Efficient Web Search on Mobile Devices with Multi-Modal Input and Intelligent Text Summarization

Eric Chang
Microsoft Research Asia
49 Zhichun Road, 5F, Sigma Center
Beijing, China
echang@microsoft.com

Helen Meng
Human-Computer Communications Laboratory, The Chinese University of Hong Kong
Shatin N. T.
Hong Kong SAR, China
hmmeng@se.cuhk.edu.hk

Yuk-chi Li
Human-Computer Communications Laboratory, The Chinese University of Hong Kong
Shatin N. T.
Hong Kong SAR, China
ycli@se.cuhk.edu.hk

Tien-ying Fung
Human-Computer Communications Laboratory, The Chinese University of Hong Kong
Shatin N. T.
Hong Kong SAR, China
tyfung@se.cuhk.edu.hk

ABSTRACT

Ease of browsing and searching for information on mobile devices has been an area of increasing interest in the World Wide Web research community [1, 2, 3, 6, 7]. While some work has been done to enhance the usability of handwriting recognition to input queries through techniques such as automatic word suggestion [2], the use of speech as an input mechanism has not been extensively studied. This paper presents a system which combines spoken query in addition to automatic title summarization to ease searching for information on a mobile device. Preliminary usability study with 10 subjects indicates that spoken queries is preferred over other input methods.

Keywords

mobile device, information retrieval, user interface, speech recognition

1. INTRODUCTION

Mobile computing has become a new trend in the computer industry. However, interface issues such as small display, tiny keypads and error-prone handwriting input are still serious barriers. Furthermore, the problems of small display sizes and cumbersome input methods compound to reduce usability. For example, in [6], Jones found that users of a small display device were much more likely to use the search functionality on a specific website. They found that 80% of small screen users began their task with the search function of the particular website and were twice as likely to use search function compared to large screen users and less likely to browse the web pages of a website to find the desired information. Clearly, improving the ease of searching for information on a mobile device can dramatically improve the efficiency of searching for information on websites. Previous works have focused on easing text input with pen using auto-completion and keyword suggestions [2]. In this study, we focus on enabling the use of spoken queries with the search engine of a particular website. Earlier work [5] has shown that spoken queries can be used for Web search on a variety of devices ranging from a browser running on a desktop computer to a PDA with a wireless connection. In this work, we focus on problems faced by users when searching for information on mobile handheld devices and propose the combination of multi-modal input and text summarization techniques to make the process more efficient.

We foresee that in the future, a multimodal approach would be most efficient for conducting searches on mobile devices. While speech can be faster than other input methods, there are also situations where other input modalities are more suitable. For example, users in a meeting may not wish to disturb the meeting with spoken queries. Furthermore, the availability of multiple input modalities would allow the user to compensate for the deficiency of one input modality with another input modality, as was demonstrated in the usability study conducted in [4]. The speech-driven search system was implemented on a text corpus which consists of the past five years of Microsoft press releases. The system accepts spoken query from users, searches through the press release database, formats the display of the retrieved articles and lists them in an HTML viewer for browsing.

2. SYSTEM DESCRIPTION

Figure 1 illustrates the architecture of the system. An application client connects to the server of search engine (IpaqRet) through a wired or wireless network. When the input query is speech, the MiPad server is triggered to perform continuous speech recognition [4]. The recognized text string is passed to the IpaqRet server to perform information retrieval on articles relevant to the input string. The titles of the relevant article are sent back to the IPAQ client, and displayed on screen in the form of article titles list. The information retrieval process for text input is similar to using speech input, except that the input text is passed to the IpaqRet server directly without triggering the MiPad server.

A listening socket waits for incoming requests from either the IPAQ client or the MiPad server. The data stream is sent to a keywords extraction module. The resulting keywords are passed to the information retrieval engine to find relevant articles from the database of indexed Microsoft press releases. The information retrieval engine returns a list of relevant document identity numbers. The titles of the corresponding articles are looked up from the title archive using the identity number. To have a better display on the small screen at the IPAQ client, the original titles of the articles are shortened using the text summarization algorithm. The search results are listed in the form of shortened title with HTML. Each title links to the corresponding press article. The result list is sent to IPAQ client for search result display.

