COMP3013 CONFERENCE COMPUTING - DISPLAYING WEB CONTENT ON SMALL DISPLAYS

Chris Lord

Southampton University
cil103@ecs.soton.ac.uk
http://chrislord.net/

ABSTRACT

The majority of web content is not suitable for viewing on small displays. In this paper, three different categories of web content manipulation to make these pages more accessible on small screens are documented, along with their strengths and weaknesses. These methods include layout manipulation, content manipulation and display techniques. Some information is provided on what methods are employed by current mobile web browsers and a suggestion for combining these ways of reformatting and display is also proposed.

Keywords

Information Visualisation, Small Screens, Web, CSS (Cascading Style Sheets), DOM (Document Object Model)

1.INTRODUCTION

Due to the increasing popularity of PDAs and convergent devices such as smart-phones, the world wide web is seeing increased access on devices with lower resolution and physical area than common desktop displays. Most content on the web was created in mind of being displayed on a desktop computer and not the smaller displays found on these devices.

Some sites choose to deal with this by detecting access by devices with constrained viewing area and redirecting the browser to a page reformatted manually and specifically for smaller screens. Other sites use proxy services and provide the main content of their site via different methods more tailored to small screens, usually by a linear transformation[3], where the content of the site is summarised and provided in a long linear list.

The problem with these previous methods is that they require the creator of the site to specifically cater for a subset of their users. There is a multitude of content on the web that is not suitable for direct migration to devices with small displays. These are rendered almost unusable due to assumptions made about the display size.

This paper intends to provide detail on some of the techniques used at client-side in browsers, such as intelligent content extraction, styling and customised display techniques, to make pages intended for use on large displays more accessible on smaller displays. It also intends to suggest what may be the ideal combination of these techniques and to provide some reasoning behind this suggestion.

2.ISSUES WITH PAGE LAYOUT AND CONTENT

There are many specific issues to overcome regarding the layout of web pages for small screens, as well as the actual content of pages. There is a finite amount of information that can be fit on a particular screen size, based on the text size that is comfortably readable[2]. Often web sites are designed with a set width and the text is set to fill this width. This becomes a problem when resolution and/or dpi (dots per inch) vary too far from that which was in mind when the page was designed, and may require the user to scroll back-and-forth, for example[4].

There is also the issue of the positioning and size of page elements. A common layout style is of a 3-columned table, where the centre column contains the main content and the surrounding columns contain relevant links and adverts. This presents a problem as screens on common portable devices often have either a square aspect ratio, or an aspect ratio that has a larger vertical component than horizontal. Images, for adverts or otherwise, are often created with the assumption of a particular resolution and dpi that is incompatible with most devices with small-screens as well.

Finally, there is the issue of the content chosen to be displayed on a page. When screen space is limited, it is important to include only the information necessary. Adverts, comments and other such irrelevances, when placed in a prominent position, can make a page difficult to navigate with a restricted input method or on a small screen.

4th Annual Multimedia Systems, Electronics and Computer Science, University of Southampton

3. Modifying Page Content

A relatively unused technique for client-side reformatting of pages is to modify the page content (as opposed to the formatting). This can take the form of contract extraction or summarisation. The reason for this lack of use may be due to the relative computational complexity of altering content as opposed to reformatting it.

3.1 DOM-based Content Extraction

An easily implementable technique for content extraction is to alter the generated DOM (Document Object Model) tree[1]. This involves using analytical techniques to identify where relevant content is located in a page. For example, by analysing the percentage of text that is a link in a particular paragraph, it can be identified whether that paragraph contains important information. Depending on the result, the paragraph can be positioned elsewhere in the document, or removed from the DOM tree entirely. Figure 1 shows an example of a page before and after using DOM-based content extraction.

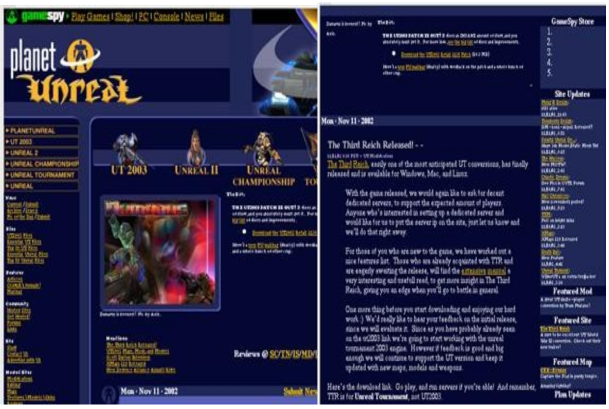


Figure 1 – Before and after of a page using DOM-based content extraction (Images from [1])

An advantage of this technique is that the original structure of the page remains, thus reducing the amount of input necessary to view content. Also, as the technique is independent of target device, it could also be implemented as a proxy service. Paper [1] chose this technique, as they believed it was an optimal method for reducing a website to only its relevant information, without altering the underlying structure of the page.

