InfoCruise: Information Navigation Using a Focus Facet Based on Context

Izumi Kohno Yoji Miyazaki Teruya Ikegami Masaki Hara Koji Kida
NEC Corporation
8916-47, Takayama-Cho, Ikoma, Nara, 630-0101, Japan
{kohno@ay, y-miyazaki@bc, t-ikegami@ct, m-hara@cw, kida@da}.jp.nec.com

Abstract

To find desired information by mobile phones, it is necessary to set search keywords easily and to explore a search if a user's information needs are not well defined. We propose an information navigation method to help users successfully find information by mobile phones. Our proposed method presents users a focus facet by analyzing the context about contents, users, and dialogs. The focus facet is presented each time users refine a search. They can select a keyword from the facet, check the search results, and then define their ambiguous needs through interaction. We compared our method with traditional search methods. InfoCruise effectively discovers desired information, because its discovery rate and the satisfaction rating are higher than those of the traditional search methods. InfoCruise efficiently searches, because both the keyword correction rate and the number of times users changed facets in it are lower than those in traditional search methods.

Keywords information retrieval, exploratory search, user interface, faceted navigation, user context

1. Introduction

With the spread of the Internet, a great deal of information about products or shops has become available online that people can locate by search systems, not only by PCs but also by mobile phones. For example, when a user wants to go out to eat, he might search for an appropriate restaurant by mobile phone. But several problems exist with mobile searches. First, since the user's search purpose is often ambiguous, he can't search well. When searching for a restaurant, many users don't know its name or fail to stumble upon the right search word. The second problem is the constraints of mobile devices. Users have difficulty viewing long lists of search results on the small screens of mobile phones. They might also feel annoyed by repeatedly setting and changing search keywords because a mobile phone's communication speed is low.

Faceted navigation is a solution when a user's search purpose is ambiguous [2], [3], [4], [5]. This search method uses the metadata of information. Metadata have several facets, which are attributes in various orthogonal sets of categories. Faceted navigation displays the aspects of a current results set, so users can switch easily between searching and browsing. It also only shows the populated facets when users drill down search results by facet. So, users can explore search results without dead-ends. Faceted navigation is a good search method when a user's search purpose is ambiguous, because users can combine querying and browsing to clarify their purposes and find information. Most faceted navigation systems show all facets and many facet values on a PC screen. They need a wide screen area. But since mobile phones have small screens, showing all facets, such as current faceted navigation systems do, seems difficult.

As a solution, we propose InfoCruise, an information navigation method to help users successfully find information by mobile phones [1]. We support searches on small screens and when the search purpose is ambiguous. Our proposed method calculates the priority of facets by analyzing the context about contents, users, and dialogs and presents keywords in a focus facet each time a search is refined.

The contents of this paper are organized as follows. Section 2 describes the kinds of contexts as factors that affect the priority of facets. Section 3 describes our navigation method's overview, and Section 4 describes our application system to search for restaurants. Section 5 describes a user study in which we evaluated our method's effectiveness and efficiency. Section 6 discusses the results of a user study, Section 7 describes related works, and Section 8 provides a conclusion.

2. Focus facet determined by context

Information has many attributes, and leveraging metadata for searching has received attention in recent years. Hearst et al. [4], [6] used metadata (facets and facet values) as a tool to guide users in formulating queries. Their interface uses metadata in a manner that allows users to both refine and expand the query. Users select some facets and decide facet values for their demand to search. But it seems difficult to show many facets on a mobile phone with a small screen, so deciding on a focus facet among all facets is crucial. We propose to select and present a focus facet based on the user context. The following three kinds of contexts as factors affect the priority of facets: contents, users, and dialogs.

2.1. Context about contents

The context about contents affects which facet will be focused on. One reason why formulating queries is difficult is a lack of knowledge about target information [7]. It is important for users to know characteristics of target information. We think characteristics of target information are the content distribution about various data attributes. For example, when a user finds a restaurant, which facet is important among style, price, or location depends on the content distribution of these attributes. If most of the restaurants around a certain location are Chinese (he might be in a Chinatown), the user should learn that Chinese is the major type of restaurant nearby. In that case, the user would rather search by the "style" facet first than by another facet.

We calculate the priority of facets using the content distribution of each facet.

2.2. Context about users

The context about the circumstances around users also affects which facet will be focused on. For example, when a user wants to find a restaurant around him, the "location" facet is important. If a user wants to find a restaurant at midnight, he would search for restaurants using "open time" first. When a user wants to find a restaurant while he is driving, he would search using "parking" first. If a user can search for restaurants with the facet that suits the user's context, he can find restaurants efficiently.

We prioritize the facet associated with user context. If we relate the "location" facet with the current position and the system detects a user's current position, the system prioritizes the "location" facet. We can relate the "open time" facet with the current time and the "parking" facet with the means of transportation.

2.3. Context about dialog

Context about dialog history also affects which facet will be focused on. If a user can't find the desired information on the first search, using a different facet from the first search might prove successful. For example, if a user can't find a good "Chinese" or "Japanese" restaurant using a "style" facet, she might have success using another facet such as "atmosphere" because she can change her aspect. We believe that context about dialog history, which means when a facet was used, is important for searching.

Instead of prioritizing the facet just used, we prioritize another appropriate facet. However, the priority of a facet used before gradually increases each time a user refines a search.

