Louvain School of Management

A Systematic Literature Review of Visual Design Metrics for Graphical User Interfaces

Author : Adrien Habay
Supervisor : Pr. Jean Vanderdonckt (UCLouvain)

Academic year : 2019 - 2020
Master [60] : Management Science

ACKNOWLEDGEMENTS

The realization of this master thesis was possible thanks to the support, time and energy of a multitude of people to whom I wish to address my most sincere thanks.

First of all, I would like to express my gratitude to my supervisor, Pr. Jean Vanderdonckt, for the advice he provided me during this thesis. I also thank him for having accepted to supervise me, for being available, for his help concerning the SLR method and for correcting this document. He always answered to my questions and encouraged me to continue my work.

I also thank Nicolas Burny for the support and follow-up of my work. He always showed interest in what I was doing because this subject concerns his Ph.D thesis and I am really happy to be able to help him on a part of his project. I would also like to thank him for agreeing to read and correct this thesis.

I would also like to thank the « kot-à-projet » Louvain-li-Nux and my girlfriend Iris for the help they gave me with the LaTeX language. I couldn't have written this thesis without their help.

Finally, I would like to thank my friends and my family for their continuous moral support during the redaction of this thesis. The coronavirus lockdown, which began in March 2020, has undoubtedly brought difficulty in carrying out this work.

ABSTRACT

Visual design of Graphical User Interfaces (GUI) aims to contribute to their usability by manipulating their visual components, such as their widgets, menus, contents and their layout by relying on a variety of techniques borrowed from general visual design and aesthetic properties. Therefore, a significant portion of GUI evaluation, but not the only one, concerns the visual aspects of GUIs. Many metrics attempt to characterize, express, and measure quality properties of the visual design of graphical user interfaces, but with different names, definitions, formulas, interpretations and empirical evidences. This thesis performs a Systematic Literature Review (SLR) for identifying and characterizing visual design metrics for graphical user interfaces on six scientific digital libraries, augmented by a snowballing procedure. We will identify a corpus of metrics that are systematically analyzed. Each identified metric will be described and accompanied by one or more formulas, the different formulas and metrics will be compared and their analysis will suggest some further consideration on how to apply and research further metrics in this area.

Contents

List of Figures

List of Tables

Chapter 1

Context

In this chapter, the reasons why computer scientists are interested in graphical user interfaces (GUIs) will be explained. These graphical interfaces have their own aesthetics features. The way by which it is possible to measure the aesthetics of these graphical user interfaces will be studied.

1.1 Problem statement

Over the last decades, digital technology has evolved until becoming an integral part of our daily lives. From media to entertainment, through recruitment, production and healthcare, the entire economy is becoming more and more digital. More than ever, companies will need to develop new business models and they will need to exploit the data they collect at large scale in order to be competitive. If they want to acquire a leadership position in the context of the digital transition, the firms will need to innovate in the digitization of their system. The digital transition is perturbing our social model and only the best adapting firms will survive. Even if it is based on old technologies, the contemporary digital economy really started only twenty years ago. It was born from the decision of the American public authorities to open the Internet to civil applications (Colin et al., 2015). With the emergence of Internet, the market of electronics devices, such as smartphones and computers, has been growing in an exponential way at the end of the 20^{th} century.

With the increasing complexity of the information we have to deal with, the usability of the user interfaces conveying information to users becomes more and more complex. Unusable product leads inevitably to dissatisfaction of the consumer (Norman, 1988). The way we use a screen and a system is affected by various factors such as the amount of information presented on the screen, its organization, its language and the distinctiveness of the screen's components (Ngo, Samsudin, et al., 2000).

1.2 Graphical User Interfaces Aesthetics

Graphical User Interface (GUI) is the most common way for an user to interact with a digital system, which may be desktop, tablet or smartphone. Graphical user interfaces are implemented by a set of software often included in operating systems. Microsoft's Windows is the most widely used operating system (around 73% share of the OS market in February 2020) (Liu, 2020), followed by macOS and iOS and then by Linux OS. Because of the importance and frequency of use of GUIs, researchers are interested in improving their quality.

In order to improve the user experience as much as possible, designers have put a significant effort into making Graphical User Interfaces usable and visually appealing (Miniukovich and De Angeli, 2014). At first, GUIs are designed to attract users eyes and then to make the users consult it easily and efficiently. Therefore, companies need to develop more attractive, intuitive and aesthetically pleasing website or application considering that GUIs aesthetics is a potential element to focus on in order to facilitate communication between device and user (Zen and Vanderdonckt, 2014). The companies interested in improving their software products can use the ISO 25000 norm to get a certification about the quality of their software.

The ISO 25000 norm, also known as SQuaRE (System and Software Quality Requirements and Evaluation), has been developed to give guidelines to website designers. This norm is an international standard for the evaluation of software quality. Software quality can be broken down into 3 different parts : software product quality, software data quality and operational quality of the software (ISO 25000, 2019). In this master thesis, we will be more interested in the software product quality. The software product quality model comprises eight quality characteristics shown in Figure 1.1:

Figure 1.1: Software product quality characteristics (ISO 25000, 2019)

In this master thesis, greater interest will be given to the question of software usability and more particularly to the user interface aesthetics. The usability is defined as « *the degree to which a product or system can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use* » and the user interface aesthetics is « *the degree to which an user interface enables pleasing and satisfying interaction for the user* » (ISO 25000, 2019). By improving the appearance of the user interface, we can increase the usability of a computer system and thus the satisfaction of the user. Aesthetics is a complex phenomenon made of culture-independent and culture-specific facets. Therefore, it can be hard to determine what a pleasing and satisfying interface is as this parameter can vary among persons. Indeed, it is often considered to be highly subjective as it depends on perception. Despite the growing interest in aesthetics, little is yet known on how to design interfaces that are indeed aesthetically pleasing (Miniukovich and De Angeli, 2014).

One of the other possible ways to improve the aesthetics of GUIs is to use visual design. Visual design of Graphical User Interfaces aims to contribute to their usability by manipulating their visual components (e.g., widgets, menus, contents, pictures, videos, banners) and their layout by relying on a variety of techniques borrowed from general visual design principles and techniques (Dondis, 1973), such as Gestalt properties (Mullet and Sano, 1995), hierarchy, visual techniques (Ngo et al., 2003), symbolic qualities (Hartono and Holsapple, 2019), and aesthetic properties (Zen and Vanderdonckt, 2014). Therefore, a significant portion of visual design, but not the only one, concerns the aesthetic aspects of GUIs.

Currently, GUI aesthetics can be evaluated using two different methods: an objective method and a subjective method. The objective method relies on complex formulas and algorithms to measure the value of a given aesthetic attribute of visual design. This method will be the one we will focus on in this work. The subjective method is based on a survey among end-users to get reviews about the value of each visual attribute for each considered interface (Zen, 2013). This last method is often used to empirically validate the objective method.

1.3 Visual Design Metrics

In order to evaluate the quality of GUIs aesthetics, computer scientists have developed metrics. Metrics are numeric values that summarize the status of specific interface attributes. Each individual metric evaluates only one isolated design aspect. By contrast, a set of metrics provides a powerful tool that evaluates multiple design issues. Metrics can thus be used to monitor some design aspects that are difficult for people to perceive accurately as well as those that are less difficult but easy to overlook (Sears, 1995).

There exists a multitude of metrics related to visual design of user interface with different names, definitions and interpretations. While many attempts have been made on proposing new metrics, they have all used different procedures and assessing tools, which do not allow clear conclusions to be drawn. This can lead to an ontological confusion. Moreover, the coverage of these metrics is not clearly defined. Metrics may be used in specific cases and several metrics can characterize the same attribute using different formulas. In addition, all metrics do not contribute equally to the overall aesthetics of the interfaces. They are very different and come from different disciplines (complexity management, layout, symmetry,..). The metrics are sometimes validated, sometimes subject to one or more different experiences and sometimes not validated at all. Therefore they have different strength of evidence.

The lack of appropriate concepts and measures of aesthetics may severely constraint future research in this area (Lavie and Tractinsky, 2004). Metrics are scattered in the literature, difficult to find and to bring together. The recommended methodology for aggregating empirical studies is a Systematic Literature Review (SLR) (Kitchenham et al., 2010). The aim of this master thesis will therefore be to carry out a SLR in order to gather all the metrics known about visual design in a single document. Each metric will be accompanied by its definition, its formula, it's strength of evidence and its context of use. The relative importance level of each metric will be studied according to the number of articles related to it.

Chapter 2

Introduction

In this chapter, the definition of a Systematic Literature Review (SLR) ans its advantages will be exposed. The way a SLR is performed will be studied. Terms such as backward or forward snowballing will be defined. The last section of this chapter will deal with visual design, its elements and principles.

2.1 Systematic Literature Review

2.1.1 Definition

A Systematic Literature Review (SLR) is a study which aims to provide a thorough overview of a research field by using a scientific and repeatable method. A systematic review attempts to collate all pieces of empirical evidence that fit pre-specified eligibility criteria to answer a specific research question (Liberati et al., 2009).

Systematic Literature Reviews are qualitative reference work which can save time for subsequent studies and provide an understanding of the existing literature on a specific topic (Casteleyn et al., 2014). The SLR method uses unambiguous and systematic procedures to minimize the occurrence of publication bias during searching, identification, synthesis, analysis and summary of studies. When it is done properly and has the minimal error, SLRs can provide reliable findings and conclusion that could help decision-makers and scientific practitioners to act accordingly (Mengist et al., 2020).

Performing a SLR is particularly important for highly multidisciplinary fields such as software engineering, for example, because many authors publish their work on multiple publishing platforms without even being aware of all its facets and contributions in the broader research field (Casteleyn et al., 2014).

A study of all available literature must be open-minded and transparent. That is why it would not suffice to merely compile a simple collection or summary of other articles. There must also be an element of analytical criticism when a SLR is performed. Also, the review cannot simply regurgitate the subject matter, but contributes to the work in its dual approach of synthesizing the available material and offering a scholarly critique of theory (Okoli and Schabram, 2010).

2.1.2 Procedure

Several studies (Casteleyn et al., 2014; Liberati et al., 2009; Villarreal-Narvaez et al., 2020) suggest that a SLR can be organized as a four-phases flow procedure consisting of Identification, Screening, Eligibility and Inclusion phases. The PRISMA diagram (Preferred Reporting Items for Systematic reviews and Meta-Analyses) shown in Figure 2.1 is a commonly used tool to represent the different phases and steps of a SLR (Liberati et al., 2009).

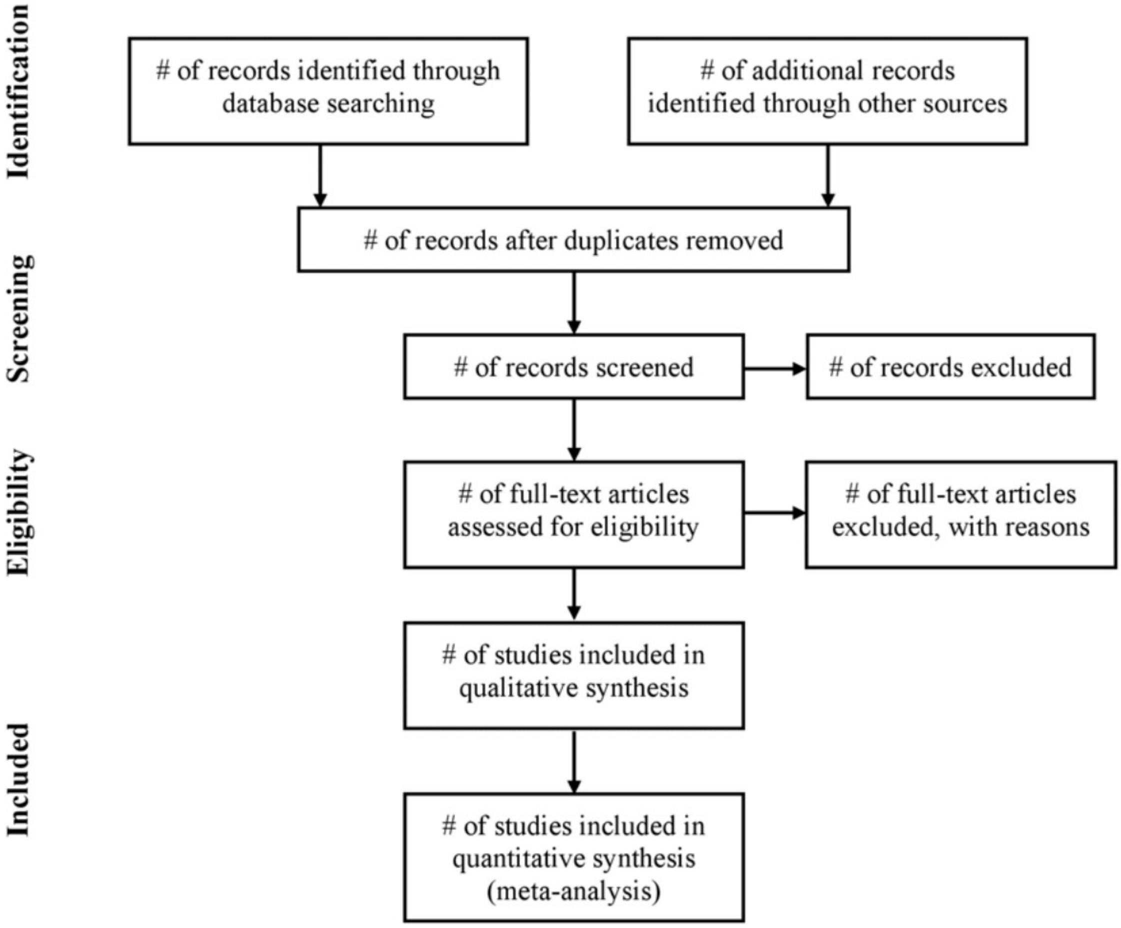


Figure 2.1: Four phase flow diagram used for a Systematic Literature Review (Liberati et al., 2009)

The aim of the identification phase is to obtain the initial set of potentially relevant primary studies, generally called « publications » (Casteleyn et al., 2014). The first step is to identify the purpose and intended goals of the review. Then, the reviewer needs to describe explicitly the details of the literature search and translate it into a semantic query. Indeed, it is necessary to define a semantic query to explain and justify how the comprehensiveness of the search is assured (Okoli and Schabram, 2010). A semantic query is introduced in a semantic Web browser or in a database to provide access to relevant publications on the topic of interest (Bast et al., 2016).

The semantic Web is a vast space for the exchange of resources between humans and machines. Using the semantic Web allows a qualitatively superior exploitation of large volumes of information. Performing a search on the semantic Web relieves users of a large part of their search tasks thanks to the increased capacity of machines to access and reason about the content of resources (Laublet et al., 2002). Indeed, the presentation of answers by the search engine is done in a way that tries to understand the user's query and the intention behind it (Amerland, 2014). However, it is interesting to note that only a decade ago, search engines were mostly lexical. Lexical means that the search engines were looking for literal matches of the query words typed by the user without making any effort to understand what the whole query actually means (Bast et al., 2016).

In order to perform an extensive search in the field of computer science, the semantic query can be introduced in the search engines of the major computer science publishing houses such as: (1) Springer (SpringerLink[1]), (2) IEEE (IEEE Xplore[2]) and (3) ACM (ACM Digital Library[3]). Then, to validate the search results, to ensure the completeness and coherence of references and to cover other publishers, two additional generic search engines can be used: (4) Google Scholar[4] and (5) DBLP CompleteSearch [5] (Casteleyn et al., 2014; Villarreal-Narvaez et al., 2020). Each search engine has its own syntax. It is therefore necessary to adapt the chosen search string according to the syntax required by the search engine.

The second part of the Systematic Literature Review is the screening phase. After the identification phase, the initial set of publications contains irrelevant results that must be removed. Firstly, all double entries are eliminated. The screening for inclusion step requires the reviewer to be explicit about what studies were considered for review and which ones were eliminated without further examination. Each publication is therefore inspected for relevance with respect to the research scope using criteria of form and contents (Okoli and Schabram, 2010).

---

[1] http://www.springerlinks.com/.

[2] https://ieeexplore.ieee.org/.

[3] http://dl.acm.org/.

[4] http://scholar.google.com/.

[5] http://www.dblp.org/search/.

The different inclusion and exclusion criteria must be clearly defined by the reviewer. This step allows to sort and keep only the articles relevant to the topic of interest (Casteleyn et al., 2014; Villarreal-Narvaez et al., 2020). Inclusion criteria on the form covers all publications in English that underwent a competitive peer-review process and where the full text is available. This includes all short and full research papers published in peer-reviewed journals, conferences, symposia, or workshops. All sources that did not go through a competitive peer-review process or are not pure research contributions (such as books, Ph.D. and master theses, patent descriptions, etc.) are thus excluded (Casteleyn et al., 2014). It is noteworthy that it is also possible to restrict the search depending on publication year.

Next, the publications that met the previous inclusion criteria are analyzed through the content criteria. To meet this criteria, publications must explicitly mention the subject (or a synonym) of the semantic research in the abstract or introduction. Indeed, it must be possible to infer from the abstract or introduction that the publication is related to the topic or to the broader research field (Casteleyn et al., 2014).

In the third phase (eligibility phase), the reviewer is free to add additional criteria required for the research. Sources that do not fulfill these criteria are considered ineligible and further excluded (Villarreal-Narvaez et al., 2020).

The last step of a Systematic Literature Review (inclusion phase) consists in verifying quantitative and qualitative aspects of the corpus of papers (Villarreal-Narvaez et al., 2020). The qualitative aspect aims at expressing the publications quality with verbal terms and the quantitative aspect use a meta-analysis to express the papers' values using charts and various types of graphs (Mengist et al., 2020).

In order to perform a quantitative analysis, the relevant data are firstly extracted and classified from the previously selected documents. This will allow additional knowledge and conclusions to be drawn. The general information of the articles include the year of publication, the name of the journal, the study type and the country where the study was conducted (Mengist et al., 2020). A meta-analysis can be performed on the data. It uses statistical techniques to integrate and summarize the results of included studies. Meta-analysis can help to generate more precise estimates on the topic under study (Mengist et al., 2020). Many systematic reviews contain meta-analyses, but not all of them (Liberati et al., 2009).

The qualitative analysis aims to answer to the formulated research question of the semantic request. It covers the explanation and narration of results, discussion, indication of the way forward for future research and the inference of a conclusion (Mengist et al., 2020).

