

Supporting Real Estate Search Through Automatic Information Suggestion

Goro Otsubo
LIFULL Co., Ltd.
Tokyo, Japan
ohtsubogoro@lifull.com

ABSTRACT

Searching real estate property is uncommon task for most of the users. As a result, the user is not familiar with the detailed search condition which is useful for search. In this paper, we propose to use voice recognition as a support for real estate property search. First user set vague search condition with GUI. Then system listens conversation between users. From the conversation, system extracts keyword and suggest detailed search condition and real estate information search results with those conditions. We will discuss system design, algorithm used to link spoken words and detailed search condition and preliminary test results.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous; See <http://acm.org/about/class/1998/> for the full list of ACM classifiers. This section is required.

Author Keywords

Voice Recognition; Multi modal; GUI

INTRODUCTION

In Japan, people will move to new real estate property only six times for the lifetime on average. In other words, searching real estate property is "An event that happens once in a decade". As a result it is very difficult to become familiar with special terms which are used to search real estate property.

In recent years real estate information sites on the Internet have been widely used. Using those sites, the user can access and search real estate information by themselves. However, search interface of those sites are like query interface of the property database. They have every function available to search the database, but they are not necessarily easy to use for all the user who has interest in real estate property.

Typical user can set the vague request to search the real estate property like the town they want to search, layout of property and price. Further, for example, it is also possible to set detailed search conditions such as "Pets allowed". However,

in most cases the user does not know what type of detailed search conditions are available and how to set them. As a result, they are not used effectively in the real estate search process.

To solve this issue, we assumed that we need support system which can recommend proper detailed search condition to assist real estate search. To do that, we have chosen to utilize the conversation between users. Asking detailed demand for new real estate property is sometimes intimidating for the user. Moreover, the user may not necessary know the proper word to search the real state information. Therefore, we assumed that the system which can extract detailed search condition from the casual user conversation would be effective to support the real estate property search.

Interfaces for searching information using voice have been studied for a long time, and commercial service which utilizes voice interface become popular in recent years due to popularization of smartphones and home devices. Recently speech recognition on the server side has become widespread, the accuracy of speech recognition has dramatically improved even for unspecified speakers.

However, in most cases speech recognition is used for simple web search and search in the app store[10]. This is due to the fact that even though the precision of speech to text conversion has improved, the next step which is "understanding meaning" remains as tough problem yet. According to the research by Luger et al.,[8] although the user's expectation for speech recognition is high, current speech recognition interface makes the user feel very stressful when using it.

Considering technical difficulty described above, we have chosen to use voice recognition as a support role. The user is not assumed to directly talk to the search interface. Instead, system listens the conversation between users. The system recognizes the keyword in the conversation, and tries to explore the appropriate detailed search condition to search the real estate property. Matching spoken word by the user and detailed search condition is the first challenge in this research. The frustration that the user feels when using the voice interface is that the system can not understand the meaning of the word they have pronounced. We need to match appropriate detailed search condition to words in users' conversation.

There are various attempt to incorporate speech interface in search as a support role for exploratory search. Andolina et al. proposed systems [4][5] that extract keywords from

users' conversations to stimulating human creative thinking. We assumed similar approach could be effective in real estate search.

Other reason why an existing system using speech recognition makes a user feel frustrated is that the system is not transparent [8]. In other words, the user has no way of knowing why the system returned a response when the misunderstood answer came back. For this reason, we need to make matching process between spoken and recognized word and recommended detailed condition as clear as possible. By showing every keywords searched during the detailed search condition and relation between them, we will be able to achieve that goal. Also, attempt to avoid keyword input by selecting and manipulating suggested keywords by touch has been proposed[7]. By using similar interaction, we assumed that we will be able to increase the effectiveness of the system. Even if system can not recognize the user's intention correctly, user may be able to find and select interesting keywords displayed on the screen and explore related detailed search condition.

We will describe the design of the system below.

SYSTEM DESIGN

Screen shot of developed system is shown in Figure 1.

