COMPUTER
ENGINEERING
PROGRAM

Middle East Technical
University
Northern Cyprus
Campus

eMINE Technical Report Deliverable 1 (D1),

August 2011

Heuristics for Visual Elements of Web Pages

Yeliz Yesilada
Middle East Technical University,
Northern Cyprus Campus,
Kalkanlı, Güzelyurt, TRNC,
Mersin 10, TURKEY

This technical report presents the heuristics that have been developed for identifying the visual elements of web pages. In order to develop these heuristics, we have used the following: our travel analysis framework which already included some guidelines, web pages that are previously used in our studies, and the related work that exists in the literature. This technical report aims to present the heuristics and detail the process followed in creating these heuristics.

METU
NCC

eMINE

The World Wide Web (web) has moved from the Desktop and now is ubiquitous. It can be accessed by a small device while the user is mobile or it can be accessed in audio if the user cannot see the content, for instance visually disabled users who use screen readers. However, since web pages are mainly designed for visual interaction; it is almost impossible to access them in alternative forms. Our overarching goal is to improve the user experience in such constrained environments by using a novel application of eye tracking technology. In brief, by relating scanpaths to the underlying source code of web pages, we aim to transcode web pages such that they are easier to access in constrained environments.

Acknowledgements

The project is supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK) with the grant number 109E251. As such the authors would like to thank to (TÜBİTAK) for their continued support.

Contents

Middle East Technical University,
Northern Cyprus Campus,
Kalkanlı, Güzelyurt, TRNC,
Mersin 10, TURKEY

Tel: +90 (392) 661 2000
http://www.ncc.metu.edu.tr/

Corresponding author:
Yeliz Yesilada
Tel: +90 (392) 661 2994
yyeliz@metu.edu.tr
http://metu.edu.tr/~yyeliz

1 Introduction

Web pages are typically designed for visual interaction. In order to support visual interaction they are designed to include a number of visual segments. These visual segments typically include different kinds of information, for example they are used to segment a web page into a number of logical sections such as header, footer, menu, etc. They are also used to differentiate the presentation of different kinds of information. For example, on the news site they are used to differentiate different news items. However, unfortunately the underlying source code is not encoded in such a way to differentiate the visual segments. They are typically encoded for visual consumption and not for machine processing. For example, on can easily look at the visual rendering and can differentiate the header which has red background colour, top menu which has black and grey background colour, headline news which has a bigger image and a larger text for the header, etc. However, when we look at the source code we cannot see such kind of clear segmentation or pattern in the source code. We can find list items, paragraphs, etc. However, identifying such segments can be very useful for different fields, for example, web pages can be properly displayed or repurposed for mobile devices, blind users can easily access them with their screen readers, search engine can use such implicit information to provide a better search result, etc.

When we look at the real world mobility studies in wayfinding and urban or architectural design show that people use a variety of objects in a physical environment to successfully complete their journeys [26, 28]. When we look at the web from this perspective and think about it as an environment where people make journeys [25], counterparts of these objects also exist on web pages [18, 37]. We have identified three broad categories of visual elements of web pages that are summarised below along with some examples [37].
(1) Way Points are points within a journey at which a decision may be made that facilitates onward movement. These include Decision Points (e.g., menu), Way Edges (e.g., colour boundaries), Navigation Points (e.g., hyperlinks), Reference Points (e.g., logo), etc.
(2) Orientation Points are used for both establishing and maintaining orientation. The knowledge about orientation suggests that such objects include Reference Point (e.g., logo), Direction (e.g., back and forward button), Location and Position (e.g., menu highlighting active item) objects, etc.
(3) Travel Assistants are different strategies used by both sighted and visually impaired travellers to re-orientate themselves when they experience problems in orientating themselves in unfamiliar or familiar environments. These strategies include usage of: Information Points (e.g., search box), Travel Aids (e.g., site map), Travel Memory (e.g., history list) and Travel Support (e.g., guided tour).

Based on this model, we have created a travel analysis framework that can be used to systematically analyse web pages for their travel support. This framework consists of two stages: (1) Inspecting a web page to identify travel objects and create a travel object inventory. The aim is to filter the page and find the objects that are useful in promoting the onward journey; (2) Classifying each travel object in the inventory according to the role it plays in the travel process (which can be greater than one).

The rest of this technical report is organised as follows: Section 2 presents detailed information about these visual elements of web pages. Section 3 presents a number of heuristics that can be used to identify such visual elements of web pages. Section 4 presents some heuristics identified in the literature and finally Section 5 presents what has been done

in the literature to identify the importance of such visual elements.