2.1.3 Snowballing

In order to fill in a Systematic Literature Review and find additional references, the snowballing method can be applied. Snowballing refers to using the reference list of a paper or the citations to the paper to identify additional papers to analyze. There are two types of snowballing: forward and backward. For backward snowballing, the process consists of going through the reference list of each source found in the SLR and including or excluding the new references according to the previously defined criteria. Forward snowballing refers to identifying new papers based on those papers citing the paper being examined. The citations to the paper being examined are studied using Google Scholar. Each candidate citing the article is examined according to the same inclusion or exclusion criteria used previously (Wohlin, 2014).

2.1.4 Virtues of a Systematic Literature Review

Performing a SLR has 3 virtues. The first one is the descriptive virtue : SLR provides the reviewer with a method for describing the underlying characteristics of any study. (Villarreal-Narvaez et al., 2020). It is therefore possible to describe each reference in the same way. The scope of the studies can be described, but also their methods, results and metadata (e.g., year of publication, name of journal, types of data sources, study site...) (Mengist et al., 2020).

The second virtue of SLR is the comparative virtue. Indeed, thanks to SLR, it is possible to compare to each other the various references found. By clearly describing those studies along various dimensions, the reviewer will be able to assess their overlapping degree, similarities and key differences. For example, it is possible to compare the number of properties affected, the type of formula used by a given reference (validated, tested, empirical or theoretical formula). The quality of the studies can be assessed on the basis of their relevance and the studies can then be ranked in order of relevance. Thanks to this virtue, it is also possible to gather and compare studies on a specific topic (Villarreal-Narvaez et al., 2020).

The last virtue of Systematic Literature Reviews is the generative virtue. SLR makes it possible to identify areas of investigation for the topic of interest that were not covered or have been little explored by previous work. In order to perform a generative analysis of the references, the reviewer can find areas of the research which need a deeper study, a verification, a validation, or a consolidation. The analysis on the basis of this virtue is often the most difficult to carry out but it allows to generate new ideas, to fill in gaps and to imagine perspectives for the project (Villarreal-Narvaez et al., 2020).

2.2 Visual Design

2.2.1 Definition

Visual design informs us how design elements go together to create well-rounded and thoughtful visuals (Gordon, 2020). It aims to improve a product's aesthetic appeal and usability with suitable images, colors, fonts and other elements. The different elements of a website are carefully placed by designers to create interfaces that optimize user experience. The first impression is made in less than 50 ms. Therefore, the interface will create an emotion for the user almost instantaneously (Siang, 2020). The objective of visual designers is to evoke emotions which result in positive user experiences. Visual designers are therefore considered as the problem solvers of the design world. In addition to creating beautiful designs aimed at generating positive emotions, they know how to explain design concepts and the decisions behind their work (Morris, 2020).

Visual design can also be seen as a strategic process. Indeed, designers have to develop a strategy to lead the user's eye to an item's functionality and to make the aesthetics consistent. It is important for designers to compose and arrange website content around each page's purpose. They have to make sure that the content gives off the right visual cues because even the most subtle details can sometimes have a major impact on what users think and feel on the web page (Siang, 2020). The visual design strategy must be established in order to present the right things in the right way and in the right place.

However, the visual design has a few cons. It is purely visual and does not take into account the semantic query, the type of data/graphic/image or the text to be displayed. Moreover, some theoretical concepts can be opposed or complicated to apply on a screen for non-professional designers.

2.2.2 Elements of Visual Design

The elements are building blocks used and combined in order to generate visual designs. The different elements listed below are used by designers to improve websites aesthetics.

1. **Lines** are the most basic element of visual design but they have different properties. They connect two points and can be used to help define shapes, make divisions and create textures. They can be regular or irregular, curved or straight, thick or thin, geometric (i.e., look like they are drawn by a ruler or compass) or organic (i.e., look like they are drawn by hand) (Siang, 2020).

2. **Shapes** are enclosed/self-contained areas. A shape has two dimensions: length and width. To define the area, the visual designer uses lines and differences in brightness, color or texture (Siang, 2020; U.S. Dept. of Health and Human Services, 2006).

3. **Volume** is used to give a three-dimensional effect on two-dimensional screens. Volume has 3 dimensions: length, width and depth. It may be created by combining two or more shapes and can be further enhanced by different tones, textures and colors (Siang, 2020; U.S. Dept. of Health and Human Services, 2006).

4. **Negative space/White space** is an important part of the layout strategy. A negative space (or white space) use the blank area around a "positive" shape to create a figure/ground effect or calm the design overall. Incorporating spaces into a design helps reduce noise, increase readability and is also able to create an illusion (e.g., the WWF logo) (Siang, 2020; U.S. Dept. of Health and Human Services, 2006).

5. **Brightness** is the balance of light or dark colors that an element has. By manipulating brightness, we can influence how easy it is to distinguish details. Indeed, if something is too bright or too dark, the visual might be washed out or hard to see (Fessenden, 2019).

6. **Colors** are used to differentiate items, to create depth, to add emphasis, to convey emotions in and add variety, to separate distinct areas of a page and to help the organization of the information (U.S. Dept. of Health and Human Services, 2006).

7. **Texture** is the surface quality of an object. It refers to how a surface feels or is perceived to feel. By repeating an element (e.g., lines), a texture will be created and a pattern formed. Depending on how a texture is applied, it may be used strategically to attract or deter attention (U.S. Dept. of Health and Human Services, 2006).

8. **Typography** refers to which fonts are chosen, their size, alignment, color and spacing (U.S. Dept. of Health and Human Services, 2006).

2.2.3 Principles of Visual Design

While the elements of visual design describe the building blocks of a product's aesthetics, the principles of design tell us how these elements can be combine and should go together for the best results. In other words, the principles can be considered as guidelines to bring effectively the elements together in a way that makes sense. According to a study performed by the Nielsen Norman Group, there are 5 visual design principles impacting the user experience (Gordon, 2020).

1. **Scale** describes the relative sizes of the elements in a design (Figure 2.2). When this principle is used properly, the most important elements in a design are bigger than the ones that are less important. Indeed, when something is big, it's more likely to be noticed. In order to achieve a visually pleasing design, it is recommended not to use more than 3 different sizes (Gordon, 2020; Siang, 2020).

The scale principle is commonly used and almost every good visual design takes advantage of it. Using this principle allows to create variety within the layout but also to establish a visual hierarchy on the website (Gordon, 2020).

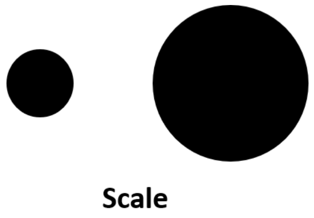
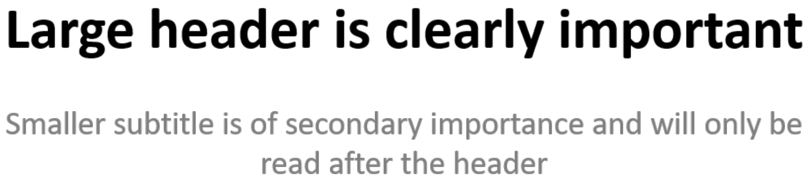


Figure 2.2: Example of scale principle (adapted from Siang, 2020)

2. **Visual hierarchy** shows the difference in importance of the elements in a design (Siang, 2020). Designers can create hierarchies through different font sizes, colors, placement on the page, spacing and a variety of other signals. Usually, items at the top are seen as the most important (U.S. Dept. of Health and Human Services, 2006). The hierarchy will allow the user to better understand the content and the layout of the page consulted. The advice provided by the experts of the Nielsen Norman Group suggests using 3 types of scale (small, medium and large) as well as playing with colors (bright colors for important items and muted colors for less important ones) in order to create a hierarchy (Gordon, 2020). Figure 2.3 shows an example of a visual hierarchy using a difference in size between the title and subtitle and a black and grey colour respectively.


Large header is clearly important

Smaller subtitle is of secondary importance and will only be read after the header

Figure 2.3: Example of visual hierarchy using different scales and colors (adapted from Siang, 2020)

3. **Balance** describes the extent to which the elements of a layout are evenly distributed. In other words, the balance principle refers to a satisfying arrangement or proportion of design elements. An interface is said to be balanced when there is an equally distributed amount of visual signal on both sides of an imaginary axis going through the middle of the screen. This axis can either be vertical or horizontal. To create a balance, the number of elements present on the interface is important but it is also necessary to pay particular attention to the surface occupied by these elements as well as their visual weight. The visual weight of an element is linked to how prominent it appears compared to everything surrounding it (Gordon, 2020).

A balanced interface appears stable while an imbalanced interface can be seen as unsustainable and unnatural (Figure 2.4). It is noteworthy that there are 2 types of balance : symmetrical balance and asymmetrical balance. Symmetrical balance concerns similar objects evenly distributed on a page or a screen, whereas asymmetrical balance concerns dissimilar objects but of equal visual weight.

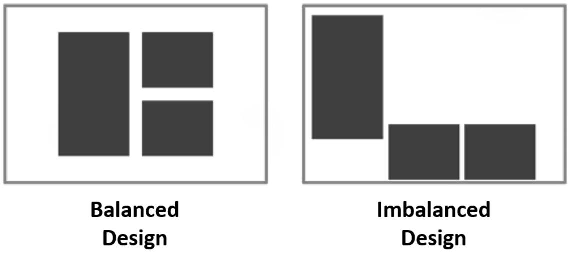


Figure 2.4: Difference between a balanced and an imbalanced design (adapted from Siang, 2020)

4. **Contrast** refers to the juxtaposition of visually dissimilar elements in order to convey the fact that these elements are different. It provides the eye with a noticeable difference (e.g., size or color) between two objects or two sets of objects. Designers often use the red color to make certain elements stand out or to provoke an emotion to the user. For example, on the iOS operating system, the red color especially signify deleting while green is often used for positive actions such as "Go" or "Accept". Designers can also play on the contrast between a text and its background in order to deemphasize less important text. Nevertheless, this can be risky because reducing the contrast of the text also reduces its legibility and could therefore make the content inaccessible (Gordon, 2020).

5. **Gestalt principles** explain how humans simplify and organize complex images that consist of many elements, by subconsciously arranging the parts into an organized system that creates a whole, rather than interpreting them as a series of disparate elements. The human eye and brain perceive an unified shape in a different way to the way they perceive the individual parts of such shapes. If the design elements are arranged properly, the Gestalt of the overall design will be very clear (Gordon, 2020; Siang, 2020). Gestalt is the reason why we can see a square, a circle and a triangle even though the lines are not complete in the figure 2.5 shown below.

Figure 2.5: Application of the Gestalt principles for the visualization of a square, a circle and a triangle instead of a set of dotted lines (Siang, 2020)

Chapter 3

Research Method: Systematic Literature Review

In this chapter, we will explain how our Systematic Literature Review dedicated to visual design metrics for Graphical User Interfaces has been performed. The final results of the Systematic Literature Review will be presented in a PRISMA diagram and in a summary table.

3.1 Identification Phase

The overall goal of this Systematic Literature Review, and more broadly of this thesis, is to provide a consolidated overview of research in the field of visual design metrics for GUIs. Particular interest will be given to references that describe visual design metrics using a mathematical formula. The SLR will allow to compare the references among themselves and also to study the relative importance of each of the formulas found in the papers. In a second phase, the metadata such as the publication venues of the references, the number of authors and the evolution over time of the subject of interest will be studied.

Initial Search

In order to find papers potentially relevant to this topic of investigation, the following semantic query has been introduced in different digital libraries:

Q = ("Visual" AND "Metric" AND "User" AND "Interface")

The semantic query was ran on both single-publisher libraries (e.g. ACM Digital Library) and multi-publisher engines (e.g. Google Scholar). The 6 digital libraries selected to carry out the search process of this SLR are detailed below, with their acronym and URL according to the format: Name [acronym]: URL.

- ACM Digital Library [ACM]: http://portal.acm.org/.
- IEEE Xplore [IEEE]: http://ieeexplore.ieee.org/.
- Science Direct [SD]: http://www.sciencedirect.com/.
- SpringerLink [SPL]: https://link.springer.com/.
- Google Scholar [GS]: http://scholar.google.com/.
- DBLP Complete Search [DBLP]: http://www.dblp.org/search/.

The selected digital libraries use their own syntax and have their own advanced search options, therefore the chosen search string has been adapted for each search engine as shown in Table 3.1.

Digital library	**Adapted query string**	**Advanced search options**
ACM	Visual AND Metric AND User AND Interface	Search items from: ACM Guide to Computing Literature Search within: Anywhere Publication date: All dates
IEEE	Visual AND Metric AND User AND Interface	Each term was searched in: Full text & metadata
SD	Visual AND Metric AND User AND Interface	Find articles with the terms of the query
SPL	Visual AND Metric AND User AND Interface	Find resources with all of the words of the query
GS	Visual AND Metric AND User AND Interface	Find articles containing all of the words of the query
DBLP	Visual Metric User Interface	Combined DBLP search

Table 3.1: Syntax and advanced search option used for each selected digital library (adapted from Ordoñez et al., n.d.)

All searches were performed between 24^{th} March 2020 and 20^{th} April 2020. The search results were sorted in decreasing order of relevance. For each search, the first 150 results were inspected to form the initial set. Supposing that the accuracy of each search engine ranking algorithm is effective, we can assume that considering the first 150 results yields a relevant and representative result set (Casteleyn et al., 2014). However, it should be noted that for the DBLP digital library, only 2 results were found which have both been included for further screening. The total number of references found after the identification phase is therefore equal to 752 (150 papers on 5 digital librairies + 2 from DBLP).

3.2 Screening Phase

The initial set of publications contains irrelevant results that must be removed. Each paper was evaluated with respect to its relevance to visual design using criteria related to form and content, as follows:

Form

Only the papers written in English that underwent a peer-review process and for which the full text was available were retained. Therefore, research papers published in peer-reviewed such as journals, conferences, symposiums and workshops were included, while books, Ph.D. theses, master theses and patents were excluded.

Content

Only the papers that explicitly introduced a metric for visual design of user interface or discussed about aesthetics and usability of GUI were retained. Publications linked to the visual aspect of the interface or to the user interface layout were also included for screening. All off-topic articles have therefore been excluded.

After excluding articles according to the form and content criteria, a total number of 670 irrelevant article were excluded, leaving $752 - 670 = 82$ papers. These 82 papers will therefore be examined in the third phase of the Systematic Literature Review.

3.3 Eligibility Phase

Among the screened papers, only those that had a formula for a metric related to visual design were retained. For this purpose, the full text of the 82 articles was reviewed in search of a formula or an algorithm. After the addition of this eligibility criterion, 59 publications that did not have a formula related to a visual design metric were therefore excluded, leaving to 23 the number of remaining publications.

At the end of this phase, duplicates coming from similar results on different digital libraries were eliminated. The number of duplicates found is 9. The final number of relevant publications identified with the SLR method is therefore equal to 14. The detailed process in which these 14 papers were found is described by a PRISMA diagram (Liberati et al., 2009) as shown in the Figure 3.1.

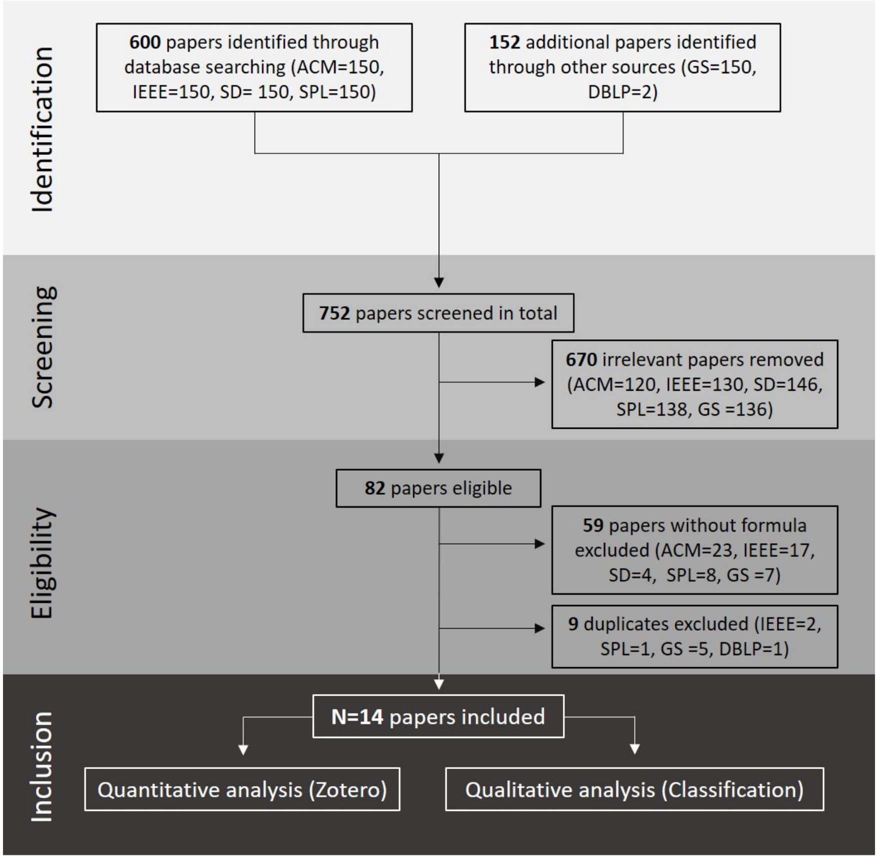


Figure 3.1: The four-phase flow PRISMA diagram of our SLR examination (based on Villarreal-Narvaez et al., 2020)

From these 14 articles, a selective backward snowballing will be carried out in order to potentially discover additional references that meet all the previously defined criteria.

3.4 Backward Snowballing

In order to complete our set of articles, a backward snowballing was carried out. Indeed, since the semantic search was Q = ("Visual" AND "Metric" AND "User" AND "Interface"), other relevant references using a synonym for visual design like aesthetics might not have been presented by the search engine algorithms. For these reasons, a selective backward snowballing was carried out by screening through the reference list of each of the 14 previously selected articles. This backward snowballing is selective because only the sources containing the word "aesthetics" have been screened.

It is noteworthy that to be rigorous, it would have been necessary to screen all the sources of our set of articles as well as the realization of a forward snowballing but this would have required a lot of time and an analysis by several people. Therefore, in this master thesis, only a selective backward snowballing will be performed.

The search for new sources on the basis of the word "aesthetics" identified 51 new articles to be screened. Among them, 36 papers were not eligible according to the conditions defined in section 3.2 and 3.3. In addition to this, 4 publications were also excluded because they were duplicates of our original articles set. Therefore, this selective backward snowballing allowed us to identify 11 new references. This step will be identified as the 1^{st} iteration of our snowballing process.

A second iteration of the selective backward snowballing has been carried out on newly discovered references. For each new article discovered, the reference list was studied in order to screen new publications containing the word "Aesthetics". With this second iteration, a total of 135 papers were screened, 51 were not relevant according to the criteria of content or form, 52 did not contain a formula, 24 duplicates were found as well as 8 new articles eligible for the study. A third iteration was done but no new papers were found. That marked the end of our process of backward snowballing.