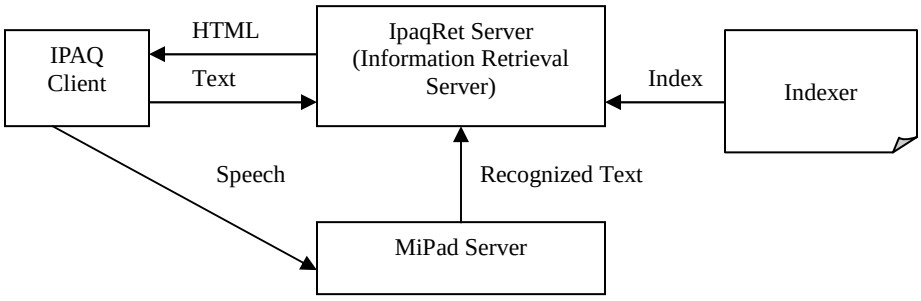

Figure 1: System block diagram.

3. HANDLING SPOKEN QUERY

All spoken queries are passed to the MiPad server [4]. MiPad currently runs on a Pocket PC connected to a Windows server where speech recognition is performed. When using the Tap and Talk technology within MiPad, users tap on a button on the screen and speak into the IPAQ's microphone (shown in Figure 2). The speech waveform is sent to the MiPad server where recognition on the input speech takes place. The recognized text strings from the speech recognizer is passed to the server through a socket. We have developed a vector space model based information retrieval engine. Stop words removal and stemming are performed during both indexing and retrieval phrases. The stop word list we used is extracted from the SMART information retrieval system [10]. The stemming is performed using the Porter stemming algorithm [8] where only the stemmed forms of the terms are stored and compared to during retrieval.

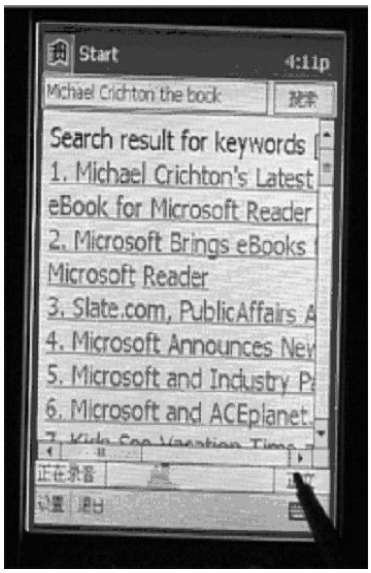

Figure 2: A frame from the demo video of the system. The volume bar on the lower part of the screen reminds the user to speak within a reasonable volume range. The user taps on the record button to start sending speech query through the mobile device.

The retrieval engine produces a list of relevant articles and the titles of the relevant articles are sent to the title extraction module. All titles are listed in a HTML file and sent back to the IPAQ client for display. Since the display is small, the task of displaying a list of original titles from the press release archive becomes a challenging task.

4. SHORTENING TEXT FOR SMALL DISPLAYS

Since the original titles for the press release are usually long (around 15 words per title), the number of titles displayed per screen is small. Therefore, horizontal scrolling is frequently needed and increases the amount of time necessary for result selection. To solve this problem, we display the search result in the form of shortened titles instead of their original titles. Our title-shortening algorithm consists of two techniques: the use of abbreviation and automatic insignificant words removal. Previously, Radev et. al. have done work related to this topic [9].

4.1 Use of Abbreviations

This technique replaces words or phrases with abbreviations or acronyms to minimize the character-length they consume. The substitution is accomplished by abbreviation dictionary lookup and maximum matching with a greedy algorithm. The abbreviation dictionary is manually edited based on the most frequently occurring words/phrases within the domain and contained 250 entries. For example, the word "demonstration" is replaced with "demo" and the phrase "Internet Explorer" is replaced with "IE".