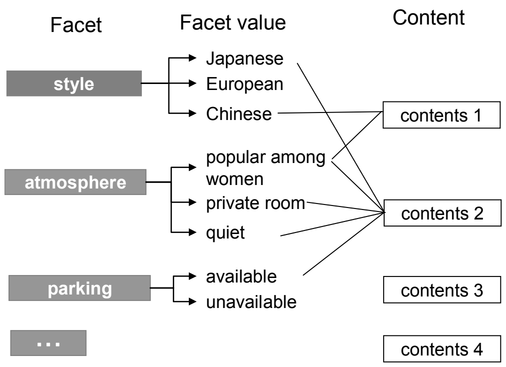

Fig. 1 Information structure

3. Navigation architecture overview

InfoCruise handles the information structure when contents have values in various orthogonal sets of categories. Fig. 1 shows an example where a restaurant has facet values in style, atmosphere, and parking facets. InfoCruise selects a focus facet among all facets and shows facet values as search keywords about the targeted information. InfoCruise also shows a dialog that explains why the facet and facet values were chosen.

Figure 2 shows the navigation architecture of InfoCruise. The system generates search keywords and dialogs in the following steps. The first step calculates the priority of facets by analyzing the context of the contents, users, and dialogs. In particular, the evaluated value of each facet is calculated using three

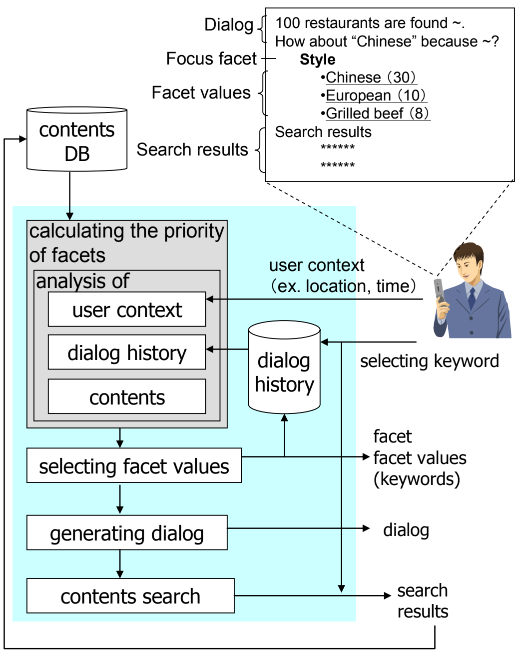

Fig. 2 Navigation architecture

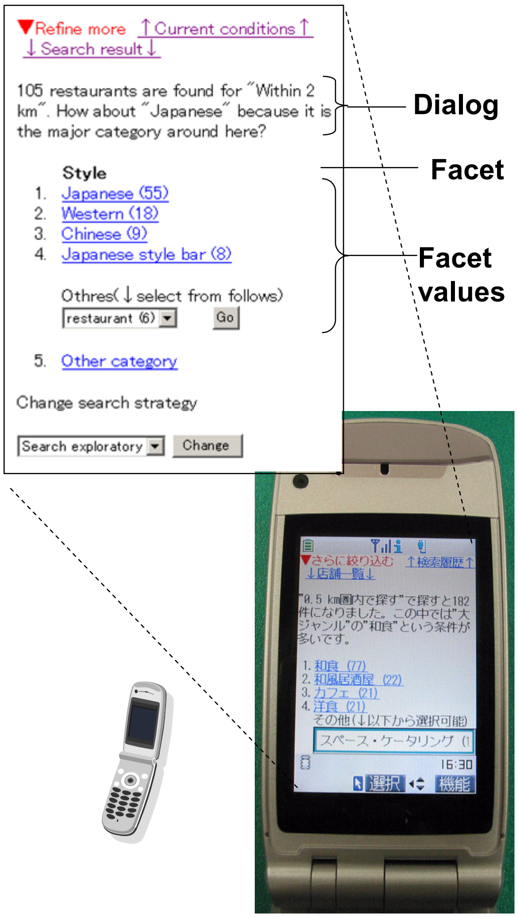

Fig. 3 Image of InfoCruise on a mobile phone

contexts, and all facets are ranked by the evaluated values. The rule for the calculation of the priority of facets is defined as the "search strategy." The second step selects facet values from the highest priority facet. The third step generates a dialog.

When a user selects a search keyword, the system refines the contents using that search keyword and generates new search keywords and dialogs for the search results once again. The contexts about the contents are updated using the refined content distribution of each facet, and the contexts about the dialogs are updated using the most recent history.

InfoCruise iteratively generates a focus facet and facet values as search keywords for the targeted information. Users can select desired search keywords, check the search results, and then define any ambiguous needs through interaction.

Our navigation architecture can define plural search strategies for some situations in an application. If users feel that the proposed search keywords aren't suitable, they can change the search strategy and select other search keywords.

4. Application system by mobile phone

To search for restaurants, we developed an application system composed of a server and a client. The client is a mobile phone. Search keywords and dialogs are created on the server computer and sent to a mobile phone. Fig. 3 is an image of InfoCruise on a mobile phone.

We used information about 4,500 restaurants in the Kansai area in Japan. The content has nine facets: location and open time, parking, price, interior design, service, cuisine, atmosphere, style, and sub-style. The facets and facet values were manually assigned. We also defined four user contexts: current position, current time, whether the user drives, and budget. These contexts will be acquired automatically in the future when the GPS function and electric money in mobile phones become widespread. However, we set these contexts manually in our prototype.