Finally, at the end of this Systematic Literature Review, $14 + 11 + 8 = 33$ references introducing a formula related to visual design of Graphical User Interfaces were identified and downloaded for further usage.

3.5 Inclusion Phase

3.5.1 Quantitative Analysis

The quantitative aspect of our corpus of papers was verified by using Zotero[6] and by generating summary statistics. Zotero is a multi-platform bibliographic management software used to manage a set of documents.

Firstly, a collection of papers were created on Zotero. Each reference has been compiled in the online library and the full-text PDF of each source has been added as an attachment. Then, the metadata from each source has been completed in order to analyze them statistically. The 33 references that have been included are available online in the Zotero library by following this link:

https://www.zotero.org/groups/2463417/visual_design_metrics/library

---

[6] https://www.zotero.org/

3.5.2 Qualitative analysis

The quality of each source will be discussed with more details in the next chapter. Different qualitative aspects will be discussed, such as the number of authors per article or the degree of validation of the different formulas presented in the papers.

3.6 Summary of the Systematic Literature Review

The evolution of the number of references according to the different steps of the SLR is presented in the summary table below (Table 3.2).

Source	**Query**	**Screened**	**Excluded**	**Duplicates**	**Included**
1. ACM	17,406	150	143	0	7
2. IEEE	23,504	150	147	2	1
3. SD	14,119	150	150	0	0
4. SPL	28,197	150	146	1	3
5. GS	248 000	150	143	5	2
6. DBLP	2	2	0	1	1
Total before snowballing	331230	752	729	9	**14**
7. Backward Snowballing 1^{st} iteration	-	51	36	4	11
8. Backward Snowballing 2^{nd} iteration	-	135	103	24	8
Total after snowballing	-	938	868	37	**33**

Table 3.2: Summary table of the number of references analyzed at the different steps of the SLR by digital libraries and by the snowballing process (adapted from Villarreal-Narvaez et al., 2020)

Chapter 4

Results and discussion

In this chapter, the 3 virtues of the Systematic Literature Review will allow us to study, characterize, compare and find perspectives for the different visual design metrics for Graphical User Interfaces found in the articles. The quality of the sources will also be discussed through a meta-analysis of the papers.

4.1 Visual Design Metrics

For each metric, the number of citations per metric (NoC) will be calculated. The number of citations will allow to rank the metrics by order of popularity in the literature. Each formula of a metric will also be accompanied by a strength of evidence level (SoE) depending on the degree of verification of the formula.

The SoE is defined below:

- If only a theoretical formula is given, SoE=1.
- If the formula is verified by at least one study with an example on a website, SoE=2.
- If the formula has been validated empirically with a correlation below 50%, SoE=3.
- If the formula has been validated empirically with a correlation between 50% and 90%, SoE=4.
- If the formula has been validated empirically with a correlation above 90%, SoE=5.

4.1.1 Balance (NoC=21)

Balance can be defined as the distribution of optical weight (the perception that some objects appear heavier than others) in a picture. In screen design, balance is achieved by providing an equal weight of screen elements, left and right, top and bottom (Ngo, Teo, et al., 2000b). In the literature, there are several balance formulas which are listed and compared here under.

4.1.1.1 First formula to measure balance (SoE=5)

A first formula of balance (BM) is given below (equation 4.1)(Altaboli and Lin, 2011; Chettaoui and Bouhlel, 2017; Kar and Zain, 2007; Ngo and Byrne, 2001a, 2001b; Ngo, Teo, et al., 2000a, 2000b, 2003; Purchase et al., 2011; Zain et al., 2011; Zen, 2013; Zen and Vanderdonckt, 2014):

$$BM = 1 - \frac{|BM_{vertical}| + |BM_{horizontal}|}{2} \in [0, 1] \tag{4.1}$$

$BM_{vertical}$ and $BM_{horizontal}$ are respectively the vertical and horizontal balances. These terms and their formulas are defined in section 6.1.1. These terms depend on the visual weight of the interface elements (Ngo, Teo, et al., 2000b). The visual weight depends on the area, color and shape of the various objects on the interface. The distribution of the visual weight in the quadrants will influence the balance of the interface.

This metric can be really useful because the balance is an essential element to focus on when visually designing interfaces. It is therefore very interesting to be able to measure it. However, it is noteworthy that balance does not necessarily induce symmetry.

4.1.1.2 Second formula to measure balance (SoE=5)

A second formula exists in order to measure the balance of Graphical User Interfaces. The formula is shown below (equation 4.2) (Ngo and Byrne, 1998; Ngo, Samsudin, et al., 2000):

$$BM = (w_L - w_R, w_T - w_B) \tag{4.2}$$

Where BM is the measure of balance, w_L is the weight on the left hand side (LHS), w_R is the weight on the right hand side (RHS), w_T is the weight on the top, and w_B is the weight on the bottom. The detailed calculations of the weights are shown in section 6.1.2. Overall, the weights depend on the area of the interface objects and the distance between their vertical/horizontal central line and the vertical/horizontal axis respectively. The balance is reached when BM = (0,0).

One advantage of this formula is the ability to compare the balance on each sides of the interface. Nevertheless, this formula has the disadvantage of only taking into account the area of the different objects. The visual weight can also depend on the color and the shape of the interface objects.

4.1.1.3 Third formula to measure balance (SoE=2)

To apply this formula, the first step is to find the center of the screen. It is located half way between the left edge of the left-most widget and the right edge of the right-most widget. The weight of either side is computed by multiplying the number of pixels used by their distance from the center. The third formula of balance is proposed by Sears, 1995 (equation 4.3):

$$Balance = 200 \times \frac{\text{weight of side one}}{\text{weight of side one} + \text{weight of side two}} \tag{4.3}$$

The interface is completely balanced if the score is equal to 100. This formula has not been empirically validated but only applied to different websites. Its strength of evidence level (SoE) is thus equal to 2.

4.1.1.4 Fourth formula to measure balance (SoE=4)

This formula is applied in a microscopic approach by comparing the balance of an image pixel by pixel. Black pixels are given a mass of one and white pixels given a mass of zero. The equation of balance (b) is given below (equation 4.4) (Bauerly and Liu, 2006):

$$b = \left(\frac{x_b}{w}, \frac{y_b}{h}\right) \tag{4.4}$$

Where b is given as a set of normalized coordinates between zero and one, w is the image width in pixels, (x_b, y_b) is the center of balance and h is the image height in pixels (cfr section 6.1.3 for the calculation of (x_b, y_b)). An image with a perfectly centered balance will have a b value equal to $(0.5, 0.5)$. This formula has been validated empirically with a correlation of 0.79.

4.1.1.5 Fifth formula to measure balance (SoE=5)

This fifth formula of balance is applied to a specific interface: text-overlaid images. Text-overlaid image consists of a large-size background image with a small number of texts overlaid on it. This formula is derived from the fourth formula of balance (equation 4.4) but takes into account the color attributes of the pixels via the formula presented in section 4.1.16.1. It is decomposed in horizontal (B_H) and vertical balance (B_V) as shown below (equation 4.5 and 4.6) (Lai et al., 2010):

$$B_H = 1 - \left|2 \times \frac{x_b}{w} - 1\right| \tag{4.5}$$

$$B_V = 1 - \left| 2 \times \frac{y_b}{h} - 1 \right| \tag{4.6}$$

B_H and B_V range between 0 and 1. The closer to 1 the value of B_H (B_V), the better the horizontal (vertical) visual balance obtained. w is the image width in pixels, (x_b, y_b) is the center of balance (depending on color) and h is the image height in pixels (cfr section 6.1.4 for the calculation of (x_b, y_b)).

4.1.1.6 Sixth formula to measure balance (SoE=4)

In this sixth formula, the number and size of objects for each quarter of the screen is calculate to obtain the Total Balance Complexity (TBC) value. The Total Balance Complexity (TBC) formula is given in equation 4.7 (Alemerien and Magel, 2014):

$$TBC = 1 - (0.5 * BQn + 0.5 * BQs) \tag{4.7}$$

Where BQn and BQs are the number and size of objects for each quarter of the screen. These terms are defined below:

$$BQn = \frac{\sum_{k=1}^{6} \frac{BQn_i}{BQn_j}}{6}$$

$$BQs = \frac{\sum_{k=1}^{6} \frac{BQs_i}{BQs_j}}{6}$$

BQn_i and BQn_j represent the number of objects in the ith and jth quarters and BQs_i and BQs_j represent the sum of sizes of objects in the ith and jth quarters. These last 4 terms vary between 0 and 1. The range of BQn is [0,1] where 0 means unbalanced and 1 means fully balanced in terms of number of objects and the range of BQs is [0,1] where 0 means unbalanced and 1 means fully balanced in terms of object size. This formula has been empirically validated with a correlation of 0.825.

4.1.1.7 Seventh formula to measure balance (SoE=5)

This balance formula is specifically applied to Android applications (Riegler and Holzmann, 2018b). The balance metric calculates the margins between GUI elements and their neighboring elements or borders. For each GUI element, the distances to the four horizontal and vertical neighbors are considered. The neighbor is defined as the closest GUI element or group border in the respective direction (i.e. left, right, top and bottom).

In order to calculate the balance value (B_s), we use the mean horizontal ($\overline{balh_s}$) and vertical margin ($\overline{balv_s}$) (cfr section 6.1.5 for their definitions) and weight them equally (equation 4.8) (Riegler and Holzmann, 2018b):

$$B_s = \frac{\overline{balh_s} + \overline{balv_s}}{2} \tag{4.8}$$

4.1.1.8 Eighth formula to measure balance (SoE=1)

The eighth formula to measure balance is proposed by Streveler and Wasserman (Streveler and Wasserman, 1984). It is computed as a comparison between the GUI objects center of mass and the screen center. Their paper is not available online so only the theoretical formula is shown below (equation 4.9) (Zen and Vanderdonckt, 2014):

$$BM = 1 - \frac{2 * \sqrt{(X_{CS} - X_{CM})^2 + (Y_{CS} - Y_{CM})^2}}{\sqrt{W^2 + H^2}} \tag{4.9}$$

Where W and H are the width and height of the interface, (X_{CS}, Y_{CS}) are the coordinates of the center of the screen and (X_{CM}, Y_{CM}) are the coordinates of the center of mass of the GUI objects.

4.1.1.9 Ninth formula to measure balance (SoE=2)

The balance metric (Ba) is inspired by the concept of an equilibrium between two areas separated by a vertical line. Balance is achieved by giving similar visual weight to both sides. Since the quantification of this weight is subjective, a heuristic has been made that considers weight as the accumulated height of the components on each side. The formula is computed as follows (equation 4.10) (López et al., 2013):

$$Ba = \frac{n_{hac}}{n_c}\left(1 - \frac{|h_l - h_r|}{h_l + h_r}\right) \tag{4.10}$$

Where n_{hac} is the number of components horizontally aligned with other components, n_c is the number of components in the interface, h_l is the accumulated height of left side components and h_r is the accumulated height of right side components.

4.1.2 Equilibrium (NoC=10)

Equilibrium on a screen is accomplished through centering the layout itself. A layout is in equilibrium when its center coincides with the center of the frame. (Ngo, Samsudin, et al., 2000; Ngo, Teo, et al., 2000b).

4.1.2.1 First formula to measure equilibrium (SoE=5)

Equilibrium (EM) is computed as the difference between the center of mass of the displayed elements and the physical center of the screen and is given by the equation 4.11 (Kar and Zain, 2007; Ngo and Byrne, 2001a, 2001b; Ngo, Teo, et al., 2000a, 2000b, 2003; Zain et al., 2011; Zheng et al., 2009):

$$EM = 1 - \frac{|EM_x| + |EM_y|}{2} \in [0, 1] \tag{4.11}$$

Where EM_x and EM_y are the equilibrium components along the x-axis and y-axis respectively (cfr section 6.2.1 for the details about the calculation of these terms).

The equilibrium components depend on the coordinates of the centers of the interface objects (x_i, y_i), the center of the frame (x_c, y_c), the surface area of the objects (a_i), the number of object on the frame (n) and the width and height of the frame (b_{frame} and h_{frame} respectively).

4.1.2.2 Second formula to measure equilibrium (SoE=5)

The second formula to measure equilibrium (EM) is given by the equation 4.12 (Ngo and Byrne, 1998; Ngo, Samsudin, et al., 2000):

$$\begin{aligned} EM &= (x_c, y_c) - (x_0, y_0) \\ &or \\ EM &= (x_c - y_0) - (y_c - y_0) \end{aligned} \tag{4.12}$$

Where (x_0, y_0) is the center of the layout, and (x_c, y_c) is the center of the frame. The layout is composed of a set of objects of areas a_1, a_2, a_3... at points $(x1, y1)$, $(x2, y2)$, $(x3, y3)$... in the xy plane (cfr section 6.2.2 for the calculation of these terms). By using this formula, the equilibrium can be reached when EM = (0,0). This formula can therefore estimate precisely if there is a vertical or horizontal disequilibrium.

4.1.3 Symmetry (NoC=17)

Symmetry can be represented as an axial duplication: when an element of screen design is present on one side of the center line, it is exactly replicated on the other side. The center line can be either vertical, horizontal or radial. Radial symmetry consists of equivalent elements balanced about two or more axes that intersect at a central point (Figure 4.1).

Horizontal symmetry **Vertical symmetry** **Radial symmetry**

Figure 4.1: Example of axes of symmetry (Walleriusz, 2015)

4.1.3.1 First formula to measure symmetry (SoE=5)

Symmetry (SYM) is the extent to which the screen is symmetrical in three directions: vertical, horizontal, and diagonal and is given by the equation 4.13 (Kar and Zain, 2007; Ngo and Byrne, 2001b; Ngo, Teo, et al., 2000a, 2000b, 2003; Salimun et al., 2010; Soui et al., 2019; Zain et al., 2011; Zheng et al., 2009):

$$\mathrm{SYM} = 1 - \frac{|\mathrm{SYM}_{vertical}| + |\mathrm{SYM}_{horizontal}| + |\mathrm{SYM}_{radial}|}{3} \in [0,1] \tag{4.13}$$

$\mathrm{SYM}_{vertical}$, $\mathrm{SYM}_{horizontal}$ and SYM_{radial} are, respectively, the vertical, horizontal, and radial symmetries (cfr section 6.3.1 for the definition of the formulas of these different variables). Overall, the symmetry depends on the coordinates on the x and y-axis of the center of the different objects in relation to the center of symmetry, the size of the objects (their width and height) and the total number of objects present in each quadrant of the interface.

4.1.3.2 Second formula to measure symmetry (SoE=5)

There is another formula for measuring symmetry. The formula of symmetry (SM) is shown below (equation 4.14) (Ngo and Byrne, 1998; Ngo, Samsudin, et al., 2000):

$$SM = \{[g_{UL} - g_{LL}], [g_{UR} - g_{LR}], [g_{UL} - g_{UR}], [g_{LL} - g_{LR}], [g_{UL} - g_{LR}], [g_{UR} - g_{LL}]\} \tag{4.14}$$

Where g_j is the weight of the quadrant j. The way of measuring the weight of a layout is explained in section 6.3.2. The vertical symmetry is obtained if: $g_{UL} - g_{UR} =$ (0,0,0,0) and $g_{LL} - g_{LR} =$ (0,0,0,0). The horizontal symmetry is obtained if: $g_{UL} - g_{LL} =$ (0,0,0,0) and $g_{UR} - g_{LR} =$ (0,0,0,0). Finally, the radial symmetry is obtained if: $g_{UL} - g_{LR} =$ (0,0,0,0) and $g_{UR} - g_{LL} =$ (0,0,0,0) (Ngo, Samsudin, et al., 2000).

The advantage of this metric is the possibility to obtain each type of symmetry individually. Thus, vertical symmetry can be obtained without necessarily having horizontal or radial symmetry for the interface.

4.1.3.3 Third formula to measure symmetry (SoE=4)

A third study defined symmetry as an analysis of the similarity of pixels on opposite sides of an axis of reflection. The study proposes an algorithm that takes a microscopic approach and compares each half of an image pixel by pixel (Bauerly and Liu, 2006). The pixels that are close to the axis of reflection have a higher influence on the overall impression of symmetry than the ones which are further away. Equation 4.15 below gives the formula of symmetry (s) (Bauerly and Liu, 2005, 2006, 2008):

$$s = \frac{2}{3mn} \sum_{i=1}^{m} \sum_{i=1}^{n} X_{ij} \left(1 + \frac{j-1}{n-1} \right) \tag{4.15}$$

Where m is the pixel length of the image dimension that is parallel to the axis of reflection, n is the number of comparisons required in each row or column of pixels, X_{ij} is a binary variable which is equal to one when the pixel pairs are the same and zero when they are opposite, i is the pixel row and j is the pixel column.

This formula has been validated empirically with a correlation of 0.78. The ratings of the subjects of the study corresponded with the s values of the images presented in the study.

4.1.3.4 Fourth formula to measure symmetry (SoE=5)

This fourth formula of symmetry is applied to an interface composed of text-overlaid images. This formula is derived from the third formula of symmetry (equation 4.15) but takes into account the color attributes of the pixels via the formula presented in section 4.1.16.1. Two pixel blocks with a small HSV color difference are considered to be more symmetrical than two with a large color difference. Therefore, when comparing two corresponding pixel blocks for measuring symmetry, the difference value of symmetry is considered to be proportional to the HSV (hue, saturation, and value) color difference between the two pixel blocks. The computation models for the horizontal symmetry (S_H), vertical symmetry (S_V), and radial symmetry (S_R) is given by equation 4.16, 4.17 and 4.18 respectively (Lai et al., 2010):

$$S_H = \frac{2}{3w(h/2)} \sum_{i=1}^{w} \sum_{j=1}^{h/2} (1 - \Delta C_{ij-i'j'}) \left(1 + \frac{(h/2)-j}{(h/2)-1} \right) \tag{4.16}$$

$$S_V = \frac{2}{3w(h/2)} \sum_{i=1}^{h} \sum_{j=1}^{w/2} (1 - \Delta C_{ij-i'j'}) \left(1 + \frac{(w/2) - j}{(w/2) - 1}\right) \tag{4.17}$$

$$\begin{aligned} S_R = \frac{8}{6wh} \sum_{i=1}^{h/2} \sum_{j=1}^{w/2} (1 - \Delta C_{ij-i'j'}) \left(1 + \left(\frac{(w/2) - j}{(w/2) - 1} + \frac{(h/2) - i}{(h/2) - 1}\right)/2\right) \\ + \sum_{i=(h/2)+1}^{h} \sum_{j=1}^{w/2} (1 - \Delta C_{ij-i'j'}) \left(1 + \left(\frac{(w/2) - j}{(w/2) - 1} + \frac{i - 1 - (h/2)}{(h/2) - 1}\right)/2\right) \end{aligned} \tag{4.18}$$

Where w and h are respectively the width and height of the image in pixel blocks, (i', j') is the coordinates of a pixel on the opposite side of a given axis of reflection for a pixel situated in (i, j) and ΔC is defined in the section 4.1.16.1. These formula have been validated with a correlation of 0.976 for S_V, 0.849 for S_H and 0.833 for S_R.

