Media Technology

Creating a better web browsing experience for visually impaired people by analysing and combining existing methods

First Bachelor Thesis

Completed by

Francesco Novy
mt101064

From the St. Pölten University of Applied Sciences
Media Technology degree course

Under the supervision of
MSc Georg Kaindl

St. Pölten, on

(Signature Author)

Media Technology

Declaration

- the attached research paper is my own, original work undertaken in partial fulfillment of my degree.

- I have made no use of sources, materials or assistance other than those which have been openly and fully acknowledged in the text. If any part of another person's work has been quoted, this either appears in inverted commas or (if beyond a few lines) is indented.

- Any direct quotation or source of ideas has been identified in the text by author, date, and page number(s) immediately after such an item, and full details are provided in a reference list at the end of the text.

- I understand that any breach of the fair practice regulations may result in a mark of zero for this research paper and that it could also involve other repercussions.

St. Pölten, on

____

(Signature Author)

Media Technology

Abstract

With the growing importance of the internet, there is also a growing responsibility to provide a good user experience for visually impaired people who browse it.
Different aids can be used for different problems, and this work wants to find out how different techniques could be utilised together to improve the quality of web browsing for sightless people. There are already various papers describing audio-only or tactile browsers or general accessibility issues, and the main point of this thesis is to collect and analyse different techniques to combine them. The capabilities of current web browsers concerning accessibility will be examined as well. At the end of this paper, a unified audio-tactile browser will be proposed by combining the previously investigated methods, based mainly on a tactile graphics tablet and a 3D audio engine.

Table of Contents

1. Introduction

Few would dispute that the World Wide Web (WWW) belongs to the most important innovations of the last 50 years. Its relevance is ever increasing, with no change of course in sight. New web technologies are constantly developed and employed, and the Internet has over the time transformed from a mainly text-based place to a multimedia-based medium.

While the WWW's importance grows, it also entails a growing responsibility to provide ways and means for visually impaired people to browse it. While there are various mainstream solutions to access information online, most effort has been done only to facilitate basic access, not to make it a joyful experience. Nowadays, the internet provides more than raw information – there are various interactive web services like Facebook (www.facebook.com) or Twitter (www.twitter.com) that play major parts in people's everyday lives. While sighted users are used to deal with a multitude of complex and dynamically changing visual information, using these pages is much more difficult for people with sight impairment.

There are plenty of papers out there which address the topic of accessibility for visually impaired people in the area of web browsing, and even more concerning blind or partially sighted people in general. General information about accessibility (in particular in web design) have been the focus of various papers (Atkinson et al. 2006; Pun et al. 2007; Takagi et al. 2007; Wang/Xu, X./Li 2007; Hamam et al. 2008; Hailpern et al. 2009; Jay/Brown/Harper 2010a). However, due to the multitude and complexity of research, it is difficult to get an orientation and an overview of the respective state of the art of science and technology.

While there are already various techniques like audio-only or tactile browsers out there, virtually no research has been done to analyse and compare these existing methods or to combine them. Therefore, this paper sets out to fill this gap. Its goal is to compile various differing methods and analyse them, as well as to try and combine them. The main focus lies not on the transmission of raw information, but on doing so in a way that is as enjoyable for visually impaired people as possible.

This leads to the following specification of the central research question:

Can existing methods for the display of web content for visually impaired people be combined to provide a web browsing experience that incorporates as many features that sighted users can use as possible?

This thesis will be separated in four main sections. First, a list of features that are currently used on common web pages will be compiled. All methods and browsers will be evaluated to find out in what extend they implement these pre-established criteria.

In section 2 (chapter 3) a short overview of mainstream-technology will be given, concentrating on the major screen reader software and their capability to convey web content to visually impaired people.

Section 3 will cover a few important methods for the transmission of web content to visually impaired people. The methods were picked by browsing through scientific online libraries like ACM (dl.acm.org) or Google Scholar (scholar.google.at). The selection of methods was generated by searching for related key words and by exploring referenced papers. While the chosen methods are not exhaustive, all major techniques that have been found have been considered.

Audible as well as tactile solutions will be examined to explore how various types of accessible interfaces and methods of non-visual rendering could be combined. These solutions for visually impaired people should be, if possible, not only instrumental but also pleasant to use. Differences between these techniques will also be described. Each solution will be reviewed regarding the pre-established criteria (chapter 4 and 5).

In the final section (chapter 6), a possible unified audio-tactile browser will be discussed. For this purpose, possibilities to combine previous findings to create such a unified browser will be analysed. The tactile graphics tablet and the 3D audio browser will form the basis of this unified browser, and general points and problems concerning such a web browser will also be discussed. While the realisation, much less the testing of this unified browser would go beyond the scope of this work, there are many findings in other papers that can be applied. The coverage of the pre-established criteria of this unified audio-tactile browser will be compared to the ones of the single methods without combination.

2. Aspects of an enjoyable web browsing experience

In order to be able to compare the different methods, a list of points will be compiled containing the features that are responsible for an enjoyable web browsing experience. The author will explain why certain aspects are important as well as ranking them in three categories: not very important (1 point), moderately important (2 points) and very important (3 points).

All following techniques will be evaluated on the basis of this list. If a method covers all the features, it can be considered as giving roughly the same experience as a sighted user would have, which would be the ideal solution. Only methods that can be used on their own will be graded, because it would be difficult to compare supplementary techniques to stand-alone solutions.

- **Accessing the basic textual information (3)**
  In general, this is the most important aspect that a browser has to cover. Textual information is the central content on most web sites, and therefore accessing it is a basic necessity.