We prepared two search strategies for the retrieval of restaurants by mobile phones. One is an exploratory search strategy used when a user wants to know the most common type of restaurant in the area. In this strategy, we prioritize the facet with the largest difference among facet values because this is the most relevant information. Users select this strategy when they don't know the area around them very well.

The second strategy is a quick search strategy used when a user wants to decide on a restaurant as soon as possible. In this strategy, we prioritize the facet in which the number of contents for each facet value is the most equal to refine the contents with certainty.

The advantage of the exploratory search strategy is that users can discover new and unexpected information by traversing the information set. But a disadvantage of the exploratory search strategy is that searches are time-consuming. In contrast, the advantage of the quick search strategy is that users can search rapidly. However users can't extract new information using the quick search strategy.

We describe the method for ranking the facets of each search strategy as follows.

4.1 "Exploratory search" strategy

In the "exploratory search" strategy, the evaluated value of each facet is calculated by Formula (1) and the facet of the highest evaluated value is selected:

$$E_i = C_i \times U_i \times D_i \quad (1)$$

E_i is the evaluated value of facet i. C_i is the evaluated value of the context about the contents of facet i. U_i is the evaluated value of the context about the users of facet i. D_i is the evaluated value of the context about the dialog of facet i.

The evaluated value of the context about the contents, C_i, is calculated using the difference of the contents distribution among the facet values. If there is the largest difference among facet values in a facet, we think the facet has some characteristic that should be pointed out to users. C_i is calculated by formula (2):

$$C_i = \sum_j^{m_i} \left(N_{i,\max} - N_{i,j}\right)^2 \Big/ m_i \qquad (2)$$
$$N_{i,\max} = \max\left(N_{i,j}\right)$$

In Formula (2), m_i is the number of facet values in facet i. $N_{i,j}$ is the number of contents of facet value j in facet i, which is normalized by the summation of the contents of all facet values in facet i. For example in Fig. 4(A), value C_{style} is higher than value $C_{atmosphere}$, because there is greater variation in restaurant styles.

The evaluated value of the context about users, U_i, is defined as follows. We prioritize the facet associated with user context. If a user's current position and time are set, the evaluated value of the "location and open time" facet is prioritized. If the user is driving, the evaluated value of the "parking" facet is prioritized. If a price range is set, the evaluated value of the "price" facet is prioritized. Specifically, $U_{location\ and\ open\ tim}$ are 1.8, $U_{parking}$ is 1.5, U_{price} is 1.3, and U_{others}=1.

The evaluated value of the context about dialog Di is calculated by Formula (3). We define the priority of a facet used just before lower than other facets, but its priority gradually increases each time a user refines a search:

$$D_i = \begin{cases} \alpha_i \times n_i & \alpha_i \times n_i < 1 \\ 1 & \alpha_i \times n_i \geq 1 \end{cases} \quad (3)$$

In Formula (3), n_i is the number of dialogs since facet i was used (when facet i was used n_i=0), and α_i is the gradient. In this system α_i =0.01. We adjusted α_i so that a facet used before appears again after several dialogs.

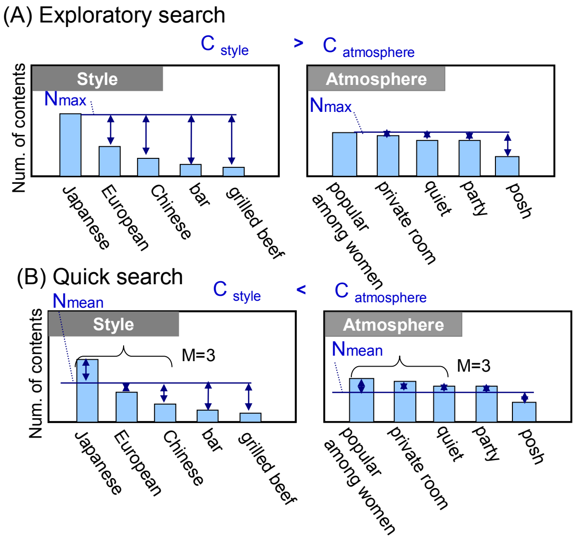

Fig. 4 Ranking based on content distribution

4.2 "Quick search" strategy

In the "quick search" strategy, the evaluated value of each facet is also calculated by formula (1) to select the facet of the highest evaluated value. The evaluated values of the context about dialog D_i, and users U_i are used in the same way as the "exploratory search."

The evaluated value of the context about contents C_i is calculated using the evenness of content distribution. A user can refine the contents with certainty if he selects a facet in which the number of contents for each facet value is equal. The facet should be pointed out to users. C_i is calculated by Formula (4):

$$C_i = \exp(-\sum_k^M \left(N_{i,mean} - N_{i,k}\right)^2 \Big/ M) \qquad (4)$$
$$N_{i,mean} = \sum_k^M N_{i,k} / M$$

In Formula (4), M is determined by the number of search keywords that can be shown on a mobile phone. We calculate the variance of the top M facet values of each facet. C_i is raised when the variance becomes small. $N_{i,k}$ is the number of contents of facet value k in facet i, which is normalized by the summation of the contents of M facet values in facet i. For example in Fig. 4(B), value $C_{atmosphere}$ is higher than value C_{style}.

4.3 Interactive navigation

Ranking facets and selecting keywords are done each time a user refines a search. If a user's current position is known, the system prioritizes the "location" facet and presents search keywords about location to refine the contents, as shown in Fig. 5. When a user selects a search keyword in the "location" facet, its priority is decreased because it was just previously used. Next, the system calculates content distribution and prioritizes the facet with greater variation such as "cuisine" under the "exploratory search" strategy.