4.1.3.5 Fifth formula to measure symmetry (SoE=3)

Miniukovich and De Angeli developed a fifth formula to measure vertical symmetry. This formula is applied in a microscopic approach (at pixel scale). Firstly, screenshots are taken on a website and an algorithm detects the image contours. Then, the algorithm keeps only the vertical component of detected contours. The number of contour pixels is reduced by taking a contour pixel and dismissing others in the 3-pixel radius and taken as key points. Further, for each key point, the algorithm look for a match in the 4-pixel radius area across the central axis. Then, the ratio of matches (k_{sym}) to all key points (k_{all}) is taken and normalized by the probability of key point match due to chance. The probability of incidental match depends on the number of all key points, the size of match-search area (S_a, which is a 4-pixel radius area) and the size of the screenshots (S_s). The normalized ratio mentioned above (Sym_{norm}) is in fact the metric of symmetry (equation 4.19) (Miniukovich and De Angeli, 2014):

$$Sym_{norm} = \frac{k_{sym}}{k_{all}} * \left(\frac{(k_{all} - 1) * S_a}{S_s}\right)^{-1} \tag{4.19}$$

This symmetry metric has been validated empirically but is moderately correlated with complexity ratings of participants and did not correlate with aesthetics ratings.

4.1.4 Sequence (NoC=11)

In visual design, the sequence refers to the arrangement of objects in a layout in a way that facilitates the movement of the eye through the displayed information. Normally the eye starts from the upper left and moves back and forth across the display to the lower right. Moreover, the eye tends to move from big objects to small objects, from bright colors to subdued colors, from color to black and white, and from irregular shapes to regular shapes (Figure 4.2) (Ngo, Teo, et al., 2000b).

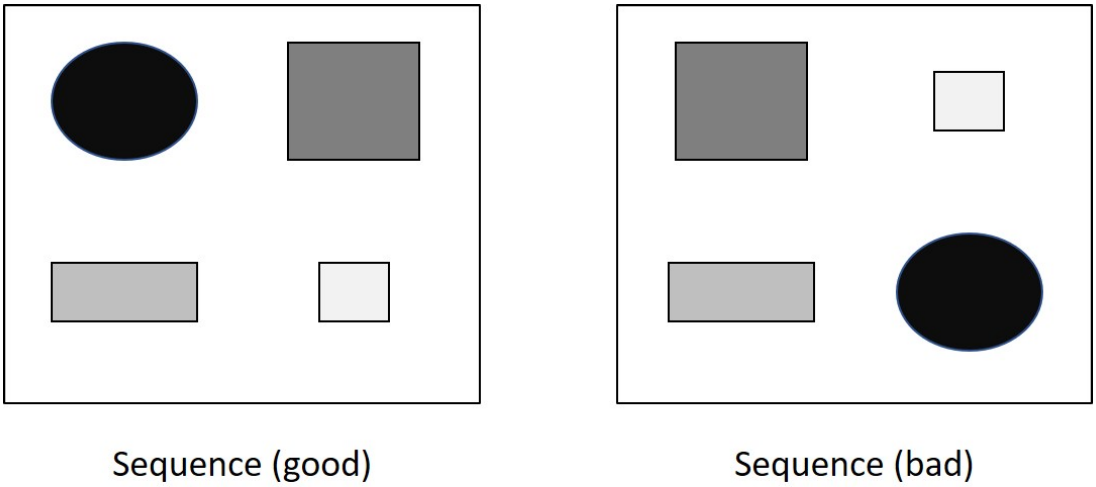


Figure 4.2: Comparison between a good and a bad sequence (Ngo, Teo, et al., 2000b)

4.1.4.1 First formula to measure sequence (SoE=5)

Sequence (SQM) is a measure of how information in a display is ordered according to a reading pattern that is common in Western cultures and is given by the equation 4.20 (Altaboli and Lin, 2011; Kar and Zain, 2007; Ngo and Byrne, 2001b; Ngo, Teo, et al., 2000a, 2000b, 2003; Salimun et al., 2010; Zain et al., 2011):

$$SQM = 1 - \frac{\sum_{k=A,C,S} \sum_{j=UL,UR,LL,LR} |q_j - v_j^k|}{8} \in [0,1] \tag{4.20}$$

with (q_{UL},q_{UR}, q_{LL}, q_{LR}) $= (4, 3, 2, 1)$: the weighting of the quadrant as it should be to have the best sequence

$$v_j^k = \begin{cases} 4 & \text{if } w_j^k \text{ is the largest in } w^k \\ 3 & \text{if } w_j^k \text{ is the } 2^{nd} \text{ largest in } w^k \\ 2 & \text{if } w_j^k \text{ is the } 3^{rd} \text{ largest in } w^k \\ 1 & \text{if } w_j^k \text{ is the smallest in } w^k \end{cases}$$

$$j = UL, UR, LL, LR \qquad k = A, C, S$$

Where UL, UR, LL, and LR stand for upper-left, upper-right, lower-left, and lower-right (the 4 quadrants of the interface). A, C and S are respectively the size, the color, and the shape of the objects. v_j^k is the effective weighting of each quadrant. Cfr section 6.4.1 for the details of the w_j^k and w^k calculations.

This metric is therefore useful to determine the sequence of an interface. A good sequenced interface will have a SQM close to 1.

4.1.4.2 Second formula to measure sequence (SoE=5)

Another study by Ngo, Samsudin, et al., 2000 has defined sequence as an arrangement of objects in a layout in a way that facilitates the eye movement. In order to improve the sequence, this study suggests to manipulate the user's attention thanks to eye attractors and visual paths (cfr section 6.4.2 for more information about attractiveness and visual paths). The formula of sequence depends on the visual paths and is given by the equation 4.21 (Ngo and Byrne, 1998; Ngo, Samsudin, et al., 2000):

$$SQM = (p, d) \tag{4.21}$$

Cfr section 6.4.2 for the definitions of p (the existence of a path) and d (the length of the path).

4.1.4.3 Third formula to measure sequence (SoE=2)

This third formula to measure sequence is used in the context of mobile applications. The proposed formula to calculate this metric is given below (equation 4.22) (Soui et al., 2019):

$$SQM = 1 - \frac{\sum_j (q_j \sum_{i=1}^{n} N_{i,j})}{4n} \in [0, 1] \qquad j = UL, UR, LL, LR \tag{4.22}$$

Where $N_{i,j}$ is the number of object on the quadrant j and n is the total number of objects on the mobile user interface. Each quadrant is given a weighting with $q_{UL} = 4$, $q_{UR} = 3$, $q_{LL} = 2$, $q_{LR} = 1$.

4.1.5 Cohesion (NoC=7)

4.1.5.1 First formula to measure cohesion (SoE=5)

The cohesion is linked to the aspect ratio. The aspect ratio refers to the relationship of width to height. In screen design, similar aspect ratios promote cohesion. Changing the aspect ratio of a visual field may affect eye movement patterns sufficiently to account for some of the performance differences. The aspect ratio of a visual field should stay the same during the scanning of a display (Ngo, Teo, et al., 2000b).

Cohesion is a measure of how cohesive the screen is and is given by the following formula (equation 4.23) (Ngo and Byrne, 2001b; Ngo, Teo, et al., 2000a, 2000b, 2003; Salimun et al., 2010):

$$CM = \frac{|\mathrm{CM}_{fl}| + |\mathrm{CM}_{lo}|}{2} \in [0, 1] \tag{4.23}$$

CM_{fl} is a relative measure of the ratios of the layout (l) and the frame of the screen (f). CM_{lo} is a relative measure of the ratios of the objects (o) and layout (l). The detailed calculations of these terms is explained in section 6.5.1.

The main variables that influence cohesion are the widths and heights of the interface objects, the layout and the frame as well as the number of objects on the frame.

4.1.5.2 Second formula to measure cohesion (SoE=4)

This second formula measure the screen layout cohesion (SLC). The screen layout cohesion is composed of 4 metrics of cohesion based on 17 aesthetic and structural attributes (GLC, UGLC, WLC and SR). The 4 metrics are presented and calculated more in details in section 6.5.2.

The formula of SLC is shown below (equation 4.24) (Alemerien and Magel, 2015):

$$SLC = \frac{GLC * W1 + UGLC * W2 + WLC + SR * W1}{2 + W1} \tag{4.24}$$

Where GLC is the Group Layout Cohesion, $UGLC$ is the UnGroup Layout Cohesion, WLC is the Widget Layout Cohesion and SR is the Semantic Relatedness (SR). The values of GLC and $UGLC$ are weighted based on the number of grouped and ungrouped widgets to the total number of widgets on a given user interface ($W1$, $W2$), respectively. In addition, SR is weighted by $W1$ because SR represents the semantic relatedness of widget groups on a given user interface. The SLC metric has been validated empirically with a correlation of 0.897.

4.1.5.3 Third formula to measure cohesion (SoE=1)

In a study performed by Kokol, Rozman and Venuti, cohesion was measured in terms of relationship between data on one screen. If all data on the screen belongs to the same entity, the cohesion is high. If data on one screen are attributes of two or more related entities, this is medium cohesion and if data do not relate directly, this is called low cohesion. This is presented in the formula below (equation 4.25) (Kokol et al., 1995):

$$COH_i = \frac{(W_e N_e + W_{Re} N_{Re} + W_u N_u)}{(N_e + N_{Re} + N_u)} \tag{4.25}$$

where COH_i is the cohesion for one entry screen, N_e is the number of data belonging to the same entity, N_{Re} is the number of data (together) which belong to related entities, N_u is the data which belongs to unrelated entities. W are the weights (arbitrary set by the researchers). Finally the value of cohesion of an entire interface is given as:

$$COH = \sum (COH)_i$$

4.1.6 Unity (NoC=9)

According to Ngo, Teo, et al., 2000b, unity allows the elements of visual design to seem like they belong together, to dovetail so completely that they are seen as one thing. Unity is achieved by using similar sizes, shapes, or colors for related information and leaving less space between elements of a screen than the space left at the margins.

4.1.6.1 First formula to measure unity (SoE=5)

Unity (UM) is the extent to which the screen elements seem to belong together and is given by the formula below (equation 4.26) (Altaboli and Lin, 2011; Ngo and Byrne, 2001a, 2001b; Ngo, Teo, et al., 2000a, 2000b, 2003; Purchase et al., 2011; Salimun et al., 2010):

$$UM = \frac{|\mathrm{UM}_{form}| + |\mathrm{UM}_{space}|}{2} \in [0,1] \tag{4.26}$$

Where UM_{form} is the extent to which the objects are related in size with:

$$\mathrm{UM}_{form} = 1 - \frac{n_{size} - 1}{n}$$

Ngo went further in the definition of UM_{form} where it depends on the size, shape, and color of objects (Ngo, Teo, et al., 2000b):

$$\mathrm{UM}_{form} = 1 - \frac{n_{size} + n_{color} + n_{shape} - 3}{3n}$$

UM_{space} is a relative measure of the space between groups and that of the margins:

$$\text{UM}_{space} = 1 - \frac{a_{layout} - \sum_i^n a_i}{a_{frame} - \sum_i^n a_i}$$

Where a_i, a_{layout}, and a_{frame} are the areas of object i, the layout, and the frame, respectively. n_{size}, n_{color}, and n_{shape} are the numbers of sizes, colors, and shapes used respectively, and n is the number of objects on the frame.

4.1.6.2 Second formula to measure unity (SoE=2)

The formula is proposed by Soui et al., 2019 in the context of mobile applications. As the screen size of the smartphone is not the same as of a computer, it is necessary to adapt the information quantities and the form of information. The unity is the extent to which the components are correlated in size and number of space between groups and margins. A good unity is obtained by using the optimum number of size component (minimize the uses of different sizes in the mobile interface) and leaving less space between objects. The formula of unity is given below (equation 4.27) (Soui et al., 2019):

$$UM = 1 - \left(0.5\left[\frac{|N_{size} - 1|}{n} + \frac{|a_{sc} + \sum_i^n a_i|}{2a_{MUI}}\right]\right) \in [0, 1] \tag{4.27}$$

Where N_{size} is the number of various sizes belong into used objects, n is the number of objects, a_{MUI} is the area of the mobile interface, a_{sc} is the area of the screen and a_i is the area of the interactive object i.

4.1.7 Proportion (NoC=5)

4.1.7.1 First formula to measure proportion (SoE=5)

For thousands of years, people and cultures have loved proportional shapes. In screen design, aesthetically pleasing proportions should be considered for major components of the screen, including windows and groups of data and text in order to improve the user experience.

The following shapes (with their proportions) are considered as aesthetically pleasing (Ngo et al., 2003):

- Square (1:1)
- Square root of two (1:1.414)
- Golden rectangle (1:1.618)
- Square root of three (1:1.732)
- Double square (1:2)

Proportion (PM) is the comparative relationship between the dimensions of the screen components and proportional shapes and is given by the equation 4.28 (Ngo and Byrne, 2001a; Ngo, Teo, et al., 2000a, 2000b, 2003):

$$PM = \frac{|\mathrm{PM}_{object}| + |\mathrm{PM}_{layout}|}{2} \in [0, 1] \tag{4.28}$$

PM_{object} is the difference between the proportions of the objects and the closest proportional shapes. PM_{layout} is the difference between the proportions of the layout and the closest proportional shape (cfr section 6.6.1 for details of the calculations of these terms). Overall, they depend on the width and height of the interface objects, the width and height of the layout and the proportion of the interface shapes.

4.1.7.2 Second formula to measure proportion (SoE=1)

This second formula of proportion (V_{ar}) proposed by Harrington et al., 2004 is given below (equation 4.29):

$$V_{ar} = 1 - \frac{|Z_i - R|}{R} \in [0, 1] \tag{4.29}$$

With Z_i: the ratio of width and height of the major content elements and groups of elements on a page and R: the golden ratio between width and height and equal to $2/(1 + \sqrt{5}) = 0.618$.

4.1.8 Simplicity (NoC=9)

4.1.8.1 First formula to measure simplicity (SoE=5)

Simplicity (SMM) is a combination of elements that results in ease in comprehending the meaning of a pattern. Simplicity in screen design is achieved by optimizing the number of elements on a screen and minimizing the alignment points.

It can be easily measured by the equation 4.30 (Chettaoui and Bouhlel, 2017; Fu et al., 2007; Ngo and Byrne, 2001a, 2001b; Ngo, Teo, et al., 2000a, 2000b, 2003; Purchase et al., 2011):

$$SMM = \frac{3}{n_{vap} + n_{hap} + n} \in [0, 1] \tag{4.30}$$

Where n_{vap} and n_{hap} are the numbers of vertical and horizontal alignment points and n is the number of objects on the frame.

4.1.8.2 Second formula to measure simplicity (SoE=4)

This second formula measure the vertical and horizontal alignment of objects in two levels: a group level (Local Alignment) and a screen level (Global Alignment). These two alignment levels are then combined to calculate the Total Screen Alignment Complexity (TAC) (equation 4.31) (Alemerien and Magel, 2014):

$$TAC = AC * \text{weight1} + SA * \text{weight2} \tag{4.31}$$

Where AC is the Alignment Complexity, SA is the Screen Alignment (cfr section 6.7.1 for the definition of AC and SA), $weight1$ is the ratio of the number of grouped-objects to the total number of objects on the screen and $weight2$ is the ratio of the number of ungrouped-objects to the total number of objects on the screen. The simplicity can be obtained my minimizing the TAC value. This formula has been empirically validated with a correlation of 0.836.

4.1.9 Density (NoC=9)

4.1.9.1 First formula to measure density (SoE=5)

Density is the extent to which the screen is covered with objects. Density is achieved by minimizing screen density levels. The density (DM) can be measured by the percentage of object on the entire frame by the following formula (equation 4.32) (Fu et al., 2007; Ngo and Byrne, 2001a, 2001b; Ngo, Teo, et al., 2000a, 2000b, 2003):

$$DM = 1 - 2\left|0.5 - \frac{\sum_i^n a_i}{a_{frame}}\right| \in [0,1] \tag{4.32}$$

Where a_i and a_{frame} are the areas of object i and the frame; and n is the number of objects on the frame.

4.1.9.2 Second formula for density (SoE=2)

A second formula to measure density in the context of mobile applications has been proposed (Soui et al., 2019). They found that an user with a low motivation to be continually interested and to make an effort to achieve a goal prefers a mobile user interface with a low-density level. Density is calculated by the equation 4.33 (Soui et al., 2019):

$$DM = 0.5\left|\frac{\sum_i^n a_i}{a_{MUI}} + \frac{a_{MUI}}{a_{sc}}\right| \tag{4.33}$$

Where n is the number of interactive objects, a_{MUI} is the area of the mobile interface, a_{sc} is the area of the screen and a_i is the area of the interactive object i.

4.1.9.3 Third formula to measure density (SoE=4)

The third formula of density developed by (Alemerien and Magel, 2014) takes into account the local density (LD_j) of a group of object j and the global density (GD) to obtain the Density-Complexity (DC) formula given below (equation 4.34):

$$DC = \left(\frac{\sum_{j=1}^{n} LD_j}{n} \right) * W1 + GD * W2 \tag{4.34}$$

Where W1 is the ratio of the area of groups to the screen area and W2 is the ratio of ungrouped area to the screen area. LD_j and GD are defined below:

$$LD_j = \frac{\sum_{i=1}^{\text{groupedobjects}} \text{Size of object } i \text{ in a group } j}{\text{Area of group } j}$$

$$GD = \frac{\sum_{k=1}^{\text{ungroupedobjects}} \text{Size of object } k}{\text{Area of ungrouped}}$$

This formula has been empirically validated with a correlation of 0.683.

