WebSight: The Use of the Grid-Based Interface to Convey Layout of Web-Pages in a Non-visual Environment

Hesham M. Kamel[1] and Halil I. Erhan[2]

[1] College of Information Technology, United Arab Emirates University, Al-Ain, UAE
[2] School of Interactive Arts and Technology, Simon Fraser University, Surrey, Canada
hesham@uaeu.ac.ae, herhan@sfu.ca

Abstract. Accessing Web content including lay out of web pages is currently limited for sight-impaired people. In general, Internet content is designed with sighted users in mind, requiring users without this ability to adapt (Edwards, 1994). The non-visual interaction methods presented by Screen Readers are often serial in nature and laborious. In this paper we introduce the design and evaluation of WebSight, a talking browser that conveys layouts of Web pages for the blind. WebSight is a plug-in for Internet Explorer and employs a universal 3X3 grid-based interface (Kamel, 2002), to assist blind people with visualizing Web content with respect to its absolute and relative positions. Each cell of the grid contains a 3X3 virtual sub-grid with nine unique positions. We conducted an experiment involving six blind and six sighted navigating a layout of a particular webpage. The study reveals that the use of absolute and relative position coupled with a grid-based interface enable blind users to build mental model of page layout at least as well as sighted users. In addition, findings of the study suggest that the grid-based interface is a universal mechanism that enhances the process of building mental models of layout designs.

Keywords: Talking browser, Blind, Grid-based interface, Web page layout, Internet, Absolute and relative positions, Visually impaired, Mental model.

1 Introduction

Cognitive and semantic aspects of a stimulus play an important role in visual and scene perception, which affects user's Web behavior [7, 15]. For visually impaired people, although there are not direct visual stimuli and channels, they can form visual mental models of a given context similar to sighted people [10, 12]. While the reliance on visual layout of hypermedia enhances information access on the Web for sighted users, the complex visual structure of pages creates challenges for visually impaired people. The difficulty arises when the implicit visual cues presented on a Web page cannot be accessed by visually impaired individuals [14]. It is often intricate for blind people to get sufficient explanations with locating specific data on a Web page without sighted-assistant. This is due to the Web technologies do not support feedback regarding visual orientation of web elements in a non-visual format.

C. Stephanidis and M. Antona (Eds.): UAHCI/HCII 2013, Part II, LNCS 8010, pp. 674–683, 2013.
© Springer-Verlag Berlin Heidelberg 2013

The design of a system should consider the intended user's needs, knowledge, and skills as well as the task environment [13, 16, 17]. A system for the visually impaired users can be informed by understanding how they use existing interfaces. For example, a clock face is a familiar interface that can be used by visually impaired users as a tool for giving the location of different types of food on a plate [10]. The potatoes, for instance, might be described as being located at "6 o'clock" (closer and in front of the blind person), where the vegetables may be located at "12 o'clock" (directly above the potatoes and farther away from the person). The clock face can become a reference when forming the mental image of different items on the plate. In this example, the analog clock interface forms a coordinate system (12 unique positions) to assist the blind user visualize absolute and relative positions.

Using of visual composition of page content including sharp layout, snappy captions and images transform plain text into a presentation that pleases sighted users [4, 8]. Nevertheless screen readers use the internal structure of such pages and "disregard" these visual enhancements, as these visual cues can equally help visually impaired users to obtain an accurate mental model of page compositions [18].

The grid-based interface solutions have been utilized in the past to assist visually impaired in using "geo-referenced" information augmented with "non-textual sounds and speech output" [19]. It has been also used in experiments to understand user performance in tracking cursor-location combining speech- and grid-based interfaces [1].

In this paper, we bring attention to why visual information given its spatial context on a Web page can be vital for blind people. Moreover, we show why some existing accessibility methods do not meet the needs of blind users and we propose a new method addressing the challenges mentioned above. Next, we give an overview of the related work in providing blind users with the ability to visualize and access information on the Web. This is followed by a description of the WebSight system. We then report the study and the results. Finally, we draw our conclusion and suggest future work.