2 Visual Elements of Web Pages

Travellers use environmental features or elements in order to make a successful journey which are called *travel objects* [26, 28]. Travellers use landmarks and memory objects to reassure themselves that they are safe to proceed and going the right way. Landmarks and memory are two main classes of travel objects, and these classes also include sub-classes which are as follows [18]:

- *Landmarks*: Alert, information point, identification point, way point and way edge.
- *Memory*: Memory, alert, information point and identification point.

These classifications are dynamic and can overlap. An identification point is both a landmark and a memory object. A landmark may be classified as an information point and a way point on closer inspection. This classification depends on the context of the travel. These travel objects are an important part of the model of travel. Fundamentally, the encapsulated process extracts travel objects from the environment in which travel takes place – a web page. Therefore, before we can extract them, we must identify such objects and specify their characteristics. These have to be in detail and must reflect the key environmental features used by travellers. The extensions to [18] have lead to three broad categories of travel objects: (1) way points, (2) orientation and (3) travel assistants. These categories also include sub-classes which are explained in the following sections. Table 1 presents these objects and examples from the real and web world.

2.1 Way Points

These are the points within a journey at which a decision may be made that directly facilitates onward movement. Below, we explain the sub-classes of way points, however other classifications may also be considered as way points depending on the journey undertaken:

Decision Points are the choice points where alternative paths of travel are possible. Travellers recall the direction they must travel to reach their destination and they change their direction of travel if necessary [24].

Way Edges are the environmental elements that are linear or continuous and act as boundaries between two areas [26]. During a journey, visually impaired travellers usually make explicit references to these elements in the environment [16].

Navigation Points provide a possible route and the traveller exercises some control by choosing to follow or not to follow it. They can be considered as decision points in wayfinding, but the traveller is not choosing from a set of options; the traveller needs to decide to follow or not to follow it. They relate directly to the *paths* that are defined as channels along which people potentially move [26].

Reference Points or Landmarks are some aspects of the environment that are unique and memorable [26]. They are defined as the most salient cues in any environment [15] and are conceptually and perceptually distinct locations [20].

	Role	Real World	web
Way Points	**Decision Points (DP)**	Junction, corridor intersections.	Menu, check boxes, and combo box.
	Way Edge (WE)	Wall, shorelines, hedges, fences, cliffs, railroad tracks [26], edge of lawns.	Colour boundaries, frame borders.
	Navigation Points (NP)	Streets, walkways, canals, railroad [26], corridors and some examples specific to buildings: stairs, escalators, elevators [28].	Hyperlinks, mail to, search box.
	Identification Points (ID)	Building name, house number, street name, campus entrance identification, floor number.	Heading, title, URL, frame title.
	Alert (AL)	Traffic lights, signs that identify hazards, bridges, stops, and curves.	Progress bar, a search field icon can change the colour and appearance to indicate that the search is in progress, displaying or animating an hour glass.
	Attention (AT)	Advertisements.	Animation, advertisement, banner.
	Reference Points (RP)	Big Ben, Precinct centre bridge on Oxford road [28], sculptures, structural or decorative elements, information booths, and etc.	Logo, title banner, some aspect of the style.
Orientation	**Location & Position (LP)**	"You are here" signs, signs that show the current position, reassurance signs on the motorway.	Hyperlink menus highlight the active page, some sites show the active page (e.g., home>products).
	Direction (DR)	Signs that show the direction to cities, towns, signs with arrows.	Scrollbar direction, Back and Forward button.
	Distance (DS)	Signs that show the distance to cities or towns.	Scrollbar (distance from top & bottom of the page).
Travel Assistants	**Information Points (IP)**	Information or help desk, police, friends.	Search Box.
	Travel Aid (TA)	Map, tactile map, road list, floor plan.	Site index, table of contents, site map, outline.
	Travel Memory (TM)	Route plan, journey plan.	Previously visited links change colour, history list, bookmark.
	Travel Support (TS)	Guided tour, visually impaired individuals may learn new routes by guiding them.	Guided tour.

Table 1: Travel objects and examples.

Increased usage of landmarks or reference points is one of the strategies used by visually impaired people to travel efficiently in their homes and communities [24]. Landmarks have to communicate some specific, identifiable features [29]. They may be primarily physical objects, but they can be sounds, odours, temperature or tactual stimuli [5, 24, 29].

Reference Point Components are directly related to the information points and are defined as two or more stimuli that, when linked, allow a traveller to determine his or her exact location [24]. A single reference point component might not be enough to identify the exact position of the traveller. Whereas, when two or more reference point components are linked, a traveller can determine his or her exact position. They are common features which

do not provide precise position. However, they might help in determining one's general position.

Identification Points are identification signs that are elementary state description of a location and usually perceived when the destination is reached [28]. These points identify an object, a place or a person in the space.