This technique can prove to be problematic for pages that make heavy use of JavaScript though, as some JavaScript pages modify the DOM themselves and may produce unexpected results if it has been modified externally. This problem is becoming increasingly important to consider as AJAX (Asynchronous Javascript And XML) becomes popular. Another danger of using contact extraction is that it may extract the incorrect data. For example, the method used in paper [1] fails on sites that make heavy use of links in the main content, as shown in figure 12 of said paper.

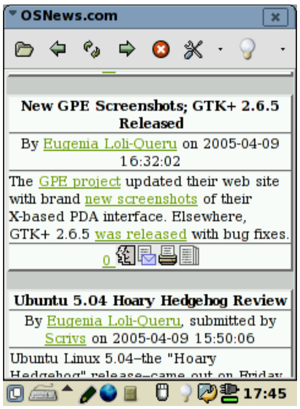
Assuming that it doesn't class it as irrelevant information, this technique would not hamper the viewing of sites based on Macromedia Flash[12], or other proprietary browser plug-ins. These types of sites have accessibility issues outside the scope of this document, however, and would need to be considered as special cases.

3.2 Summarisation

Another technique, covered in the paper *Text compaction for display on very small screens*[5], is to summarise certain data on a page. This technique works by contextually matching certain phrases, words and numbers and providing shorter versions as replacements. The particular implementation in [5] is applied to e-mails and can, for example, replace the text '2nd November 2005' with '2/11/05'. Further reductions can be accomplished by omitting unimportant words from sentences, however, algorithms must be carefully tested so as not to alter the meaning of the content. A particular downfall of this technique is that it is locale-specific (specialised for a particular region) and if employed incorrectly, can actually make content harder to read.

4. Modifying Page Layout

The most common method of altering content to be viewable on devices with small screens is to alter the layout of the page, usually by employing custom CSS. This is the method employed by the Minimo browser[6], pictured in figure 2.



New GPE Screenshots; GTK+ 2.6.5 Released

By Eugenia Loli-Queru on 2005-04-09 16:32:02

The GPE project updated their web site with brand new screenshots of their X-based PDA interface. Elsewhere, GTK+ 2.6.5 was released with bug fixes.

0

Ubuntu 5.04 Hoary Hedgehog Review

By Eugenia Loli-Queru, submitted by Scrivs on 2005-04-09 15:50:06

Ubuntu Linux 5.04–the "Hoary Hedgehog" release–came out on Friday

Figure 2 – The Minimo (Mini Mozilla) browser

By examining the CSS that Minimo uses[7], it can be seen exactly what methods it uses to reformat pages. The Minimo CSS works by overriding properties of the style of the page in question by marking those properties as important. It cancels all explicit size settings, removes all padding and margins, removes floating elements, flattens tables, sets the maximum width of all elements to the width of the browser to remove the horizontal scrollbar and highlights all links. This gives websites a 'stacked' appearance, similar to the proprietary method of the Opera[4] browser and is a form of linear transformation. Research papers [9][10] show that it is desirable to eliminate horizontal scrolling to reduce the amount of user interaction required to understand the content of a page.

This technique does not reorganise the content in any order of importance, however, so sites with large amounts of irrelevant content before the main content will still be difficult to navigate.

5. Display Techniques

The final technique that will be documented is that of modifying the way the content is interpreted when it is displayed, without actually modifying the content, contrary to the previous two categories of techniques.

This is an area that is being very actively researched, however, there have yet to be any commercial web browsers that exploit the following covered techniques.

5.1The Gateway

The Gateway[8] is a method of displaying a full page on a small screen without employing any modification of page layout or content (an overview transformation[3]). The advantage of this is that there is no possibility of making page content inaccessible or altering its meaning.

The Gateway works by scaling a full rendering of the web page to the dimensions necessary to fit in a single screen on the target display. The page remains static and a user can select parts of the content that would like to view in additional detail, at which point it provides a full-scale view of that particular component. Figure 3 shows The Gateway in use.