User can set the search conditions which are area to search, layout and price of real estate property via GUI. After setting the search condition, users have free conversation about their intention for new house between themselves. System continuously monitors their conversation and recognized text is shown in the lower part of the screen. System automatically and continuously extract keywords from conversation and tries to find the related detailed search condition such as "Within 800 meters from convenience store", "Pet allowed". System also display the search result of the real estate property specifying each detailed condition. Link is shown between detailed search condition and searched property. There are two major challenges in developing the proposed system. First, we needed to develop algorithm which search detailed search condition using user's conversation data. Second, we need interaction interface which will effectively support user to search real estate information even if the system does not recognize the user's intention correctly. We will discuss these challenges next.

MATCHING ALGORITHM BETWEEN DETAILED SEARCH CONDITION AND SPOKEN WORDS

One of the major frustrations felt using existing speech interface is that the system recognizes only programmed keywords while there is no clue for the user about which word to speak. As a result, quite often system does not recognize the word that the user pronounced. Even though the user pronounces the word which has similar meaning to the keyword that system recognizes, system can not understand the similarity between those words.

Similar difficulty is recognized in the field of question and answer retrieval (herein Q&A retrieval) task[6]. Major challenge for Q&A retrieval is word mismatch between the user's question and the question-answer pairs in the archive [11]. To

solve the word mismatch problem, many different approaches have been proposed.

In this research, we tried to utilize Word2Vec [9]. By using Word2Vec, each word is vectorized and similarity between words can be calculated. Basic idea in this research is that we tried to match spoken words and detailed search condition not with simple word match, but with match between related words expanded from spoken words and detailed search condition using Word2Vec. Following is the algorithm we used.

1. For each detailed search condition, manually set two to five related keywords.
2. For each manually set keyword, search related words using Word2Vec, and record them as "expanded related keywords".
3. Convert conversation voice to text string using speech to text conversion function.
4. Extract noun and verb from the recognized text, and record them as "spoken words"
5. For each spoken words, search related words using Word2Vec, and record them as "expanded related keywords".
6. Tries to find same word from expanded keywords from spoken words, and expanded keywords from manually set keywords. If same word can be found, put link between spoken word and detailed search condition.

Example of matching is shown in Figure 2.

Next we will discuss the training data set for Word2Vec. Related words extracted using Word2Vec depends on nature of training data set. As an example, we will show the result of retrieved related words of the word "Noise" using Wikipedia[3] as training data in Table 1. (Original words are in Japanese) Generally, words used in connection with "Noise" are lined up, but they are different from what the user associates when searching for a real estate property. Users who are concerned about "Noise" may choose "Top floor" detailed search condition if they are concerned about the noise from the floor above. Users who care about the noise from the roads may choose "Higher than the second floor". In either case, those detailed search conditions has no relations to the words extracted using Wikipedia as a training data.

We have also gathered text data from the web site called All About Japan [1] which has large amount of text related to real estate property. We found out that related words extracted is more suitable for user's intuition when searching the real estate property. Retrieved related words of the word "Noise" using All About Japan as training data is also shown in Table 1. "Sound leak" is what user may be interested in when searching real estate property.

However, number of words in All About Japan is not necessary enough. As a result, in many cases we could not extract related words because spoken word does not exist in All About Japan data set.