Identity is what makes one part of an environment distinguishable from another; it is a characteristic that allows the traveller to differentiate parts of the environment [2]. Travellers can use identification points to validate their arrival at the destination ("this is it").

Attention These are the objects that attract traveller's attention and may change the traveller's focus. They may be used for observations that may lead to interesting discoveries but yield most initiation control to the environment [27].

Alert These objects alert the traveller to a change in the environment or control of the journey [18]. Like the attention objects, they also attract the traveller's attention, but they usually notify approaching action or danger.

2.2 Orientation Points

Orientation is defined as the knowledge of one's direction and distance relative to things observed or remembered in the surroundings and keeping track of these spatial relationships as they change during locomotion [4, 5]. The concepts[1] of position or location, directionality and laterality are important cognitive components for orientation during mobility [24]. Moreover, as one moves towards a desired goal, establishing orientation and maintaining orientation are critical components of successful travel.

The knowledge about orientation suggests that a person needs information about location, distance and direction in order to be oriented in a journey. Landmarks are used to give a sense of location [19] and are defined as spatial anchors since they provide precise information about one's location [24]. Landmarks are also important for the orientation of visually impaired travellers [4].

Direction Directional information is essential to the navigator's ability to remain oriented within the environment [14, 26]. A sense of direction that is an ability to maintain direction while moving, is usually equated with a sense of orientation [28]. Directional information can be provided through the directional signs that designate direction towards a place, an object or an event in form of a name, symbol or pictograph and an arrow. They may also show which direction the traveller is moving along.

Distance The ability to make accurate distance estimations facilitates establishment and maintenance of orientation [24]. Objects that provide distance information may indicate distance from the traveller's starting position or from the traveller's destination. They may also show where one is with respect to nearby objects and the target location.

Location or Position Landmarks provide implicit location or position information, because

[1] Concepts are defined as mental representations, images or ideas [24].

different travellers may need varying amounts and type of information about landmarks. However, there may be some objects in the environment which provide location or position information explicitly. These location or position objects are directly related to reassurance signs and are defined as checkpoints which are used by travellers to reassure themselves that they are on the right track [28].

2.3 Travel Assistants

Sighted or visually impaired travellers may all experience problems in orienting themselves from time to time in an unfamiliar or familiar environment. They use different strategies for solving this problem including consulting a map; exploring the space systematically, either alone or with a guide; or following verbal or written directions [24].

Information Points By using these objects, a traveller can directly request information. The traveller controls the type and amount of information requested and supplied, so they are an active information supply [24].

Travellers can update their spatial information by interacting with other people while travelling. This is one of the strategies used by travellers for re-orienting themselves [24]. A supportive environment can be thought of in terms of information points at frequent and regular intervals. Particularly, information points may be important for visually impaired travellers, since they compensate for not having access to distant cues that are so useful to the sighted travellers [30].

Travel Aids provide an overview of the environment. They usually place the entire environment within the traveller's view. We refer to them as passive information supply, because, unlike information points, the traveller does not control the type and amount of information. They can also be considered as secondary sources, which can be used for spatial knowledge acquisition. They may help travellers to determine their position in the environment, their direction of travel, and the relative position of other objects or places in the environment [15].

Travel Memory holds information about where the traveller has been and provides means to get back there. It can be considered as an external memory aid to supplement internal memory [18].

Travel Support A traveller may make a journey without controlling all the details of the journey, that is to say, the traveller may not actively control the journey undertaken. Travellers may make a journey by actually being guided throughout. This could be a strategy for travellers to learn the spatial relationships in an unfamiliar environment [15].

3 Heuristics for Identifying and Classifying Visual Elements of web Pages

This framework is based on the model of travel. It aims to capture a systematic method to discover objects and their roles as explained in the previous section. Fundamentally, the

framework is comprised of two stages:

1. Inspecting a web page in order to create a travel object inventory;
2. Classifying each travel object in the inventory according to the role it plays in the travel process.

3.1 Identification of Visual Elements

If we look at the web landscape from the real world travel perspective, travel objects also exist in the web landscape[2] (see Table 1). For example, if we consider a typical web page. The logo at the top left corner acts as a *reference point*; it is a unique and memorable feature of the page. The colour boundary on the side is a *way edge*; it is used to separate the "side content" from the "main". The heading "personal banking" is used as an *identification point* and used to identify that part of the page. As these examples demonstrate, the travel objects play an important role in the mobility of the web users. Since in the visually impaired users' web landscape these objects are not presented appropriately, their mobility is reduced. Travel objects should be presented in a way that they can fulfill their intended roles and ease travel on the web.