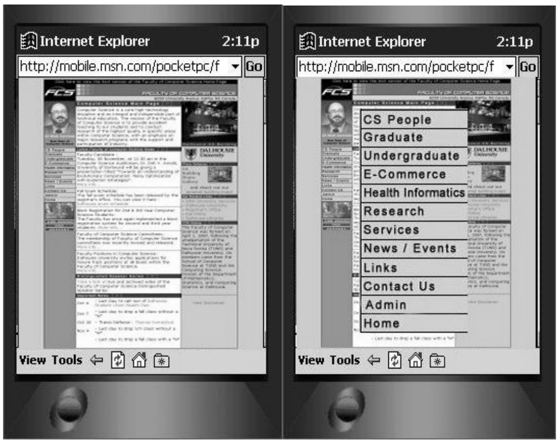


Figure 3 – A demonstration of The Gateway (images from [8])

Initial research shows that users prefer using the Gateway to approaches that use linear migration[3]. The research in paper [3] also suggests that users preferred direct migration to linear transformation. This research only had ten participants however, so no valid conclusions can be drawn from this. It is unlikely that direct migration is a more popular method of display than linear transformation, due to its usage and showcase in the popular Opera web-browser[4] and the Minimo web-browser[6].

5.2Transparent Widgets

Transparent widgets are less of a display technique and are more involved with the interface in which web content is displayed. It is suggested[2] that available display space can be maximised if control widgets (e.g. the back and forward buttons in a browser) are displayed semi-transparently over page content.

The added complexity of navigation this technique introduces will rarely warrant the display space it affords, however, the idea of transparent widgets may be useful if implemented in a different way. The paper[2] admits a learning period and high initial error-rates with the chosen implementation, which would not be acceptable in a commercial product.

5.3Flip Zooming

Flip Zooming[11] is a technique very similar to The Gateway[8]. It employs a zooming view of the document and also removes all scrollbars, similar to The Gateway. To achieve this, pages are broken up into pieces that fit on the target device display. The currently focused piece is displayed at the original zoom level in the centre of the display, surrounded by the other page pieces, displayed at a much lower zoom level. Shortcuts are provided for switching between the next and previous pieces, as well as being able to choose a piece directly by selecting it with the pointing device. Figure 4 shows Flip Zooming being used on a prototype browser with a 160x160 pixel display, vs. direct migration on the same display.

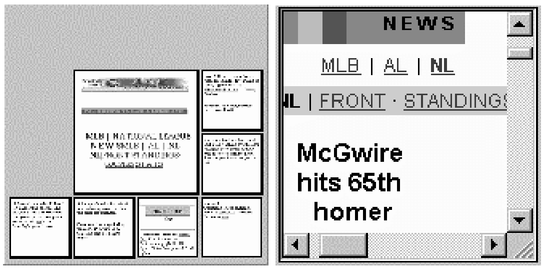


Figure 4 – Flip Zooming vs. Direct migration at a resolution of 160x160 pixels (images from [11])

The process of splitting a web page into cards requires some analysis for those pieces to contain relevant information, and thus for the technique to be useful. The implementation in paper [11] used a simple lexical analysis algorithm. It was found that the unfocused pieces were too small to be of use unless a user had previous knowledge of the layout of the page, and so further techniques were used extract content, such as those discussed earlier in this paper.

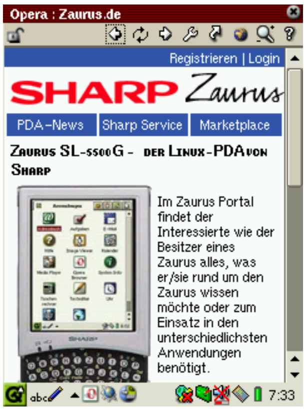
A disadvantage of this technique, as with using transparent widgets, is that there is a considerable familiarisation period. An ideal solution would be intuitive enough to be obvious on first glance how things work, or require very minimal discovery. Some of the problems covered in paper [11], such as unfocused pieces being unreadable, may not be valid on a display with a higher resolution, as are becoming common in recent hand-held devices.

6. A Combination of Techniques

Based on the information gathered on the previously examined processes and techniques, I suggest that a combination of these may be the best way of making general web content more accessible on small screens.

Firstly, I suggest that DOM manipulation be used to assign priority to paragraphs of text. I suggest that if a paragraph with significantly lower priority than other paragraphs is found, it be reordered towards the end of the page. These paragraphs could be identified by the link-text to standard text ratio, or by using some simple lexical analysis, such as counting the amount of prepositions or calculating the space to word ratio. Most lexical analysis techniques have the side-effect of making any process locale-specific, so if possible, it is to be avoided, or made optional.

Secondly, I suggest that a stylesheet be used to flatten and resize content, and to highlight links. At this point, any given web-page would be rendered similarly to Opera's small screen rendering[4], as shown in figure 5.



Zaurus SL-5500G - der Linux-PDA von Sharp

Im Zaurus Portal findet der Interessierte wie der Besitzer eines Zaurus alles, was er/sie rund um den Zaurus wissen möchte oder zum Einsatz in den unterschiedlichsten Anwendungen benötigt.