4. 4 Example of operation

We now explain an example of our system operation. In Fig. 6(A), a user sets his current position, budget, date, and time as user contexts. He also selects his favorite search strategy from "search exploratory" or "search quickly." The concept and purpose of each search strategy were explained to users in advance. His current position and time are set, so the evaluated value of the "location and open time" facet is prioritized. In Fig. 6(B), the system presents such dialogs to users as "If you select 'within 2 km' of Kyoto, 105 restaurants are available" and search keywords about distance. The system automatically searches for restaurants open at the current time and calculates the distance from the current position to each shop. The system defines the presentation number of shops and distance such as "105" restaurants and "2" km. When a user selects a search keyword "within 2 km," the system generates a focus facet and facet values once again to refine 105 contents. In Fig. 6(C), the system presents to a user a focus facet "style" with the largest difference among facet values under the search exploratory strategy. The system presents a dialog such as "105 restaurants are found within 2 km. How about Japanese because it is the most common type around here?" When a user selects "Japanese" as a search keyword, the system presents a focus facet "cuisine" with the largest difference among facet values to refine the 55 contents. Sometimes after a user

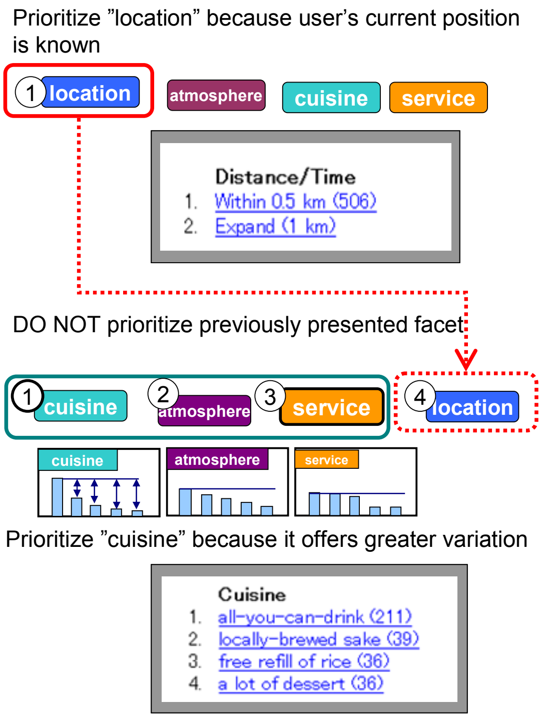

Fig. 5 Ranking a facet iteratively

selects keywords and the number of contents are reduced well (e.g., under five contents), the system stops proposing search keywords, as shown in Fig. 6(E).

If the user feels that the proposed search keywords aren't suitable, she can change the search strategy, for example, from an exploratory to a quick search strategy (Fig. 6(C)). The system newly generates a focus facet and facet values to refine the same 105 contents and presents an "atmosphere" focus facet in which the number of contents for each facet value is the most equal. In Fig. 6(F), the system presents search keywords about atmosphere.

5. User study

We evaluated our information navigation method from the viewpoints of effectiveness and efficiency. We compared InfoCruise to two experimental systems using traditional search methods. One system searches by category (facet) using a pull down menu. All categories and search conditions are shown fixedly at all times. The other resembles traditional faceted navigation systems in which only one facet is shown and the appearance order of the facets is fixed, for example alphabetically.