2 Related Work

Goble et al [6] have addressed the possibility of including travel and mobility techniques in the usability metrics of web design. Their goal was to maximize cues and minimize obstacles to give high Internet mobility to blind users. Goble has presented a framework for identifying travel objects and registering them as either cues to aid travel or obstacles that hinder travel for visually impaired users.

Francisco-Revilla and Crow [5] conducted a study on how people interpret modular layouts of news and shopping Web pages. They found when the layout complexity increases, the interpretation process gets longer and the reading more varied. It was noted that before looking at the main content, subjects first "framed" the Web page by looking for familiar structural elements that can be used as "point of reference.

Strain et al [18] created a grid-based interface consisting of 9 squares and rendered information using multimodal feedback. Each cell and user's mouse positions on the grid are conveyed via speech in relation to the entire grid. A distinctive and periodic wave along with a pulse effect is played to convey the border location of each cell.

WebSight addresses important needs that are not included in Strain's system. For example: it employs a rendering technique which conveys directional orientation; allows users to randomly access a particular cell; integrates a virtual grid technique; enables users to turn grid on/off; and renders each cell by its position.

3 WebSight System Design

WebSight divides the viewing area of a web page into nine cells, employing a 3X3 grid-based interface [1, 9, 10, 11, 18]. The grid introduces nine fixed regions with unique center points, where one is the top left corner and nine is the bottom right corner (see Figs. 1 and 2). As the user navigates the on-screen content, WebSight renders the spatial orientation of each element via voice synthesis. The function keys (B, X or Shift-B, Shift-L or Shift-X) allow the user to navigate forward or backward through buttons, links or text respectively. If a link spans two adjacent cells, that link by default belongs to the first cell of the two.

WebSight Internal Data Organization: WebSight organizes key HTML elements on a Web page into two data structure lists:

- Content-based: A list that is based on the semantics of the raw HTML tags.
- Cell-based: A list whose HTML element order is primarily based on the cell that the element lies in within the web page. The cell-based list can also be considered as a collection of content-based lists that contain HTML elements. These elements lie within the cell that the list represents. To determine which cell an HTML element belongs to, WebSight uses the element's position relative to the top-left corner of the Web page.

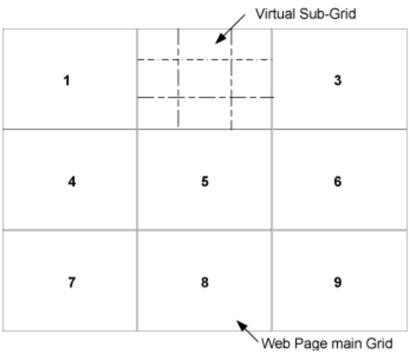

Fig. 1. Layout of Main Grid and Virtual Sub-Grid

A user can navigate these lists in one of three basic modes. The first mode is content-based, navigating via the inner organization of the HTML tags. The two other modes rely on the cell-based lists: continuous and internal. The continuous cell-based navigation is moving from one cell to the next (e.g., navigating from Cell 1-2, where all links in Cell 1 is rendered before rendering the first link that appears Cell 2). The internal navigation is moving within a given cell until the user prefers to move to another cell. Consequently, for any HTML element, the user can navigate to the next element that follows it content-wise or cell-wise. This navigational method utilizes the user's preference based on which part of the screen is desired. It provides the blind user with a flexible non-visual technique to dynamically navigate a Web page.

Content-Based Navigation: Once this mode is selected, WebSight renders the information as the blind user navigates the page, based on the HTML organization and in a non-sequential manner. Note that the visual ordering of the links does not necessarily coincide with the structure in the corresponding HTML tags. For example, if the cursor is on the link "My sessions" in Cell 2, pressing the next link function key 'L', WebSight announces the next link as "Cell 3, Right, Link, About My Web Canine" instead of "Down far left, Link, Feedback" (Fig. 2).