4.1.9.4 Fourth formula to measure density (SoE=5)

This fourth formula of density is specifically applied to Android applications. The density metric (D_s) shows how cluttered or empty the user interface is. The equation 4.35 shows that the areas $A_{u,s}$, of all GUI elements U_s are summed up and related to the area A_s of the entire screen s, resulting in the density D_s (Riegler and Holzmann, 2018b):

$$D_s = \frac{1}{A_s} \sum_{u=1}^{U_s} A_{u,s} \tag{4.35}$$

4.1.10 Regularity (alignment) (NoC=11)

Regularity is an uniformity of elements based on some principle or plan. Regularity can be achieved by establishing standard and consistently spaced horizontal and vertical alignment points for screen elements and by minimizing the alignment points (Ngo, Teo, et al., 2000b). Experience in practical comparisons of real-world designs suggests that height and width, as well as bottom and right edge alignments are also factors in subjective judgments of visual regularity. A way to improve the usability is to improve the regularity because highly chaotic arrangements are more difficult to use (Noble and Constantine, 1996).

4.1.10.1 First formula to measure regularity (SoE=5)

It is possible to measure how regular the screen is thanks to the formula of regularity (RM) given below (equation 4.36) (Ngo and Byrne, 2001b; Ngo, Teo, et al., 2000a, 2000b, 2003; Salimun et al., 2010; Zen, 2013):

$$RM = \frac{|RM_{alignment}| + |RM_{spacing}|}{2} \in [0, 1] \tag{4.36}$$

$RM_{alignment}$ is the extent to which the alignment points are minimized and $RM_{spacing}$ is the extent to which the alignment points are consistently spaced (cfr section 6.8.1 for details of the calculations of these terms). Overall, they depend on the numbers of vertical and horizontal alignment points, the number of distinct distances between column and row starting points and the number of objects on the frame.

4.1.10.2 Second formula to measure alignment (SoE=2)

A second formula to measure regularity (RM) is given, in the context of mobile applications, by the equation 4.37 (Soui et al., 2019):

$$RM = 1 - \frac{N_{av} + N_{ah} + N_{sp}}{3} \in [0, 1] \tag{4.37}$$

Where N_{av} is the number of vertical alignment points (number of rows), N_{ah}: the number of horizontal alignment points (number of columns), N_{sp} is the number of distinct distances between column and row starting points and n is the number of component of the mobile user interface.

4.1.10.3 Third formula to measure regularity (SoE=2)

Regularity is a structural metric that measures the uniformity or orderliness of the user interface layout. It is based on the principle that ease of use is hampered by very chaotic arrangements. The uniformity of the layout focuses on the spatial arrangement of the interface components. An user interface with a regularity score of 0.6 to 0.9 is considered as a good interface (see equation 4.38) (Lau and Wilson, 1998):

$$\text{Regularity} = 1 - \frac{(N_{size} + N_{top} + N_{left}) - A}{3C - A} \tag{4.38}$$

Where C is the number of components on the screen, N_{size}, N_{top} and N_{left} are the number of different sizes, top edge alignments and left edge alignments of components respectively. $A = [2\sqrt{C}] + 1$ is an adjustment for the minimum number of possible alignments and sizes.

4.1.10.4 Fourth formula to measure regularity (SoE=2)

The fifth formula to measure regularity is expressed as the layout uniformity of the interface in the paper of Noble and Constantine (Noble and Constantine, 1996). In fact, the layout uniformity is a synonym of regularity because they measure the same things (alignment and uniformity/orderliness of an user interface layout). They defined the layout uniformity (LU) as follows (equation 4.39) (Noble and Constantine, 1996):

$$LU = 100 * \left(1 - \frac{(N_h + N_w + N_t + N_l + N_b + N_r) - M}{6 * N_c - M}\right) \in [0, 100] \tag{4.39}$$

Where N_c is the total number of components on the screen, dialogue box, or other interface composite, and N_h,N_w, N_t, N_l, N_b and N_r are, respectively, the number of different heights, different widths, different top edge alignments, left edge alignments, bottom edge alignments, and right edge alignments of visual components. M is an adjustment for the minimum number of possible alignments and sizes that makes the value of LU range from 0 to 100:

$$M = 2 + 2 * \left\lceil 2\sqrt{N_{total}} \right\rceil$$

with "$\lceil \quad \rceil$" denoting « smallest integer greater than ». Layout Uniformity is based on the argument that highly chaotic arrangements are more difficult to use. However, complete regularity is not the goal. Indeed, moderately high regularity characterizes good designs (LU between 0.5 and 0.9) (Noble and Constantine, 1996).

4.1.10.5 Fifth formula to measure alignment (SoE=4)

In their paper, Riegler and Holzmann describe the misalignment metric (Riegler and Holzmann, 2018b). This metric formula is specifically applied to Android applications.

In order to calculate the misalignment metric M_s, the mean number of horizontal ($\overline{numh_s}$), vertical ($\overline{numv_s}$) and central alignments ($\overline{numc_s}$) of pairs of neighboring elements of a screen s is calculated as follows (equation 4.40) (Riegler and Holzmann, 2018b):

$$M_s = 1 - (\overline{\mathrm{numh}_s} + \overline{\mathrm{numv}_s} + \overline{\mathrm{numc}_s}) \tag{4.40}$$

From this formula it is therefore possible to measure alignment or regularity, by calculating $1 - M_s$. The calculation of the terms of M_s is described below:

$$\overline{\mathrm{numh}_s} = \frac{1}{A_s} \sum_{g=1}^{G_s} A_{g,s} \frac{\mathrm{numh}_{g,s}}{\mathrm{num}_{g,s}}$$

Where $A_{g,s}$ is the area of a group g of GUI elements on the screen s, A_s is the area of the screen, G_s is the weighted arithmetic mean of the width and height elements over all groups, $\mathrm{numh}_{g,s}$ is the number of horizontal alignments in a group g and $\mathrm{num}_{g,s}$ is the number of possible alignments (horizontal, vertical and central) in a group. The mean number of vertical and central alignments $\overline{\mathrm{numv}_s}$ and $\overline{\mathrm{numc}_s}$ are calculated in a similar way.

4.1.10.6 Sixth formula to measure regularity and alignment (SoE=2)

Examples of alignment formulas are given in a book written by López S., Simarro and López P. The formulas consist of 3 different alignment formulas which calculates the horizontal, vertical and perpendicular alignment separately and one formula of regularity. The first one, called linearity, measures the horizontally aligned arrangement of the element in the interface. Equation 4.41 shows the linearity metric (Li) (López et al., 2013):

$$Li = \frac{n_{hac} \times n_{va}}{n_c^2} \tag{4.41}$$

Where n_{hac} is the number of components horizontally aligned with other components, n_{va} is the number of different vertical alignments and n_c is the number of components in the interface.

The second one, called sequentiality (Se), is the vertically analogous metric to linearity (equation 4.42) (López et al., 2013):

$$Se = \frac{n_{vac} \times n_{ha}}{n_c^2} \tag{4.42}$$

Where n_{vac} is the number of components vertically aligned with other components,n_{ha} is the number of different horizontal alignments and n_c is the number of components on the interface.

The last formula of alignment proposed is called orthogonality. It considers the perpendicular alignment of elements of different shapes and sizes and ensures that every component is horizontally and vertically aligned. The expression of this metric (Or) is shown below (equation 4.43) (López et al., 2013):

$$Or = \frac{n_{vac} \times n_{hac}}{n_c^2} \tag{4.43}$$

Where n_{hac} is the number of components horizontally aligned with other components, n_{va} is the number of components horizontally aligned with other components and n_c is the number of components in the interface.

The formula of regularity proposed is shown below (equation 4.44) (López et al., 2013):

$$Re = \frac{n_{vac} \times n_{hac} \times (n_c - n_s + 1)}{n_c^3} \tag{4.44}$$

Where n_{vac} is the number of components vertically aligned with other components, n_{hac} is the number of components horizontally aligned with other components, n_c is the number of components in the interface and n_s is the number of different shapes.

4.1.11 Economy (NoC=5)

4.1.11.1 First formula to measure economy (SoE=5)

The economy consists in using as few display elements as possible. The objective is to convey a message in the simplest possible way. This can be done with minimal use of colors, technical displays and font style or visual design elements (Ngo, Teo, et al., 2000b).

The formula of economy (ECM) is given below (equation 4.46) (Ngo and Byrne, 2001a, 2001b; Ngo, Teo, et al., 2000a, 2003; Salimun et al., 2010):

$$ECM = \frac{1}{n_{size}} \in [0, 1] \tag{4.45}$$

Ngo went further in the definition of economy (ECM) with this formula (Ngo, Teo, et al., 2000b):

$$ECM = \frac{3}{n_{size} + n_{color} + n_{shape}} \in [0, 1] \tag{4.46}$$

with: n_{size}, n_{color}, and n_{shape}, the numbers of sizes, colors and shapes used respectively.

4.1.12 Homogeneity (NoC=5)

4.1.12.1 First formula to measure homogeneity (SoE=5)

The relative degree of homogeneity of a composition is determined by how evenly the objects are distributed among the four quadrants of the screen. The degree of evenness is a matter of the quadrants that contain nearly equal numbers of objects (Ngo, Teo, et al., 2000b).

Homogeneity (HM) is a measure of how evenly the objects are distributed among the quadrants and is given by the formula below (equation 4.47) (Ngo and Byrne, 2001b; Ngo, Teo, et al., 2000a, 2000b, 2003; Soui et al., 2019):

$$HM = \frac{W}{W_{max}} \in [0,1] \tag{4.47}$$

W is the number of different ways a group of n objects can be arranged for the four quadrants when n_j is the total number of objects in quadrant j.

$$W = \frac{n!}{\prod_{j=UL,UR,LL,LR} n_j} = \frac{n!}{n_{UL}!n_{UR}!n_{LL}!n_{LR}!}$$

W is maximum when the n objects are evenly allocated to the various quadrants of the screen and therefore:

$$W_{max} = \frac{n!}{\frac{n}{4}!\frac{n}{4}!\frac{n}{4}!\frac{n}{4}!} = \frac{n!}{(\frac{n}{4}!)^4}$$

with n_{UL}, n_{UR}, n_{LL}, and n_{LR} the numbers of objects on the upper-left, upper-right, lower-left, and lower-right quadrants, respectively and n the number of objects on the frame (Ngo, Teo, et al., 2000b).

4.1.13 Rhythm (NoC=6)

4.1.13.1 First formula to measure rhythm (SoE=5)

Rhythm in visual design refers to regular changes in the elements used. It combines order with variation to make the appearance exciting. Rhythm is accomplished through variation of arrangement, dimension, number and form of the elements.

Rhythm (RHM) is the extent to which the objects are systematically ordered and is given by the equation 4.48 (Kar and Zain, 2007; Ngo and Byrne, 2001b; Ngo, Teo, et al., 2000a, 2000b, 2003; Zain et al., 2011):

$$RHM = 1 - \frac{|\mathrm{RHM}_x| + |\mathrm{RHM}_y| + |\mathrm{RHM}_{area}|}{3} \in [0,1] \tag{4.48}$$

The rhythm components are defined more in details in section 6.9.1. They depend on the coordinates of the center of the frame, the coordinates of the centers of the interface objects, the area of the interface objects and the total number of objects on the quadrant.

4.1.14 White-space fraction (NoC=1)

A study by Harrington et al. has defined a design rule: white space (including margins) should total about half of the total page area. The non white-space area can be estimated by totaling the areas of the content objects (Harrington et al., 2004).

4.1.14.1 First formula to measure the white-space fraction (SoE=1)

The formula proposed by Harrington et al., 2004 to measure the white-space fraction (V_{ws}) is given below (equation 4.49):

$$V_{ws} = 1 - 4\left(\frac{\sum A_i}{A_p} - 0.5\right)^2 \tag{4.49}$$

Where A_i is the area of object i and A_p is the area of the page.

4.1.15 Composition (NoC=3)

4.1.15.1 First formula to measure composition (SoE=2)

This metric is used in the context of mobile applications and serves to increase the visual clarity of the GUI by the meaningful arrangement of the interface components. The objective of this metric is the visual and semantic grouping of the components that are related in the same border (the line, the color, the shape, etc.). This metric counts the number of objects that have a clear boundary by the group. In general, a composition close to 1 helps younger or older users with low computer skills to interact with the system. To measure this, we use the following formula (equation 4.50) (Soui et al., 2019):

$$COM = 1 - \frac{G + UG}{2n} \in [0, 1] \tag{4.50}$$

Where G is the number of groups with clear boundary by line, background, color, or space, UG is the number of ungrouped objects and n is the number of component of the mobile user interface.

4.1.15.2 Second formula to measure composition (SoE=3)

This metric of composition, called « Grouping metric» in the paper written by Alemerien and Magel, measures the number of objects that have a clear boundary by line, background, color, space, or size (Alemerien and Magel, 2014). The formula of composition calculates the percentage of grouped and ungrouped objects (G and UG respectively).

Finally, the grouping complexity (GT) is obtained as shown below (equation 4.51) (Alemerien and Magel, 2014):

$$GT = UG + GC \tag{4.51}$$

with :

$$UG = 1 - \frac{\sum_{i=1}^{N} GW}{N}$$
$$GC = \frac{G}{M} * \text{Weight}$$

Where N is the total number of objects on the screen, G is the ratio of the number of different object types to the total number of objects (M) in all groups. Weight is the ratio of total number of grouped objects to total number of objects on the screen. GW is a variable that indicates whether the object is grouped or not. The value of GW equals 1 if the object exists in a group, otherwise GW equals 0. The formula of the grouping complexity has been empirically validated with a low correlation of 0.462.

4.1.15.3 Third formula to measure composition (SoE=3)

Under the grouping technique, elements with similar functions or information are surrounded by a boundary by line, background color and/or space. Grouping makes it easier for the users to extract the information assigned to the group. The method for calculating the grouping complexity is computed as follows (equation 4.52) (Fu et al., 2007):

$$CG = \frac{g_i}{g} \in [0, 1] \tag{4.52}$$

Where g_i is the number of groups with clear boundary by line, background, color, or space and g is the total number of groups. This formula has been empirically validated but no correlation value was given. The study only concluded that the proposed formula is useful.

4.1.16 Color contrast (NoC=2)

Contrast refers to the juxtaposition of visually dissimilar elements in order to convey the fact that these elements are different. It provides the eye with a noticeable difference between two objects or two sets of objects. In this case, we will talk of the color contrast which can provoke an emotion to the user or emphasize/deemphasize text elements.

4.1.16.1 First formula to measure color contrast (SoE=5)

The color model used to compute this metric is the HSV model (hue, saturation, and value). Hue varies from 0° to 360° while the corresponding colors vary from red through yellow, green, cyan, blue, magenta, and back to red. Saturation varies from 0 (unsaturated) to 1 (fully saturated). Value varies from 0 to 1 with the corresponding colors becoming increasingly brighter (Figure 4.3).

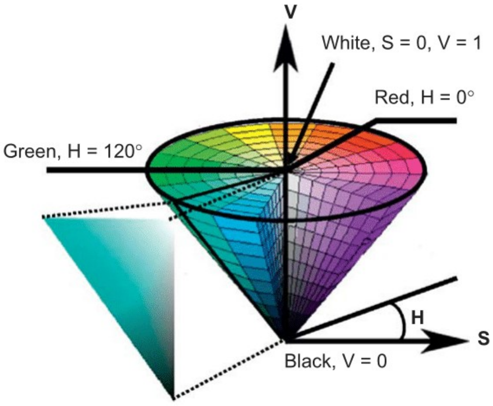


Figure 4.3: The coordinate system for the HSV color space (Lai et al., 2010)

The color difference between two colors, (H1, S1, V1) and (H2, S2, V2), is therefore given as (equation 4.53) (Lai et al., 2010):

$$\Delta C = \frac{1}{\sqrt{4}}\sqrt{(V_1+V_2)^2+(V_1S_1cosH_1-V_2S_2cosH_2)^2+(V_1S_1sinH_1-V_2S_2sinH_2)^2} \tag{4.53}$$

A human perceives quite differently two colors with a ΔC close to 1 while the ΔC of two perceptually similar colors is close to 0. Therefore, a pixel block with a large ΔC against the color of the background region should appear to be more contrasted and visually heavier than one with a small ΔC.

4.1.16.2 Second formula to measure color contrast (SoE=2)

In their paper, Miniukovich and De Angeli calculated the figure-ground contrast. They defined this term as the difference in color or brightness between two adjacent areas. This difference forms an edge and the magnitude of this difference defines the strength of the edge (Miniukovich and De Angeli, 2014). They used a microscopic approach (pixel scale) to generate a formula for this metric.

In their study, they analyzed screenshots of web pages. The edges of two adjacent areas were detected by using an algorithm which requires two input thresholds in the range of 0 to 1: low and high. Lower input thresholds detects more edges while higher thresholds detects fewer edges. In the paper of Miniukovich and De Angeli, the threshold varied from 0.1 to 0.7 with a step of 0.1, which gave seven levels (l) of edge strength (E_l). The edge pixels were counted for each level and the difference between successive levels was computed. Then, the differences were weighted and summed up. The weakest edges received the highest weight, since they contribute the most to the difficulty of visual differentiation. The last step was to normalize the sum by the total number of edge pixels. The figure-ground contrast metric (E_{norm}) is finally computed as follows (equation 4.54) (Miniukovich and De Angeli, 2014):

$$E_{norm} = \frac{\sum_{l=1}^{6}((E_l - E_{l+1}) * (1 - \frac{(l-1)}{6}))}{E_1 - E_6} \tag{4.54}$$

This formula has been empirically validated with a low correlation (0.46).

4.1.17 Color complexity (NoC=1)

The use of color in an user interface has been shown to improve performance and visual search time, to facilitate organizing information and to aid memory. Color can also create positive user perceptions and is preferred to monochromatic screens. A colored interface is seen less monotonous, reduce eye strain, and is more globally more pleasant (Riegler and Holzmann, 2018b). Research has found that if the number of colors on a display is increasing, the response time to a single color rises and color confusion is more likely (Luria et al., 1986). A study found that the maximum number of colors that a person can handle is in the range of 4 to 10, while lower numbers should be encouraged (Brooks, 1965).

4.1.17.1 First formula to measure color complexity (SoE=3)

This metric of color complexity is applied specifically to Android applications. The metric is based on the number of dominant colors. Dominant colors are occurring more frequently on a pixel-basis than non-dominant colors. The algorithm also computes the difference in terms of HSV value for the dominant colors to obtain the formula of color complexity shown in equation 4.55 (Riegler and Holzmann, 2018b):

$$C_s = \frac{\frac{numc_s}{numcmax} + (1 - \overline{ccm_s}) + c_{h_s,h_{s+1}}}{3} \tag{4.55}$$

Where *numcmax* denotes the maximum number of dominant colors on a certain screen of the whole user interface, $numc_s$ is the number of dominant colors c_1,c_2,.. on screen s, $\overline{ccm_s}$ is the average color combination match of the dominant colors and $c_{h_s,h_{s+1}}$ is the color combination match of the current screen's color histogram h_s and the succeeding screen's color histogram h_{s+1} (cfr section 6.10.1 for more details on the calculation of the variables). This formula was empirically validated with a low correlation (around 40% on three different applications).