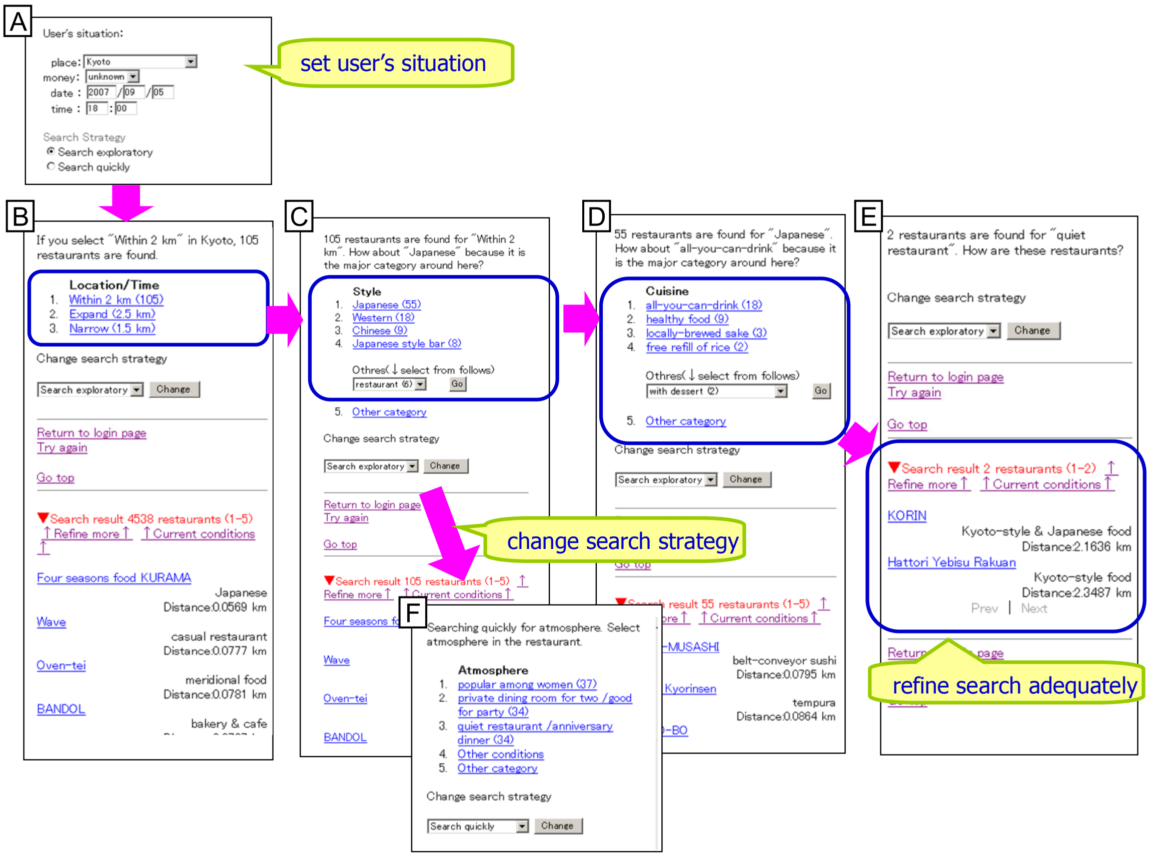

Fig. 6 Example of search flow

5.1. Measures

Some measures were defined to evaluate our system's effectiveness and efficiency.

5.1.1. Effectiveness. Since InfoCruise is an exploratory search system, it has difficulty defining what contents a user wants before a search. Since we don't know the correct answer, we can't use such traditional search measures as relevance or recall ratios. Therefore, we measured the discovery rate and the satisfaction rating to evaluate effectiveness. The discovery rate is the number of tasks in which users found the content versus the number of all tasks. It does not matter whether the content found by the user is what the user wanted before the search. The satisfaction rating is a subjective ranking in which a score of 5 is defined as satisfied and 1 as dissatisfied. If a user can't find a restaurant, the rating is 0.

5.1.2. Efficiency. InfoCruise is an exploratory search system in which users repeat a search and confirm its results until they get their desired information. They clarify their needs through a process, so reducing the total operation time is not important. We used two indexes to measure whether futile operations decreased. First, we measured the correction rate to evaluate the efficiency. InfoCruise shows search keywords and the number of search results for each keyword to refine the current contents to help users preview their search results. Since users can select appropriate keywords using InfoCruise, so they won't need to correct keywords during the search. The correction rate is the number of corrected search keywords versus the number of search keywords set in one search. Second, we measured the number of times users changed facets for InfoCruise compared to traditional faceted navigation methods. InfoCruise analyzed the facet priority using context. This method will probably reduce the times users changed facets.

5.2. Experiment

Two kinds of experiments were conduced to compare InfoCruise with traditional search systems.

5.2.1. Experiment 1: comparison with search by category system. We evaluated InfoCruise's effectiveness for a traditional search by a category system using the discovery rate and the satisfaction rating and its efficiency using the correction rate. The two systems had the same database and a similar UI (Fig. 7). The contents were located at the top of the screen and the search conditions at the bottom. The compared search system fixedly showed its categories and search conditions at all times. Users selected conditions using a pull-down menu and a "GO" button.

The 16 participants, who ranged in age from 20 to 40, commonly used mobile phones but they were not PC or Internet experts.

Participants searched for an appropriate restaurant under a designated situation. The tasks were prepared in five abstraction levels for search purposes to evaluate the InfoCruise's effectiveness when user needs were ambiguous. As shown in Fig. 8, each task defines a situation that designated some search conditions. Task T1 is the most abstract, which means the user needs were ambiguous. The situation is only designated by the location and date conditions. T5 is the most concrete, so its situation is designated by location, date, and two search conditions.

In our experiment, users executed five tasks (one for each task level) for InfoCruise and the compared system. Users pushed the "complete" button when they found a desired restaurant or users pushed the "give up" button when they wanted to quit. To measure the discovery rate, users could quit whenever they chose. After finishing one task, they completed a questionnaire containing satisfaction ratings and comments and then executed the next task. After finishing all five tasks of one experimental system, they moved to the other experimental system. The users executed five different tasks on the two experimental systems.

5.2.2. Experiment 2: comparison with faceted navigation system. We evaluated InfoCruise for a traditional search system like faceted navigation. InfoCruise presents user search keywords in a focus facet based on context. In contrast, the compared system shows search keywords in one facet whose appearance order is fixed alphabetically. If the ranking method of facet in InfoCruise is useful, the times that users change facets may be reduced. The main purpose of the experiment is to evaluate efficiency using the times users changed facets. We also measured the discovery rate, the satisfaction rating, and the correction rate for the traditional search system like faceted navigation.