In the first stage of the framework, web pages are analysed to find out the provided travel objects and create a travel object inventory. The aim of identification is to filter the page and find the objects that are useful in promoting the onward journey. These objects are the regions or portions of the page i.e., an HTML element, collections of elements or parts of elements depending on the rendering. In order to identify these objects, a journey should be made from top to bottom by noting which parts of the page (rendering or underlying code) are useful in promoting the onward journey. These then become travel objects.

A set of guidelines are also developed for identifying objects. These guidelines are established by investigating a large number of web pages. The aim is to make the identification process systematic and consistent. If the process is proved to be so, the guidelines can form the basis of heuristics for travel object identification within an automated process. These heuristics can also evolve throughout the application. The guidelines are principally grouped into four. They are summarised as follows:

- **Extracting travel objects from a page** These are the fundamental strategies for extracting travel objects, and the important aspects about the environment and travel objects. E.g.,
  - *A bird's-eye view* of a page may help to spot visual groupings and draw a sketch of the page. Then the sketch can be extended by zooming in and out from these groupings and by considering their relationships.
  - *Granularity* A travel object may be atomic or composite (i.e., composed of other travel objects).
  - *Content or Context Change* When a page is analysed linearly (from top to bottom), any context or content change can be considered as a clue to differentiate objects.

[2]The web landscape is defined as the combination of the page and the agent (e.g., browser) [18].

- **HTML source code** Some details are hidden in the source code and some are in the rendering, so it is important to inspect both. E.g.,
  - *An image map* When the rendering of a page is analysed, it may be difficult to realise the image maps, whereby analysing source code may help to elucidate the details of the image maps.
  - *Layout tables* The details of invisible layout tables can be obtained from the source code. Sometimes groups of objects are located in a cell of the layout table, so it may help to check whether or not they can be considered as a single travel object.
  - *Image with a link* It might not be easy to recognise whether there is a link on an image or not so it might help to check the source code.
- **Using elements of a document** The general knowledge about the structure of a standard document can be used in finding out the travel objects (the role of a part of a document can be considered). For example, sections, paragraphs, titles, headings, bulleted or numbered lists are all different travel objects. Similarly, the HTML elements can also be considered, for instance, links. E.g.,
  - *Sections and headings* The heading and the content part (section) should be considered separately because it is likely that they have different roles during a journey. The headings can be obtained by checking whether the source code contains H1 through H6 tag set. However, this may not be enough because not all headings are explicitly specified by using this tag set. Different typefaces may be used to indicate headings; thus it is important to inspect both the rendering and the underlying source code.
  - *Boundaries* (e.g., line, colour and space) are used to visually divide information or sections. For example, a line boundary can be created by using a HR tag. Besides helping in recognising the context division in a page, they may also be considered as travel objects.
  - *Links* All the links on a page are candidates for being travel objects despite the fact that they could be grouped together with other objects depending on the context.
  - *Logo* If an image or icon is used to identify that particular page, then it can be considered as a travel object.
  - *Animations and Adverts* can be considered as single objects as they usually have different context.
- **Neighbourhood objects** Objects that are grouped together to provide a common function can be considered as a single travel object. E.g.,
  - *Functional dependency* If the functionality of an object depends on anther object then they can be considered as a single travel object. For instance, the search capability in a page is usually provided by a search box, a ‘go’ button and a label. These three objects can be considered as a single travel object because the functionality of the ‘go’ button depends on the search box and cannot be used on its own.

- *Single destination* If a group of objects (e.g., an image and a paragraph) links to the same page, then that group can be considered as a single object.
- *Common style* If objects with different content are presented with the same style (e.g., menu items), then it is likely that those objects form a composite object.
- *Interaction* If the traveller needs to interact differently with consecutive objects then they can be considered as different objects.

3.2 Classification of Visual Elements

This second stage aims to classify the extracted travel objects. The main use of this classification is to discover the roles of each travel object in the inventory. Every travel object has at least one role during a journey and depending on the journey, it may have more than one role or it may have different roles in different journeys. Since we cannot consider all the different possible journeys, we only consider the possible roles of travel objects in a general context.

The classification process consists of a series of questions (a questionnaire) that have to be answered for every object in the created inventory. These questions aim to capture the definitions of the classifications of travel objects (Table 1). For example, a *Travel memory* is defined as "an object that holds information regarding where the traveller has been and provides means to get back". Therefore, questions like "does this object show where the traveller has been?", "does it show previously visited places", etc., are included.

The expected answers to the questionnaire are *yes* or *no*. This is to simplify the process, make it systematic and have it in a form which is easy to automate. The results are then evaluated to infer the possible roles of the travel objects. The aim of asking every question to every object is to try to decrease the subjectivity of the approach and to provide a systematic rationale to classification.