- **Exploring multimedia content (3)**
  While a matter of course for sighted users, graphical information can pose a major problem for blind or partially sighted people. Nearly all modern websites use graphics, some more excessively than others, and without a way to explore them a lot of information can get lost. Videos are another type of content that can be found on the internet and that is difficult to access for visually impaired people.

- **Knowing their position in the document (1)**
  Because visually impaired people cannot check their position in the document with a quick glance at the scrollbar, other means have to be found to find out where in the web page they currently are. Although this feature can often be useful, it is not really essential for a good web browsing experience.

- **Getting styling information (2)**
  The use of Cascading Style Sheets (CSS) is ever growing, because it enables web developers to realize complex designs and layouts. To convey the styling information to people with sight loss would approximate their experience to the one of sighted users and enable them to perceive a web page in the way the web designer intended it to be perceived.

- **Knowing where hyperlinks are leading to (2)**
  While sighted users can find out if a certain web site is interesting for them by quickly scanning it, visually disabled people lack this possibility. If a technique is implemented to compensate the lack of this ability, blind or partially sighted users could save a lot of time by avoiding unnecessary exploration of non-relevant web pages.

- **Price and Availability (3)**
  Because most individuals who are blind or partially sighted live below the poverty line (Headley/Pawluk 2010), the price for browsing solutions is an important aspect that should not be overlooked. Of course, potential pricing information has to be taken with a grain of salt, because real world costs could possibly differ a lot if a product would go

into mass production compared to the cost of a prototype. Apart from the price, the percentage of visually impaired people that could use a specific method will also be taken into account: For example, a technique that relies on Braille can only be used by people with Braille literacy.

- **Getting an overview of a web page (1)**
  Modern websites use a multitude of layouts, which sighted users can easily digest. However visually impaired people have no easy way of seeing a web page as a whole. This feature would allow them to perceive the overall picture of a website, or at least to get a feeling of what elements can be found around the user's current position.

- **Handling dynamic content changes (2)**
  The ability to perceive dynamic content updates is often needed in modern websites. Some kind of technique has to be used to deliver these content changes to visually impaired people, while avoiding to overwhelm them with too much simultaneous information.

If all proposed criteria are met, a maximum of 17 points can be reached. If a feature is only partially covered, a part of the points can be awarded. The following table shows this list with shortened table headers, in the same order as listed above:

	Text	Media	Position	Styling	Links	Avail.	Overview	Dynamic	Total
Maximum	3	3	1	2	2	3	1	2	17

Table 1: Evaluation schema (Own compilation)

This list lays no claim to be exhaustive, but the author believes that, while others might put differing emphasis on the different features, overall it reflects the most important points that account for a pleasant web browsing experience.

3. A short introduction to screen readers

A screen reader is "a text-to-speech system, intended for use by blind or low-vision users, that speaks the text content of a computer display." (Anon. n.d.) Alternatively, modern screen readers can also write in Braille on special output devices called Braille lines (Rotard/Knödler/Ertl 2005).

Usually, screen reader users use Microsoft Windows operating system because it offers the biggest choice of screen reader software – the majority of visually impaired people do not use Windows' built-in accessibility features. However, the number of Mac OS X users has recently increased because of improvements of Apples built-in VoiceOver screen reader (Borodin et al. 2010).

Most other screen reader software only runs on Microsoft Windows-operating system. The most used screen readers, according to a survey made by WebAIM.org in May 2012, are JAWS, Window-Eyes, NVDA and System Access (To Go) (WebAIM.org 2012).

According to a practical screen reader comparison made by the Iceland National Institute for the Blind, Partially Sighted and Deaf Blind (INIB), the screen readers JAWS, HAL and NVDA fared best when browsing the web, while VoiceOver and Orca (an open source screen reader (Orca Website n.d.)) showed some shortcomings (Hreinsson 2011) – Though it has to be noted that they tested performance in Internet Explorer 8, Firefox 3.5 and an older Safari version only, so there may have been improvements since then. Their final verdict is:

> "Overall we definitely found that the applications do pretty well with allowing users to browse the Internet. We would recommend any of the Windows Screen Readers but reserve a bit of judgment on the Apple approach to web browsing, noting that there may be revolutionary interface innovations in the near future if some of the touchpad features found in iOS are ported to their desktops or laptops. We found navigating the web with Firefox and Orca slow and cumbersome and not up to our expectations, though we certainly know there is a lot of work going into improved user experience with Orca." (ibid.)

Some screen readers are for free while others cost several hundred Euros (Freudenfeld 2010).

There are a few basic problems with screen readers and likewise with most web displaying methods for people with sight loss, the main one being dynamic content. To tackle this problem, the W3C introduced the Accessible Rich Internet Applications Suite (ARIA). ARIA "…defines a way to make Web content and Web applications more accessible to people with disabilities. It especially helps with dynamic content and advanced user interface controls developed with Ajax, HTML, JavaScript, and related technologies." (W3C 2011) However, most of the dynamic content available today does not implement the ARIA standard (Borodin et al. 2010), because it would mean considerable additional work and therefore expense.

While most traditional screen reader applications like JAWS maintain a static representation of web content for the user, which do not update except when a new page loads, other aids like the FireVox browser (ibid.) or the SASWAT audio web browser

(Jay/Brown/Harper 2010b) announce content updates as they occur. This approach, while seeming better, entails an entirely new problem, namely that without correct ARIA attributes, web browsers are unable to filter out irrelevant updates and therefore often interrupt and distract the user from more important information.

Another fundamental problem with screen readers is that many users use them with audio output, which leads to problems when there is sound from media in the background that is masking it. This makes visually impaired people unable to hear the screen reader and causes them to lose control of the page. A solution for this problem would be to add a shortcut to the screen reader to target page-audio volume separately from the global (screen reader) audio (Miyashita et al. 2007).