The InfoCruise interfaces are shown in Fig7 (a). The interfaces of the compared system were the same as InfoCruise, but the appearance order of the facets was fixed alphabetically. If the users don't like the

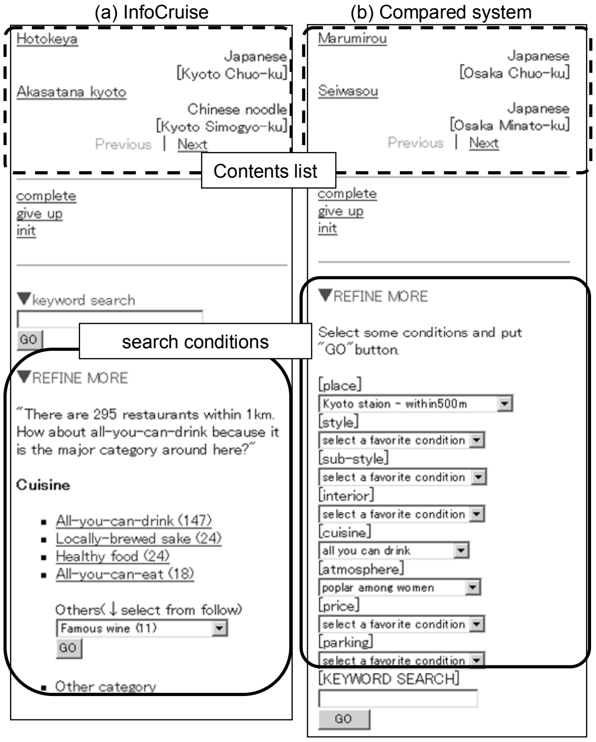

Fig. 7 Interfaces of two experimental systems

Abstraction level

abstract

Task level	Designated condition	Example (terms underlined here only)
		Find your favorite restaurant under this situation.
T1	Location, Date (near)	It is <u>9:30 pm</u> on <u>Friday</u> now. You are in <u>Osaka.</u>
T2	Location, Date (distant)	You will go to <u>Kyoto</u> <u>at lunchtime</u> <u>next Sunday</u> .
T3	Location, Date, Driving	It is <u>12:30 am</u> on <u>Sunday</u> now. You are driving in <u>Nara</u>.
T4	Location, Date, One search condition	You will go to <u>Osaka</u> <u>next Friday evening</u>. Your budget is <u>under 5000 yen</u>.
T5	Location, date, Two search conditions	You will go to <u>Kobe</u> <u>next Saturday evening</u> <u>for a party.</u> Many people will come. There will be some <u>children</u>.

concrete

Fig. 8 Task examples of each abstraction level for search purposes

presented search conditions in one facet, they can change the facet by clicking the “Other category,” in both InfoCruise and the compared system.

The procedures of Experiment 2 were the same as in Experiment 1, but only tasks 1, 3, and 5 were used: the five participants were in their 20s and 30s.

5.3. Results

We describe the results of the two experiments from the viewpoint of effectiveness and efficiency.

5.3.1. Effectiveness. Table 1 shows the results of Experiment 1 comparing InfoCruise to a traditional search system showing all categories any time fixedly. The discovery rate and the average satisfaction rating in all tasks and at each task level are shown in Table 1. The discovery rate of InfoCruise in all tasks (87%) was 10% higher than the compared system showing all categories any time (78%). The satisfaction rating of InfoCruise in all tasks (3.42) was 0.6 point higher than the compared system (2.8). The difference between the two systems was significant ($p<0.05$). The results suggest that InfoCruise is effective for desired information discovery.

Based on the discovery rate of each task level in Table 1, InfoCruise’s ratings were better than the compared system in all task levels except Task 2. The difference between InfoCruise and the compared system in Task 1 was 25%, but the difference in Task 5 was only 4%. Furthermore, based on the satisfaction rating of each task level in Table 1, InfoCruise’s rating was only significantly better than the compared system in T1. This means that InfoCruise is more effective under T1 than under T5. T1 is the most abstract task. The result suggests that InfoCruise is especially effective when user needs are ambiguous.

Table 3 shows the results of Experiment 2 comparing InfoCruise to a traditional search system like faceted navigation in which the appearance order of facets is fixed alphabetically. The discovery rate of InfoCruise in all tasks (100%) was higher than the compared system (87%). The satisfaction ratings between two systems were not different.

5.3.2. Efficiency. Table 2 shows the correction rate of Experiment 1 comparing InfoCruise to a traditional search system showing all categories any time. InfoCruise’s correction rate, which was also calculated by each search strategy, was much smaller in all tasks (8.9%) than the compared system (37.4%). The correction rate of the “search exploratory” strategy was 5.9%, and the correction rate of the “search quickly” strategy was 16.7%. The average task completion time was 255 seconds for InfoCruise. In contrast, it was 291 seconds for the compared system. This means that since users can select appropriate keywords using InfoCruise, they don’t need to correct keywords during the search and they can find contents more quickly than using the compared system.

Table 3 shows the number of times that users changed facets and the correction rate of Experiment 2 comparing InfoCruise to a traditional search system in which the appearance order of facets is fixed alphabetically. The number of changing facets for InfoCruise (0.8) was smaller than the compared system

Table 1 Discovery rate and satisfaction rating in Experiment 1

Discovery rate

Task level	T1	T2	T3	T4	T5	All tasks
InfoCruise	94%	88%	100%	81%	75%	87%
Compared	69%	94%	88%	64%	71%	78%

Satisfaction rating

Task level	T1	T2	T3	T4	T5	All tasks
InfoCruise	3.69	3.31	3.93	3.33	2.87	3.42
Compared	2.38	3.44	3.19	2.36	2.57	2.80
p	0.04	0.77	0.23	0.10	0.76	0.03

Table 2 Correction rate in Experiment 1

	Correction rate
InfoCruise	8.9%
Compared	37.4%

Search strategy	Correction rate	Num. of tasks
exploratory	5.9%	42
quick	16.7%	27
exploratory +quick	15.6%	6

Table 3 Times user changed facets, correction and discovery rates, and satisfaction rating in Experiment 2

	Num. of changing facets	Correction rate	Discovery rate	Satisfaction rating
InfoCruise	0.8	7.3%	100%	3.73
Compared	1.7	18.4%	87%	3.47

(1.7). The operation load for changing facets was reduced. InfoCruise's correction rate was also smaller than the compared system.