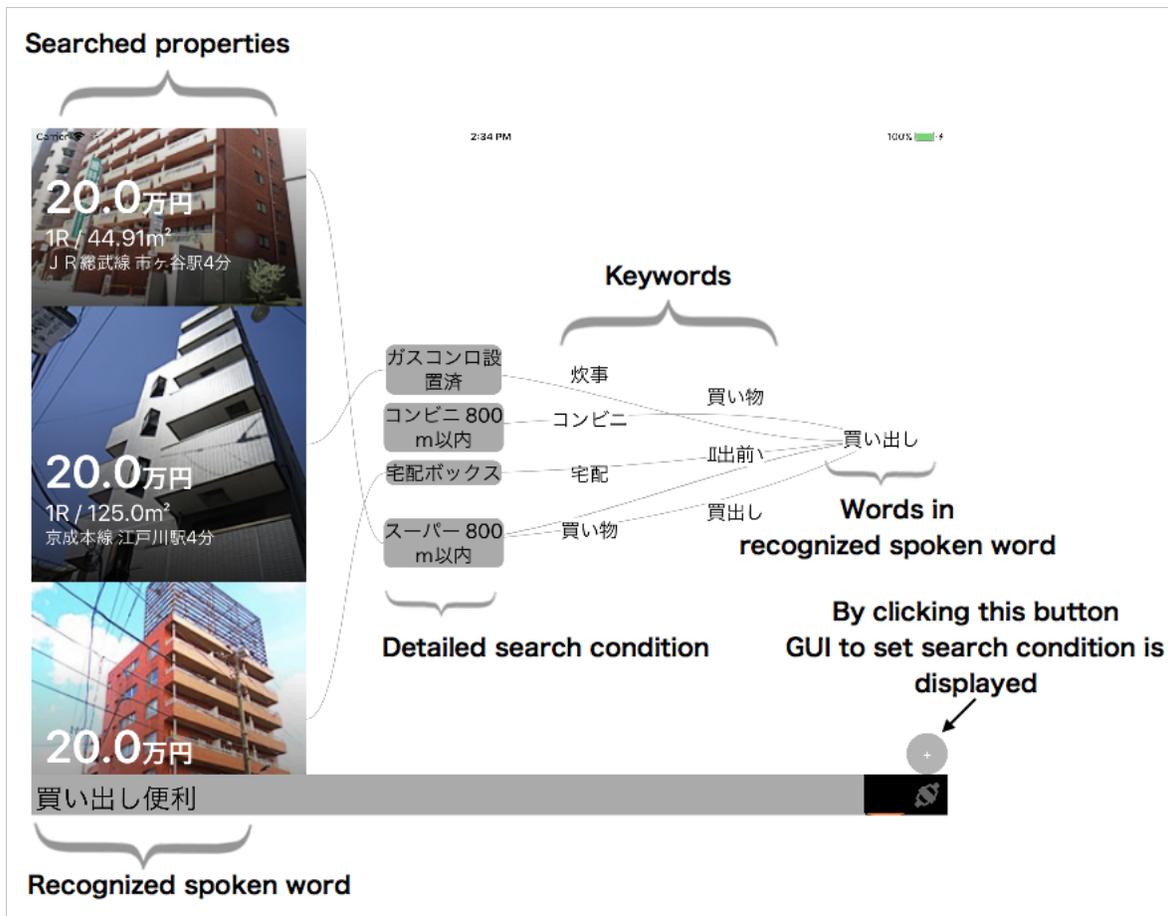


Figure 1. Screen shot of developed system

Rank	Wikipedia	All About Japan
1	Soot	Water leakage
2	Air pollution	Shaking
3	Salinity	Suspicious person
4	Exhaust gas	Sound leak
5	Interference	Exhaust gas

Table 1. Related words of "Noise" searched using different training data

Considering characteristics of each data set described above, we decided to use both data set for keyword extraction. First we tries to extract keywords from All About Japan data set. If no matching word is found, we tries to extract keywords using Wikipedia data set.

USER INTERACTION

By using the algorithm described above, we can expect that we will be able to search the detailed search condition better than simple word match. However, still there are errors and we don't expect that we can search the detailed search condition with high precision.

The other aspect that we should consider is transparency of the system which is described earlier. To ease user frustration, we need to make inference process clear to the user.

Considering these factors, we designed the interaction interface shown in Figure 1. Not only detailed search condition and spoken words are displayed, but also manually set keywords and expanded keywords are shown on the screen. All the words displayed on the screen can be used as an search keyword by dragging and dropping the word into the text area shown in the lower part of the screen.

As stated before, precision of voice-to-text speech and search of detailed search condition is not necessary high. In that case, user may be frustrated if we only display "no results found" or totally irrelevant result. In this research we tried to display as much words as possible on the screen. By viewing those words, it it probable that some words on the screen might be interesting for the user. If so, user can start new search via interaction on the display, not by voice recognition.