4 Related Work – Heuristics in the Literature

In the literature, some work also looks at understanding the role and the functions of the segments. In our work, we consider these segments as visual elements. Some segments in a web page has specific roles, for example some of them are used as headers, footers and some are used as the main content or headline in a web page. In this section, we review the literature and identify the work that has been done to understand the role of the segments.

- [39] proposes an algorithm based on random walks that classifies elements of web pages into five categories which are Content (C), Related Links (R), Navigation and Support (N), Advertisement (A) and Form (F). Their algorithm automatically categorises web elements by developing five graphs, one for each functional category, with the basic elements in the web page as vertices. Each graph is specifically designed such that most of the probability of stationary distribution of a random walk is concentrated in the nodes that belong to its corresponding category. They are focusing on a very small set of roles/functions of web elements. However, in reality web elements have many more different roles.

- [1] uses heuristics to recognise the major sections of a web page which are: 1) top, 2) main content, 3) left and right menus, 4) bottom and 5) clutter such as advertisement.
  - Top includes the title of the page and a menu bar – Heuristic$_{top}$: "If a list of hyperlinks (i.e., a menu bar) or a table including a list of hyperlinks is placed within the top 200 pixels of the page [21], it is considered to be the top section."
  - Left and right menu includes navigation links – Heuristic$_{menu}$: "If a list of hyperlinks or a table including a list of hyperlinks is placed on the left (right) side of the page, occupying up to 30 percent of the page width [18], and its upper boundary is below the top section and its lower bound is above the bottom section, then it is considered to be the left (right) menu section.";
  - The main content is included in the center – Heuristic$_{main}$:"The remaining area (see Heuristic$_{top}$, Heuristic$_{bottom}$ and Heuristic$_{menu}$) is considered to be the main area";
  - Clutter usually contains images that are located in the bottom or side of a page.
  - Bottom – Heuristic$_{bottom}$: "If a table is placed within the lowest 150 pixels of the page [21], it is considered to be the bottom section.".

  [1] are very good and promising, however they are based on a specific model of web pages which is the header is at the top, the menus ate on the sides and the bottom of the page is clearly marked. Although this covers a wide variety of web pages, there are so many different structures on the web.

- [23] uses VIPS algorithm [10, 9] to first identify the blocks in a web page. By using heuristics these blocks are then categorised into three: Navigation bar, navigation list and content block. These three categories are then used to filter the unnecessary content (blocks which are not in these three categories). The following heuristics are used:
  - Navigation bar – Heuristic$_{bar}$: if the blocks have links, and if does not have another domain and if the over half of the content is links and if average link length is more than 10 and it does not have contents other than links then that block is a navigation bar.
  - Navigation list – Heuristic$_{list}$: if the blocks have links, and if does not have another domain and if the over half of the content is links and if average link length is more than 10 and it does have contents other than links then that block is a navigation list.
  - Contents – Heuristic$_{content}$: 1) If a block has links and the number of words is more than 100, then it is a content block. 2) If a block has links and it does not have another domain name and it does not have over half of links then it is a content block.

  It is again very difficult to see how these heuristics can be generalised, and also the role of the blocks are only focuses on these tree items which do not reflect the actual roles of elements in a web page. In summary, they do not really focus on identifying the role of these blocks but they rather focus on grouping these blocks into three categories.

- [34, 35] differentiate link blocks from content blocks. Even though they do not try to understand the role of structural elements, their blocking algorithm does differentiate between content and link blocks.

- [13, 12] aims to identify high-level content blocks which are header, footer, sidebars (left and right) and the body. The authors also refer to these layout components altogether as semantic structure. They assume that header and footer typically has a flat shape, and header is located at the top and the footer is located at the bottom.

  - Header – Since the header block locates on the top of the page, they define a threshold N and let the upper N pixels of a web page to be the header region. They propose the formula: N=bas_threshold+F(heightwidth) where F(x) = a(b*x+c) where base_threshold, a,b and c are constants. Their experiments show that base_threshold=160, a=40, b=20 and c=1 are
  - Footer – Similar approach to header is used for the footer.
  - Left and Right Sidebar – 1/4 part of a web page to be the left sidebar and 1/4 of part of the page to be the right sidebar.

  Even though the main higher-level content blocks are covered, the way they are identified are focused on a specific web side template. Header at the top, footer at the bottom, left and right bars on each.