Figure 5 – Opera small screen rendering (Image from www.linux-user.de)

At this point, I suggest that paragraphs over a particular length be collapsed into an expandable link made up of keywords from the paragraph, or from a header, if such a header can be extracted from the main body of the page. For example, on a news site, news stories may be collapsed under an expandable link of their heading. To avoid becoming locale-specific when using keyword extraction, these words could be identified via length, or the link could be made from the first few words of the paragraph, followed by an ellipsis. When these links are activated, the full paragraph could be expanded underneath them. If the paragraph is already expanded, it would be collapsed. This could be achieved by modifying the CSS visibility property in the DOM tree and adding a new text block at the same level in the tree that contains the link. Alternatively, it could be done at the renderer level, without modifying the DOM tree, although this may prove to be too complex a modification to perform in this manner.

This is similar to The Gateway and Flip Zooming in that it puts the focus on a single part of the page, but differs in that it can be implemented in a generic way in any browser, using a combination of JavaScript and CSS. The combination of these methods should eliminate horizontal scrolling, minimise vertical scrolling and highlight the user's attention to the important parts of the page, without too heavily altering the content and running the risk of making any information inaccessible, or worse, inaccurate.

7. Conclusion

It has been shown that the only assured method to make general web content accessible on a device with a small display is to manually reformat said content. As portable devices start to converge and become more common, this may happen more often. An intermediary, automated method, however, is required if today's devices are to be used efficiently to browse the web. A number of methods for achieving this have been examined, with varied amounts of success. Perhaps the most promising techniques are those that provide an entirely new way of representing a document, such as The Gateway[8], or Flip Zooming[11], however, none of those covered in this document have been developed to an extent where someone might use them today.

The most popular techniques being used in browsers today involve styling (used in the Minimo[6] and Opera[4] browsers) and DOM modification, either at the client-side or by special proxies (used by the Opera[4] browser on mobile phones). As portable devices become cheaper and more commonplace, it is likely that more and more websites will provide special stylesheets for small screens, or even recreate their content specifically for small devices (via methods such as news feeds or alternative locations).

As with any emerging technology, there will likely be many new developments in the years to come.

8. References

[1] Suhit Gupta, Gail Kaiser, David Neistadt and Peter Grimm, *DOM-based Content Extraction of HTML Documents*, Dept. of Comp. Sci., Columbia University. International World Wide Web Conference, 207-214 (2003).

[2] Tomonari Kamba, Shawn A. Elson, Terry Harpold, Tim Stamper and Piyawadee “Noi” Sukaviriya, *Using small screen space more efficiently*. Conference on Human Factors in Computing Systems, 383-390 (1996)

[3] Bonnie MacKay, Carolyn Watters and Jack Duffy, *Web Page Transformation when Switching Devices*, Faculty of Computer Science, Dalhousie University. MobileHCI 2004, 228-239 (2004).

[4] *Opera's Small-Screen Rendering(tm)*, Opera Software, http://www.opera.com/products/mobile/smallscreen/, last accessed 19th November 2005

[5] Simon Corston-Oliver, *Text compaction for display on very small screens*, Microsoft Research (2001).

[6] *Minimo Project*, Mozilla Foundation, http://www.mozilla.org/projects/minimo/, last accessed 19th November 2005

[7] *Minimo small screen rendering stylesheet*, Mozilla Foundation,

https://bugzilla.mozilla.org/attachment.cgi?id=140182, last accessed 19th November 2005

[8] Bonnie McKay, *The Gateway: A Navigation Technique for Migrating Small Screens*, Dalhousie University. Doctoral Consortium, CHI 2003, 684-685 (2003).

[9] Matt Jones, Gary Marsden, Norliza Mohd-Nasir, Kevin Boone and George Buchanan, *Improving Web interaction on small displays*, Interaction Design Centre, School of Computing Science, Middlesex University. 8th international conference on World Wide Web, 1129-1137 (1999).

[10] Matt Jones, Goerge Buchana and Harold Thumbleby, *Sorting out Searching on Small Screen Devices*, University of Waikato, Middlesex University, University College London. 4th international symposium on mobile human-computer interaction, 81-94 (2002).

[11] Staffan Björk and Johan Redström, *An Alternative to Scrollbars on Small Screens*, PLAY: Applied Research on Art and Technology, Viktoria Institute. CHI '99 Extended Abstracts, 316-317 (1999).

[12] *Macromedia Flash Player*, Macromedia, http://www.macromedia.com/software/flashplayer/, last accessed 16th December 2005