Hands-Free Searching Using Google Voice

Hamzeh Al_bool[1], Maytham Ali[2], Hamdi Ibrahim[3], and Adib M.Monzer Habbal[4]

InterNetWorks Research Group, UUM College of Arts and Sciences,
University Utara Malaysia, 06010 UUM, Sintok, MALAYSIA

Abstract
Google is one of the most popular information retrieval systems which it has a fast increasing in the number of users. It can provide a foundation for improving the learning and teaching environments. However, not all potential users have the capabilities that allow them to use Google without impediments. This is specifically true for motor-disabled users. Therefore, for this category of users a special approach of interaction with Google interface is necessary. The missing capabilities of this category of users should be substituted by capabilities that these users have in order to deal with Goggle. This study entails the development of Hands-Free voice recognition Google Search Engine to allow physically disabled the opportunity to operate a Google and browse the result of search without using a keyboard or mouse. The implementation is carried out with Java programming language and the result of usability testing shows that 93% of the participants found the system usable.

Key words:
Voice Recognition, motor-disabled users, Google, Java Programming Language

1. Introduction

Recent advances in computer technology have enabled users of voice recognition products to achieve desirable results which were previously impossible with the largest mainframe computers or workstations. One of the most important reasons that led to the production a large number of voice input systems is the ability of the voice input systems to remove a lot of literacy barriers in communicating with the instructions of the computer.

Besides, the inventions of the technological innovations have suggested the emerging knowledge society, in which all our engagements in life will be focused on having knowledge of something. Such knowledge can be acquired in many ways and in different forms. On the other hand, many search engines have been developed to facilitate information search across the globe via the World Wide Web (www). Google is one of the most popular search engines and information retrieval systems. It is a huge, global web-based application with search engine. Furthermore, it is very easy to use [17] with very simple interface. On top of that, it is linked with more than hundreds of millions of web pages involving similar numbers of different conditions. It has been recorded with over ten millions of queries daily [3].

Unfortunately not all potential users have all capabilities that allow them to use Google information retrieval system without impediments. There are many people who have injuries and disabilities which prevent many of them to deal with computer applications to suit their conditions. It has been revealed that more than 50% of newly hired computer users reported musculoskeletal with symptoms like eyestrain, neck and shoulder pain, low back pain, elbow pain tendonitis), forearm pain (muscles) and nerve entrapments [6]. Additionally, there are also disability constraints in using the traditional keyboard input system. Hence, another form of input is necessary, such as voice in recognition systems.

The development of information retrieval system with voice recognition products provides a foundation for improving the learning and teaching environments through technology. Voice recognition might be able to help the disabled users to explore information through a set of voice commands that provide easy ways and controls mechanism for searching information. It is expected that voice will replace the physical manipulation as the dominant input modality. This shift will dramatically alter input needs, and the way users interact with computers.

This study is to introduce a new service which is not currently available in Google search engine (GSE) [18]. It proposes the implementation of voice recognition input in GSE. The voice recognition input will entail the development of Hands-Free GSE Using Voice Recognition to provide physically challenged individuals

Manuscript received August 5, 2010
Manuscript revised August 20, 2010

especially the visually-impaired with the opportunity to operate the GSE and browse the result of searching without using keyboards or mouse. Furthermore, it will assist people who normally have problem with musculoskeletal disorders when working with computers via typing of instruction.

This paper is organized into six main sections as follows: It starts with this section that introduces the study and this paper. It is followed with a section that discusses some related works. Section 3 follows next, discussing on the concept of hands-free voice recognition GSE. The implementation and evaluation are provided in Sections 4 and 5, which are finally followed with a concluding section.

2. Related Works

In the quest to assist the disabled individuals, in terms of their mode of interactions with computer-enhanced technologies, [5], proposed an assistive system utilizing voice recognition input technology. In the application, the voice commands are generated to assist the disabled people to operate basic electronic equipments without necessarily touching the equipments. In addition, the authors used mouse emulators and keyboard emulators for interaction with computer for different computing purposes.

Later, for physically disabled people to enjoy part of the dividends of the 21st century technological innovations, [1] propose a framework for controlling computer desktop with issuance of voice commands.