4.1.18 Typography complexity (NoC=1)

4.1.18.1 First formula to measure typography complexity (SoE=5)

Typography plays an important role in application design. Smaller fonts lead to a higher fixation duration and thus reduce the reading performance. On the other side, choosing a larger font size increases legibility. The color contrast can also influence the legibility of the text (Riegler and Holzmann, 2018b).

The formula of typography complexity presented below is applied in the context of Android applications. Typographic aspects that are analyzed in this formula include the number of different text sizes, or heights of text elements, the average text size and the average text foreground/background color combination match. The formula of typography complexity (T_s) is given below (equation 4.56) (Riegler and Holzmann, 2018b):

$$T_s = \frac{\frac{numt_s}{numtmax} + \left(1 \frac{\overline{t_s}}{tmax}\right) + (1 - \overline{ccm_s})}{3} \tag{4.56}$$

Where $numt_s$ is the number of different text sizes, or heights of text elements, $numtmax$ denotes the maximum number of text sizes and $tmax$ denotes the maximum text size, both on a certain screen of the whole user interface. $\overline{t_s}$ is the average text size which is determined by calculating the average height of all text elements found on the screen s and $\overline{ccm_s}$ is the average text foreground/background color combination match (cfr section 6.10.1 for the calculation of this term).

4.1.19 Essential efficiency (NoC=1)

4.1.19.1 First formula to measure essential efficiency (SoE=3)

A well designed user interface should have as few steps as possible to achieve the user's goal. Essential efficiency is the ratio of the number of steps in the essential use case to those that are required to perform the task with a given user interface. It measures how close a given design is to the essential use case model, in terms of the number of steps involved to perform a task or collection of tasks (equation 4.57) (Lau and Wilson, 1998):

$$\text{Essential Efficiency} = \frac{\text{Essential Length}}{\text{Operational Length}} \tag{4.57}$$

Where Essential Length represents the essential minimum number of steps from users' perspective. Essential efficiency computed for a given user interface design with a value of 1 or close to 1 indicates that the design is efficient in performing the given tasks (Lau and Wilson, 1998).

4.1.20 Size complexity (NoC=2)

4.1.20.1 First formula to measure size complexity (SoE=5)

The size complexity can influence the global complexity of the interface. Alemerien and Magel, 2014 have developed a formula to measure the Object Size Complexity (SC_k) and the overall size complexity (SC). The formula is given below (equation 4.58) (Alemerien and Magel, 2014):

$$SC = \frac{\sum_{k=1}^{W_i} SC_k * \text{Weight(k)}}{W_i} \tag{4.58}$$

with:

$$SC_k = \frac{\sum_{j=1}^{N} S_j}{N}$$

Where N is the number of objects in type kth and S_j is the number of different sizes, S_j is 1 if the object size is not counted before and 0 if the object size is counted. W_i is the number of objects types and $Weight(j)$ is the number of objects in category jth divided by the total number of objects on the screen. The formula of the overall size complexity has been empirically validated with a correlation of 0.943.

4.1.20.2 Second formula to measure size complexity (SoE=3)

Size complexity involves the categorization of elements into groups according to actual physical size and variation in those sizes. It is based partially on information order, and is calculated for each type of information so that the complexity don't increase significantly when different information appear in different sizes. In terms of the magnitude size of the size complexity (CS), objects are related to size as follows (equation 4.59) (Fu et al., 2007):

$$SC = 1 - \frac{\sum_{I}^{type}(n_{size} - 1)}{n} \in [0, 1] \tag{4.59}$$

Where n_{size} is the number of sizes and n is the total number of objects. This formula has been empirically validated but no correlation value was given. The study only concluded that the proposed formula is useful.

4.1.21 Visual coherence (NoC=2)

4.1.21.1 First formula to measure visual coherence (SoE=3)

Visual Coherence (VC) measures how closely visual organization matches the semantic relationships among concepts associated with components. It is based on the principle that well-structured interfaces group semantically related components together.

Visual Coherence of a visual group is the ratio of the total « relatedness » of pairs of visual components to the total number of pairs. Total VC for a screen or dialogue box is the sum of the VC for each level of grouping (equation 4.60) (Constantine, 1996; Noble and Constantine, 1996):

$$VC = 100 * \left(\frac{\sum_{\forall l} G_l}{\sum_{\forall l} N_l * (N_l - 1)/2} \right) \tag{4.60}$$

with $G_l = \sum_{\forall i,j | i \neq j} R_{i,j}$

Where N_l is the number of components in group i, $R_{i,j}$ is the semantic relatedness between components i and j. In practice, semantic relatedness can be simplified to $R_{i,j} = 1$ if components i and j are « substantially related », $R_{i,j} = 0$ otherwise. This formula correctly favors organizing user interface components into subgroups, but only so long as the groups make sense, that is, they enclose substantially related components (Noble and Constantine, 1996).

This metric is one of the few that takes into consideration the semantic aspect and is therefore very interesting.

4.1.22 Element smallness (NoC=1)

4.1.22.1 First formula to measure element smallness (SoE=5)

E_s, the screen smallness, is a measure of the size of GUI elements on a certain screen s. This formula is specifically applied to Android applications. Firstly, the screen is divided into groups. A group g is defined as a region of interest on the screen, which contains one or more GUI elements belonging together. There are various criteria for the formation of groups such as common background color, position on the screen or functionality. Then, the mean element width and height of each group g of a screen s ($\overline{width_s}$ and $\overline{height_s}$ respectively) is calculated from the $\text{width}_{u,g,s}$ and $\text{heigth}_{u,g,s}$ of each element u in this group (cfr section 6.11.1 for more details). Finally, $\overline{width_s}$ and $\overline{height_s}$ are combined to a single metric : E_s, the screen smallness (equation 4.61) (Riegler and Holzmann, 2018b):

$$E_s = \frac{\left(1 - \frac{\overline{\text{width}_s}}{\text{width}_s}\right) + \left(1 - \frac{\overline{\text{height}_s}}{\text{height}_s}\right)}{2} \tag{4.61}$$

Where $width_s$ and $height_s$ are the width and height of the screen s. Smaller objects tend to be harder to identify and interact with, so this metric needs a higher score for smaller elements and smaller score for higher elements.

4.1.23 Edge congestion (NoC=1)

4.1.23.1 First formula to measure edge congestion (SoE=4)

Edge congestion relies on the notion of critical spacing – minimal distance between objects at which the user starts having difficulties differentiating the objects (Miniukovich and De Angeli, 2014).

Miniukovich and De Angeli have developed an algorithm to detect edges on a screenshot of a web page. This metric is applied in a microscopic approach by comparing the position of the different pixels of the screenshot. If at least two pixels of two different edges are in the same 20-pixel vicinity, they are marked as congested. The marking is done in both horizontal and vertical directions. Finally, the edge congestion metric is defined as the ratio between congested pixels (p_c) to all edge pixels (p_a) (equation 4.62) (Miniukovich and De Angeli, 2014):

$$cong = \frac{p_c}{p_a} \tag{4.62}$$

This metric has been validated empirically with a low correlation of 0.53.

4.1.24 Concentricity (NoC=1)

4.1.24.1 First formula to measure concentricity (SoE=1)

The concentricity metric measures how the interface objects are gathered in the center of the interface or rather kept in the corners (Zen and Vanderdonckt, 2014). Equation 4.63 shows the formula of concentricity (C) (Zen and Vanderdonckt, 2014):

$$C = \frac{\overline{d}}{d_{diag}} \tag{4.63}$$

It can be measured using the average of distances between the interface center and each object center (Zen and Vanderdonckt, 2014).

4.1.25 Order and complexity (NoC=13)

4.1.25.1 First formula to measure order and complexity (SoE = 5)

By compiling some of the formulas described above, it is possible to create a scale of order or complexity with extreme complexity at one end and minimal complexity (order) at the other (Ngo, Teo, et al., 2000b).

The order and complexity scale (OM) is measured as shown below (equation 4.64) (Kar and Zain, 2007; Ngo and Byrne, 2001a, 2001b; Ngo, Teo, et al., 2000a, 2000b, 2003; Zain et al., 2011):

$$OM = \frac{\sum_i^{13} \alpha_i M_i}{13} \in [0,1], \qquad 0 \leq \alpha_i \leq 1 \tag{4.64}$$

with

$\mathrm{M} = (M_1,M_2,M_3,M_4,M_5,M_6,M_7,M_8,M_9,M_{10},M_{11},M_{12},M_{13})$

= (BM, EM, SYM, SQM, CM, UM, PM, SMM, DM, RM, ECM, HM, RHM)

Where BM, EM, SYM, SQM, CM, UM, PM, SMM, DM, RM, ECM, HM and RHM is given by equation 4.1, 4.11, 4.13, 4.20, 4.23, 4.26, 4.28, 4.30, 4.32, 4.36, 4.46, 4.47 and 4.48 respectively. The weighting component α is assumed to be a constant. Aesthetic measure M_i has its own weighting component α_i. The choice of weights is an optimization problem that needs to be solved with mathematical algorithms (Ngo, Teo, et al., 2000b).

The strength of evidence level is 5 for this formula. Each metric listed above has been calculated and applied to more than 50 different screens. Then, an empirical study has been performed with a high correlation (0.930) between the relative ratings by the viewers and those obtained using the computational measures. A further study (Ngo and Byrne, 2001b) showed that BM, SQM, UM, PM, DM and HM account for most of the variability in overall ratings, but the others metrics have less impact.

4.1.25.2 Second formula to measure order and complexity (SoE=5)

This second formula of order and complexity is derived from the Birkhoff's formula of the measure of order and complexity (M): $M = O/C$. Where O is the order and C the complexity (Birkhoff, 1933).

The formula proposed by Ngo computes 4 different metrics defined previously and is given by the equation 4.65 (Ngo and Byrne, 1998; Ngo, Samsudin, et al., 2000):

$$M = \frac{O_{BM} + O_{EM} + O_{SM} + O_{SQM}}{C} \in [0,3] \qquad (C > 0) \tag{4.65}$$

Where O_{BM} is a measure of balance, O_{EM} is a measure of equilibrium, O_{SM} is a measure of symmetry, and O_{SQM} is a measure of sequence. The complexity C of a layout is defined as the number of its components (cfr section 6.12.1 for more details on these terms).

This formula has been verified empirically by comparing the order value obtained from the formula 4.65 to the order perceived by the subjects of the study. The correlation was very high (0.97).

4.1.25.3 Third formula to measure complexity (SoE=3)

Guangfeng Song has developed an algorithm to measure complexity and visual segmentation of Web pages (cfr section 6.12.2 for the explanations of the algorithm). The formula of complexity is given below (equation 4.66) (Song, 2007):

$$C = \frac{\sum C_k S_k}{S_k} \tag{4.66}$$

Where C_k and S_k are the complexity and size of sub-block k.

Overall, the algorithm analyzes the interface to generate blocks that are hierarchically arranged. The formula is then used in the hierarchy of blocks from bottom to top. The lowest level blocks are assigned a complexity equal to the number of objects they enclose. The estimated complexity of the whole web page is finally calculated with the complexity of all first level blocks (Song, 2007).

This formula has been validated empirically. However, the correlation of the estimated complexity of the first level of blocks with the perceived complexity was low (0.433).

4.1.25.4 Fourth formula to measure complexity (SoE=3)

A study, written by Soui et al. in the context of mobile applications, proposes that novice users usually prefer an interface with a low level of complexity while users having higher education level prefer Mobile User Interfaces with a high level of complexity. The complexity metric is therefore calculated as follows (equation 4.67) (Soui et al., 2019):

$$CM = \frac{N_{vap} + N_{hap}}{2n} \in [0, 1] \tag{4.67}$$

Where N_{vap} is the number of vertical alignment points, N_{hap} is the number of horizontal alignment points and n is the number of object on the mobile user interface.

4.1.25.5 Fifth formula to measure complexity (SoE=4)

In the paper of Alemerien and Magel, 5 metrics contributing to the overall complexity of the screen layout were defined (Alemerien and Magel, 2014). These 5 metrics are the Total Balance Complexity (TBC) (equation 4.7), the Total Screen Alignment Complexity (TAC) (equation 4.31), the Density Complexity (DC) (equation 4.34), the Grouping Complexity (GT) (equation 4.51) and the Size Complexity (SC) (equation 4.58). The Overall Screen Layout Complexity (LC) formula is therefore given below (equation 4.68) (Alemerien and Magel, 2014):

$$LC = ((TAC * w1 + TBC * w2 + DC * w3 + SC * w4 + GT * w5)/5) * 100\% \tag{4.68}$$

All the complexity metrics have been empirically validated and a weight for each metric was calculated based on the ratings of the study participants. w1, w2, w3, w4 and w5 are the weight for the TAC, TBC, DC, SC, and GT metrics respectively. The values of those weights are 0.84, 0.76, 0.80, 0.72, and 0.88, respectively. The correlation between the user ratings and the computational value of the overall screen layout complexity is 0.804.

4.1.25.6 Sixth formula to measure complexity (SoE=4)

The sixth formula of complexity proposed by Stickel et al. is called the XAOS metric. The formula for visual complexity (X) is shown below (equation 4.69) (Stickel et al., 2010):

$$X = A * O * S \tag{4.69}$$

A is the number of possible interactions, which can be considered as functional elements or just actions. O is the number of higher level structures or gestalt groups, in short organizational elements and S is the summed entropy of RGB values. This formula has been empirically validated with a correlation of 0.77.

4.1.26 Inconsistency (NoC=1)

4.1.26.1 First formula to measure inconsistency (SoE=5)

This metric is applied in the specific case of Android applications. According to Nielsen, consistency is one of the most important characteristics of usability in user interfaces (Nielsen, 1989). Consistency refers primarily to action sequences, terms, units, layouts, colors, typography, etc within an application state. A clear and clean organization that uses consistency makes it easier to recognize the main objects on the screen and to ignore secondary information when necessary.

The inconsistency metric defined below aims to determine if the entire user interface is visually aesthetic. Therefore, it computes the metric values of the :

1. Number of GUI elements
2. Element smallness (cfr section 4.1.22.1)
3. Misalignment (cfr section 4.1.10.5)
4. Density (cfr section 4.1.9.4)
5. Imbalance (cfr section 4.1.1.7)
6. Color complexity (cfr section 4.1.17.1)
7. Typographic complexity (cfr section 4.1.18.1)

These metric values are calculated for each screen s and compared by calculating the standard deviation. In order to normalize the result in a [0,1] range, the result of the standard deviation calculation is divided by $xmax$, which denotes the maximum value for a certain metric over all screens in order to obtain the inconsistency (I) as shown below (equation 4.70) (Riegler and Holzmann, 2018b):

$$I = \frac{1}{N} \sum_{n=1}^{N} \frac{1}{xmax_n} \sqrt{\frac{1}{J-1} \sum_{j=1}^{J} (x_{j_n} - \overline{x_n})^2} \in [0,1] \tag{4.70}$$

Where x_{j_n} denotes the metric value x on screen j and metric n. N is the number of metrics per screen and $\overline{x_n}$ is the mean complexity value for metric n over all screens. This formula has been empirically validated with a high degree of statistical significance ($P \leq 0.05$ for 3 different applications).

4.2 Qualitative analysis

Now that the different metrics and their corresponding formulas have been presented, a qualitative analysis of the papers found in the SLR will be performed via a meta-analysis of the publications and a comparison of the papers.

4.2.1 Authors

Figure 4.4a shows the number of author/publication. Figure 4.4b shows the number of authors who conducted a study on visual design metrics between 1995, the year when the first papers describing visual design metrics were published (Kokol et al., 1995; Sears, 1995), and 2020.

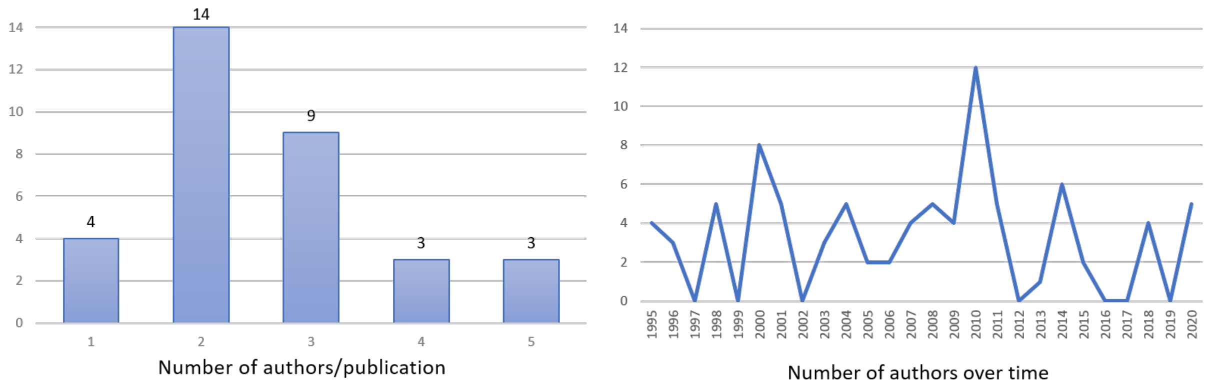


Figure 4.4: a) Number of authors per paper and b) number of author over time (based on Villarreal-Narvaez et al., 2020)

A typical paper on visual design metrics has two authors (14 studies, representing 42.4%) or 3 authors (9 studies, representing 27.3%) (Figure 4.4a). Four studies were conducted by single authors (Constantine, 1996; Sears, 1995; Song, 2007; Zen, 2013) while others are conducted by 4 or 5 authors. Figure 4.4b suggests that a first peak of interest was reached in 2010, the year when the highest number of authors were writing formulas about visual design metrics. It should be noted, however, that new formulas for visual design metrics appear sporadically and that a new peak could therefore be reached in the future.

4.2.2 Venues

In terms of publication venues, authors publish almost as much conference papers (52%) as they write in journals articles (48%). There are no preferred journal or conference proceeding. Indeed, all the publications are dispersed which add difficulty to locate specific information.

4.2.3 Hot topics

Table 4.1 lists the five most influential studies on visual design metrics in decreasing order of their Google Scholar number of citations (dated July 28, 2020).

