Nonvisual Presentation and Navigation within the Structure of Digital Text-Documents on Mobile Devices

Martin Lukas Dorigo[1], Bettina Harriehausen-Mühlbauer[2], Ingo Stengel[1], and Paul S. Haskell-Dowland[1]

[1] Plymouth University, Plymouth, United Kingdom
[2] University of Applied Sciences Darmstadt, Darmstadt, Germany
martin.dorigo@plymouth.ac.uk

Abstract. This paper introduces a novel concept for an assistive technology in support of blind and visually impaired persons for nonvisual presentation and navigation within the structure of digital text-documents on mobile devices (smart phones, internet tablets, etc.) which enables them to get a fast overview over the structure of an entire document. The advantages compared to state of the art screen readers are that it enables them to identifying the type, position, length and context of each logical structure element, the current cursor position within the document as well as to accessing any structure element at any time in an arbitrary order. For the nonvisual presentation of the document structure elements auditory icons, tones and vibration feedback are used. Navigation and cursor routing is provided by performing gestures on the touch screen. There is no expensive special hardware required.

Keywords: Assistive technology, nonvisual, presentation, navigation, cursor routing, mobile devices, smart phone, internet tablet, auditory icon, tone, vibration feedback, touch screen, gestures, logical structure, overview, skimming and scanning, document accessibility, blind, visual impairment.

1 Introduction

There are thousands of digital text-documents such as websites or PDF files available on the Internet. Especially for large and highly structured documents like newspaper articles or scientific literature the structure is very important. If the reader has on overview over the structure of the entire document in mind, he is able to efficiently skimm and scann over the document and to quickly find relevant information within the document's content, he is searching for.

On visual media like screen or print, each logical structure element is physically presented by a specific visualisation, allowing sighted readers to recognise the structure at a glance. This structure can be quite complex. Therefore most user agents for the reading of digital text-documents provide a visual overview function, where the structure of the document is only (without the content) visually presented at a higher level of abstraction. The content (text and graphics) is substituted by place holders like dots, lines and thumbnails, indicating the length and size of it [Fig. 1].

Blind and visually impaired persons on the other hand can not see this visual physical presentation of the logical document structure.
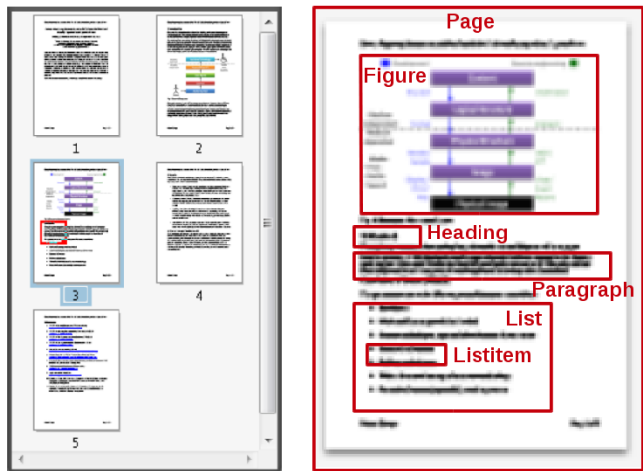

Fig. 1. Visual presentation and navigation over the structure of digital text-documents

2   State of the Art

Blind and visually impaired persons are using screen readers like Apple VoiceOver [1], Google TalkBack [2] or JAWS for Windows [3] for the nonvisual reading of digital text-documents [Fig. 2.].
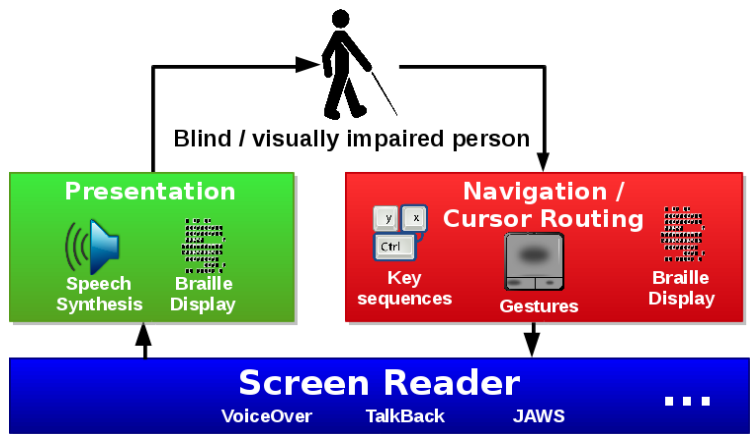

Fig. 2. State of the art screen reader

The screen reader uses speech synthesis or a braille display for presenting the text to the user. These output methods are mainly purposed for presenting text (document

content) but not for the presentation of non-text data like the logical document structure.