EVALUATION

We have conducted two types of evaluation so far. First, we evaluated how effective search algorithm is for various user input. Since we would like to evaluate effectiveness of our

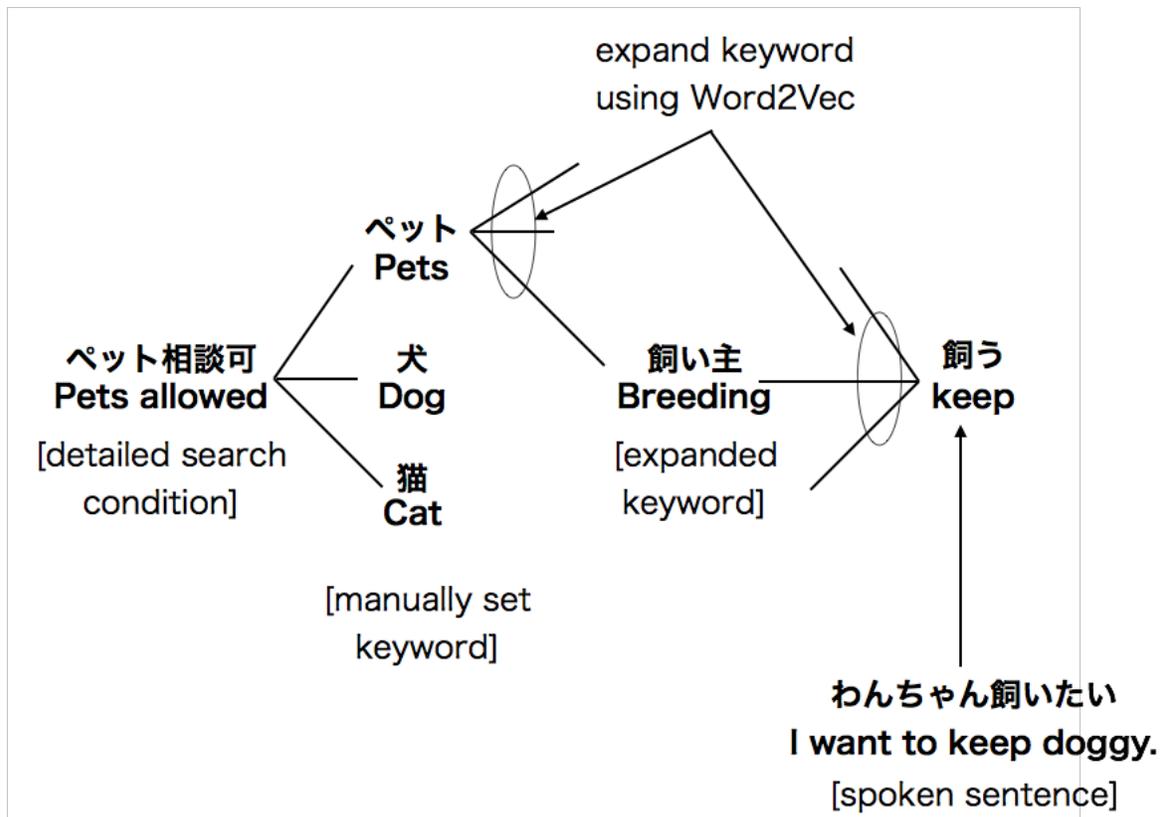


Figure 2. Concept diagram to match spoken words and detailed search condition

algorithm, we used text input rather than voice input to avoid error caused by speech recognition.

We have chosen fifty sentences from Q&A site about real estate search[2] such as "Good view and well-ventilated". All the sentences do not include the exact word in detailed search condition. Therefore, none of the conditions can be searched using simple word match algorithm. For each sentence, we have selected corresponding detailed search condition. System displays up to four detailed search condition. If at least one search condition displayed is related to input sentence, we evaluated it as a success.