Authors	**Study**	**Year**	**Citations**
1. Ngo et al., 2003	Modelling interface aesthetics	2003	184
2. Sears, 1995	AIDE: a step toward metric-based interface development tools	1995	122
3. Bauerly and Liu, 2008	Effects of Symmetry and Number of Compositional Elements on Interface and Design Aesthetics	2008	110
4. Bauerly and Liu, 2006	Computational modeling and experimental investigation of effects of compositional elements on interface and design aesthetics	2006	105
5. Zheng et al., 2009	Correlating low-level image statistics with users-rapid aesthetic and affective judgments of web pages	2009	88

Table 4.1: The five most influential studies on visual design metrics according to the number of Google Scholar citations

The most influential study is the study of Ngo: Modelling interface aesthetics. Indeed, in this article there are 14 different formulas of metrics (53% of the total number of metrics (26) described above). The metrics presented in this study have been empirically validated with very high correlation and they have also been applied on different screens. The second most influential study is the study of Andrew Sears. He is one of the first to have developed a visual design metrics, which is why his paper is often cited. Then, there are the studies by Bauerly and Liu, in which they developed and validated metrics to measure balance and symmetry at the microscopic scale. The formulas are computed at the pixel scale. Finally, the study by Zheng et al., 2009 empirically validated the formula of balance, symmetry and equilibrium proposed by Ngo.

4.2.4 Number of formulas for visual design metrics

The evolution of the number of formulas for visual design metrics mentioned in the literature is shown in Figure 4.5. This graph shows that the peak of citations was in 2000. In fact, Ngo's formulas were written in 2000 and the validation of his formulas was done in 2000 and 2001. Although many authors were writing about visual design metrics, very few formulas were mentioned this year.

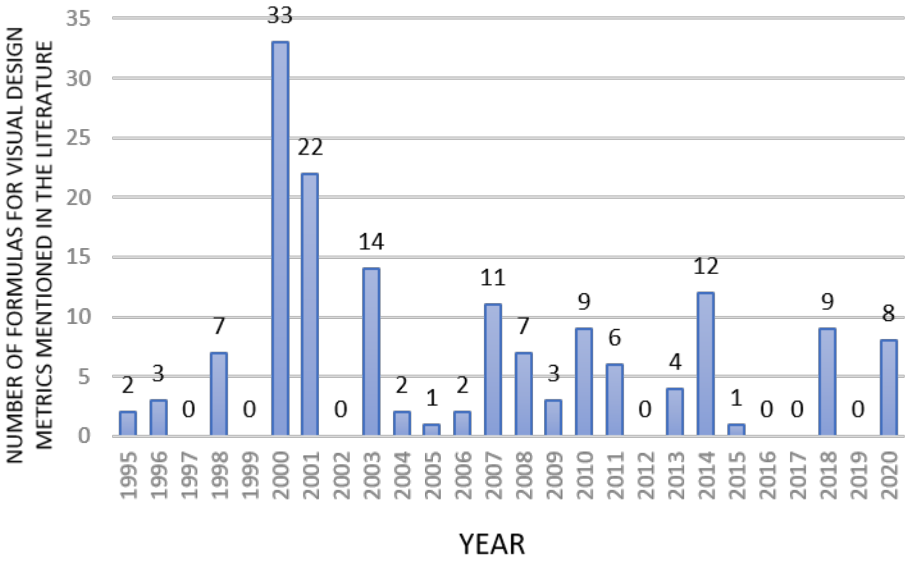


Figure 4.5: Number of visual design metrics mentioned in the literature between 1995 and 2020

4.2.5 Scope of application of the formulas

All the formulas mentioned before 2018 were interested in the visual design of Graphical User Interfaces in general. It is only in 2018 and 2020 that formulas specifically adapted to mobile applications were developed (Riegler and Holzmann, 2018b; Soui et al., 2019). It is therefore conceivable that in the future, this subject will gain more and more interest as the designers will have to develop more aesthetically pleasing Mobile User Interfaces.

Chapter 5

Conclusion and perspectives

In this master thesis, a Systematic Literature Review related to visual design metrics for Graphical User Interfaces was performed, resulting into a corpus of N=33 studies. In these studies, a total of 26 metrics, for which a formula was developed, were collected. 63 formulas have been gathered in this work to measure these different metrics. Among all the metrics, the most popular one is the balance (mentioned 21 times), followed by the symmetry (mentioned 17 times), the order and complexity (mentioned 13 times), the regularity and sequence (mentioned 11 times) and the equilibrium (mentioned 10 time). The most popular formulas for these metrics were the formulas of Ngo which were empirically validated with a high correlation.

In their publications, several authors seem to agree that balance is one of the most important visual technique to achieve (Bauerly and Liu, 2006; Lai et al., 2010; López et al., 2013; Ngo and Byrne, 1998; Vanderdonckt and Gillo, 1999. This metric is so important that 9 different formulas were developed to measure it. Symmetry is also consequential but less than balance. The reason is that symmetry is a more stringent requirement than balance and these two metrics are in some ways related. In their book, López et al., 2013 described the use and benefit of balance as high. Two other metrics were given these attributes: regularity and sequence. Regularity is important because chaotic arrangements are more difficult to use (Noble and Constantine, 1996). A lot of formula were also developed to measure alignment and regularity.

Measuring order and complexity is also crucial for usability and user experience, as it influences emotions, understanding and ease of task execution (Alemerien and Magel, 2014; Stickel et al., 2010). Color-related metrics are also of great consequence because colors add dimension, realism, and attracts an user's eye. Moreover, the color contrast can influence the legibility and can be used to emphasize/deemphasize interface objects.

Throughout this document, we could have seen that each metric can be calculated in different ways by computing a set of visual properties that can be quantified with a fair degree of accuracy. The difficulty to provide an acceptable model of aesthetics measurements resides in the fact that most of the metrics are highly correlated because they are often the resulting transformation of the same inputs attributes (Zen and Vanderdonckt, 2014). This phenomenon was seen through the different formulas proposed by the authors which often included the same variables.

Poorly-documented metrics were also collected (like white-space fraction or concentricity for example). Very few metrics took the semantic aspect into account. In the future, some metrics should be developed for text design in particular (text font, text size or text color). It might also be interesting to investigate certain metrics more deeply in the future.

It is however important to note that there are still a multitude of metrics for which no formula has yet been developed (Vanderdonckt and Gillo, 1999). Among these metric without formula, we can, for example, mention the horizontality/verticality of the interface, the understatement/exaggeration, the neutrality, the singularity, the transparency, the distribution of the interface objects among the four quadrants of the layout, the predictability of the presentation order of the interface objects or the continuity (this list is not exhaustive) (Vanderdonckt and Gillo, 1999). These different metrics without formulas will also need to be studied in greater depth in the future.

Chapter 6

Appendix 1: Formulas of Visual Design Metrics for Graphical User Interfaces

6.1 Balance

6.1.1 First formula to measure balance

A first formula of balance (BM) is given below (equation 6.1)(Ngo, Teo, et al., 2000b):

$$BM = 1 - \frac{|BM_{vertical}| + |BM_{horizontal}|}{2} \in [0, 1] \tag{6.1}$$

$BM_{vertical}$ and $BM_{horizontal}$ are, respectively, the vertical and horizontal balance with:

$$BM_{vertical} = \frac{w_L - w_R}{max(|w_L|, |w_R|)}$$

$$BM_{horizontal} = \frac{w_T - w_B}{max(|w_T|, |w_B|)}$$

with w_j: the visual weight of the different object of the interface. The visual weight depends on the area, color, and shape of the various objects on the interface. L, R, T and B stand respectively for left, right, top and bottom.

$$\begin{aligned} w_j &= \sum_{i}^{n_j} d_{ij} \left(\frac{a_{ij}}{a_{max}} + |c_{ij} - c_{frame}| + s_{ij} \right) \\ & 0 \leq (c_{ij}, c_{frame}, s_{ij}) \leq 1 \\ & j = L, R, T, B \end{aligned}$$

a_{max} is the area of the largest object on the frame with:

$$a_{max} = max(a_{ij}, i = 1, 2, ..., n_j, j = L, R, T, B)$$

Where a_{ij}, c_{ij}, and s_{ij} are, respectively, the area, color and shape of object i on side j. d_{ij} is the distance between the central lines of the object and the frame and n_j is the total number of objects on the side. Each color is given a weighting between 0 (white) and 1 (black) (Ngo, Teo, et al., 2000b).

Another definition of w_j is defined in Ngo and Byrne, 2001a:

$$w_j = \sum_{i}^{n_j} a_{ij} d_{ij} \qquad j = L, R, T, B$$

6.1.2 Second formula to measure balance

A second formula exists in order to measure the balance of Graphical User Interfaces (equation 6.2) (Ngo, Samsudin, et al., 2000):

$$BM = (w_L - w_R, w_T - w_B) \tag{6.2}$$

Where BM is the measure of balance, w_L is the weight on the left hand side (LHS), w_R is the weight on the right hand side (RHS), w_T is the weight on the top, and w_B is the weight on the bottom.

The weight of a layout is the algebraic sum of the weight of its components:

$$w = \sum_{i} a_i d_i$$

Where a_i is the area of an object and d_i is the distance between its vertical central line and the vertical axis.

If BM = (0, 0), then the layout is in a state of balance. If $w_L - w_R > 0$, then the LHS is heavier than the RHS. If $w_L - w_R < 0$, then the RHS is heavier than the LHS. If $w_T - w_B > 0$, then the top is heavier than the bottom. If $w_T - w_B < 0$, then the bottom is heavier than the top.

6.1.3 Fourth formula to measure balance

The balance point, b, is the Cartesian coordinate at the center of the visual mass of the image. This center can easily be found once individual masses are assigned to each pixel. Black pixels are given a mass of one and white pixels given a mass of zero. Equation 6.3 and 6.4 give the center of balance as (x_b, y_b) where w is the image width in pixels, h is the image height in pixels and W, the visual weight, is the summation of black pixels in each pixel column (equation 6.3) or row (equation 6.4). For the experimental procedure and analysis, b is given as a set of normalized coordinates between zero and one (equation 6.5) (Bauerly and Liu, 2006):

$$\sum_{x_i=1}^{w} W_x(x - x_b) = 0 \tag{6.3}$$

$$\sum_{y_i=1}^{h} W_y(y - y_b) = 0 \tag{6.4}$$

$$b = \left(\frac{x_b}{w}, \frac{y_b}{h}\right) \tag{6.5}$$

6.1.4 Fifth formula to measure balance

This formula is derived from the fourth formula of balance (equation 6.5) but takes into account the color attributes of the pixels (equation 4.53). Balance is decomposed in horizontal (B_H) and vertical balance (B_V) as shown below (equation 6.6 and 6.7) (Lai et al., 2010):

$$B_H = 1 - \left|2 \times \frac{x_b}{w} - 1\right| \tag{6.6}$$

$$B_V = 1 - \left|2 \times \frac{y_b}{h} - 1\right| \tag{6.7}$$

B_H and B_V range between 0 and 1. The closer the value of B_H (B_V) is to 1, the better the horizontal (vertical) visual balance obtained. w is the image width in pixels, h

is the image height in pixels and (x_b, y_b) is the center of balance. The center of balance, taking into account the pixel color, is calculated below:

$$W_i = \sum_{j=1}^{h} \Delta C_{ij-B}$$

$$W_j = \sum_{i=1}^{w} \Delta C_{ij-B}$$

Where ΔC is calculated via equation 4.53 with (H_{ij}, S_{ij}, V_{ij}) the HSV color of a pixel block situated at (i, j) and (H_B, S_B, V_B) the background color.

$$\sum_{i=1}^{w} W_i(i - x_b) = 0$$

$$\sum_{j=1}^{h} W_j(j - y_b) = 0$$

Where w is the image width in pixels, h is the image height in pixels and W is the visual weight, depending on the pixel color.

6.1.5 Sixth formula to measure balance

For each GUI element u of each group g on a screen s, the horizontal margins (i.e. left and right margin, respectively $marl_{u,g,s}$ and $marr_{u,g,s}$) and vertical margins (i.e. top and bottom margin, respectively $mart_{u,g,s}$ and $marb_{u,g,s}$) to the neighboring GUI elements and borders are measured. Afterwards, the maximum horizontal and vertical margin among all margins within a group is determined. The mean horizontal margin per group is calculated and related to the maximum horizontal or vertical margin in the group, resulting in $balh_{g,s}$. This is done analogously with the vertical margins, resulting in $balv_{g,s}$.

$$balh_{g,s} = \frac{1}{U_{g,s}} \sum_{u=1}^{U_{g,s}} \frac{\mathrm{marl}_{u,g,s} + \mathrm{marr}_{u,g,s}}{2 * \mathrm{marhmax}_{g,s}}$$

Where $U_{g,s}$ is the number of GUI elements of group g on screen s. The mean balance is weighted with the group area $A_{g,s}$ to obtain $\overline{balh_s}$ in this case. This is done analogously with the vertical margins:

$$\overline{balh_s} = \frac{1}{A_s}\sum_{g=1}^{G_s} A_{g,s} * balh_{g,s}$$

In order to calculate the balance value (B_s), we use the mean horizontal ($\overline{balh_s}$) and vertical margin ($\overline{balv_s}$) and weight them equally (equation 6.8) (Riegler and Holzmann, 2018b):

$$B_s = \frac{\overline{balh_s} + \overline{balv_s}}{2} \tag{6.8}$$

6.2 Equilibrium

6.2.1 First formula to measure equilibrium

The first formula to measure equilibrium is given by the equation 6.9 (Ngo, Teo, et al., 2000b):

$$EM = 1 - \frac{|EM_x| + |EM_y|}{2} \in [0, 1] \tag{6.9}$$

where EM_x and EM_y are the equilibrium components along the x-axis and y-axis respectively.

$$EM_x = \frac{2\sum_i^n a_i(x_i - x_c)}{nb_{frame}\sum_i^n a_i}$$

$$EM_x = \frac{2\sum_i^n a_i(y_i - y_c)}{nh_{frame}\sum_i^n a_i}$$

The equilibrium components depend on the coordinates of the centers of the interface objects (x_i, y_i), the center of the frame (x_c, y_c), the surface area of the objects (a_i), the number of object on the frame (n) and the width and height of the frame (b_{frame} and h_{frame} respectively) (Ngo, Teo, et al., 2000b).

6.2.2 Second formula to measure equilibrium

The second formula to measure equilibrium is given by the equation 6.10 (Ngo, Samsudin, et al., 2000):

$$\begin{aligned} EM &= (x_c, y_c) - (x_0, y_0) \\ &or \\ EM &= (x_c - y_0) - (y_c - y_0) \end{aligned} \tag{6.10}$$

Where (x_0, y_0) is the center of the layout, and (x_c, y_c) is the center of the frame. The center of a set of objects of areas a_1, a_2, a_3... at points $(x1, y1)$, $(x2, y2)$, $(x3, y3)$... in the xy plane is:

$$(x_0, y_0) = \left(\frac{\sum_i a_i x_i}{\sum_i a_i}, \frac{\sum_i a_i y_i}{\sum_i a_i} \right)$$

The equilibrium is reached when EM = (0,0).

6.3 Symmetry

6.3.1 First formula to measure symmetry

Symmetry (SYM) is the extent to which the screen is symmetrical in three directions: vertical, horizontal, and diagonal and is given by the equation 6.11 (Ngo, Teo, et al., 2000b):

$$\text{SYM} = 1 - \frac{|\text{SYM}_{vertical}| + |\text{SYM}_{horizontal}| + |\text{SYM}_{radial}|}{3} \in [0, 1] \tag{6.11}$$

SYM$_{vertical}$, SYM$_{horizontal}$ and SYM$_{radial}$ are, respectively, the vertical, horizontal, and radial symmetries with:

$$\text{SYM}_{vertical} = \frac{\begin{array}{l} |X'_{UL} - X'_{UR}| + |X'_{LL} - X'_{LR}| + |Y'_{UL} - Y'_{UR}| + |Y'_{LL} - Y'_{LR}| + \\ |H'_{UL} - H'_{UR}| + |H'_{LL} - H'_{LR}| + |B'_{UL} - B'_{UR}| + |B'_{LL} - B'_{LR}| + \\ |\Theta'_{UL} - \Theta'_{UR}| + |\Theta'_{LL} - \Theta'_{LR}| + |R'_{UL} - R'_{UR}| + |R'_{LL} - R'_{LR}| \end{array}}{12}$$

$$\mathrm{SYM}_{horizontal} = \frac{\begin{array}{l}|X'_{UL} - X'_{LL}| + |X'_{UR} - X'_{LR}| + |Y'_{UL} - Y'_{LL}| + |Y'_{UR} - Y'_{LR}| + \\ |H'_{UL} - H'_{LL}| + |H'_{UR} - H'_{LR}| + |B'_{UL} - B'_{LL}| + |B'_{UR} - B'_{LR}| + \\ |\Theta'_{UL} - \Theta'_{LL}| + |\Theta'_{UR} - \Theta'_{LR}| + |R'_{UL} - R'_{LL}| + |R'_{UR} - R'_{LR}|\end{array}}{12}$$

$$\mathrm{SYM}_{radial} = \frac{\begin{array}{l}|X'_{UL} - X'_{LR}| + |X'_{UR} - X'_{LL}| + |Y'_{UL} - Y'_{LR}| + |Y'_{UR} - Y'_{LL}| + \\ |H'_{UL} - H'_{LR}| + |H'_{UR} - H'_{LL}| + |B'_{UL} - B'_{LR}| + |B'_{UR} - B'_{LL}| + \\ |\Theta'_{UL} - \Theta'_{LR}| + |\Theta'_{UR} - \Theta'_{LL}| + |R'_{UL} - R'_{LR}| + |R'_{UR} - R'_{LL}|\end{array}}{12}$$

Where X'_j, Y'_j, H'_j, B'_j, Θ'_j and R'_j are, respectively, the normalised values of X_j, Y_j, H_j, B_j, Θ_j and R_j with:

$$X_j = \sum_i^{n_j} |x_{ij} - x_c| \qquad j = UL, UR, LL, LR$$

$$Y_j = \sum_i^{n_j} |y_{ij} - y_c| \qquad j = UL, UR, LL, LR$$

$$H_j = \sum_i^{n_j} h_{ij} \qquad j = UL, UR, LL, LR$$

$$B_j = \sum_i^{n_j} b_{ij} \qquad j = UL, UR, LL, LR$$

$$\Theta_j = \sum_i^{n_j} \left| \frac{y_{ij} - y_c}{x_{ij} - x_c} \right| \qquad j = UL, UR, LL, LR$$

$$R_j = \sum_i^{n_j} \sqrt{(x_{ij} - x_c)^2 + (y_{ij} - y_c)^2} \qquad j = UL, UR, LL, LR$$

Where UL, UR, LL, and LR stand for upper-left, upper-right, lower-left, and lower-right, respectively. (x_{ij},y_{ij}) and (x_c,y_c) are the coordinates of the centers of object i on quadrant j and the frame. b_{ij} and h_{ij} are the width and height of the object and n_j is the total number of objects on the quadrant (Ngo, Teo, et al., 2000b).