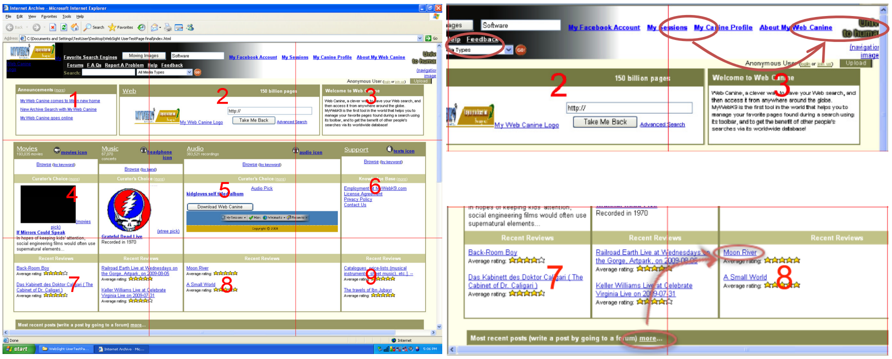

Fig. 2. Test Web page with Grid (Left); Content-Based navigation between links in Cell 2 and 3 (Right upper); Continuous-Based navigation between Cell 7 and 8 (Right lower)

Continuous-Based Navigation: In this mode, users navigate through the information of each cell in a top-down left-right orientation. Since the link "More" is the last link-element in Cell 7, WebSight announces, "Cell 8, Up Right, Moon River" as the next link-element (see Fig. 2).

Internal-Based Navigation: In this mode once more, WebSight renders the information in a top-down left-right orientation. However, when the user reaches the last element in a given cell, WebSight loops back and announces the first link in that cell as the next element.

Random Navigation: Jumping to a particular cell allows users to locate a precise content (e.g., the username and password fields). Random cell selection can be used with any of the three navigation modes mentioned above. If a user is navigating forward cells, WebSight places the user's cursor at the first element of that cell. On the other hand, if a user is navigating backwards, WebSight places the cursor at the last element of that cell.

Navigation and Visualization through Virtual Sub-Grid: WebSight virtually integrates another grid inside each of the nine cells [15], and then converts the sub-grid positions to convey the visual point of reference for every HTML element. However, the sub-grid is not visible on the interface, while the system still utilizes its functional effect. This visual feature improves screen readability for communication between sighted and blind users. As the user navigates from element A to element B, WebSight renders two types of information: relative and absolute positions.

The Method of Relative Positions: The directional information is given in reference to the current cursor position on the sub-grid whether the cursor moves forward or backward. In addition, WebSight reports when the directional travel favors one cardinal direction more than the other. When traveling from the end of one line to the beginning of the next line, from a technical viewpoint, the cursor has traveled in a down-left manner. WebSight renders the relative direction change to be “down, far left” instead of “down, left”(see Fig. 3-Left). Note that the link positions are based on the first character that appears in a given link and not the length of the link itself.

Fig. 3. (Left) In Cell 5, moving from link “Audio Pick” to link “kid gloves self-titled album”, the user hears “Down Far Left”; (Right) The letter ‘H’ in link ‘Help’ starts in Cell 1, WebSight reports its location to be in Cell 1 and not Cell 2

The Method of Absolute Positions: WebSight reports the current position of the cursor using the cell on the main grid, followed by the location on the virtual grid within that cell. The function key ‘C’ gives the absolute position of the current cursor location. The directions reported are the eight typical compass directions (e.g., center, top middle, bottom middle, left middle, right middle, top left corner, bottom left corner, top right corner and bottom right corner). For example pressing the function key C, when the cursor is at the link “Favorite Search Engines, WebSight renders “Cell 1, Start at Top Middle, Link, Favorite Search Engines” (see Fig. 2).