In regard to the pre-established criteria, all screen readers fare nearly identical. Therefore, they will be represented as a single item "screen readers". They can accurately convey the textual information found on websites. Images, on the other hand, are only represented via simple descriptions, which is not really helpful most of the time. While some screen readers are quite expensive, the free alternatives are sufficiently capable too. Therefore, and because they work with Braille as well as with audio, 3 points are awarded to availability. The final score for screen readers is 6 points.

	Text	Media	Position	Styling	Links	Avail.	Overview	Dynamic	Total
Maximum	3	3	1	2	2	3	1	2	17
Screen readers	3	0	0	0	0	3	0	0	6

Table 2: Evaluation of the major web browsers (Own compilation)

4. Haptic aids

Touch is an essential sense for visually impaired people, so it comes as no surprise that there are several solutions for blind and partially sighted people that are based on haptic features. In this chapter, Braille as the basis of reading for visually challenged people will be covered shortly, followed by an introduction of a few selected haptic displaying methods and a short analysis of their problems and advantages.

4.1. A short introduction to Braille

Braille is "a system of writing or printing, devised by L. Braille for use by the blind, in which combinations of tangible dots or points are used to represent letters, characters, etc., that are read by touch.” (Anon. n.d.) It is widely used for visually impaired people to be able to read textual information. While a statistic by the National Federation of the Blind shows that Braille-reading is an essential skill to improve the quality of live for people with sight impairment, it also notes: “Fewer than 10 per cent of the 1.3 million people who are legally blind in the United States are Braille readers. Further, a mere 10 per cent of blind children are learning it.” (Jernigan Institute 2009) A survey showed that only 28 per cent of screen reader users use their Braille output at all (WebAIM.org 2012). Other numbers concerning Braille are no better: Fewer than 1 per cent of the two million blind or partially sighted people in the UK are users of Braille (Rose 2012), and while the number of working people who are registered as blind in the UK is significantly higher at 20 per cent, it is still not very high. Of the blind and partially sighted population of the UK of five to sixteen years old, around 4 per cent are Braille users (The Office for Disability Issues 2009).

In this context, it has to be noted that all tactile reading techniques rely heavily on the user’s ability to read Braille. As long as the Braille literacy is not increased dramatically, these solutions will not be able to gain widespread success. Therefore, a major step for the improvement of the web browsing experience of visually impaired people would be to drastically raise the Braille literacy. As long as this doesn’t happen, further enhancement in the tactile area will not have any considerable effect, because a large part of the methods described in this paper rely heavily on Braille. Users with no Braille skills can hence only use auditive aids.

4.2. Tactile graphic tablets for text and image display

Tactile graphics displays are special devices that use a pin matrix device with a certain resolution (for example 120x60 pins), which are set electromagnetically.

Providing different versions of web pages for visually disabled people with a reduced layout is not a satisfactory solution, because visually impaired people want to browse on the same web pages, use the same links and read the same content in the same layout that sighted people do. Special versions of a web pages help to access the textual content, but they don’t convey the entire content and the feeling of the website.

Martin Rotard, Sven Knödler and Thomas Ertl from the Visualization and Interactive Systems Institute from the University of Stuttgart proposed a tactile web browser for visually disabled people that utilises a tactile graphic tablet (Rotard/Knödler/Ertl 2005).

They proposed a tactile web browser for XHTML documents that renders text, graphics and mathematical content in a way that conveys the whole content of a website. The idea is that screen readers, which extract the textual information which is displayed on the screen and linearize it, are only able to get the written information across to the user, but not the graphical information. Especially in scientific education it is necessary to have access to images, diagrams and formulas, because without this information visually impaired students may not get a deep understanding of the learning materials.

They use a transformation schema that can process XHTML documents. The tactile browser extracts the styling information from the Cascading Style Sheets (CSS) and recalculates the layout of the document for the reduced resolution of the tactile graphics display. Various methods are used for the styling of the content, like the automatic inclusion of special tags for text attributes like colour, links or underlined text or the implementation of a transformation of hexadecimal colour values into the nearest colour keyword name.

There are, of course, various problems that arise when dealing with tactile graphics displays. Because of high cost, the tablet's resolution has to be kept at a low level. Images have to be scaled down to fit the lower resolution, but it is important to retain the ratio between the size of the image and the size of the user screen, because a small image should continue to appear small. Many HTML elements like tables, lists or frames require some kind of rethinking or restyling to be presented in a better way.

There also has to be a dedicated key to cycle through the links and images that are currently displayed on the tablet, to prevent the user from losing the context like it would be the case when using the standard TAB-key to cycle through interactive elements. By selecting an image, a special exploration mode is started, in which the selected image is shown full-screen. There are several complex methods used to convert images in a way that makes it possible to display them on a tactile graphics tablet.

