NEW WWW BROWSER FOR VISUALLY IMPAIRED PEOPLE USING INTERACTIVE VOICE TECHNOLOGY

Yasuo Horiuchi  Fujisawa Atsushi  Akira Ichikawa

Chiba University
1-33 Yayoi-cho, Inage-ku, Chiba-shi, Chiba, Japan
hory@ics.tj.chiba-u.ac.jp

Abstract
It is important for visually impaired people to be able to access information via the WWW, however, this is currently difficult since conventional web browsers are based on the "GUI". In this paper, we introduce new web-browser technologies based on the concept of an interactive voice system technology. First, it is suggested that the guide method with a logical-tree structure is superior to the display-based method. Second, a new web browser which utilizes this logical guide method is introduced and a prototype web browser is developed. Based on subjective evaluations, it is demonstrated that this browser is useful for visually impaired people. Lastly, the method of guiding visually impaired people through forms and tables is introduced and its effectiveness is demonstrated by subjective evaluations.

1 INTRODUCTION
It is obvious that the WWW (World Wide Web) would be useful for visually impaired people to access a variety of information. Unfortunately, it is very difficult for them to use the WWW, because currently, conventional web browsers are based on the GUI (Graphical User Interface). Now, Screen Reader technology, which can read the CUI (Character User Interface) display out loud using a voice synthesizer, has been developed for visually impaired people[1]. Using this and the CUI WWW browser “Lynx”, visually impaired people can access the WWW. However, Lynx is not equipped to guide visually impaired people since it was designed only for sighted people. In this study, some preliminary experiments were performed and new web-browser technologies, based on the concept of an interactive voice system technology are proposed. Also, the prototype system was developed for visually impaired people who are novice PC users.

2 GUI BROWSING BY VISUALLY IMPAIRED PEOPLE
Sighted people can directly view the information on the display. However, visually impaired people can not, so the information must be expressed via Braille display or through the use of voice synthesizers. The use of voice synthesizers is preferred because Braille display is costly. To enable visually impaired people to use a computer, it must be able to reply by synthesized voice when they type on the keyboard. It is important that when something is typed the computer replies interactively, since it is difficult for visually impaired people to know the state of the display. Therefore, the computer must guide these users by voice synthesis.

2.1 Two Guide Methods on GUI
There are two typical guide methods (a direct-guide method based on a graphical structure and a guide method based on a logically structured tree) which enable visually impaired people to use the GUI. In the direct-guide method based on a graphical structure, users can operate the GUI directly using several keys. However, in the guide method based on a logically structured tree, graphical information on the GUI is transformed into a logically structured tree.

As an example, consider the situation, where two applications (“Calendar” and “Write”) are running on the GUI display (Figure 1) and the user wants to change the font style on the window of the “Write” application which is running in the background. Using the direct-guide method, which is based on a graphical structure, users
would have to carry out the following operations, in order:

1. Press the appropriate key to change applications
2. Press the appropriate key to select Menu
3. Press the right cursor key twice (File → Edit → Style)
4. Press the down cursor key twice (Plain → Bold → Italic)
5. Press the enter key (Execute the font style change to Italic)

However, using the guide method based on a logically structured tree, the graphical information on the GUI is transformed into a logical tree structure, as shown in Figure 2. Users can navigate this tree using the cursor key. For the case of the above example, at first, the user is at the “Client Region” of the application “Calendar” and wants to select the “Italic” font style in the application “Write”. In this case, users would have to carry out the following operations:

1. Press the left cursor key (Move to “Calendar” application)
2. Press the down cursor key (Move to “Write” application)

3. Press the right cursor key (Select “Write” application and the focus becomes “Menu”)
4. Press the right cursor key (Select “Menu” and the focus becomes “File”)
5. Press the down cursor key twice (Move to “Style”)
6. Press the right cursor key (Select “Style” and the focus becomes “Plain”)
7. Press the down cursor key twice (Move to “Italic”)
8. Press the enter key (Execute “Italic”)

Thus, the logical-tree-based guide method uses a smaller variety of keys, while the display-based guide method has fewer key strokes.

2.2 Comparison of the Two Methods

To determine which method is more useful for visually impaired people, subjective evaluations were performed. Four visually impaired people who are novice PC users were asked to execute simple tasks such as the above example using the two types of guide methods. Figure 3 shows the relationship between the number of key strokes and the total time required to accomplish the task. As the result shows, the logical-tree-based guide method requires more key strokes, but its total task time is less than that required by the display-based guide method. It seems reasonable to assume that the logical-tree-based method is superior to the display-based method. Furthermore, users could perform operations smoothly using the logical-based method and preferred it without exception. The seven-step MOS (mean-opinion score; 1 = cannot be used, 7 = excellent) of the logical-based method was assigned a value of 6.25 by the four visually impaired people who tested the method.
3 WEB BROWSER FOR VISUALLY IMPAIRED PEOPLE

3.1 Web Browser Based on Logical Structure

Hypertexts on the WWW are linked with other texts. Figure 4 shows an example of how these texts are linked. This structure is represented as a tree structure in figure 5.