4.2 Insignificant Words Removal

This technique retains the phrase(s) of the article title that contains most significant words (title terms) and removes other phrases. The titles are compressed independently for each article, and the significance of the title terms are considered within the single article. The significance of a title term is directly related to its frequency of occurrence in the title and the number of paragraphs containing it. A TERMSCORE is used to model the significance. The higher the TERMSCORE a term has, the more important the term is.

The TERMSCORE is calculated based on the well-known Term Frequency/Inverse Document Frequency (TF/IDF) measure [10]. However, since we are measuring the significance of a title from a single article but not the whole collection, we define

the meaning of our "IDF" as how often the title terms appear in the paragraphs of the article. The stemmed forms of the terms are considered when counting term frequencies. A term will have zero TERMSCORE if it is a stop word.

The title is separated into phrases with phrase boundary defined at the position of punctuations and prepositions. We choose to also segment the title at prepositions because very often we cannot find pauses within a title. The significance of phrases is determined by the summation of the TERMSCOREs. The sum of the TERMSCORE of all the words in a phrase is the PHRASESCORE. The phrases are ranked according to the PHRASESCORE, and the highest scoring phrases are considered to be the most significant phrases. To fit the title into the PDA screen, we limit the length of each title to 35 characters. A subset of all phrases will be selected based on their PHRASESCORE ranking within the limitation of 35 characters length. Such phrase selection approach can optimize the use of the limited display space and allow keywords to be included as much as possible. Some examples of title-shortening are:

"Windows XP Ushers In New Era of Communications" => "Win XP Ushers In New Era of Comm"

"Sidewalk, the Personalized City Guide to Entertainment, Goes Live in San Francisco" => "The Personalized City Guide"

5. CONCLUSION

In this paper, we have presented methods of enabling users to search for information efficiently on mobile devices. To allow better search results, information is extracted from the corpus to be searched. The important words which occur in the corpus are used to improve the effectiveness of the speech recognition algorithm, while text summarization techniques are applied to improve the efficiency of browsing through search results. Preliminary usability study with 10 subjects indicates that spoken queries is preferred over handwriting input. A more detailed usability study will be completed in the future.

6. ACKNOWLEDGEMENT

This work was carried out while the 3rd and the 4th author were visiting students at MSR Asia.

7. REFERENCES

1. G. Buchanan and M. Jones, "Search Interfaces for Handheld Mobile Devices", *Poster Proc. 9th Int. World-Wide Web Conf.*, 2000, http://www9.org/final-posters/48/poster48.html.
2. O. Buyukkokten, H, Garcia-Molina, A. Paepcke, "Focused Web Searching with PDAs", *Proc. 9th Int. World-Wide Web Conf.*, 2000.
3. O. Buyukkokten, H, Garcia-Molina, A. Paepcke, "Seeing the Whole in Parts: Text Summarization for Web Browsing on Handheld Devices", *Proc. 10th Int. World-Wide Web Conf.*, 2001.
4. X. D. Huang et. tal. "Mipad, A Multimodal Interaction Prototype", *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2001.
5. H. Gu, J. Li, B. Walters, and E. Chang, "Spoken Query for Web Search and Navigation", *Poster Proc. 10th Int. World-Wide Web Conf.*, 2001.
6. M. Jones, G. Marsden, N. Mohd-Nasir, K. Boone, G. Buchanan, "Improving Web Interaction on Small Displays", *Proc. 8th Int. World-Wide Web Conf.*, 1999, http://www8.org/w8-papers/1b-multimedia/improving/improving.html.
7. O. Kaljuvee, O. Buyukkokten, H, Garcia-Molina, A. Paepcke, "Efficient Web Form Entry on PDAs", *Proc. 10th Int. World-Wide Web Conf.*, 2001.
8. M. F. Porter, "An algorithm for suffix stripping", *Program*, 14(3): 130-137, 1980.
9. D. Radev and W. Fan, "Automatic summarization of search engine hit lists", *Proc. ACL Workshop on Recent Advances in NLP and IR.*, Hong Kong, October 2000.
10. G. Salton, "The SMART document retrieval project", *Proc. Fourteenth International ACM/SIGIR Conference on Research and Development in Information Retrieval*, 1991.