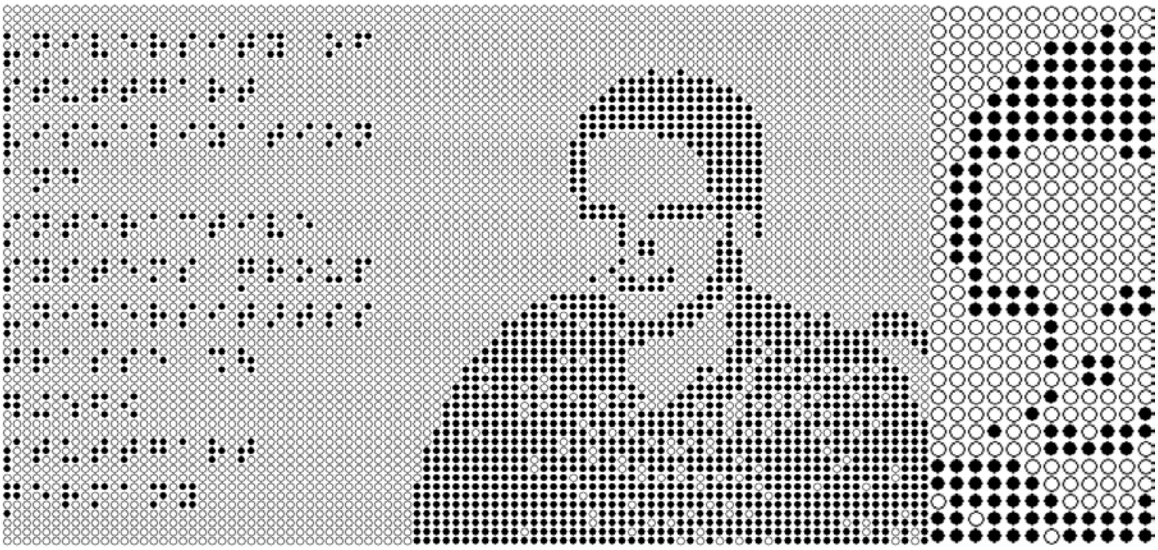

Figure 1: A simple web page displayed on a tactile graphics tablet (simulation on the left and magnified area on the right). (Rotard/Knödler/Ertl 2005)

In general, these tablets provide a very good way for visually impaired people to explore websites. The biggest draw-back is the low resolution and resulting problems with complex, wide websites. Since 2005, websites have become more and more sophisticated, while current tactile graphics tablets still move at around 120x60 pins (Headley/Hribar/Pawluk 2011). Additionally, they usually employ a comparatively big pin spacing (i.e., 2.5 mm for the hyperBrailler by Metec AG), which is limiting for presenting small graphics. Modern LCD displays use resolutions of up to 1920x1200 pixels, outpacing tactile graphics tablets by a factor of over ten.

Another problem with this technique is the missing possibility to display text on top of images, making the display of background graphics impossible. To counter this problem, a multi-amplitude pin layout can be used (ibid.), although this would lead to unreasonably high priced devices.

The price of such tablets is, while not unreasonably high, not favourable for people with little money to spare: For example, the Talking Tactile Tablet 2 by Touch Graphics Inc (which does not provide a web browser, but a 120x60 pin layout similar to the one specified by Martin Rotard and Sven Knödler) is currently available online for $799.00 (Touch Graphics Inc. n.d.). There are also other approaches which stray from the pin-layout, like the TeslaTouch, which works with electric signals that get processed by a special wristband which amplified the signal and delivers it to the finger. While this makes it possible to provide higher resolutions, because no mechanical pieces are necessary, it also results in a significantly slower reading speed (Xu, C. et al. 2011), which is not feasible for comfortable web browsing.

The last flaw these tablets have is their inability to adapt to dynamically changing content. While sighted users would see a popup at the top of the screen and can then choose to react to it, blind or partially sighted users will not notice content changes without further notice. If dynamic content change happens below the current reading position, the user will not even know that there has actually ever been a different content there, and if it occurs before, he will probably not even notice it at all. And this is under the premise that the tablet and respectively the browser software even detect dynamic changes and execute them on the tactile surface.

Overall, these tablets are very good to explore textual or graphical content. They cannot play back videos; therefore they only get 2 points awarded for the multimedia-feature. Because of their relatively high price and their dependence on braille literacy, no point is given for availability. In total, a score of 8 points is reached.

	Text	Media	Position	Styling	Links	Avail.	Overview	Dynamic	Total
Maximum	3	3	1	2	2	3	1	2	17
Tactile graphics tablet	3	2	0	2	0	0	1	0	8

Table 3: Evaluation of the tactile graphics display (Own compilation)

4.2.1 Movable, mouse-like tactile displays

Instead of displaying a website on a (comparatively) big “full page” pin display, a small, mouse-like tactile display can be used. This device works by showing the user what is currently “under the mouse” on a small (i.e., 2x4 pins) display. By moving the mouse around, another part of the screen is shown on the pin matrix (Headley/Hribar/Pawluk 2011). Headley, Hribar and Pawluk’s device works with a four-stage multi-amplitude pin layout that works by showing Braille text with full amplitude and background images with approximately 50 per cent amplitude, to enable users to differentiate between them.

They also note that while their technique seems to work well enough, the average time to find and read only five Braille labels in one diagram was, according to their test, 9.4 minutes – a time span that may be acceptable for a scientific diagram or a map, but which seems unreasonable when concerning webpages. Furthermore, while this method works well for the display of maps and smaller images, it is uncertain how well it would work with websites, which cover a much bigger space, and where an instant overview over the whole page is very important.

Expanding the multi-amplitude pin matrix to a bigger screen would increase the devices price to very high levels, considering that for a 2x4 pin device the price is "…under $500 for the prototype" (Headley/Pawluk 2010) – especially in regard to the fact that most blind or visually impaired people live below the poverty line (ibid.).

The evaluation of the movable tactile display is basically similar to the bigger tactile graphics display. The main difference is the lack of overview for the user. The author has also deducted one point for textual information and one point for multimedia, because reading is noticeably slower than on the larger displays. The total score is 5 points.