These two results suggest that users can search efficiently using InfoCruise.

5.3.3. User Comments. We gathered user comments by questionnaires. There were four possible answers to the question, "In what situation might you use InfoCruise?" The answers were: "when I want to quickly decide on a restaurant" (5 users), "when I don't have any idea where to go" (3), "when I want to pick a restaurant as close as possible" (1), and "when I have to decide on a restaurant in a new place" (1). Furthermore, there were four answers to the question, "What are the good points of InfoCruise?" The answers were: "I could find new information" (3), "it is easy to use on a mobile phone" (3), "I enjoyed the search process" (2), and "I got various kinds of information" (2).

6. Discussion

First, we discuss the effectiveness. InfoCruise's discovery rate was higher than the compared system showing all categories any time. Based on the user comments, users want to use InfoCruise when they don't have any idea where they will go or when they have to decide on a restaurant in a new place. In this situation, selecting keywords is difficult using such a traditional search method as the compared system because the user information needs are ambiguous. On the other hand, users can select keywords from presented keywords, define their needs through interaction, and find a restaurant using InfoCruise in this situation, so the discovery rate might increase.

InfoCruise's satisfaction rating was also higher than the compared system showing all categories any time. Based on user comments, they enjoyed finding new information and the search process itself. Even if the restaurant found was unexpected, the user might be satisfied with it through the process. On the other hand, when empty result sets were shown using the compared system, users felt like they had hit a dead end. After such a dead end, users seemed to settle on a kind of restaurant that they really didn't want. So the satisfaction rating of the compared system might be lower than InfoCruise.

Next, we discuss the efficiency. The correction rate of InfoCruise was much smaller than the compared system showing all categories any time. Fig. 9 shows an example of a user's operation on both systems. The facets, the selected search conditions, and the number of search results are shown. In the compared system, a user changed the keyword set before on the same facet such as "location" and "style", perhaps because the number of search results was too great or too small. There was waste in the user operations, and the correction rate of the compared system was high. Otherwise, in InfoCruise, a user added keywords from different facets and gradually refined the contents. The user alternately selected a keyword from the location, cuisine, sub-style, and atmosphere facets. Fig. 10 shows the presented facet and search keywords in this experiment. The user's selection order followed the one presented by the system. Users can add keywords gradually after checking the search results, so the correction rate is low.

In addition, users can narrow down restaurants from an unexpected view with InfoCruise. For example in the experiment, one user, who selected a keyword in an atmosphere facet on the fourth search (Fig. 9), commented that "I don't have any restaurant preference." If the user hadn't searched for restaurants using InfoCruise, she wouldn't have selected a keyword from an atmosphere facet. The presented focus facet on InfoCruise inspired new interest in the user such as atmosphere.

Based on the correction rate of each search strategy, the correction rate for the "quick search" strategy was a little higher than for the "exploratory search" strategy. The search keywords presented for "quick search" might not fit the user's intention so well.

In the comparison between InfoCruise and the system in which the appearance order of facets is fixed alphabetically, the times users changed facets in InfoCruise was about half that of the compared system. So the algorithm for ranking facets might be useful to reduce user operation loads. If users aren't satisfied with the presented facet, they can change it. But in our experiment, many users tentatively selected a keyword from the facet presented by the system and corrected their selected keywords later. So the correction rate of the compared system was higher than InfoCruise.

In conclusion, since users enjoyed the search process and found a restaurant using InfoCruise even when their needs were ambiguous, both InfoCruise's discovery rate and satisfaction rating might increase. The result suggests that InfoCruise is effective for desired information discovery when a user's information needs are not well defined. In InfoCruise, a user added keywords from different facets and gradually refined the contents, so the keyword correction rate was low. The algorithm for ranking facets was also useful to reduce user operation loads. The result suggests that InfoCruise is efficient for searching for information.

Compared system

Facet / Search	Location	Style	Num. of search results
1st	**Osaka Within 1 km**	**European**	19
2nd	**Osaka Within 500 m** change	European	12
3rd	**Osaka Within 100 m** change	European	0
4th	**Osaka Within 500 m** change	**Japanese** change	101
5th	Osaka Within 500m	**European** change	12

InfoCruise

facet / Search	Location	Cuisine	Sub-style	Atmosphere	Num. of search results
A 1st	**Osaka Within 1 km**				708
B 2nd	Osaka Within 1 km	**All-you-can-drink**			349
C 3rd	Osaka Within 1 km	All-you-can-drink	**Japanese style pub**		69
D 4th	Osaka Within 1 km	All-you-can-drink	Japanese style pub	**Popular among women**	17

Fig. 9 Example of user's operation in our experiment

A

If you select "Within 1 km" in Namba, 708 restaurants are found.
Location/Time
- Within1 km（708）
- Expand（1.5 km）
- Narrow（0.5 km）

B

708 restaurants are found for "Within 1 km". How about "all-you-can-drink" because it is the major category around here?
Cuisine
- all-you-can-drink（349）
- locally-brewed sake（71）
- all-you-can-eat（44）