Navigation and cursor routing commands are provided by performing key sequences on the keyboard, using the cursor routing keys of the braille display or by performing gestures on a multitouch pad or touchscreen. These commands enable the user to identify which element types and number of each type are contained within the document for e.g. by browsing the list of headings, list of graphics, list of hyperlinks, etc. But they do not provide any information about the position, length and context of a structure element and the current cursor position within the document. Furthermore it is not possible to access any structure element at any time in arbitrary order.

In addition to these commands, VoiceOver allows the user to explore which element of the currently visible portion of the document is visually physically rendered at which position on the screen by touching the touchpad or touchscreen at an arbitrary position and by sliding horizontally or vertically. But the knowledge of this physical structure does not necessary implicate the understanding of the logical structure behind it.

Several research projects aim to present various types of structured non-text data in a nonvisual way: Presenting HTML structure in audio [4], haptic cues to aid nonvisual structure recognition [5], nonvisual overviews of complex data sets [6], document structure presentation and navigation using a braille display [7], spatial exploration of web pages [8], intelligent non-visual navigation of complex HTML structures [9] and navigation of hierarchical structures [10].

But at the moment there is no solution for nonvisual presentation and navigation within the structure of digital text-documents on mobile devices.

3   Preliminary Results

My relevant survey [11] among 205 blind and visually impaired persons proved the need for a novel concept for nonvisual presentation and navigation within the structure of digital text-documents on mobile devices.
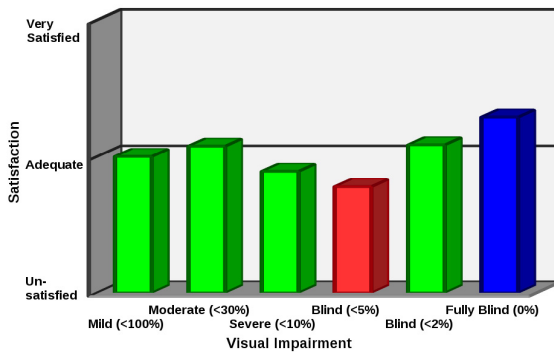

Fig. 3. Satisfaction with their current situation as to the reading of digital text-documents

At the moment all categories of blind and visually impaired persons are not very satisfied with their current situation as to the reading of digital text-documents [Fig. 3]. There is serious need for action required.

The logical document structure is very important for them as to the reading and navigation of a text-document [Fig. 4]. All of the participants would wish to have a better overview over the structure.
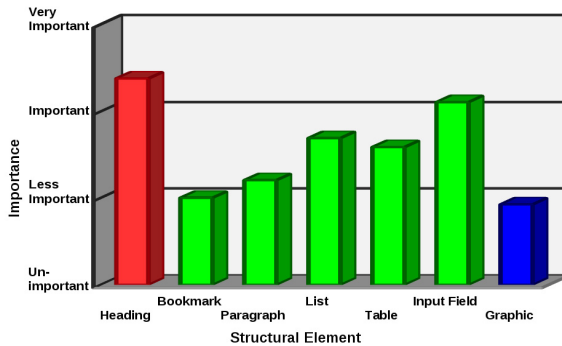

Fig. 4. Importance of logical structure elements for blind and visually impaired persons

Keyboard and mouse (for the categories who are not fully blind) are the most used input methods for the reading and navigation within digital text-documents [Fig. 5]. The participants would wish to have a more easy, intuitive and standardised method.
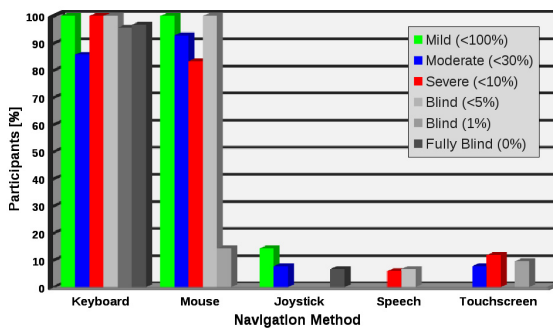