	Text	Media	Position	Styling	Links	Avail.	Overview	Dynamic	Total
Maximum	3	3	1	2	2	3	1	2	17
Movable tactile display	2	1	0	2	0	0	0	0	5

Table 4: Evaluation of the movable, mouse like tactile display (Own compilation)

4.3. Haptic wearables for videos

Haptic wearables are garments that have tactile devices included. The tactile device can be of any shape or size and serve for varying purposes, like, amongst others, to store haptic information regarding a video. For this purpose, a rectangular layout like a video screen can be used (Rahman et al. 2010).

Rahman, Alkhaldi, Cha and El Saddik from the University of Ottawa propose a jacket and a haptic arm band. Through a dedicated online service, tactile authors wearing the wearables can add tactile properties to a certain scene of variable duration and store the tactile content on a web server. Users can then choose from all videos with haptic annotation and access them, causing the wearables to render the feelings (ibid.)

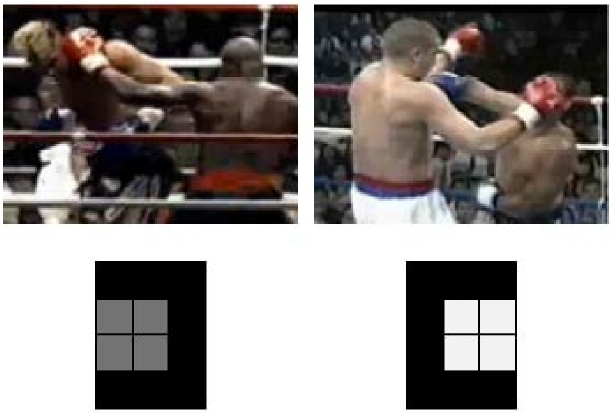

Figure 2: A content example showing left-hand and right-hand punches (Rahman et al. 2010).

With this technique, it is possible to convey basic feelings to visually impaired users. However, voice over or similar means are still necessary to convey real information from a video. The wearables are quite complex and contain many electronic parts, which, in general, does not bespeak exceptional durability or a low price.

Additionally, this method can only be useful if the type of video suits it: The experience would probably not be greatly enhanced when watching a romantic comedy or report.

Another problem is the sources: In their research, they used the Youtube API to access video material, but without an author who adds the haptic information there is nothing to watch. Because of this, haptic wearables can only become a feasible aid for blind or partially sighted people if widespread adoption takes place and enough people start authoring content.

A big benefit of this technique is that it is, except for the need of specific annotation, completely agnostic of most other methods. It can be used with a screen reader or an audio browser as well as with a tactile graphics tablet. If the necessary equipment would go in mass production and therefore be sold at a reasonable price, it could eventually reach a critical mass and generate enough haptic annotation. Such wearables can also be an enjoyable addition for sighted users, creating a more immersive experience for everybody. This fact could help garner more users and consequently more content creators.

Because these haptic wearables cannot be used without some other kind of web browsing technology, no score will be given.

5. Auditory aids

Hearing is the other sense with increased importance for people with sight loss. In comparison to haptic information, audio has greatly differing qualities. In this chapter, some audio-based techniques beyond current screen readers will be presented.

5.1. Audio enriched hyperlinks

It is not an unreasonable effort for sighted people to explore undescriptive hyperlinks like “click me” or “follow this link”, because they can quickly scan the new site and decide if it is of importance to them. However, their visually impaired counterparts do not have this luxury, because common screen readers can only depict content in a linearized way from the top to the bottom of the page.

To solve this problem, Peter Parente developed audio previews for links, giving users the ability to get a quick overview of the page a hyperlink is pointing to (Parente 2003). His method pre-fetches HTML sites and parses basic information, like the title of the page or the type of content that forms the majority of the page (e.g. tables, images …).

In general, audio enriched hyperlinks are a very simple and effective way to provide visually impaired people with a “glance” at a following site. The usefulness is of course dependent on the information that can be gathered about a website: If this information is too long, blind or partially sighted people lose the main advantage of saving time, if it is too short, they will have to visit the site anyway because they cannot be sure what it is about.

Parente’s current implementation of Audio Enriched Links (AEL) is available as an extension for JAWS at no cost. However, development has not continued since 2005 (Parente 2005).

As with the haptic wearables, audio enriched hyperlinks are only supplementary to other techniques, and cannot be used on their own. Because of this, no score will be given for them. Adding them would be an easy way to let visually impaired people know where hyperlinks are leading to.

5.2. VoiceXML and automatic semantic partitioning

VoiceXML is a web standard defined by the W3C to describe interactive voice dialogues between a human and a computer. It is specified as:

> "VoiceXML is designed for creating audio dialogs that feature synthesized speech, digitized audio, recognition of spoken and DTMF key input, recording of spoken input, telephony, and mixed initiative conversations. Its major goal is to bring the advantages of Web-based development and content delivery to interactive voice response applications." (W3C 2004)

VoiceXML can be useful for visually impaired users; however it would be extremely time-consuming for web developers to maintain their content in various formats (i.e., HTML and VoiceXML). Because of this, Ramakrishnan, Stent and Yang from the Stony Book University developed the HearSay Audio Browser (Ramakrishnan/Stent/Yang 2004), which automatically parses web sites and partitions them through a structural and semantic analysis.

HearSay creates a semantic partition tree and then generates a speech dialog interface to this tree. Such a dialog could proceed as follows:

Computer:	Please choose one of these: Top-Navigation 1, Top-Navigation 2, Headline News, Exit
User:	Top-Navigation 1
Computer:	Please choose one of these: Menu Item 1, Menu Item 2, Menu Item 3, Back
User:	Menu Item 3

…

In this way, blind or partially sighted people can browse web pages in a more natural way. It can also be used by sighted users in situations when they cannot interact visually with a web browser, for example while driving. However, the effectiveness of this method heavily relies on the quality of the partitioning that has been done: If the site cannot be analysed correctly, the resulting dialog may not be very satisfactory for the user. It can also only work on information-heavy websites, web services like Facebook or Twitter are hard to be partitioned in a reasonable way.