Because of this tree structure, when visually impaired people want to browse the WWW, the logical-based guide is effective, as has been discussed. In this study, the prototype web browser (named the Audio Browser), based on a logically structured tree was developed and attached to Internet Explorer for Windows 95.

In the Audio Browser, users can browse the WWW using a numeric keypad. The key map is similar to Lynx's and is modified for visually impaired people. The key map is shown in figure 6. Key “4” corresponds to the “Back Button” on conventional browsers. When key “6” is pressed, the browser will follow the link to the new page. When key “7”, “1”, “8”, “2”, “9” or “3” is pressed, the related item is selected and read aloud by the voice synthesizer. When key “5” or “.” is pressed, the current link or current sentence are read aloud, respectively. When users want to stop the voice synthesis while it is activated, they can press the “0” key.

3.2 Evaluation of Audio Browser

To evaluate the Audio Browser, subjective evaluations were performed. The subjects were three visually impaired and six sighted people who were blindfolded. They were asked to answer certain questions on WWW pages. To choose the answer, they had to select the link which they thought was the correct answer. When they chose the wrong answer, they could return to the previous page containing the question. As a result, all the subjects answered all the questions correctly and there were few mistakes as they could move back correctly. These results suggest that the Audio Browser would be useful for visually impaired people.

4 FORMS AND TABLES

4.1 Forms

In this paper, we introduce a method to enable visually impaired people to deal with forms

---

1 This structure is a network in the strict sense.

2 Sentence means a part of text ending with a period or divided by <h1>, <hr> or <p> tags etc.
(“Radio Button”, “Check Box”, “Select Menu” or “Popup Menu”, and “Text Input Field”). For example, in the “Lynx” browser, “Radio Button” and “Check Box” look similar. Sighted people can identify them by reading the corresponding text. However, visually impaired people can only listen to the current item. Therefore it is necessary for the voice synthesizer to read aloud what kind of button it is that users want to know.

In our new method, the browser changes form tags into links (represented with texts) to allow users to listen to the description of tags. The translations of the tags are:

- “Radio Button” → the link shown as “Radio Button On/Off”
- “Check Box” → the link shown as “Check Box On/Off”
- “Select Menu” → the link shown as “Select Menu ITEM”, where ITEM is the message shown in the display
- “Text Input Field” → the link shown as “Text Input Field ITEM”, where ITEM is the current texts which were input previously

For example, when a link shown as a radio button is OFF and a user chooses the link, the system says aloud “Radio Button Off” and if the user presses key “6” (to go to the link), then the button becomes ON, the system says aloud “Radio Button On” and another button which was previously ON is changed to OFF.

When a link of “Select Menu ITEM” or “Text Input Field” is selected by pressing the key “6”, the browser shows another page which directs users to select an item or to input text data. In these situations, it is important to jump to another page. If a user can input data in the same page and unintentionally presses one of the numeric keypad or a short cut key during typing some characters in the text field, he would jump somewhere and become lost. Therefore, in our method, users can select an item or input data in another page and they can go back to the original page when they are finished. From subjective evaluation, it is shown that most subjects can use forms correctly with our method.

4.2 Tables
Tables contain two-dimensional information and are difficult for visually impaired people to browse. In this paper, we introduce how the browser can read aloud tables. Tables, on the WWW, consist of “Table Header” (<TH>) and “Table Data” (<TD>) tags. In order to recognize tables, the table header is important. Users need to know the contents of each row or each column, from the first cell to the last. Therefore, in our method, the following commands are introduced to allow visually impaired people to browse interactively:

- Move the current cell Left, Right, Up or Down
- Read aloud the current cell
- Read aloud the table header of the row containing the current cell
- Read aloud the table header of the column containing the current cell
- Read aloud the row containing the current cell
- Read aloud the column containing the current cell
- Read aloud the size of the table
- Read aloud the title of the table
- Read aloud the position of the current cell

To evaluate this method, subjective evaluations by four blind users and 11 sighted people who were blindfolded, were performed. The subjects were asked to answer some questions while browsing the tables. We obtained good results with over 90% accuracy. Thus, the interactive method acts effectively to consolidate contents of tables for the user.

5 CONCLUSION
In this paper, we introduce new web-browser technologies, based on the concept of an interactive voice system technology which enable visually impaired people to access the WWW. From subjective evaluations, it was concluded that (1) A logical structure is most effective. (2) The prototype web browser which was developed is useful. (3) Our method, to enable visually impaired people to understand contents of forms and tables, is also effective. In the future, we would like to extend these concepts not only for visually impaired people but also for people with upper limb disabilities as well as able-bodied people.

6 REFERENCES