Fig. 5. Input methods used for the reading and navigation within text-documents

In addition their assistive technology should be mobile and not bound to a fixed location because mobile devices are the most used hardware by blind and visually impaired persons for reading digital text-documents [Fig. 6].
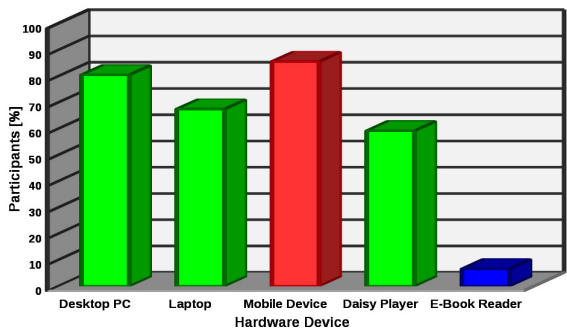
Fig. 6. Hardware devices used by blind and visually impaired for reading text-documents

4   Novel Concept

The following concept for nonvisual presentation and navigation within the structure of digital text-documents on mobile devices has been developed as a result of the state of the art and the initial survey [Fig. 7]:
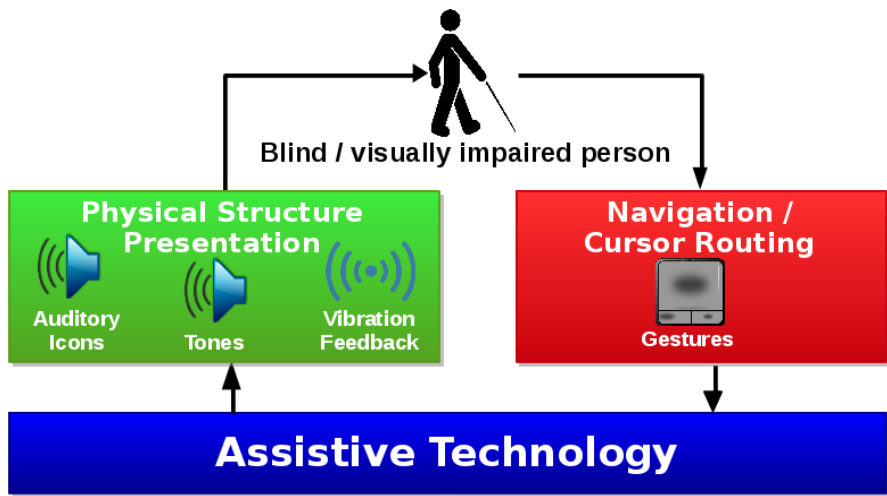

Fig. 7. Concept for nonvisual presentation and navigation within the structure of digital text-documents on mobile devices

4.1   Presentation

The logical document structure is physically presented in a nonvisual way serving the aural and the tactile human sense.

For each logical structure element type which contains content (text or graphic) an auditory icon as well as a tone has been defined. The standard elements, valid for any type of text-document, are: Heading, Paragraph, Graphic, Hyperlink, List-item, Table-row and Table-cell. In addition, context specific elements may be defined for specific types of documents. For e.g. in case of scientific literature these additional elements are: Abstract, Keywords, Summary, Conclusions, References, etc.

The auditory icon of an element is rendered when the cursor (finger) enters or passes by the element. The appropriate tones are simultaneously rendered for all nested structure elements at the current position of the cursor as long as the cursor is within an element.

The pitch at with these auditory icons and tones are rendered describes the hierarchy level of an element. Level 1 starts at musical note C4 (261.63 Hz). For each level, the pitch is increased by one note until B4. This allows the presentation of a maximum of 7 (resp. 12 if half tones are included) distinguishable hierarchy levels.

As with the aural medium, for each logical structure element containing content a vibration pattern has been defined. The appropriate vibration patterns are simultaneously rendered for all nested structure elements at the current position of the cursor as long as the cursor is within an element.

The speed at which the vibration pattern is rendered describes the hierarchy level of the element.

4.2   Navigation

Navigation and cursor routing are provided by performing gestures on the touch screen. The structure elements of the entire document are laid-out on the touchscreen in reading order from top to the bottom where the height represents the length of the content included within an element.

By putting down one finger, the cursor is set to the specific position and the appropriate auditory icons, tones and vibration patterns are rendered. By sliding up and down with one finger, the user is able to navigate within the document's structure in reading order forward and backward at an arbitrary speed. By lifting up the finger, the cursor is removed and rendering stops.

4.3   Cursor Routing

By single tapping with one additional finger anywhere on the touchscreen during the first finger (cursor) is down and pointing at a specific structure element, this specific element can be selected and the system cursor routed to it. This can be used for e.g. for reading the content contained in the selected element (text or textalternative for graphics) to the person using speech synthesis.