As a result, 34 out of 50(68%) sentences can be evaluated as success. 9 out of 34 sentences which are evaluated as success include manually defined extend keywords. If we remove them from the result, 25 out of 41 (61%) sentences can be searched successfully using proposed algorithm.

Second, using current system, we conducted simple user evaluation. We have explained system's concept and function, and put the system beside, we had conversation about what type of real estate property the user is interested in. Five users have participated in the test. During the user test, we got consistent response from the user. Every participants see the importance of system's assistant role. In some cases, the user can find the interesting results. Four out of five participants see detailed search condition that they have never searched real estate property with before. We also observed that the user quite often try

to use the voice interface as an main interface. Even though current system only extracts and search with detailed search condition, the user tried to search with voice sentence like "Search property in Tokyo area". Even though we designed the system so that the user can set those search condition with GUI, the user often forget that.

From the result of user test, we realized that interaction design should be improved so that user does not misunderstand that system can recognize every request that user might have. In the current design, search condition setting GUI is hidden while the system listen the conversation. In this case, the user expects system can understand any word what they say. To avoid such a misunderstanding, we need to show search condition setting GUI upfront. When the user speaks, and do not operate the GUI, we can show current voice recognition interface over the search condition setting GUIs.

CONCLUSION AND FUTURE DIRECTION

We have developed the real estate search system which recommends detailed search conditions from users' conversation. Evaluation result shows the mixed results. We could confirm the potential of proposed algorithm. However, we also recognized that user interface design should be improved. Based on the evaluation results, we will redesign the system interaction, and will conduct further user test to evaluate how effective system can support the search of real estate information.

REFERENCES

1. 2017. All About Japan : House,Real estate property. https://allabout.co.jp/r_house/. (2017).
2. 2017. OKWeb-Sumai(Housing). <https://okwave.jp/c622.html>. (2017).
3. 2017. Wikipedia (Japanese). <https://ja.wikipedia.org/>. (2017).
4. Salvatore Andolina, Khalil Klouche, Diogo Cabral, Tuukka Ruotsalo, and Giulio Jacucci. 2015. InspirationWall: Supporting Idea Generation Through Automatic Information Exploration. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition (C&C '15)*. ACM, New York, NY, USA, 103–106. DOI : <http://dx.doi.org/10.1145/2757226.2757252>
5. Salvatore et al Andolina. 2015. IntentStreams: Smart Parallel Search Streams for Branching Exploratory Search. In *Proceedings of the 20th International Conference on Intelligent User Interfaces (IUI '15)*. ACM, New York, NY, USA, 300–305.
6. Jiwoon Jeon, W. Bruce Croft, and Joon Ho Lee. 2005. Finding Similar Questions in Large Question and Answer Archives. In *Proceedings of the 14th ACM International Conference on Information and Knowledge Management (CIKM '05)*. ACM, New York, NY, USA, 84–90. DOI : <http://dx.doi.org/10.1145/1099554.1099572>
7. Khalil Klouche, Tuukka Ruotsalo, Diogo Cabral, Salvatore Andolina, Andrea Bellucci, and Giulio Jacucci. 2015. Designing for Exploratory Search on Touch Devices. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 4189–4198. DOI : <http://dx.doi.org/10.1145/2702123.2702489>
8. Ewa Luger and Abigail Sellen. 2016. "Like Having a Really Bad PA": The Gulf Between User Expectation and Experience of Conversational Agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 5286–5297. DOI : <http://dx.doi.org/10.1145/2858036.2858288>
9. Thomas Mikolov. 2013. word2vec:Tool for computing continuous distributed representations of words. (2013). <https://code.google.com/word2vec/>.
10. Verto. 2017. What's the Future of Personal Assistant Apps? (2017). <http://research.vertoanalytics.com/what-the-future-of-personal-assistant-apps-webinar-deck>.
11. Xiaobing Xue, Jiwoon Jeon, and W. Bruce Croft. 2008. Retrieval Models for Question and Answer Archives. In *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '08)*. ACM, New York, NY, USA, 475–482. DOI : <http://dx.doi.org/10.1145/1390334.1390416>