Strengths of Hearsay are its ability to implement dynamically changing web content by providing an interface for reviewing the changes and its free availability (Borodin 2012).

A VoiceXML browser with automatic semantic partitioning would therefore reach 8 points of the pre-defined criteria: It can express textual information and beyond that can handle dynamic content updates. Because it is free and is not dependent on Braille, 3 points are awarded for availability.

	Text	Media	Position	Styling	Links	Avail.	Overview	Dynamic	Total
Maximum	3	3	1	2	2	3	1	2	17
VoiceXML browser	3	0	0	0	0	3	0	2	8

Table 5: Evaluation of the VoiceXML browser (Own compilation)

5.3. 3D-audio web browsers

Stuart Goose and Carsten Möller proposed a 3D audio only interactive web browser, which uses spatialization to convey the structure of HTML documents (Goose/Möller 1999). They collected various methods to adapt audio browsers to the increasing complexity of website GUIs. As most other approaches, they considered the inability to get an overview of a web document as the biggest problem for blind or partially sighted users. Their 3D audio web browser analyses a standard HTML page and then produces an audio rendering which combines the use of discerning text-to-speech synthesizer-voices and of earcons (distinctive sounds that are used to convey specific information).

During various preliminary experiments, they were able to find out that humans can only accurately identify auditive positions on the x-axis, while on the y- and z-axis position detection was only working more roughly. This is a big drawback for 3D audio systems, because it significantly reduces the usable room for auditive signals. For their experiments, they used a semi-circle around the front of the listener's head (compare Figure 3).

The main content is read from a static position right in front of the user, while a second, different voice periodically announces the current position through the document as a percentage, originating at a position along the arc relative to the position through the document. Even though these two voices are deployed simultaneously, they can still be distinguished because of the so-called "cocktail party effect" (which describes the ability of a person to focus their listening attention primarily upon a single talker while surround by other voices or noise) and by employing different sounding voices and different volume levels.

When encountering hyperlinks, a specific earcon is sounded at the relative position along the arc, therefore enabling the user to know approximately where in the document they are. They take this approach even further by playing take-off, flying and landing sounds when following intra-document hyperlinks, moving these sounds from the position of the hyperlink to the destination position in the document. Similarly, when following links to other documents, they use a sound that creates the effect of a spaceship being launched high into orbit and then descending on the desired position in the target document. These methods, while quite simple, provide the user with quick overviews of where in a website they are currently positioned, therefore substituting a scrollbar. Because they use sounds that everybody can relate to, no learning phase is required to make sense of the used earcons.

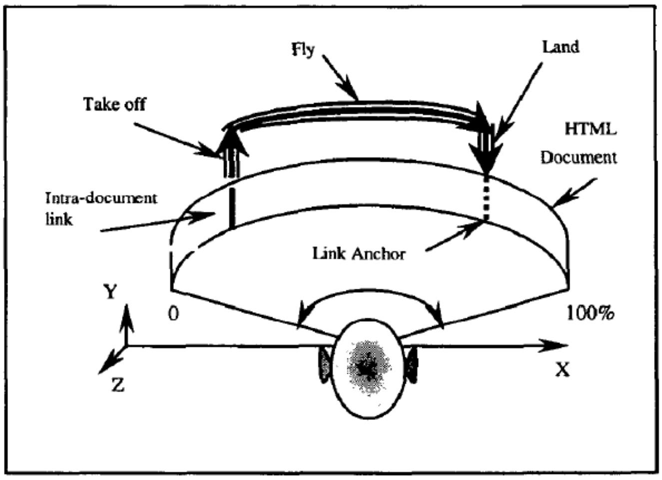

Figure 3: 3D-audio intra-document traversal (Goose/Möller 1999)

They also implemented a so-called “sound survey” feature. The sound survey allows the user to get a quick overview of specified HTML elements, like hyperlinks, around his or her current position in the document. It takes the momentary position in the document and magnifies the area immediately before and after it and projects it onto the entire arc, allowing for a more accurate rendering.

Most, if not all, of the features implemented in the interactive 3D audio web browser can be easily combined with other auditive features like the aforementioned Audio Enriched Hyperlinks. At a basic level, it only distributes all information it can gather from a mono- (or basic stereo-) signal into a three-dimensional space. This makes it a great extension for all audio based browsers.

There are virtually no drawbacks when compared to more traditional audio-only web browsers, but all the same 3D audio browsers have still not replaced them – which illustrates one of the main problems in the area of accessibility, namely the realisation of forward-looking research. The 3D audio browser as proposed by Stuart Goose and Carsten Möller also includes audio enriched hyperlinks, although their implementation is not as sophisticated as Peter Parentes. Only the title and the time required for listening are provided for the user.

Because of its ability to convey the current position in the document, as well as its implementation of the so-called sound survey, which can give a basic overview of the user's current position in the document, a total of 9 points is reached. Only one point is awarded for the link previews, because the implementation by Stuart Goose and Carsten Möller is only a basic one.

The author could not find any pricing information regarding 3D audio browsers, but as it requires no special equipment, this paper assumes that pricing would be reasonable. As it does not depend on Braille, a broad base of people could use it. Therefore, 3 points for availability have been awarded.