4.4   Advantages

The advantages of this proposed concept compared to state of the art screen readers are that it enables blind and visually impaired persons to:

1. Identify the **type** of each structure element and number of each type contained within the document
2. Identify the **position** of each structure element within the document
3. Identify the **length** of the content contained in each structure element
4. Identify the **context** of each structure element
5. Identify the current **cursor position** within the document
6. **Access** any structure element at any time in an arbitrary order

There is no expensive special hardware required. All features used are included in mobile devices out of the box. However the concept is not limited to mobile devices. It can also be used on a regular laptop or desktop PC by connecting a standard multi-touch pad and vibration feedback device to it.

In addition to blind and visually impaired the concept also benefits persons without disability in situations where the visual sense is in use by other activities and since the displays of mobile devices are much smaller than on regular laptop and desktop PCs.

Furthermore the concept is not limited to the context of text-documents. It can be used for nonvisual presentation and navigation within any hierarchical structure containing any kind of content (for e.g. program source code or file systems).

5 Evaluation Methodology

As proof of concept, a prototype App has been implemented for conducting a quantitative analysis. A pilot test with the implemented prototype is going on among blind and visually impaired participants online via Internet.

5.1 Research Questions

The hypotheses in [Chapter 4.4] lead to the following research questions. Does the proposed novel concept enable blind and visually impaired persons to:

1. Identify the type of each structure element and number of each type contained within the document ?
2. Identify the **position** of each structure element within the document ?
3. Identify the **length** of the content contained in each structure element ?
4. Identify the **context** of each structure element ?
5. Identify the current **cursor position** within the document ?
6. **Access** any structure element at any time in an arbitrary order ?
7. Receive **additional value** compared to using state of the art screen readers ?
8. Improve their **satisfaction** as to the reading of digital text-documents compared to the initial situation [Chapter 3] ?

5.2 Procedure

After the participants have downloaded, installed and started the prototype App on their own mobile device, the following procedure will be initiated [Fig. 8].
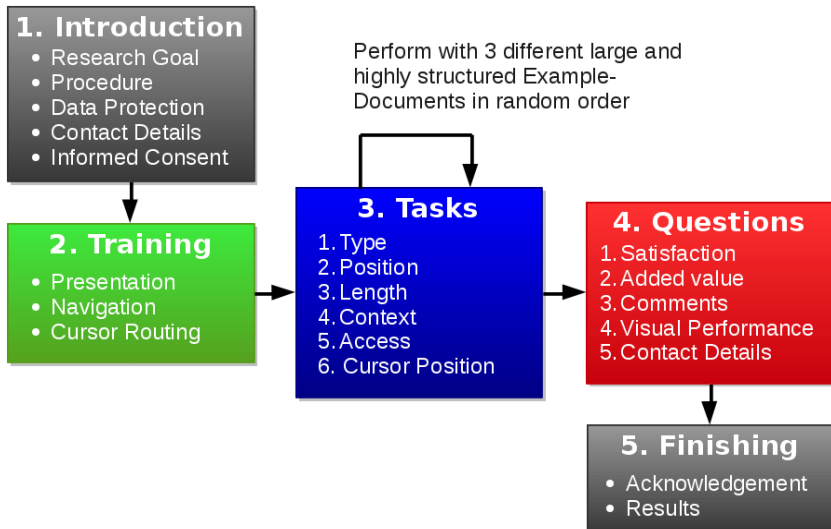
Fig. 8. Evaluation Methodology Procedure

At the beginning an introduction gives detailed description of the research containing all relevant information, a person requires to decide, if the person would like to participate within the pilot test or not. This includes: Research goal, procedure, data protection, contact details of the researcher as well as of the supervisory team and ethical approval. If the person decides to participate into the pilot test, he or she gives informed consent and confirming the minimal age 18 years.

The introduction is followed by a training of the novel auditory icons, tones and vibration pattern for the aural and tactile presentation of each structure element as well as the gestures for navigation and cursor routing.

After the training the participant is asked to perform the following tasks. This part will be repeated with 3 different large and highly structured example documents in random order. In addition, the time as well as finger movements performed are recorded.

1. Identifying which element **types** and number of each element type are contained within the document.
2. Routing the cursor to the **position** of a specific element within the document
3. Routing the cursor to the **longest** as well as to the **shortest** element
4. Finding the element within a specific **context**. For e.g. the graphic within a third subsection followed by a list
5. Locating the current **cursor position** within the document
6. **Accessing** different structure elements in an arbitrary order. For e.g. selecting an element in the middle, afterwards one at the end and finally one at the beginning of the document.