- [11] proposes a Function-Based Object model (FOM Model). According to this model, a web page consists of basic (smallest information body that cannot be further divided, only as a whole can perform certain functions) and composite objects (is a set of objects (both basic and composite) that perform certain functions together). They also propose a number of object categories:

  1. Information object, for content information
  2. Navigation object, provides navigation guide. Further divided as 1) navigation bar (provides global navigation), 2) navigation list (provides local navigation), 3) independent navigation guide (to provide navigation guide to certain piece of information)
  3. Interaction object, provides user side interaction
  4. Decoration object, serves only decoration purpose
  5. Special function object, performs special function such as advert, logo, contact, copyright, reference, etc.
  6. Page object, serves as the basic document of a web site, can be index page or content page.

  The authors also further define a number of rules for basic objects which include: 1) presentation property (media type, encoding formatting, layout information), 2) semanteme property, 3) navigation property (destination of a hyperlink), 4) decoration property (e.g, background colour), 5) interaction property (button, input, etc). Even though the proposed model is very comprehensive, it is very technical and does not take into account the users' understanding of the content. It does not also capture

the detailed understanding of the document. For example, how would you categorise header, footer, etc?

- [21] aims to segment a web page into a set of coherent objects. The overall aim is to build a representation for a web page in which objects are placed into well-defined tree hierarchy according to where they belong in an HTML structure of a web page. They then focus on recognising some common areas on web pages such as header, footer, left and right menus, and the center of the page. They mainly define a set of heuristics to identify these common areas. These heuristics are based on a model where there is a rigid abstraction of the visual representation of the page.

- This is not so much related to classification of web page segments but it is about the classification of images [17]. By analysing images from different kinds of pages such as business, governmental, education, news, etc. and they propose the following three categories: 1) unlisted images: which are standalone images that appear anywhere in the page; 2) listed images: are two or more images that are systematically ordered with a web page; and 3) semi-listed images are visually similar to listed images. These groups are also differentiated based on their DOM tree.

- [33] defines a web page as a composition of basic visual blocks and separators – they indicate that visual blocks are visual parts in a web page that cannot be divided further. They have a very simple classification of blocks. They mainly classify them as nontext blocks (buttons, images, inputs, etc) and text blocks (is the area containing a paragraph of text, except text on forms). This is a very technical classification of content and has nothing to do with the role of the objects.

- [22] defines two types of blocks: informative and redundant content blocks. Informative content blocks include content which is semantically meaningful to users and redundant content blocks include redundant data such as company logos, navigation panels, advertisement banners, etc. Their focus is information retrieval therefore they focus on eliminating the intra-page redundancy.

- [6, 7, 8] indicates that web pages includes a lot of additional information besides the main content such as advertisements, copyright notices, etc that would effect the quality of data mining. Therefore, they propose an algorithm to identify the segments and based on the features of these segments to identify the main content. They mainly focus on identfying the following classes of objects: h1 (main artcile heading), h2 (second-level heading in the article), subtitle (the subtitle of the page), perex (the leading parag of the article), paragraph (an ordinary paragraph), data (publication date), author (author name), authordate (other object that belongs to the article), aobject (other object that belongs to the article), and none. Even though some of these are related to the role of segments in the page, it seems like they are mixed – some are related to the role of blocks in the page for example, h1, h2 etc and some are related to the semantics of the content for example author, authordate, etc. This is mainly because they are related to data mining. Furthermore, these are not systematically classified, they look like an add-hoc classification of the role of elements.

- [38] aims to eliminate blocks that contain noisy information with respect to data mining. Even though they do refer to different content blocks for example navigation

panels, copyright, privacy notes, etc, they do not aim to identify the role of information blocks in a web page.

- [3, 32] aim to identify the fragments in a web page such that the page can be transcoded to better support accessibility for blind users. [3, 32] proposes different roles for these fragments which include: proper content, updated index, general index, no-role, header, footer, advertisement, delete, layouttable. Even though this looks like a comprehensive list, there still some roles that are missing in this group of roles. They have also not done systematically, it seems like the authors came up with a number of roles that are important from accessibility perspective.

5 Related Work – The Importance of the Visual Elements

There are also some work that aims to identify the importance of blocks/segments of web pages:

- [31] uses VIPS algorithm [10] to segment a web page and then proposes two learning algorithms to be used for identifying importance of these blocks. Block importances are originally identified with four levels: *level-1:* noisy information such as advertisement, copyright, decoration; *level-2:* useful information, but not very relevant to the topic of the page, such as navigation, directory, etc; *level-3:* relevant information to the theme of the page, but not with prominent importance, such as related topics, topic index, etc.; *level-4:* the most prominent part of the page such as headlines, main content, etc. However, after their experiment they proposed to have the following levels [31, 36] which means combining the second and third group: *level-1:* noisy information such as ads, copyright, decoration, etc.; *level-2:* Useful information, but not very relevant to the topic of a page, such as navigation, directory, etc. or relevant information to the theme of a page, but not with prominent importance, such as related topics, topic index, etc.; *level-3:* The most prominent part of a page, such as headlines, main content, etc.