4 WebSight User Study

We conducted a within-subject design experiment to understand the effectiveness of the grid-based Interface in forming mental model of web page layouts.. We had the participants listen to specific content of a web page, then draw a line(s) on a sheet of paper representing the location of each element. The page used for testing is designed particularly for this study by the authors and not available online (see Fig. 2).

We tested the system with two groups (visually impaired and sighted). The first group of participants was at the Orientation Center for the Blind (OCB), Albany, California. This group consisted of six partially sighted students (BPs). Their age ranged from 21 to 45. All of the six participants had some degree of vision ranging from light perception to identifying figures with no details. We asked this group to draw their mental reflection of the tasks using a Sewell raised line drawing kit. This kit includes a hard clipboard that is coated with a rubber surface. A sheet of plastic is fixed with clamps on top of the board. The users drew on the plastic sheet by pressing an inkless stylus into the rubber to create a raised impression on the sheet (Fig. 5).

Fig. 4. Sewell raised line drawing board

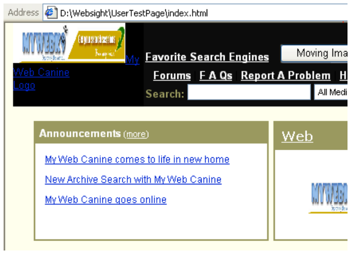

Fig. 5. The links shown in Cell 1

The second group consisted of six sighted participants (SPs) ages ranged from 41 to 65 years of age. These participants are university teachers and were not blindfolded. They used white sheets of paper and a pen for drawing. All of the participants had experience in using computer for at least five years. In addition, the visually impaired participants had experience in using a talking browser for the same length of time. The purpose of the study was to explore the advantages/drawbacks of navigating 2D presentation of web information in a non-visual format. We hypothesized that web navigational methods adopted by screen readers for blind users were not adequate enough to grasp web page layout. Therefore, our goal was not to discover if blind individuals could navigate the Internet, rather it was to see if they are able to properly visualize elements of the page layout(s) through the use of the grid-based Interface.

4.1 Method

The testing sessions were divided into two parts: Training and task performance session. In the training session, each participant had approximately 20 minutes of tutorial on how to use the system and hands-on practice using the interface. In addition, during the tutorial and task performance, we simulated the non-visual environment by dimming and covering the computer screen with opaque paper. Both groups relied on the system auditory feedback to build a mental model of the Web page layout. The participants were allowed as much time as desired to complete each task. No hints were given at any time. If the participants asked for assistance during testing, we referred them to an electronic version of the functions key list.

After completing the tasks, each participant gave feedback on the system usability. The sighted participants were asked to look at their drawings and compare it with the locations of the Web page content. Since the blind participants could not perform the visual comparison, we asked them to give us their verbal feedback regarding the usability of the system. We used both participants' performance and verbal responses to draw conclusions about the system's usability and effectiveness. All of the participants' output and comments were audio recorded for further analysis. We compared the drawings with the content of the actual Web page. The participant's drawing of a link was considered correct if the marking on the drawing was in the general area of the sub-grid (e.g. Cell Five, top-left, top-right, bottom, center, right-center etc.).

Task 1: We selected five specific links in Cell 1 (Fig. 6): “Favorite Search Engines”, “Help”, “MyWeb Canine comes to life in new home”, “MyWeb Canine goes online” and “Web”. We instructed the participants to navigate to each one using the function keys L, R and C, and then draw their corresponding locations as lines on the sheet of paper. We asked them to treat the drawing area of the paper as Cell 1 and not as the whole screen. Our goal was to see the initial effect of the WebSight feedback without prior knowledge of the information. 4/6 of the BPs and SPs got this task correct. The remaining 2 in each group had a single issue with their drawing (e.g., drawing an extra line or a line that was not quite in the right general location). All of the BPs, seemed to take a longer time to familiarize themselves with the dimension of the paper (e.g., BP4), where SP3 drew a grid to visually assist him. BP6 was able to use the first link as a point of reference to draw perspectives of the other links. SP2 commented “it would be better if the sheet of paper was a square shape” (see Fig 6).