In fact, efforts on utilizing voice recognition for searching purposes are not new. In relation, a voice recognition system has also been designed for visually impaired people [7]. Although the system is specific to people with vision problem, the developed voice recognition system in their work is more general. If not for the provision of alternative input system, these categories of challenged individuals might not have the opportunity to survive in this knowledge world.

On the other hand, voice recognition has also been employed as security information in place of the conventional knowledge-based authentication approach [4]. This implies that voice message is a promising stress-free approach to replace the conventional typing-based input system.

On top of the robustness of a computer implementation, its friendly usage experience is considered more important. Voice recognition system is not an exception in this regard. The usability of any voice recognition system is very important and is proven by [5]. Also, they suggest that human factors should surround the use of voice recognition system.

However, it was also revealed that using voice recognition to replace the traditional keyboard input system requires additional support [8]. This implies that there is a need to take care of the sensitivity of the interaction mode in terms of noise management and the like exists.

The related works are reviewed to help deeply understand the missing and the impediments capabilities of motor-disabled users that should be substituted by capabilities that these users have in order to deal with Google. Mainly, the studies in the literature come up with important information that is needed to be more aware of each of end users for Hands-Free searching using voice recognition GSE. The importance of developing such system is to help the motor-disabled users to explore information by developing a set of voice commands that provide easy way and control mechanism for searching information over Google information retrieval system.

Unlike other types of computer system, developing a voice recognition system is better done using dynamic programming approach [9]. This serves as the justification for choosing Java programming language for system development in this study.

3. Hands-Free Searching Using Google Voice Concepts

The concept behind the proposed voice recognition system over GSE is based on a native method, robot class; Grammar file, sphinx-4, and Wiener filter as discussed in the following paragraphs.

Native method *(a.k.a Exec Method):* it is primarily designed to enhance the interaction between the program and the operating system. It is a way of combining the power of C or C++ programming with Java [13]. In this study, such advantage is utilized to be able to run the internet browser application and also for easy display of website addresses.

Robot Class in java*:* This class is used to generate native operating system input events for the purposes of testing automation, self-running demos, and other applications where control of the mouse and keyboard are needed [2]. The primary purpose of Robot is to facilitate automated testing of Java platform implementations. It is employed to

deal with voice commands by taking over the control of the keyboard and mouse commands. The type of command method in this case is voice command.

Grammar File: for the grammar file, a context free grammar is specified as the input is capable of accepting the voice input and produces a java language functions that recognizes the appropriate instances of the grammar. The implication of this is that instances are created for all possible grammars.

Sphinx-4*:* it is a state-of-the-art voice recognition system written entirely in Java, and is considered as a suitable framework for Voice recognition to adopt. The main blocks are the FrontEnd, the Decoder, and the Linguist. Supporting blocks include the Configuration Manager and the Tools blocks. The communication between the blocks, as well as communication with an application, is described in [16]. Sphinx-4 is used to analyze the uses of voice commands as shown in figure 1.

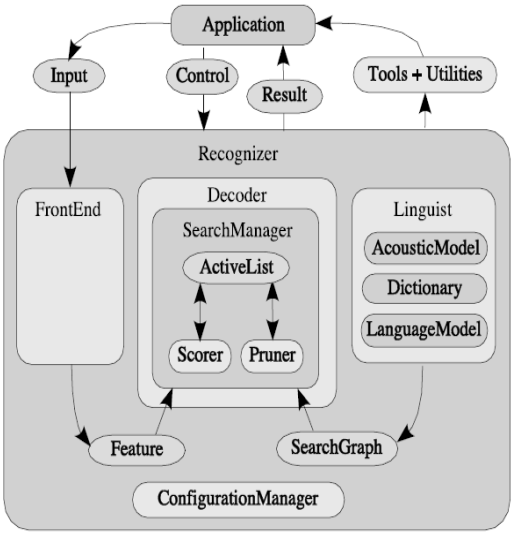

Fig. 1 Sphinx-4 Decoder Framework.

Wiener filter**:** it is based on tracking a priori SNR using Decision-Directed method, proposed by [19]. The two-step noise reduction (TSNR) technique removes the annoying reverberation effect while maintaining the benefits of the decision-directed approach. However, classic short-time noise reduction techniques, including TSNR, introduce harmonic distortion in the enhanced speech. To overcome this problem, a method called harmonic regeneration noise reduction (HRNR) is implemented in order to refine the a priori SNR used to compute a spectral gain able to preserve the speech harmonics.