	Text	Media	Position	Styling	Links	Avail.	Overview	Dynamic	Total
Maximum	3	3	1	2	2	3	1	2	17
3D audio browser	3	0	1	0	1	3	1	0	9

Table 6: Evaluation of the 3D audio browser (Own compilation)

6. A unified audio-tactile browser for the visually impaired

In this final section the feasibility of a combined audio-tactile browser will be discussed. The author believes that trying to create a browser that could actually be used in a real world scenario is the best way to demonstrate possibilities and problems that could arise when existing technology is combined.

Because it fares best in visual areas, the tactile graphics tablet will be used as the basis for a potential unified web browser for the visually impaired. It does not rely on any additional effort on the side of web developers; therefore it can always return an at least basic reproduction of a website.

As listed in the previous chapter, the main flaws with these tablets are their inability to deal with dynamic content updates and their small size. While the latter can hardly be addressed without a lot of additional research, the former can be complemented by adding an audio engine to the browser.

Therefore, the author proposes the addition of a 3D-audio engine with limited features. While common textual content would not be read by the audio engine because it would overlap with reading on the tactile tablet, the audio engine could be used to proclaim content updates. If the dynamic content change would occur before the current reading position, the browser could announce a content change and read the new content out. If the update happens after the current position, there would only be a short notification that the content has changed, letting the user read it when he comes to it.

Another function of the audio engine would be to provide audio enriched hyperlinks, therefore helping partially sighted users to get a quick understanding of following websites.

In this state, the unified audio-tactile browser could already provide a fairly good browsing experience to visually impaired users – but it would omit the vast majority of blind or partially sighted people who are not able to read Braille writing. While the most important thing to do would be to increase Braille literacy, it would do no good to develop an advanced web browser for the visually impaired if only a fraction of them could actually use it. Therefore, a Braille-free mode should be implemented in the browser, making it useful for everybody.

This Braille-free mode would use the audio engine to read textual content to the user, while they could still explore images on the tactile display. While this Braille-free mode would omit styling information displayed on the tactile display, there would also be a Learning mode which would actually read everything that is visible on the tablet, encouraging users to read along as they listen to the audio and therefore enabling them to slowly get accustomed to Braille. While the learning effect of this technology would have to be analysed separately, it could at least show them the additional information they could access if they learned Braille.

While the dialog driven VoiceXML-technique can be useful sometimes, its use would collide with the other techniques mentioned before, therefore it would not be implemented. The same goes for the moveable tactile mouse, which has the unique ability to explore sophisticated graphics and maps, but falls short in most other areas. The best solution, of course, would be to implement a multi-amplitude pin layout for the larger tactile graphic displays, but at the moment this would make it financially impossible to reach a broad base of blind or partially sighted users.

Haptic wearables are a useful addition to visually disabled people who want to watch a lot of videos – but it is heavily restricted because of the necessity to manually annotate videos. Therefore, it could only expand its usefulness if it reaches a broad user base, which in turn is difficult because the available content is lacking. This vicious circle could only potentially be broken if an automated annotation technology would be developed, but it is unclear if and how this could work.

If this unified audio-tactile browser would be implemented as proposed, it would reach a score of 15 points. Based on the price for tactile graphic tablets, its price could pose a problem for some people. Because it would not depend completely on Braille, 2 points for availability have been awarded.

	Text	Media	Position	Styling	Links	Avail.	Overview	Dynamic	Total
Maximum	3	3	1	2	2	3	1	2	17
Unified browser	3	2	1	2	2	2	1	2	15

Table 7: Evaluation of the proposed unified audio-tactile browser (Own compilation)

The only major feature that would still be omitted is the display of videos or moving images in general. To solve this problem, additional research would be necessary. If graded by the pre-established criteria, the score of the unified browser would be significantly higher than the ones of the different techniques when used on their own, as can be seen in Table 9.

	Text	Media	Position	Styling	Links	Avail.	Overview	Dynamic	Total
Maximum	3	3	1	2	2	3	1	2	17
Screen readers	3	0	0	0	0	3	0	0	6
Tactile graphics tablet	3	2	0	2	0	0	1	0	8
Movable tactile display	2	1	0	2	0	0	0	0	5
VoiceXML browser	3	0	0	0	0	3	0	2	8
3D audio browser	3	0	1	0	1	3	1	0	9
Unified browser	3	2	1	2	2	2	1	2	15

Table 8: Final evaluation roundup (Own compilation)

7. Conclusions

7.1. Implications

There is a lot of research in the area of accessibility for visually impaired people and there are lots of techniques used to enable them to browse the web. While it is impossible to consider all of them, there are certainly possibilities to combine some of them in meaningful ways. If this is done, an overall better web browsing experience can be accomplished.

The proposed unified audio-tactile browser, which is the result of this paper, depends on nearly no additional research besides the implementation of existing technology. Therefore, nothing stands in the way of realization and further testing. When reviewed according to the features as presented in chapter 2, the unified audio-tactile browser stands out as the one which covers the most desirables features.

The author believes that it is possible to greatly enhance the web browsing experience with the technologies that are already researched. Therefore, he believes that more work should go into refining them and bringing them to market, because in the end only finished products that can be sold at a reasonable price can really help visually impaired people when dealing with the problems the World Wide Web can pose for them.

7.2. Future work

While a lot of research has already been done in this area, there are still many possibilities for further work.

The effectiveness of tactile graphics displays could be improved by researching ways to increase their size or lowering their cost. In general, many of these technologies have been available for some time, but few of them have really been further developed for end user use.

For the haptic wearables, an automatic annotation algorithm could be researched which would make them considerably more useful.

The most obvious further research would be to actually build the proposed unified audio-tactile web browser and test it in real-world scenarios.