Task 2: The participants were asked to locate all the links in Cell5, via the L and R function keys. However, we specifically instructed them to not use the absolute position rendering function key C. Once more, we asked them to treat the whole sheet of paper as Cell 5 and not the whole screen. In this task, our interest was to find out the effect on the lack of providing the absolute position information in relation to the virtual sub-grid in Cell 5: 4/6 of the BPs and 6/6 SPs got this task wrong, because there were missing information. When we were asked how to go forward without the missing information, we told them to use their best judgment. All of the 12 participant’s drawing did not match the actual layout of Cell 5. Since much of the layout orientation was “down and to the left”, most of the participants interpreted the verbal description “left” to be at the left edge of the paper. 3/6 BPs and 5/6 SPs pushed the links either to the bottom left or right corner, where SP3 ran out of drawing area. BP5 had the best imagination, relying only on the relative position information (Fig. 7).

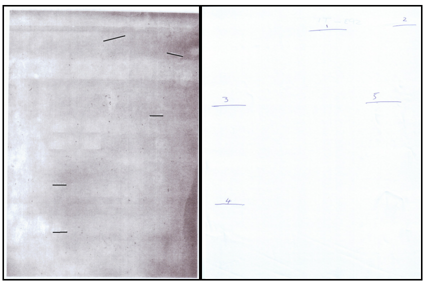

Fig. 6. Task 1 by BP5 (left) and by SP3 (right)

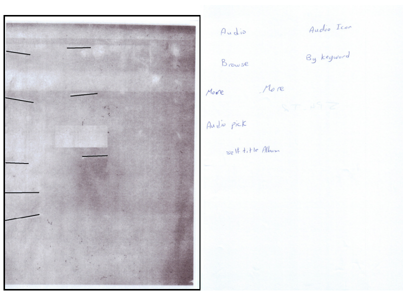

Fig. 7. Task 2 by BP5 (left), and SP4 (right)

Task 3: This task is similar to Task 2. However, this time around we allowed participants to use the absolute position function key C. All of the BPs and SPs was relieved when we told them that they can use the function key C. 3/6 BPs and 3/6 SPs were able to successfully draw the links in cell five. Comparing the results of task 2 and 3, we clearly notice the importance of conveying the absolute position information for locating on screen objects (see Fig 8).

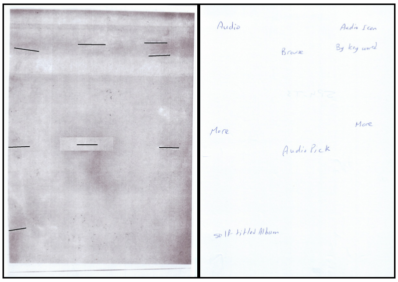

Fig. 8. Task 3 by BP5 (left), and by SP4 (right)

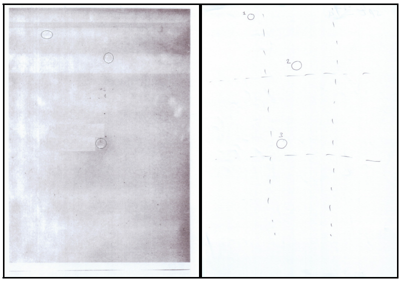

Fig. 9. Task 4 by BP1 (left) and by SP2 (right)

Task 4: We asked the participants to locate all the buttons on the web page, and then draw their corresponding locations on a sheet of paper. This time, we asked them to treat the sheet as the whole screen and not as a one single cell. In this task, we were interested in finding out if the visualization of the virtual sub-grid differs from that of the whole page. 2/6 SPs seemed to have a little trouble visualizing the virtual sub-grid within each cell and the main grid, where none of the BPs reported having any problems. It is interesting to note that 5/6 SPs drew a grid to accomplish the task. BP1 performed as good as SP2 who drew a grid as a visual guidance (see Fig 9).