4. Implementation

A system has been implemented with the technique discussed in the previous section using Java based interface. The command and control grammar-based recognition is implemented to make the system robust and speaker independent. At present, the system includes two types of functions. Firstly, through a microphone that is connected to a computer, the user can search information by using Google search engine through a write function method voice commands. Write function is a method that is built to allow users to enter any sentences they want to search for, in Google search text box by voice. Secondly, the user can access all Google search engine functions, such as Web, Images, Maps, Books and deal with all sub functions Google provides through using mouse voice commands and keyboard voice commands. Table 1 shows all the voice commands with resulting events to be carried out by the commands.

Table: 1 Voice Command Description

Voice Commands Name	**Command Events**
Search	To put the mouse pointer in the Google text box.
Go	To run the Google process.
Up, Down, Right, and Left.	To control the mouse in four trends
Open	To click on any button or link in which it depends on the mouse position
Tab	To move to the next section of a Google page.
Enter	To begin the desired process, which is usually an alternative over the pressing an OK button.
Back	To return to the previous page.
Home	To return to the Google Home Page.

Having specified the commands, with specified desired actions as listed in Table 1 this study proceeds with initializing the steps in implementing the system. Figure 2 shows the steps involved in the system implementation by opening the Google page. The diagram explains that users can utter any sentence that they want to search for. If the sentence is not included in the grammar file then the system will show expiration message that the sentence is not included in the grammar file. But if the sentences are in the grammar file the sentences will be displayed in Google text box search.

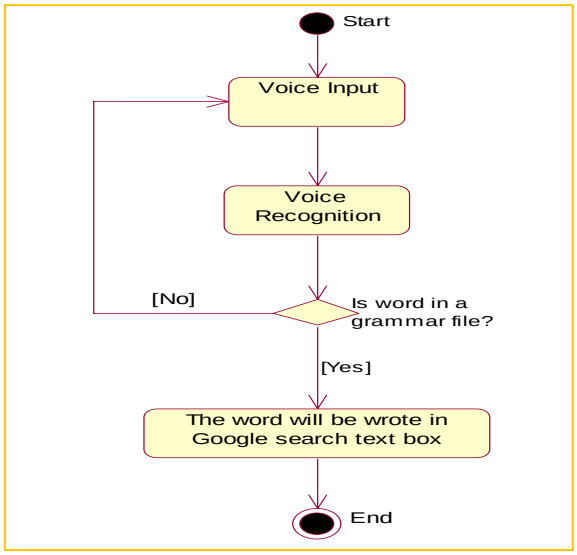

Fig. 2 Opening Google page.

Further, this study also emphasizes on using voice to control the input and navigate in a web page. Figure 3 illustrates the algorithm for the function. The diagram explains that users can use the voice command shown in Table 1 to control the search process and explore the search result. In addition, users can control mouse movements and keyboard commands with voice to move to any place in the Google web page and to access any link or function.

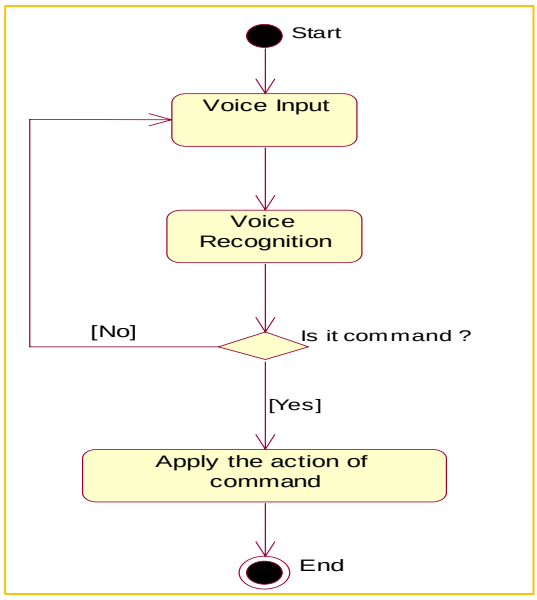

Fig. 3 Using voice commands to control search process.

