Blind Browser Using Web Extraction

Akash P. Ahire[1], Poonam S. Gavali[2], Sapna G. Tamboli[3], Kiran K. Patil[4]

[1] *Student, Department. of Computer Engineering, P.V.G's COE, Nashik, Maharashtra, INDIA*
akashahire2011@gmail.com
[2] *Student, Department of Computer Engineering, P.V.G's COE, Nashik, Maharashtra, INDIA*
punamgavali95@gmail.com
[3] *Student, Department of Computer Engineering, P.V.G's COE, Nashik, Maharashtra, INDIA*
sapna.tamboli101@gmail.com
[4] *Assistant Professor, Department of Computer Engineering, P.V.G's COE, Nashik, Maharashtra, INDIA*
kirankhagote@gmail.com

ABSTRACT

Over the years, the World Wide Web (WWW) has been an excellent medium for a number of benefits of mankind. It has facilitated communication between people located in any part of the world. It has provided access to unlimited information in various formats. Apart from a number of services, numerous kinds of entertainment factors can also be found on the Web. Though all these facilities have been extremely profitable to mankind, sadly, the visually handicapped people face great difficulties in accessing these benefits. Web browsers for the visually handicapped people in the past have been limited to converting documents to Braille or speech, or extracting text and filtering. However, the human aspects of web surfing for blind people have not been adequately addressed. This paper presents an architecture of an open source, light weight web browser that makes it easy for the visually handicapped people to surf the web. The proposed architecture allows a blind person to navigate any web content through simple speech commands and voice feedback to any keyboard operation. The browser will have an integrated text extraction engine that inspects the content of the page to construct a structured representation. The internal nodes of the structure represent various levels of abstraction of the content.

Keyword : - *Web Browser for blind, content reorganization, speech, blindness, assistive devices, text reading, hand-held object, OCR.*

1. INTRODUCTION

Internet has brought about an incredible improvement in human access to knowledge and information. However, blind people face difficulties in accessing these text materials. Web browsers for the visually handicapped people in the past have been limited to converting documents to Braille or speech, or extracting text and filtering. However, the human aspects of web surfing for blind people have not been adequately addressed. This paper presents an architecture of an open source, light weight web browser that makes it simple for the visually impaired people to surf the web. The proposed architecture allows a blind person to navigate any web content through simple speech commands and voice feedback to any keyboard operation. The browser will have an integrated text extraction engine that inspects the page to construct a structured representation. The internal nodes of the structure represent various levels of abstraction of the content. This helps in cushy and flexible navigation of the page so as to rapidly home into objects of interest. Finally, the browser is integrated to an automatic Text-To-Speech and Text-To-Braille transliteration engine that outputs the selected text in the form of speech and/or Braille.

1.1 Motivation

Most of the web browsers do not support any automatic speech recognizer that will enable blind users to navigate a web page through speech commands. Most of the web browsers do not support any speech output to any navigation related operations performed by a blind user. Most of the web browsers for blind are integrated with text to speech engines, however, very few browsers support text to Braille representation that may be used by blind users to archive the web document in printed form. Although many web browsers are integrated with Text-to Speech technology. However, most of these browsers readout the information on a web page to the user in a sequential pattern. This creates problems during navigation to a blind person. In order to provide proper information access and to bridge the communication slot between the visually impaired and the sighted community, the need to build some advance technologically supported systems that will allow a blind person to access web contents easily and efficiently is indispensable. We have decided to developed web browser for visually impaired people which will allow them to access web content and perform web related tasks more easily and efficiently. The presentation of each web page should be structured properly. This system will be used by not only blind people but also illiterate people. The user will have access to web contents through key logger and voice commands. Visually impaired people will be having automated mail sending facility.

Automatic Text Extraction

In order to handle complex backgrounds, we propose two novel feature maps to extracts text features based on stroke orientations and edge distributions respectively. Here stroke is defined as a uniform region with bounded width and significant extent. These feature maps are combined to build an Adaboost-based text classifier.