- [22] aims to discover content blocks and in order to find the important blocks. They classify the blocks into redundant and informative content blocks.

- [6, 7, 8] aims to also identify the main content block. Therefore, they implicitly consider the importance of the blocks.

- [3, 32] also aims to annotate the fragments in a web page and their importance. The importance of a fragment is assigned with a real number which is between -1 and 1, 0 is the default role. The importance of the block is then used when the page is transcoded. The annotator can also specify the importance of a fragment and based on the importance value when the page is transcoded the fragment is located respectively. For example, if a fragment is annotated as an advertisement, then it receives an importance value which is -0.8 and then when the page is transcoded it is placed at the bottom of the page.

6 Summary

This technical report presents a detailed information about the visual elements of web pages. It mainly presents a set of heuristics that can be used to identify visual elements on web pages and also discusses some heuristics that have been proposed in the literature. Finally, it also presents some works in the literature that proposes to identify the importance of visual elements of web pages.

References

[1] Hamed Ahmadi and Jun Kong. Efficient web browsing on small screens. In *Proceedings of the working conference on Advanced visual interfaces*, AVI '08, pages 23–30, New York, NY, USA, 2008. ACM.

[2] P. Arthur and R. Passini. *Wayfinding: People, Signs, and Architecture*. McGraw-Hill, 1992.

[3] C. Asakawa and H. Takagi. Annotation-based transcoding for nonvisual web access. In *ASSETS'00*, pages 172–179. ACM Press, 2000.

[4] B.B. Blasch, W.R. Wiener, and R.L. Welsh, editors. *Foundations of Orientation and Mobility*. AFB Press, American Foundation for the Blind, second edition, 1997.

[5] M. Brambring. Mobility and orientation process of the blind. In *Proceedings of the NATO Advanced Research Workshop*, pages 493–508, 1984.

[6] R. Burget. Layout based information extraction from html documents. In *Document Analysis and Recognition, 2007. ICDAR 2007. Ninth International Conference on*, volume 2, pages 624 –628, sept. 2007.

[7] Radek Burget. Automatic document structure detection for data integration. In *Proceedings of the 10th international conference on Business information systems*, BIS'07, pages 391–397, Berlin, Heidelberg, 2007. Springer-Verlag.

[8] Radek Burget and Ivana Rudolfova. Web page element classification based visual features. In *2009 First Asian conference on Intelligent Information and Database Systems*. IEEE Computer Society, 2009.

[9] Deng Cai, Shipeng Yu, Ji-Rong Wen, and Wei-Ying Ma. Extracting content structure for web pages based on visual representation. In *Proceedings of the 5th Asia-Pacific web conference on Web technologies and applications*, APWeb'03, pages 406–417, Berlin, Heidelberg, 2003. Springer-Verlag.

[10] Deng Cai, Shipeng Yu, Ji-Rong Wen, and Wei-Ying Ma. Vips: a vision based page segmentation algorithm. Technical Report MSR-TR-2003-79, Microsoft Research, 2003.

[11] J. Chen, B. Zhou, J. Shi, H. Zhang, and Q. Wu. Function-based object towards website adaptation. In *Proceedings of the Tenth International World Wide Web Conference*, Hong Kong, 2001. ACM.

[12] Y. Chen, W.Y. Ma, and H.J. Zhang. Detecting web page structure for adaptive viewing on small form factor devices. In *Proceedings of the Twelfth International World Wide Web Conference*, 2003.

[13] Yu Chen, Xing Xie, Wei-Ying Ma, and Hong-Jiang Zhang. Adapting web pages for small-screen devices. *IEEE Internet Computing*, 9:50–56, January 2005.

[14] R. Darken and J. Sibert. Navigating large virtual spaces. *International Journal of Human-Computer Interaction*, 8(1):49–72, 1996.

[15] R.P. Darken and B. Peterson. Spatial orientation, wayfinding and representation. In K.M. Stanney, editor, *Handbook of Virtual Enviroment Technology*, chapter 28. Lawrence Erlbaum Associates, Inc, 2002. `http://vehand.engr.ucf.edu/handbook/Chapters/Chapter28/Chapter28.html`.

[16] R. Edwards, S. Ungar, and M. Blades. Route descriptions by visually impaired and sighted children from memory and from maps. *Journal of Visual Impairment and Blindness*, 92(7):512–521, 1998.

[17] Fariza Fauzi, Jer-Lang Hong, and Mohammed Belkhatir. Webpage segmentation for extracting images and their surronding contextual information. In *MM '09: Proceedings of the seventeen ACM international conference on Multimedia*, pages 649–652, 2009.