5 Results and Discussion

It was noted that all of the 12 participants did not rely much on the relative position information. Knowing the absolute positions of the current and previous elements were enough to determine the current position in relation to the virtual sub-grid as well as build a mental model of the relative data of these elements.

The absolute position information contains the relative position with respect to the virtual sub-grid. Therefore the navigational access and building a mental model, explicitly rendering the relative position might not be necessary. These findings clearly show in task 2 when participants were asked to visualize element relations in cell 5 without using the absolute position functionality.

As for rendering elements that reside at the four corner regions of a cell, users automatically imagined them to be at extreme far top (for the top corners) or extreme far bottom (for the bottom corners), though they were in the correct general area. This was noted in tasks 1 and 3, where users drew the links that were rendered at the bottom left corner little close to each other.

Using the expression “up right” feedback to mean “up and to the right” was a little confusing to 4/6 BPs and 1/6 SP. Their interpretation of the “up right” was an upright angle and not up and to the right. Using “top middle” for top center, “bottom middle” for bottom center etc. was a little ambiguous to 3/6 BPs and 1/6 SP. They suggested that the use of the word “center” is a better choice to use instead of “middle’. These findings suggest that the language used to auditory convey spatial information should be carefully chosen. The fact that some of the BPs had the need to use different methods to aid their drawing (e.g., draw a grid on the plastic sheet), it dictates that visually impaired are able to have mental models but look for the proper tool to support its building process.

Overall, it should be noted that the sighted participants might have been a little faster than the blind participants in task completion time, since they were able to see the continuous visual feedback of their drawing. Task completion time was not a factor of our evaluation performance. Some of the spoken words were not clear due to the nature of the voice synthesis. However, none of the participants we had with different nationalities including Americans complained about the British female accent of Audrey that WebSight employs.

6 Conclusion and Future Work

In this paper, we presented WebSight, a system to render spatial organization of web content. The system employs a grid-based interface along with utilizing absolute and relative positions to convey layout of web pages in a non-visual environment. Given the two schools of thoughts on whether the blind has a mental model similar to the sighted or not, we have implemented WebSight to investigate which of these two schools is relatively more adaptable for building interfaces for the blind. In adapting the Interface to meet the needs of the visually impaired people, we were able to show that blind users are capable of retaining layout of web pages at least as well as sighted users. Moreover, controlling the visual access to the layout, the study shows that WebSight's grid-based Interface is a practical method to aid the same process of constructing mental models for blind and sighted users. In addition, the choice of vocabulary used with voice synthesis to convey directional information in a non-visual environment could affect (alter) the development of building mental models.

For future work, we are planning to add a function for toggling relative position information to study its effect on the process of building mental models of page layouts. In addition, we would like to investigate if users will favor more the relative or absolute position feedback. Also, we like to include the virtual sub-grid as part of the information rendering technique (e.g., starting at left-center, sub cell 4, link, MyWebK9). It is needed to understand which message would give a more effective meaning to the user, referring to the virtual sub-grid position as numbers or as locations e.g., "element is at left-middle" or "element is at sub-cell 4".

Acknowledgments. The authors would like to thank Michael Le at the dental school of UCSF for his hard work on the implementation of WebSight. Also, the authors would like to extend their appreciation to Dr. Emad Mohamed and Jihad Aravassery at the College of IT for their "visual" assistance with this paper. The author's appreciation is also extended to Raahim Beg for his help in editing the paper.