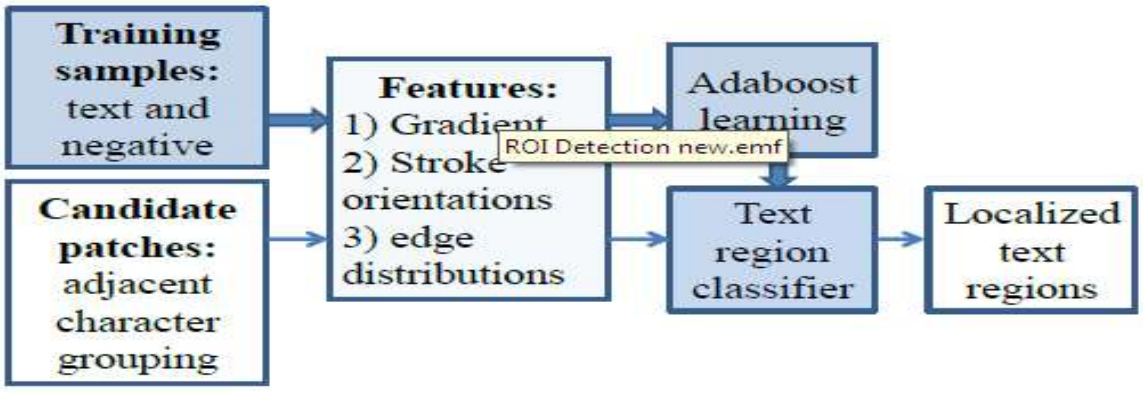

Fig. 2. Diagram of the proposed Adaboost learning based text region localization algorithm by using stroke orientations and edge distribute

2. RELATED WORK

A number of attempts were made worldwide to build web browsers and enhance usability of web for the visually handicapped people. In this paper we have thoroughly studied some of the most widely used web browsers for the blinds. Further, we have compared each of the browsers based on their key GUI features. Based on our study we consider the following GUI features to be essential for a web browser for the blinds.

- Open source architecture
- Text to speech output

- Navigations through voice feedbacks
- Speech recognition
- Text extraction and representation
- Text to Braille conversion
- Voice feedback for keyboard operations
- Mouse based gesture recognition

Next, we will discuss some of the commonly used web browsers for the visual impaired people. Among the early attempts, voice input and input for surfing [1] was adopted for the visually handicapped people. IBM's Home Page Reader [2], among others, presents the web page in an easy-to-use interface, and converts the text to speech, having different gender voices for reading texts and links. But the disadvantage of this is that the developer has to design a complex new interface for the complex graphical web pages to be browsed and for the screen reader to recognize.

Features	WebbIE	eGuide Dog	Shruti Drishti	IBM Home Page Reader
Open Source	NO	YES	YES	NO
Speech input method	NO	NO	NO	NO
Mouse gesture recognition	NO	NO	NO	NO
Voice feedback for keyboard operations	NO	NO	NO	NO
Text extraction and filtering	YES	YES	YES	YES
Hierarchical representation	YES	YES	NO	NO
Text to Braille	NO	NO	YES	NO
text to speech	YES	YES	YES	YES

Table 1. Comparison between Different Web Browsers for Visually Handicapped people

3. LITERATURE SURVEY

Depending on the purpose and content of a web page, these can be classified in various ways. It is important to understand and capture these because the presentation of a page to a blind person would depend on its type. We give below a broad classification with this in mind. This classification is based on the various formats followed by web page designers to design the pages based on what they want to present.

- Single article page
- Multiple article page
- E-mail
- Search Engine

• Portals
• Blogs
• Forms
• Social Networking
• Forums
• Online shops and auction web site

It may be noted that all these pages may contain multiple frames, tables, JavaScript, servlets, etc. These also have to be Identified and handled when the content of the web page is presented.