References

- Anon. (n.d.). the definition of screen reader. *Dictionary.com*. Retrieved August 23, 2012 a, from <http://dictionary.reference.com/browse/screen+reader?s=t>.
- Anon. (n.d.). the definition of Braille. *Dictionary.com*. Retrieved August 22, 2012 b, from <http://dictionary.reference.com/browse/Braille>.
- Atkinson, M.T./Gucukoglu, S./Machin, C.H.C./Lawrence, A.E. (2006). Making the mainstream accessible. n.p.: ACM Press, pp. 21–28.
- Borodin, Y. (2012). Bridging Web Accessibility Divide with the HearSay Browser. *SlideServe*. Retrieved October 27, 2012, from <http://www.slideserve.com/tudor/bridging-web-accessibility-divide-with-the-hearsay-browser>.
- Borodin, Y./Bigham, J.P./Dausch, G./Ramakrishnan, I.V. (2010). More than meets the eye. n.p.: ACM Press, p. 1.
- Freudenfeld, M. (2010). Screenreader im Internet 2010.. Retrieved from <http://www.incobs.de/produktinfos/screenreader/webtest_2010/artikel.php>.
- Goose, S./Möller, C. (1999). A 3D audio only interactive Web browser. n.p.: ACM Press, pp. 363–371.
- Hailpern, J./Guarino-Reid, L./Boardman, R./Annam, S. (2009). Web 2.0. n.p.: ACM Press, p. 821.
- Hamam, A./Eid, M./El Saddik, A./Georganas, N.D. (2008). A quality of experience model for haptic user interfaces. n.p.: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), p. 1.
- Headley, P.C./Hribar, V.E./Pawluk, D.T.V. (2011). Displaying braille and graphics on a mouse-like tactile display. n.p.: ACM Press, p. 235.
- Headley, P.C./Pawluk, D.T.V. (2010). A low-cost, variable-amplitude haptic distributed display for persons who are blind and visually impaired. n.p.: ACM Press, p. 227.
- Hreinsson, H. (2011). Practical Screen Reader Comparison.
- Jay, C./Brown, A./Harper, S. (2010a). A “visual-centred” mapping approach for improving accessibility to Web 2.0 for people with visual impairments.
- Jay, C./Brown, A./Harper, S. (2010b). Evaluation of the SASWAT Web browser.
- Jernigan Institute (2009). The Braille Literacy Crisis in America.
- Miyashita, H./Sato, D./Takagi, H./Asakawa, C. (2007). Making multimedia content accessible for screen reader users. n.p.: ACM Press, pp. 126–127.
- Orca Website (n.d.). Orca - A free, open source, flexible and extensible screen reader. *Orca - A free, open source, flexible and extensible screen reader*. Retrieved October 28, 2012, from <http://projects.gnome.org/orca/>.
- Parente, P. (2003). Audio enriched links. In: *ACM SIGACCESS Accessibility and Computing*, , no. 77-78, p. 2. doi:10.1145/1029014.1028633.

- Parente, P. (2005). Audio Enriched Links 1.42a. Retrieved August 24, 2012, from <http://www.oatsoft.org/Software/audio-enriched-links-1/downloads/1.42a>.
- Pun, T./Roth, P./Bologna, G./Moustakas, K./Tzovaras, D. (2007). Image and video processing for visually handicapped people. In: *J. Image Video Process.*, vol. 2007, no. 5, pp. 4:1–4:12. doi:10.1155/2007/25214.
- Rahman, A./Alkhaldi, A./Cha, J./El Saddik, A. (2010). Adding haptic feature to YouTube. n.p.: ACM Press, p. 1643.
- Ramakrishnan, I.V./Stent, A./Yang, G. (2004). Hearsay. n.p.: ACM Press, p. 80.
- Rose, D. (2012). Braille is spreading but who’s using it?. Retrieved from <http://www.bbc.co.uk/news/magazine-16984742>.
- Rotard, M./Knödler, S./Ertl, T. (2005). A tactile web browser for the visually disabled. n.p.: ACM Press, p. 15.
- Takagi, H./Saito, S./Fukuda, K./Asakawa, C. (2007). Analysis of navigability of Web applications for improving blind usability. In: *ACM Transactions on Computer-Human Interaction*, vol. 14, no. 3, p. 13–es. doi:10.1145/1279700.1279703.
- The Office for Disability Issues (Caxton House) (2009). Braille and Moon. Retrieved October 27, 2012, from <http://odi.dwp.gov.uk/inclusive-communications/alternative-formats/braille-and-moon.php>.
- Touch Graphics Inc. (n.d.). Talking Tactile Tablet 2. *Talking Tactile Tablet 2 (TTT) Bundle*. Retrieved October 27, 2012, from <http://touchgraphics.com/OnlineStore/index.php/ttt-hardware/talking-tactile-tablet-2/talking-tactile-tablet-2-ttt.html>.
- W3C (2011). WAI-ARIA Overview. Retrieved August 24, 2012, from <http://www.w3.org/WAI/intro/aria.php>.
- W3C (2004). Voice Extensible Markup Language (VoiceXML) Version 2.0. Retrieved August 24, 2012, from <http://www.w3.org/TR/voicexml20/#dml1.1>.
- Wang, Z./Xu, X./Li, B. (2007). Enabling seamless access to digital graphical contents for visually impaired individuals via semantic-aware processing. In: *J. Image Video Process.*, vol. 2007, no. 3, pp. 3:1–3:14. doi:10.1155/2007/18019.
- WebAIM.org (2012). Screen reader survey.. Retrieved from <http://webaim.org/projects/screenreadersurvey4/>.
- Xu, C./Israr, A./Poupyrev, I./Bau, O./Harrison, C. (2011). Tactile display for the visually impaired using TeslaTouch. n.p.: ACM Press, p. 317.

List of Figures

List of Tables