Subsequent to the tasks, the participants are asked to answer the following questions:

1. Does the proposed concept provide an **additional value** for them compared to using state of the art screen readers ?
2. How **satisfied** are the participants with their current situation as to the reading of digital text-documents using this novel concept ?
3. Additional **comments** like advantages, disadvantages and ideas for further improvement ?
4. Which **visual performance** do the participants have ?
5. Optionally **contact details** like e-mail, phone number or address. This is used for further questions and for sending the published results.

In the end, the participants are acknowledged for their participation in this pilot test and they are informed how they can obtain the published results. After that the App will be automatically terminated and the result will be sent back via Internet.

5.3   Prototype App

A prototype App, that implements the presented test procedure, has been developed. This App is freely available for anyone who is interested in testing this novel concept and in participating in this pilot test using their own mobile device. It can be downloaded from the Apple App Store on iTunes for iOS devices as well as from Google Play for users of the Android operating system.

5.4   Pilot Test

At the moment a pilot test is going on among 163 blind and visually impaired persons using the implemented prototype App online via Internet using their own mobile devices. First results successfully proved that the proposed concept provides a potential solution and may be a great contribution to the field of document accessibility. The results of this pilot test will be published in further publications.

6   Conclusions

The logical structure of large and highly structured text-documents is very important for the reader but blind and visually impaired persons can not see the visual physical structure presentation. At the moment, there is no solution for nonvisual presentation and navigation within the structure of digital text-documents on mobile devices. Therefore a novel concept has been developed using auditory icons, tones and vibration feedback for presentation and gestures on the touch screen for navigation and cursor routing. First results of an ongoing pilot test among blind and visually impaired participants using an implemented prototype App successfully proved that the proposed concept provides a potential solution and may be a great contribution to the field of document accessibility.

References

1. Apple, VoiceOver for iOS,
   `http://www.apple.com/accessibility/voiceover` (last accessed: January 28, 2013)
2. Google TalkBack: An Open Source Screenreader for Android,
   `http://www.google.com/accessibility/products/` (last accessed: January 28, 2013)
3. Freedom Scientific, JAWS for Windows Screen Reading Software,
   `http://www.freedomscientific.com/products/fs/`
   `jaws-product-page.asp` (last accessed: January 28, 2013)
4. James, F.: Presenting HTML Structure in Audio: User Satisfaction with Audio Hypertext. In: Proceedings of the International Conference on Auditory Display, ICAD 1996, Palo Alto, CA, USA, November 4-6 (1996)
5. Jay, C., Stevens, R., Hubbold, R., Glencross, R.: Using haptic cues to aid nonvisual structure recognition. Journal ACM Transactions on Applied Perception (TAP) 5(2), Article No. 8 (2008)
6. Kildal, J., Brewster, S.A.: Non-visual overviews of complex data sets. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, CHI 2006, Monréal, Québec, Canada, April 22-27 (2006)
7. Minatani, K.: Development of a DAISY Player That Utilizes a Braille Display for Document Structure Presentation and Navigation. In: Miesenberger, K., et al. (eds.) ICCHP 2012, Part I. LNCS, vol. 7382, pp. 515–522. Springer, Heidelberg (2012)
8. Petit, G., Dufresne, A., Robert, J.-M.: Introducing tactoWeb: A tool to spatially explore web pages for users with visual impairment. In: Stephanidis, C. (ed.) Universal Access in HCI, Part I, HCII 2011. LNCS, vol. 6765, pp. 276–284. Springer, Heidelberg (2011)
9. Pontelli, E., Gillan, D., Gupta, G., Karshmer, A., Saad, E., Xiong, W.: Intelligent nonvisual navigation of complex HTML structures. In: Universal Access in the Information Society, vol. 2(1), pp. 56–69 (2002)
10. Smith, A.C., Cook, J.S., Francioni, J.M., Hossain, A., Anwar, M., Fayezur Rahman, M.: Nonvisual tool for navigating hierarchical structures. In: Proceedings of the 6th International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS 2004, New York, New York, USA, October 18-20 (2004)
11. Dorigo, M., Harriehausen-Mühlbauer, B., Stengel, I., Dowland, P.S.: Survey: Improving document accessibility from the blind and visually impaired user's point of view. In: Stephanidis, C. (ed.) Universal Access in HCI, Part IV, HCII 2011. LNCS, vol. 6768, pp. 129–135. Springer, Heidelberg (2011)