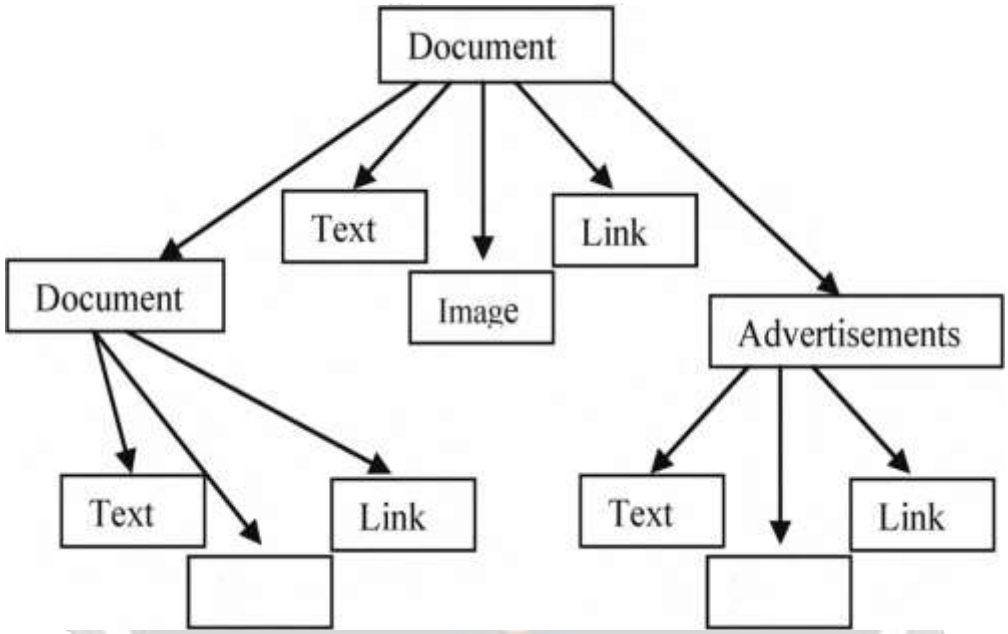

Figure 3: Hierarchical Representation of a Web Page

Web pages are almost always specifically designed for sighted people. The main aim of web page designers is to convey information to Web surfers in a manner that is both attractive and convenient. However such organization of a page is not necessarily appropriate for a blind person. The main motivation of the architecture is to address this issue. The objective of designing a special browser is to develop a framework, with the required toolset, to enable sightless people to browse the web. It is necessary to integrate multimodal input mechanisms with the system to simplify input mechanisms.
Output should be converted to speech or Braille.

4. PROPOSED SYSTEM

Our proposed system essentially consists of three different modules:

(1) User Input Module,

(2) Text Extraction and re-organization, and

(3) Output Representation.

4.1. User Input Module

The user input module deals with the different input methods. Operations on the browser, which are usually performed by mouse clicks, can be performed by speech commands. A speech recognizer is integrated with the browser, which is trained to recognize all the commands. The system is speaker independent. Apart from browser operations, different navigational operations can also be performed by voice commands. For ease in data input using keyboard, an optional voice feedback for every key operation is provided by the system. Figure 4 shows the input module.

4.2. Text Extraction and Re-organization

The extraction of text and presenting it to a visually handicapped person has many difficult aspects to it. With innumerous web pages present in the web, there is a varied diversity in the type of the pages. A prototype hierarchical structure of a web page is illustrated in figure 3. From the figure we can observe that, a web page may contain more than one kind of contents like, links, images, advertisements, and animations. These contents may not provide valuable information to a visually impaired person. Further, the document structure of an email page is also different from other page