5. Test and Evaluation

A. User test

The user test has been carried out. It was heavily based on the users' own words. This study chooses not to collect any quantitative data or define goals for the tests to reach in order to pass as successful or not. The reason for this is the same as made us go with a qualitative investigation in the first place. The users had informative opinions of the program after the tests. The evaluation include direct questions on whether each user thinks he or she will be using voice recognition input-based interaction in GSE situations in the future, and if the tested program is a real alternative in this sense.

In terms of procedure, this study invited users personally. In the invitation, users were provided with the system and a suggestion on the details to be tested with each user. This study intends to test different aspects of the reliability and easy to use concepts with different users. The reason is that it is more effective to have the test users inspired and performs the tasks spontaneously, than to make them all try every detail. The invitation also involves educating the users in terms of dictation software as a way of preparing them for the actual tests that are following. Along with the invitation process, the program will be customized in order to fit each task and user.

When this study has made-up decisions with the selected participants in the test and evaluation process, the investigation were decided to be on reliability, easy to use, and satisfactions of Hands-Free searching using voice recognition GSE. Therefore this study concluded that the users should all be familiar of using GSE. Furthermore, this study thought it would be rewarding to have at least a couple of the users already familiar with normal speech recognition. Finally, the participants were of total thirty, divided as follows:

Table 2: The participants of user test

Number of Users	Users description	Age average	Gender
Two	programmers/engineers working with speech recognition development	42.5	M
Five	students already familiar with normal speech recognition	36	M/F

Fiftee n	Students that using Google search a lot but without experience of speech recognition.	27.2	M/F
Two	Users have motor disabilities.	23.8	M
Six	Newly hired computer users	26.6	M/F

All participants were gathered from University Utara Malaysia, and users who have motor disabilities from outside the university. They were volunteered to perform the user test sessions.

Besides, the vocabularies that are stored in the grammar file consist of more than 3000 words. These vocabularies were collected manually from different sources; also they are related to learning and education area. This study gives special attention to the participants about the domain of words that are availably to be search about, in which they are only related to English education and learning area.

The user test was focused on three aspects which are: (1) how would users describe good reliability of our techniques? (2) how would users describe good usability of our techniques?; and (3) are users satisfied?.

Therefore, the valuation of the program in the context scenarios had come in order to identify the main three aspects of this study as follows:

Firstly, we introduced Hands-Free GSE Using Voice Recognition program to all participants and explained the main functions of our system to increase the awareness among the participants on what this system can do and how they can deal with it.

Secondly, we installed our prototype software on thirteen systems and we divided participants into randomly groups, except the users they have motor disabilities we put them in an independent group to be more focused. Also, programmers/engineers working with speech recognition development they do not have any group where those played the oversight role on all groups within test and evaluation process.

Thirdly, before we started we distributed a set of documents to participants, which it contains a list of one hundred words and keyphrases organized for the purposes of the search processes. these documents also contain the voice commands and the description of these commands where the main purpose to provide these voice commands to make participants more familiar about the voice commands that are available and more clearly about the functions of these commands in order to browse Google interfaces.

Finally, we asked the participants to use the system and browse the result of searching without using keyboards or mouse.

The results of this laboratory experiment have been analyzed and presented in the next section.

B. Result

There are two obvious problems found in the developed system. Firstly, technical problems such as voice recognition limitations in the current voice input systems. Secondly, the limitations of the grammar file size where it can only contain 120,000 words. These problems reduce the reliability of gathered result.

In the study, the observations reveal that users deal directly with Google interface. This study interprets that the developed system is easy to use. This is partly because of the interface of GSE itself, which is very simple and minimal.

In addition, from the qualitative observation, this study quantifies the data for better view of the results. As stated earlier, this study investigates the ease-of-use, reliability, and satisfaction among users in the user test. Hence, the results were gathered accordingly, and are presented in Figure 4.