[18] Carole Goble, Simon Harper, and Robert Stevens. The travails of visually impaired web travellers. In *HT' 00*, pages 1–10. ACM Press, 2000.

[19] R. Ingram and S. Benford. Improving the legibility of virtual environments. In *Proceedings of the Second Eurographics Conference on Virtual Environments*, 1995.

[20] S. Jul and G.W. Furnas. Navigation in electronic worlds: a CHI 97 workshop. *ACM SIGCHI Bulletin*, 29(4):44–49, 1997.

[21] Milos Kovacevic, Michelangelo Diligenti, Marco Gori, Marco Maggini, and Veljko Milutinovic. Recognition of common areas in a web page using visual information: a possible application in a page classification. In *Second IEEE International Conference on Data Mining (ICDM'02)*, page 250, 2002.

[22] Shian-Hua Lin and Jan-Ming Ho. Discovering informative content blocks from web documents. In *KDD '02: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 588–593, New York, NY, USA, 2002. ACM.

[23] Bing Liu, Chee Wee Chin, and Hwee Tou Ng. Mining topic-specific concepts and definitions on the web. In *Proceedings of the 12th international conference on World Wide Web*, WWW '03, pages 251–260, New York, NY, USA, 2003. ACM.

[24] R.G. Long and E.W. Hill. Establishing and maintaining orientation for mobility. In B.B. Blasch, W.R. Wiener, and R.L. Welsh, editors, *Foundations of Orientation and Mobility*, chapter 2, pages 39–59. AFB Press, American Foundation for the Blind, second edition, 1997.

[25] D. Lowe and W. Hall. *Hypermedia and the Web: An Engineering Approach*. John Wiley and Sons Ltd, 1998.

[26] K. Lynch. *The Image of the City*. The MIT Press, 1960.

[27] G. Marchionini. *Information Seeking in Electronic Environments*. Cambridge Series on Human-Computer Interaction. Cambridge University Press, 1995.

[28] R. Passini. *Wayfinding in Architecture*. Van Nostrand Reinhold, New York, 1984.

[29] R. Passini, J. Delisle, C. Langlois, and G. Proulx. Wayfinding information for congenitally blind individuals. *Journal of Visual Impairment and Blindness*, pages 425–429, 1988.

[30] R. Passini and G. Proulx. Wayfinding without vision; an experiment with congenitally totally blind. *Environment and Behaviour*, 20(2):227–252, 1988.

[31] Ruihua Song, Haifeng Liu, Ji-Rong Wen, and Wei-Ying Ma. Learning block importance models for web pages. In *Proceedings of the 13th international conference on World Wide Web*, WWW '04, pages 203–211, New York, NY, USA, 2004. ACM.

[32] H. Takagi, C. Asakawa, K. Fukuda, and J. Maeda. Site-wide annotation: Reconstructing existing pages to be accessible. In *ASSETS'02*, pages 81–88. ACM Press, 2002.

[33] Peifeng Xiang and Yuanchun Shi. Recovering semantic relations from web pages based on visual cues. In *Proceedings of the 11th international conference on Intelligent user interfaces*, IUI '06, pages 342–344, New York, NY, USA, 2006. ACM.

[34] Yunpeng Xiao, Yang Tao, and Qian Li. Web page adaptation for mobile device. In *Wireless Communications, Networking and Mobile Computing*, 2008.

[35] Yunpeng Xiao, Yang Tao, and Wenji Li. A dynamic web page adaptation for mobile device based on web2.0. In *Proceedings of the 2008 Advanced Software Engineering and Its Applications*, pages 119–122, Washington, DC, USA, 2008. IEEE Computer Society.

[36] Xing Xie, Gengxin Miao, Ruihua Song, Ji-Rong Wen, and Wei-Ying Ma. Efficient browsing of web search results on mobile devices based on block importance model. In *Proceedings of the Third IEEE International Conference on Pervasive Computing and Communications*, pages 17–26, Washington, DC, USA, 2005. IEEE Computer Society.

[37] Yeliz Yesilada, Robert Stevens, and Carole Goble. A foundation for tool based mobility support for visually impaired web users. In *WWW2003*, pages 422–430, 2003.

[38] Lan Yi, Bing Liu, and Xiaoli Li. Eliminating noisy information in web pages for data mining. In *KDD '03: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 296–305, New York, NY, USA, 2003. ACM.

[39] Xinyi Yin and Wee Sun Lee. Understanding the function of web elements for mobile content delivery using random walk models. In *Special interest tracks and posters of the 14th international conference on World Wide Web*, WWW '05, pages 1150–1151, New York, NY, USA, 2005. ACM.