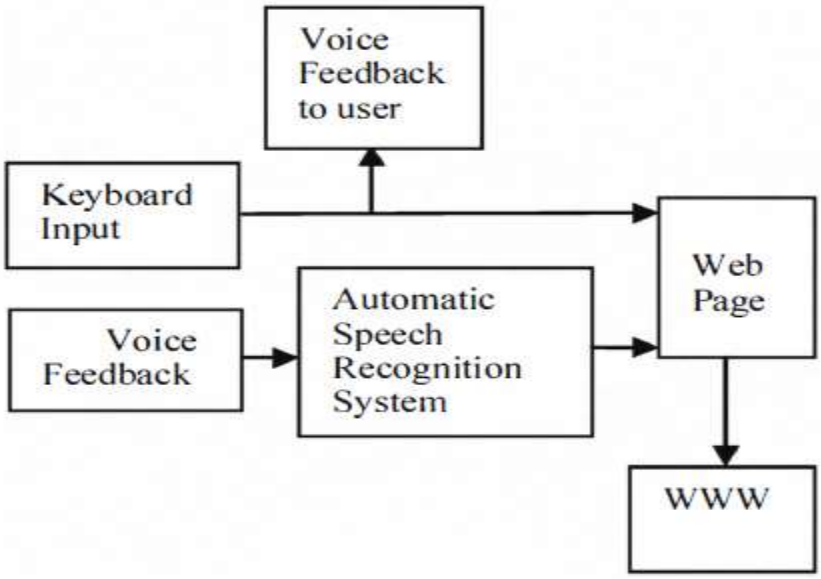

Figure 4: Block Diagram of the Input Module

In order to handle such kind of diversity in web pages, it is required to do a set of pre-processing on the corresponding pages before presenting it to the user. For that, the whole structure of the page along with its contents has to be extracted. This can be performed using an HTML parser. Using the parser, the syntactic information of a page is extracted from the html tags. For example, the header tags give information about the content following it and divider tags divide the pages into section which can be used to isolate different portions of the page as convenient. After getting an idea about the format of the page, the whole information in the page can be divided into sections and then presented to the user in a format convenient to them.

Facilities provided By this Browser:

1. Parsing
2. Browsing
3. Newspaper
4. Magnifier
5. Online Radio
6. Gmail

Algorithms Used:

1. Dom Parser – FivaTech Algorithm
2. Text-to-Speech - FLITE Algorithm

5. CONCLUSION

In this paper, the architecture of a special browser for the visually handicapped people has been described. The browser is currently under construction. The output module, i.e., Text-to-Speech and Text-to-Braille and the voice feedback for keyboard operation is complete. An Automatic Speech Recognition system is currently being trained with the voice commands. Being open source, future developments on the browser by different groups would be easy. The browser is expected to be more user-friendly and effective for the visually handicapped people and hope fully reducing the gap of information availability between the sighted and the visually impaired people.

6. REFERENCES

[1] http://www.synapseadaptive.com/wynn/IBM_Home_Page_Reader.htm I Indian Standard Code for Information Exchange.

[2] Dasgupta T., Basu A., "A Speech Enabled Indian Language Text to Braille Transliteration System", International Conference on Information and Communication Technology for Development (ICTD),2009,

[3] Chu, c.: Two Dimension Interactive Voice Browser for the VisuallyImpaired. In: K. Miesenberger eds. ICCHP 2004, LNCS 3118. Berlin: Springer,(2004),721-724.Qatar.

[4] Chu, c.: Two Dimension Interactive Voice Browser for the Visually Impaired. In: K. Miesenberger eds. ICCHP 2004, LNCS 3118. Berlin: Springer, (2004), 721-724.

[5] Liang, Z., Song, X., Zhu, Z., Liu, R.: Design and Implementation of a New Browser for Blind Person. Computer Engineering and Applications. (2004), 14, 106-108. [8] Mercedes Macias, Fernando Sanchez, Improving Web Accessibility for Visually Handicapped People Using KAI, Proceedings of the 3rd International Workshop on Web Site Evolution (WSE.OI)

[6] A. Huete, J. Victores, S. Martinez, A. Gimenez, and C. Balaguer. Personal Autonomy Rehabilitation in Home Environment by a Protable Assistive Robot. In *IEEE Trans. Systems, Man, and Cybernetics, Part C: Applications and Reviews*, Vol. 42, No. 4, pp. 561-570, 2011.

[7] ICDAR 2011 Robust Reading Competition: A. Shahab, F. Shafait, and A. Dengel. ICDAR Robust Reading Competition Challenge 2: Reading Text in Scene Images. In Proceedings of ICDAR, pp. 1491-1496, 2011.

[8] K. Kim, K. Jung, and J. Kim, "Texture-based approach for text detection in images using SVM and continuously adaptive mean shift algorithm," In *IEEE Trans. on PAMI*, Vol.25, No.12, pp.1631-1639, 2003.
[9] "KReader Mobile User Guide", knfb Reading Technology Inc, http://www.knfbReading.com
[10] S. Kumar, R. Gupta, N. Khanna, S. Chaudhury, and S. D. Joshi, "Text Extraction and Document Image Segmentation Using Matched Wavelets and MRF Model," In *IEEE Trans on Image Processing*, Vol. 16, No. 8, pp. 2117-2128, 2007.
[11] J. Liang, D. Doermann, and H. Li, Camera-based analysis of text and documents: a survey. In *International Journal of Document Analysis and Recognition (IJDAR),* No. 2-3, pp. 84-104, 2005.
[12] R. Liu and Y. Wang, "Auditory Feedback and Sensory Substitution During Teleoperated Navigation," In *IEEE/ASME Transactions on Mechatronics*, Vol. 17, No. 4, pp. 680-686, 2011.