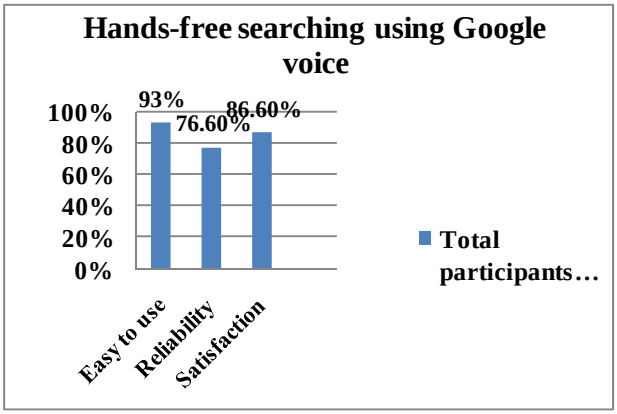

Fig. 4 The result of the participants satisfaction on the Voice-based Google search Instruction

From the graph in Figure four, this study interprets that users found that the system is easy to use. In addition, the results of search activities were found highly reliably. With the high reliability and ease of use, the users were found highly satisfied with system.

Two different noise suppression methods were tested in Matlab [19]. They were noise suppression using Spectral enhancement and noise suppression using Wiener filter. A wiener filter of the method proposed by Plapous was tested. Furthermore, a better approach will be adding an implementation of Wiener filter in Java to the front end of Sphinx-4, as the front end of Sphinx-4 is flexible and pluggable.

This study has been introduced a new service which is not currently available in Google search engine (GSE) [18]. Furthermore, to the best of our knowledge, there is no similar or close study to ours in order to compare our approach with other studies approaches or to comparison between our results with other studies results.

6. Conclusion

Previous studies have shown that there is a need to provide an alternative to the conventional keyboard and mouse input system, not only because of disability on the part of the users but also, the associated stress with the modes of interaction among normal people. This has led to several techniques in the area of voice recognition. In this age, the world is being ruled with knowledge throughout the potentials of technological innovations. Voice input system will remove all barriers imposed by the traditional input system and will enable the physically challenged individuals access to necessary information as long as they can communicate with audible voice. In this study, the developed voice-based GSE is found usable. It is therefore recommended that future researches should look at the integration of language translation that can make such a system to be able to work with variety of world languages. However, there are some limitations in the existing voice input systems. One of which is the recognition accuracy. Sphinx-4 is written in java and it does not have any noise suppression module. It is very important to have one to improve its performance in the presence of noise. It will be interesting to implement the recommended Wiener filter implementation in java. It can be easily added to the Front end block of Sphinx-4.

7. References

[1] Abdeen, M., Mohammad, H. & Yagoub, M. C. E. An Architecture for Multi-lingual Hands-free Desktop Control System for PC Windows. *IEEE Xplore*, 1747- 1750. (2008).

[2] Baldwin. RIntroduction to the Java Robot Class in Java, retrieved, August 10, 2010, from http://www.dickbaldwin.com/java/Java1472.htm. . (2003)

[3] Brin S., Page L. The anatomy of a large-scale hypertextual Web search engine* 1. *Computer Networks and ISDN Systems 30*, 107-117. . (1998).

[4] Cui, B. & Xue, T. Design and Realization of an Intelligent Access Control System Based on Voice Recognition, *International Conference on computing, communication, Control, and Management, ISECS, IEEE*, 229 – 232. . (2009).

[5] Gao, X. T., Ong, S. K., Yuan, M. L. & Nee, A. Y. C. Assist Disabled to Control Electronic Devices and Access Computer Functions by Voice Commands, *Communication of the ACM*, 37 – 42. (2007).

[6] Gerr F., Marcus M., Ensor C., Kleinbaum D., Cohen S., Edwards A., Gentry E., Ortiz D., Monteilh C.A prospective study of computer users: I. Study design and incidence of musculoskeletal symptoms and disorders. *American Journal of Industrial Medicine 41`*. 221-235. (2002).

[7] Halimah, B. Z., Azlina, A., Behrang, P. & Choo, W. O. Voice Recognition System for the Visually Impaired: Virtual Cognitive Approach, *IEEE Xplore*. (2008).

[8] Mills, S., Saadat, S. & Whiting, D. Is Voice Recognition the solution to keyboard-Based RIS?, *IEEE Xplore*, 1 -6, (2006).

[9] Ney, H. & Ortmanns, S. Dynamic Programming Search for Continuous Voice recognition, *Signal Processing Magazine, IEEE Xplore*, 16(5), 64 – 83. (1999).

