples of the paths used in the experiments are shown in Figures 6 and 7, and the experimental results are shown in Figure 8. The n-MSSP path and the optimization path yielded superior results compared with the baseline method, the shortest path. Theoretically, the global optimization method should exhibit the best results, but for some paths, the simpler minimum stroke path was superior. This result suggests that the evaluation formula does not work effectively for some paths, and that improvements can be made.

V. CONCLUSION

In this paper, we proposed a method for summarizing voice directions. Using the n-MSSP (n-min stroke shortest path) algorithm, multiple candidate routes were generated that required progressively fewer right or left turns than the shortest path. We regarded the street names as a key point in generating voice guidance; hence, the routes were simplified based on the street names. Furthermore, the best route and its guide text are determined from multiple candidate routes using the evaluation formula for the degree of summarization of the route.

Through evaluation experiments, a total of six combinations of two different routes and three different guide text generation methods were evaluated. The experimental results showed that the proposed method yielded a higher route recall agreement rate in participants. In particular, the combination of the $\lceil n/2 \rceil^{th}$ n-MSSP and the vertex based path simplifier function was effective. Although the optimization method was effective, the evaluation formulas did not work well for some paths. This suggests that the evaluation formula requires improvements.

In the future, we plan to implement a new summarization method that better combines vertex- and street name-based summarizations, and to improve map rendering to highlight road names and other key points of the route.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Numbers JP19H04115 and JP21K19766.

REFERENCES

[1] C. A. Rizzardo, H. A. Colle, E. A. McGregor, and D. Wylie, “Using a single map display both for navigational planning and for turn-by-turn vehicle guidance: Configural spatial

knowledge acquisition," Journal of Experimental Psychology. Applied, Vol. 19, No. 4, pp. 301-319, 2013.

[2] E. W. Dijkstra, "A note on two problems in connexion with graphs," In Numerische Mathematik, Vol. 1, pp. 269-271. 1959.

[3] K. Oura, D. Yamamoto, I. Takumi, A. Lee, and K. Tokuda, "On-Campus, User-Participatable, and Voice-Interactive Digital Signage," Journal of Japanese Society for Artificial Intelligence, Vol. 28, No. 1, pp. 60-67, 2013. (in Japanse)

[4] D. Yamamoto et al., "Development of talking bus route guidance system and its evaluation by demonstration experiment," DICOMO, Vol. 2018, pp. 1872-1878, 2018. (in Japanese)

[5] R. Hara, (supervisor: D. Yamamoto), "A proposal of voice navigation method by simplifying routes and guidance sentences for a signage-type voice guidance system," Graduation thesis, Nagoya Institute of Technology, 2020. (in Japanese)

[6] R. C. Thomson and R. Brooks, "Efficient generalisation and abstraction of network data using perceptual grouping," Proceedings of the 5th international conference on GeoComputation, pp. 23-25, 2000.

[7] R. C. Thomson and R. Brooks, ""Good continuation" principle of perceptual organization applied to the generalization of road networks," Proceedings of the 19th international cartographic conference, pp. 1215-1223, 1999.

[8] Q. Zhang and P. Fisher, "Road network generalization based on connection analysis," in Proceedings of the 11th Interantional symposium on Spatial Data Handling, pp. 343-353, 2005.

[9] D. Yamamoto, M. Murase, and N. Takahashi, "On-Demand Generalization of Road Networks based on Facility Search Results," IEICE Transactions on Information and Systems, Vol. E102-d, No. 1, pp. 99-103, 2019.

[10] M. Murase, D. Yamamoto, and N. Takahashi, "On-demand generalization of guide maps with RoadNetworks and category-based web search, results," in Proceedings of the 14th international symposium on Web and Wireless Geographical Information Systems (W2GIS 2015), Vol. 19, pp. 53-70, 2015.

[11] D. Yamamoto, S. Ozeki, and N. Takahashi, "Focus+Glue+Context: An improved fisheye approach for web map services," in Proceedings of the ACM SIGSPATIAL GIS 2009, Seattle, Washington, pp. 101-110, 2009.

[12] K. Fu, Y. C. Lu, and C. T. Lu, TREADS: "A safe route recommender using social media mining and text summarization," in Proceedings of the 22nd ACM SIGSPATIAL international conference on Advances in Geographic Information Systems, pp. 557-560, 2014.

[13] A. Asahara, S. Shimada, and K. Maruyama, "Macroscopic structural summarization of road networks for Mobile traffic information services," in Proceedings of the 7th international conference on Mobile Data Management, pp. 42-49, 2006.

[14] Y. Li, H. Su, U. Demiryurek, B. Zheng, K. Zeng, and C. Shahabi, "PerNav: A route summarization framework for personalized navigation," in Proceedings of the 2016 international conference on management of data, pp. 2125-2128, 2016.

[15] H. Su et al., "Making sense of trajectory data: A partition-and-summarization approach," in Proceedings of the 2015 IEEE 31st international conference on Data Engineering, pp. 963-974, 2015.

[16] A. Lee, K. Oura, and K. Tokuda, "MMDAgent—A fully open-source toolkit for voice interaction systems," in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 2013, pp. 8382-8385, 2013.

[17] R. Nishimura, D. Yamamoto, T. Uchiya, and I. Takumi, "Web-based environment for user generation of spoken dialog for virtual assistants," EURASIP Journal on Audio, Speech, and Music Processing, Vol. 2018, No. 1, article number 17, 2018.

[18] Y. Hiura, (supervisor: D. Yamamoto), "Proposal of an efficient nth min stroke shortest path search method," Master's thesis, Nagoya Institute of Technology, 2020. (In Japanese)

[19] T. L. Saaty, "A scaling method for priorities in hierarchical structures," Journal of Mathematical Psychology, Vol. 15, No. 3, pp. 234-281, 1977.

[20] M. Visvalingam and J. D. Whyatt, "Line generalisation by repeated elimination of points," The Cartographic Journal, Vol. 30, No. 1, pp. 46-51, 1993.

 ISBN: 978-1-61208-983-6