C

349 restaurants are found for "all-you-can-drink". How about "Japanese style pub" because it is the major category around here?
Sub-Style
- Japanese style pub（69）
- restaurant bar（23）
- Italian（21）

D

69 restaurants are found for "Japanese style pub". How about "good for party" because it is the major category around here?
Atmosphere
- good for parties（36）
- popular among women（17）
- private dining room for two（9）

Fig. 10 Presenting facets and search keywords in the experiment

7. Related work

Our purpose is to support searches on small screens and when the user's search purpose is ambiguous. Faceted navigation is an existing solution for ambiguous search purposes. Most faceted navigation systems show all facets and many facet values on a screen [3], [4], [5]. If a screen is big enough such as a PC, these techniques are very effective for searches. However, mobile phones have small screens, and showing all facets such as existing faceted navigation systems seems difficult. Our method, which selects a focus facet for which users can choose a search keyword easily, is effective on the limited screens of mobile phones. FaThumb [2] is faceted navigation for mobile searches in which a number keypad navigates the metadata. Fathumb needs to make hierarchical metadata that include up to nine facets at one hierarchy. Our method sets the priority facets using context. If the number of facets increases, our method can automatically present a refined facet.

InfoCruise has two main features. One is presenting keywords to lighten the load for generating search queries. The other is interactive navigation. We discuss the relation between our method and other works by their features.

From the viewpoint of the presenting keyword function, Google Suggest, one technology that provides additional keywords after user input [8], presents keywords related to the input keyword by analyzing co-occurrence keywords selected from user logs. InfoCruise also presents keywords in a focus facet. But our method generates keywords based on the context about the users and contents. We can present keywords adapted to user circumstances or content distribution.

From the viewpoint of easing the load for generating search queries, research exists on query-free retrieval that automatically gathers information without user input. Reference [9] selects information based on text being written or read in Emacs. Reference [10] selects information based on the user's physical context such as his location, the people in the area, or the date. These systems automatically retrieve contents; otherwise InfoCruise presents keywords to refine them. We believe when systems infer using context, presenting keywords and confirming interactively might provide satisfying search results.

From the viewpoint of interactive navigation, relevance feedback [11] repeatedly performs a search and locates results close to a user's demand by generating queries from the search results users judged as matching their demands. The content of search results has many features, so the system doesn't know which feature is important from the user's evaluation of the content. A user selects a search keyword made directly from the metadata of the contents in InfoCruise. Our method could bring search results closer to a user's demand more directly.

8. Conclusion

We proposed InfoCruise, an information navigation method that presents keywords in a focus facet by analyzing the context about the contents, users, and dialogs. Our method calculates the priority of facets by analyzing three kinds of contexts. We developed an application system to search for restaurants by mobile phone and evaluated our method compared to traditional search methods. In our experiments, both the discovery rate and the satisfaction rating of our system are higher than those of the traditional search methods, so InfoCruise is effective to discover desired information. Both the keyword correction rate and the number of times users changed facets in our method are lower than those in the traditional search methods, so users can search efficiently with InfoCruise.

As future work, developing metadata assigned technology is important. In this paper, we used manually assigned metadata. Keywords must be automatically extracted that describe content features as well as people do. Moreover, keyword categorization technology is needed to make pairs of facets and facet values. In our method, dialog history affects the priority facets, but it is only a short-term history in one search. Long-term dialog history must also be treated because it reflects user preference. Our future work also involves calculating the priority facets based on context containing long-term dialog history.

References

[1] Izumi Kohno, Yoji Miyazaki, Masaki Hara, and Teruya Ikegami, InfoCruise: Information Navigation Presenting a Focus Facet Based on Context, IARIA, International Conference on Advances in Computer-Human Interaction, ACHI 2008, pp. 45-52, 2008.

[2] Amy K. Karlson, George G. Robertson, Daniel C. Robbins, Mary P. Czerwinski, and Greg R. Smith, FaThumb: a facet-based interface for mobile search, Proc. SIGCHI, ACM Press, pp. 711-720, 2006.

[3] Ka-Ping Yee, Kirsten Swearingen, Kevin Li, and Marti Hearst, Searching and organizing: Faceted metadata for image search and browsing, Proc. SIGCHI, ACM Press, pp. 401-408, 2003.

[4] Marti Hearst, Ame Elliot, Jennifer English, Rashimi Sinha, Kirsten Swearingen, and Ka-Ping Yee, Finding the flow in web site search, Comm. of the ACM, 45, 9, pp. 42-29, 2002.

[5] Endeca. http://endeca.com/

[6] Marti Hearst, Clustering versus faceted categories for information exploration, Comm. of the ACM, 49, 4, pp. 59-61, 2006.

[7] Ryen W. White, Bill Kules, Steven M. Drucker, and M. C. Schrafel, Supporting exploratory search, Comm. of the ACM, 49, 4, pp. 36-39, 2006.

[8] Google Suggest: http://www.google.com/

[9] Bradley Rhodes and Thad Starner, The Remembrance Agent: A continuously running automated information retrieval system, The Proceedings of The First International Conference on The Practical Application of Intelligent Agents and Multi Agent Technology (PAAM '96), pp. 487-495, 1996.

[10] Bradley Rhodes, Using Physical Context for Just-in-Time Information Retrieval, IEEE Transactions on Computers, Vol. 52, No. 8, pp. 1011-1014, 2003.

[11] Rocchio J., Relevance feedback in information retrieval: In the SMART Retrieval System, Experiments in Automatic Document Processing, pp. 313-323, 1971