[10] Rashid, R. A., Mahalin, N. H., Sarijari, M. A. & Abdul Azizi, A. A. Security System Using Biometric Technology: Design and Implementation of Voice Recognition System (VRS), *International Conference on Computer and Communication Engineering, IEEE*, 898 – 902. (2008).

[11] Sergey B., Lawrence P. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems* 30:107-117.(1998).

[12] Shrawankar U., Thakare V. Feature Extraction for a Voice Recognition System in Noisy Environment: A Study, *Second International Conference on Computer Engineering and Applications (ICCEA), 2010, 1*. 358-361. (2010).

[13] Vairale V., Honwadkar K. Wrapper Generator using Java Native Interface. *International Journal of Computer Science, 2*(2). 126-139. (2010).

[14] Yuan, K., Hou, W. & Zhao, Y. Human Factor Research of Voice Search on Mobile Internet, *IEEE Xplore.* (2008).

[15] Vaishnavi, V., & Kuechler, W. *Design research in information systems*, retrieved, July 19, 2010, from http://www.isworld.org/Researchdesign/drisISworld.htm. (2005).

[16] Walker W., Lamere P., Kwok P., Raj B., Singh R., Gouvea E., Wolf P., Woelfel J. Sphinx-4: A flexible open source framework for voice recognition. *Sun Microsystems, Inc. Mountain View, CA, USA*:18. (2004).

[17] SeniorNet. Lesson 4. retrave in 12 augest, 2010 from http://www.seniornet.org/php/default.php?ClassOrgID

=5337&PageID=5920Hamzeh Boul is typing a message. (2004).
[18] Google, ANNUAL REPORT PUR-SUANT TO SECTION 13 OR 15(d) OF THE SECURITIES EXCHANGE ACT OF 1934 For the fiscal year ended December 31, 2008. retrieved, July 19, 2010, from http://investor.google.com/ (2009).
[19] Plapous, C.; Marro, C.; Scalart, P. Im-proved Signal-to-Noise Ratio Estima-tion for Speech Enhancement, IEEE Transactions on Audio, Speech, and Language Processing, 14 . 2098 -2108. (2006).

Hamzeh Mohammad Al_Abool received the B.S. degree in Software engineering from Al-Balqa' Applied University, Jordan, in 2008. Currently, he doing Master studies in Information Technology (IT) at University Utara Malaysia and expected to graduate by this year. He has participated as presenter in the national conference titled " National Conference on Rural ICT Development (RICTD)" held on 2009, also he participated in the ITU-UUM Asia Pacific Centre of Excellence Training Workshop on " Wireless communication System: Wireless Network Fundamentals" held on 2009.

Maytham Abdulhussein Ali received the B.S. degree in Computer Science from The University of Mustansiriya of Iraq in 2007. Currently, he doing Master studies in Information Technology (IT) at University Utara Malaysia and expected to graduate by this year. He has participated in the national conference titled " National Conference on Rural ICT Development (RICTD)" held on 2009, also he participated in the ITU-UUM Asia Pacific Centre of Excellence Training Workshop on" Wireless communication System: Wireless Network Fundamentals" held on 2009.

Hamdi Khalifa Ibrahim received the B.S. degree in Computer Science from Sebha University, Libya, in 2004. Currently, he doing Master studies in Information Technology (IT) at University Utara Malaysia and expected to graduate by this year. He got on International Computer Driver License (ICDL) certificate in 2006. He has been working for two years as a teacher in computer institutes and programmer in the General Company of Electricity for four years. He has participated in The 2010 International Conference on Education Technology and Computer.

Adib M.Monzer Habbal received his degree in computer engineering and post graduate Diploma in Informatics engineering from Aleppo University, Syria, in 2003 and 2005 respectively. In March 2007, he received the Master's degree in Information Technology from University Utara Malaysia, Malaysia. Currently, He is a lecturer at UUM, Malaysia. His main research interests include Internet performance engineering, network security, network application, TCP and congestion control in Mobile Ad-hoc Networks. He serves on the Technical program committee of many international conferences, such as NETAPPS 2010, ICCAIE 2010, and ISCI 2011. Also he is a technical reviewer for IEEE TENCON 2009 and IEEE WiMob 2010.

#