References

1. Dai, L., Goldman, R., Sears, A., Lozier, J.: Speech-based cursor control: a study of grid-based solutions. SIGACCESS Access. Comput (77-78), 94–101 (2003)

2. Darken, R., Sibert, J.: Way finding strategies and behaviours in large virtual worlds. In: Proc ACM CHI 1996, pp. 132–149 (1996)
3. Edwards, A., Edwards, A., Mynatt, E.: Tutorial: Enabling Technology For Users With Special Needs. In: Proceedings of Human Factors in Computing Systems: CHI 1994, pp. 405–406 (1994)
4. Faraday, P., Sutcliffe, A.: Making contact points between text and images. In: Proceedings of the 6th ACM International Conference on Multimedia, New York, pp. 29–37 (1998)
5. Francisco-Revilla, L., Crow, J.: Interpreting the layout of web pages. In: Proceedings of the 20th ACM conference on Hypertext and hypermedia (HT 2009), pp. 157–166. ACM, New York (2009); Revilla, L.,Crow, J. Interpreting the Layout of Web Pages
6. Goble, C., Harper, S., Stevens, R.: The travails of visually impaired web travellers. In: Proceedings of the 11th International ACM Conference on Hypertext and Hypermedia, pp. 1–10. ACM, New York (2000)
7. Henderson, J.M., Weeks, P.A., Hollingworth, A.: The effects of semantic consistency on eye movements during complex scene viewing. J. Exper. Psych. Hum. Percept. Perform. 1, 210–228 (1999)
8. Hoffos, S., Sharpless, G.: The CD-I Designers Guide. McGraw-Hill, New York (1991)
9. Kamel, H.M., Landay, J.A.: The Integrated Communication 2 Draw (IC2D): A Drawing Program for the Visually Impaired. In: Proceedings of ACM Conference on Human Factors in Computing Systems: CHI 1999, Pittsburgh, PA, United States, pp. 222–223 (1999)
10. Kamel, H.M., Landay, J.A.: A study of blind drawing practice: creating graphical information without the visual channel. In: Proceedings of the Fourth International ACM Conference on Assistive Technologies: Assits, Arlington, VA, United States, pp. 34–41 (2000)
11. Kamel, H.M., Landay, J.A.: Sketching images eyes-free: a grid-based dynamic drawing tool for the blind. In: Proceedings of the Fifth International ACM Conference on Assistive Technologies: Assits 2002, Edinburgh, Scotland, pp. 33–40 (2002)
12. Kamel, H.M., Landay, J.: Constructing moving pictures eyes-free: an animation tool for the blind. In: CHI Extended Abstracts, Minneapolis, United States, pp. 772–773 (2002)
13. Norman, D.A.: The Design of Everyday Things. Doubleday, New York (1990)
14. Plessers, P., Casteleyn, S., Yesilada, Y., Troyer, O.D., Stevens, R., Harper, S., Goble, C.: Accessibility: a web engineering approach. In: Proceedings of the 14th International Conference on WorldWideWeb (WWW 2005), pp. 353–362. ACM, New York (2005)
15. Rayner, K.: Eye movements in reading and information processing: 20 years of research. Psych. Bull. 124(3), 372–422 (1998)
16. Sharp, H., Rogers, Y., Preece, J.: Interaction design: beyond human-computer interaction. Book, p. xxvi, 773 p. Wiley (2007)
17. Shneiderman, B.: Designing the User Interface: Strategies for Effective Human-Computer Interaction. Addison-Wesley Longman Publishing Co., Inc., Boston (2009)
18. Strain, P., McAllister, G., Murphy, E., Kuber, R., Yu, W.: A grid-based extension to an assistive multimodal interface. In: CHI 2007 Extended Abstracts on Human Factors in Computing Systems, p. 2680 (2007)
19. Zhao, H., Plaisant, C., Shneiderman, B., Lazar, J.: Data Sonification for Users with Visual Impairment: A Case Study with Georeferenced Data. ACM Trans. Comput.-Hum. Interact. 15(1), Article 4, 28 pages (2008)