6.3.2 Second formula to measure symmetry

There is also a different formula for measuring symmetry. The second formula of symmetry (SM) is given below (equation 6.12) (Ngo, Samsudin, et al., 2000):

$$SM = \{[g_{UL} - g_{LL}], [g_{UR} - g_{LR}], [g_{UL} - g_{UR}], [g_{LL} - g_{LR}], [g_{UL} - g_{LR}], [g_{UR} - g_{LL}]\} \tag{6.12}$$

Where g_j is the weight of the quadrant j. The weight of a layout is the algebraic sum of the weight of its components:

$$g = \sum_i (|x_i - x_c|, |y_i - y_c|, l_i, b_i)$$

l_i is the length of an object, b_i is the breadth of the object, (x_i, y_i) is the center of the object, and (x_c, y_c) is the center of the frame.

6.4 Sequence

6.4.1 First formula to measure sequence

Sequence (SQM) is a measure of how information in a display is ordered in relation to a reading pattern that is common in Western cultures and is given by the equation 6.13 (Ngo, Teo, et al., 2000b):

$$SQM = 1 - \frac{\sum_{k=A,C,S} \sum_{j=UL,UR,LL,LR} |q_j - v_j^k|}{24} \in [0, 1] \tag{6.13}$$

with (q_{UL},q_{UR}, q_{LL}, q_{LR}) $= (4, 3, 2, 1)$ the weighting of the quadrant as it should be to have the best sequence

$$v_j^k = \begin{cases} 4 & \text{if } w_j^k \text{ is the largest in } w^k \\ 3 & \text{if } w_j^k \text{ is the } 2^{nd} \text{ largest in } w^k \\ 2 & \text{if } w_j^k \text{ is the } 3^{rd} \text{ largest in } w^k \\ 1 & \text{if } w_j^k \text{ is the smallest in } w^k \end{cases}$$

$$j = UL, UR, LL, LR \qquad k = A, C, S$$

v_j^k is the effective weighting of each quadrant with:

$$w_j^A = q_j \sum_i^{n_j} a_{ij} \qquad j = UL, UR, LL, LR$$

$$w_j^C = q_j \sum_i^{n_j} |c_{ij} - c_{frame}| \qquad j = UL, UR, LL, LR$$

$$w_j^S = q_j \sum_i^{n_j} s_{ij} \qquad j = UL, UR, LL, LR$$

$$w^k = (w_{UL}^k, w_{UR}^k, w_{LL}^k, w_{LR}^k)$$

Where UL, UR, LL, and LR stand for upper-left, upper-right, lower-left, and lower-right (the 4 quadrants of the interface). A, C and S are respectively the size, the color, and the shape of the objects. a_{ij}, c_{ij}, and s_{ij} are, respectively, the area, color, and shape of object i on quadrant j (Ngo, Teo, et al., 2000b).

6.4.2 Second formula to measure sequence

Another study by Ngo, Samsudin, et al., 2000 has defined sequence as an arrangement of objects in a layout in a way that facilitates the eye movement. In order to improve the sequence, this study suggests to manipulate the user's attention thanks to eye attractors and visual paths.

1) The attractiveness of an object depends on its weight (Ngo, Samsudin, et al., 2000):

$$a \times q_j, \qquad j = UL, UR, LL, LR$$

Where a is the area of the object and q is the value of the quadrant in which the object resides with $(q_{UL}, q_{UR}, q_{LL}, q_{LR}) = (4, 3, 2, 1)$. The greater the weight is, the more attractive the object is.

2) The use of visual paths will improve the sequence because it will lead the eye from left to right and from top to bottom. Indeed, with this configuration, a faster performance is likely, compared to the case where the layout requires numerous jumps between widely separated parts of the display. A complete path is a path that exists from the upper-left quadrant of the screen to the bottom right quadrant. By definition, the length of a path is the number of edges and a short path is more desirable than a longer one (Ngo, Samsudin, et al., 2000).

To find a visual path in a layout:

1. Start from an eye attractor
2. Move to an adjacent attractor on the same level or the level below
3. Repeat 2 until the last node is reach

3) The formula of sequence is given by the equation 6.14 (Ngo, Samsudin, et al., 2000):

$$SQM = (p, d) \tag{6.14}$$

with the following definitions:

$$p = \begin{cases} 1 & \text{If a path exists} \\ 0 & \text{Otherwise} \end{cases}$$

$$d = n_e$$

Where SQM is the measure of sequence, p is the existence of a complete path, d is the length of the path, and n_e is the number of edges. This study suggests that an optimal layout should have a SQM with p = 1 and $d \leq 3$

6.5 Cohesion

6.5.1 First formula to measure cohesion

Cohesion is a measure of how cohesive the screen is and is given by the following formula (equation 6.15) (Ngo and Byrne, 2001b; Ngo, Teo, et al., 2000a, 2000b, 2003; Salimun et al., 2010):

$$CM = \frac{|CM_{fl}| + |CM_{lo}|}{2} \in [0, 1] \tag{6.15}$$

CM_{fl} is a relative measure of the ratios of the layout (l) and the frame of the screen (f) with:

$$CM_{fl} = \begin{cases} t_{fl} & \text{if } t_{fl} \leq 1 \\ \frac{1}{t_{fl}} & \text{Otherwise} \end{cases}$$

CM_{lo} is a relative measure of the ratios of the objects (o) and layout (l) with:

$$CM_{lo} = \frac{\sum_i^n f_i}{n}$$

with:

$$t_{fl} = \frac{\frac{h_{layout}}{b_{layout}}}{\frac{h_{frame}}{b_{frame}}}$$

$$f_i = \begin{cases} t_i & \text{if } t_i \leq 1 \\ \frac{1}{t_i} & \text{Otherwise} \end{cases}$$

with:

$$t_i = \frac{\frac{h_i}{b_i}}{\frac{h_{layout}}{b_{layout}}}$$

Where b_i and h_i, b_{layout} and h_{layout}, and b_{frame} and h_{frame} are the widths and heights of object i, the layout, and the frame, respectively and n is the number of objects on the frame (Ngo, Teo, et al., 2000b).

6.5.2 Second formula to measure cohesion

The 4 components of Screen Layout Cohesion (SLC) are described below (Alemerien and Magel, 2015).

1) Group Layout Cohesion (GLC)

This metric takes into account the relationships (R_{ik}) between widgets (w_i and w_k) based on 17 attributes that are associated with each pair of widgets (Alemerien and Magel, 2015). R_{ik} represents the similarity relatedness between the widgets in terms of structural and aesthetic attributes. Therefore, when the values of R_{ik} are computed, the GLC metric can be calculated, where the value of GLC is the cumulative sum of the weighting cohesion values for all groups on a given user interface as shown in equation 6.16 (Alemerien and Magel, 2015):

$$GLC = \sum_{t=1}^{Groups} \left(\frac{\sum_{i=1}^{w} \sum_{k=1}^{w-i} R_{ik}}{(w*(w-1))/2} \right) * Gw_t \tag{6.16}$$

The ratio of the number of widgets in a specific group (G_t) to the total number of grouped widgets on the screen is considered as the weight of each group (Gw_t). w is the number of widgets, and R_{ik}, the relationship between widgets on a screen is given by equation 6.17:

$$R_{ik} = \frac{\begin{array}{c} R_h + R_w + R_I + R_r + R_b + R_t + R_{ft} + R_c + R_{fs} + \\ R_f + R_{br} + R_{fc} + R_{x1} + R_{x2} + R_{y1} + R_{y2} + R_{al} \end{array}}{\text{Number of relationships}} \tag{6.17}$$

Where h is the height, w the width, l the left margin, r the right margin, b the bottom margin, t the top margin, c the backcolor, fs the font size, f the font style, ft the font name, br the border style, fc the forecolor, $x1$ the x1_location, $x2$ the x2_location, $y1$ the y1_location, $y2$ the y2_location and al the text alignment.

2) UnGroup Layout Cohesion (UGLC)

This metric uses the value of R_{ik} that is computed in equation 6.17, but R_{ik} is the similarity relatedness of ungrouped widgets. The value of $UGLC$ metric is the cumulative sum of the R_{ik} of ungrouped widgets on a given user interface as shown below (equation 6.18) (Alemerien and Magel, 2015):

$$UGLC = \frac{\sum_{i=1}^{w} \sum_{k=1}^{w-i} R_{ik}}{w * (w-1)/2} \tag{6.18}$$

3) Widget Layout Cohesion (WLC)

The value of WLC metric is the cumulative sum of the weighting cohesion values for all widget types on a given user interface as shown below (equation 6.19) (Alemerien and Magel, 2015):

$$WLC = \sum_{t=1}^{Types} \left(\frac{\sum_{i=1}^{w} \sum_{k=1}^{w-i}}{w * (w-1)/2} \right) * Tw_t \tag{6.19}$$

The value of R_{ik} are calculated from equation 6.17. The ratio of the number of widgets from the same type (T) to the total number of widgets on the screen is considered as the weight of each widget type (Tw_t).

4) Semantic Relatedness (SR)

Firstly, the S_{ik} value is calculated, which is the similarity relatedness between two widgets (w_i and w_k) that exist in a group. The value of the SR metric is the cumulative sum of the values of weighting semantic relatedness for all groups on a given user interface as shown below (equation 6.20) (Alemerien and Magel, 2015):

$$SR = \sum_{t=1}^{Groups} \left(\frac{\sum_{i=1}^{w} \sum_{k=1}^{w-i} S_{ik}}{(w * (w-1))/2} \right) * Sw_t \tag{6.20}$$

The ratio of the number of widgets in a specific group (G_t) to the total number of grouped widgets on the screen is considered as the weight of each group (Sw_t).

6.6 Proportion

6.6.1 First formula to measure proportion

The following shapes (with their abbreviation and proportions) are considered as aesthetically pleasing (Marcus, 1992):

- Square (sq) (1:1)
- Square root of two (r2) (1:1.414)
- Golden rectangle (gr) (1:1.618)
- Square root of three (r3) (1:1.732)
- Double square (ds) (1:2)

The formula of the proportion metric (PM) is given below (equation 6.21):

$$PM = \frac{|\mathrm{PM}_{object}| + |\mathrm{PM}_{layout}|}{2} \in [0,1] \tag{6.21}$$

PM_{object} is the difference between the proportions of the objects and the closest proportional shapes:

$$\mathrm{PM}_{object} = \frac{1}{n} \sum_{i}^{n} \left(1 - \frac{min(|p_j - p_i|, j = sq, r2, gr, r3, ds)}{0.5} \right)$$

PM_{layout} is the difference between the proportions of the layout and the closest proportional shape:

$$\mathrm{PM}_{layout} = 1 - \frac{min(|p_j - p_{layout}|, j = sq, r2, gr, r3, ds)}{0.5}$$

with:

$$(p_{sq}, p_{r2}, p_{gr}, p_{r3}, p_{ds}) = \left(\frac{1}{1}, \frac{1}{1.414}, \frac{1}{1.618}, \frac{1}{1.732}, \frac{1}{2} \right)$$

$$p_i = \begin{cases} r_i & \text{if } r_i \leq 1 \\ \frac{1}{r_i} & \text{Otherwise} \end{cases}$$

$$p_{layout} = \begin{cases} r_{layout} & \text{if } r \leq 1 \\ \frac{1}{r_{layout}} & \text{Otherwise} \end{cases}$$

with:

$$r_i = \frac{h_i}{b_i}$$

$$r_{layout} = \frac{h_{layout}}{b_{layout}}$$

Where b_i and h_i are the width and height of object i. b_{layout} and h_{layout} are the width and height of the layout and p_j is the proportion of shape j (Ngo, Teo, et al., 2000b).

6.7 Simplicity

6.7.1 Second formula to measure simplicity

equation 6.22 calculates the alignment for each group of object (GA_i) and the equation 6.23 calculates the Alignment Complexity (AC) for all groups on the screen (Alemerien and Magel, 2014):

$$\text{Group Alignment } (GA_i) = \frac{\sum_{i=1}^{K}(V_p + H_p)}{2K} \in [0,1] \tag{6.22}$$

Where V_p is the number of vertical alignment points, H_p is the number of horizontal alignment points and K is the number of grouped objects on the screen.

$$AC = \sum_{i}^{m} GA_i * \text{Weight(i)} \tag{6.23}$$

Where the weight is the number of objects in a group (i) divided by the total number of grouped objects and m is the number of groups on the screen.

Then, the global alignment is calculated as follows (equation 6.24):

$$\text{Screen Alignment } (SA_i) = \frac{\sum_{i=1}^{N}(V_p + H_p)}{2N} \in [0,1] \tag{6.24}$$

Where V_p is the number of vertical alignment points, H_p is the number of horizontal alignment points and N is the number of ungrouped objects on the screen.

Finally, the Total Screen Alignment Complexity combining the global alignment and the group alignment is calculated below (equation 6.25) (Alemerien and Magel, 2014):

$$TAC = AC * \text{weight1} + SA * \text{weight2} \tag{6.25}$$

Where $weight1$ is the ratio of the number of grouped-objects to the total number of objects on the screen whilst $weight2$ is the ratio of the number of ungrouped-objects to the total number of objects on the screen.

6.8 Regularity or alignment

6.8.1 First formula to measure regularity

It is possible to measure how regular the screen is thanks to the formula of regularity (RM) given below (equation 6.26) (Ngo and Byrne, 2001b; Ngo, Teo, et al., 2000a, 2000b, 2003; Salimun et al., 2010; Zen, 2013):

$$RM = \frac{|RM_{alignment}| + |RM_{spacing}|}{2} \in [0, 1] \tag{6.26}$$

$RM_{alignment}$ is the extent to which the alignment points are minimised with:

$$RM_{alignment} = \begin{cases} 1 & \text{if } n = 1 \\ 1 - \frac{n_{vap}+n_{hap}}{2n} & \text{Otherwise} \end{cases}$$

$RM_{spacing}$ is the extent to which the alignment points are consistently spaced with:

$$RM_{spacing} = \begin{cases} 1 & \text{if } n = 1 \\ 1 - \frac{n_{spacing}-1}{2(n-1)} & \text{Otherwise} \end{cases}$$

Where n_{vap} and n_{hap} are the numbers of vertical and horizontal alignment points. $n_{spacing}$ is the number of distinct distances between column and row starting points and n is the number of objects on the frame.

6.9 Rhythm

6.9.1 First formula to measure rhythm

Rhythm (RHM) is the extent to which the objects are systematically ordered and is given by the equation 6.27 (Kar and Zain, 2007; Ngo and Byrne, 2001b; Ngo, Teo, et al., 2000a, 2000b, 2003; Zain et al., 2011):

$$RHM = 1 - \frac{|\mathrm{RHM}_x| + |\mathrm{RHM}_y| + |\mathrm{RHM}_{area}|}{3} \in [0,1] \tag{6.27}$$

with:

$$\mathrm{RHM}_x = \frac{\begin{array}{c}|X'_{UL} - X'_{UR}| + |X'_{UL} - X'_{LR}| + |X'_{UL} - X'_{LL}| + \\ |X'_{UR} - X'_{LR}| + |X'_{UR} - X'_{LL}| + |X'_{LR} - X'_{LL}|\end{array}}{6}$$

$$\mathrm{RHM}_y = \frac{\begin{array}{c}|Y'_{UL} - Y'_{UR}| + |Y'_{UL} - Y'_{LR}| + |Y'_{UL} - Y'_{LL}| + \\ |Y'_{UR} - Y'_{LR}| + |Y'_{UR} - Y'_{LL}| + |Y'_{LR} - Y'_{LL}|\end{array}}{6}$$

$$\mathrm{RHM}_{area} = \frac{\begin{array}{c}|A'_{UL} - A'_{UR}| + |A'_{UL} - A'_{LR}| + |A'_{UL} - A'_{LL}| + \\ |A'_{UR} - A'_{LR}| + |A'_{UR} - A'_{LL}| + |A'_{LR} - A'_{LL}|\end{array}}{6}$$

where X'_j, Y'_j, and A'_j are, respectively, the normalised values of X_j, Y_j, and A_j with:

$$X_j = \sum_i^{n_j} |x_{ij} - x_c| \qquad j = UL, UR, LL, LR$$

$$Y_j = \sum_i^{n_j} |y_{ij} - y_c| \qquad j = UL, UR, LL, LR$$

$$A_j = \sum_i^{n_j} a_{ij} \qquad j = UL, UR, LL, LR$$

Where UL, UR, LL, and LR stand for upper-left, upper-right, lower-left, and lower-right, respectively. (x_{ij}, y_{ij}) and (x_c, y_c) are the coordinates of the centers of object i on quadrant j and the frame. a_{ij} is the area of the object and n_j is the total number of objects on the quadrant.

6.10 Color complexity

6.10.1 First formula to measure color complexity

This metric of color complexity is applied specifically to Android applications. Firstly, the number of dominant colors for screen s, $numc_s$, has to be determined. Dominant colors are colors that are occurring more frequently on a pixel-basis than non-dominant colors. The dominant colors are then compared pairwise to calculate all possible color combination matches. The complexity value of the color combination match (ccm_s) is calculated as a weighted difference of the H, S and V values of the two colors (equation 6.28). The final color combination match metric, $\overline{ccm_s}$, is calculated as the average of these matches. The more different the dominant colors on a screen s are, the bigger the value for $\overline{ccm_s}$.

$$ccm_{c1,c2} = 0.70 * \left(\frac{\delta H_{c1,c2}}{H_{max}}\right) + 0.15 * \left(\frac{S_{max} - \delta S_{c1,c2}}{S_{max}}\right) + 0.15 * \left(\frac{V_{max} - \delta V_{c1,c2}}{V_{max}}\right) \tag{6.28}$$

For each screen, the number of used colors $numc_s$ as well as the mean color combination match $\overline{ccm_s}$ is determined. For an app-wide analysis, the color histogram h_s of each screen s is computed. The comparison of the current with the following screen's histogram is performed using the Bhattacharyya distance. If the histograms are similar regarding the distribution of the used colors, the distance is small. That means, that even if all single screens have visually pleasant color combinations, but the used colors are not consistently used throughout the app, this problem is detected (Riegler and Holzmann, 2018b).

The formula of color complexity shown in equation 6.29 is finally obtained (Riegler and Holzmann, 2018b):

$$C_s = \frac{\frac{numc_s}{numcmax} + (1 - \overline{ccm_s}) + c_{h_s,h_{s+1}}}{3} \tag{6.29}$$

Where $numcmax$ denotes the maximum number of dominant colors on a certain screen of the whole user interface, $numc_s$ is the number of dominant colors c_1,c_2,..on screen s, $\overline{ccm_s}$ is the average color combination match of the dominant colors, the bigger the value f and $c_{h_s,h_{s+1}}$ is the color combination match of the current screen's histogram h_s and the succeeding screen's histogram h_{s+1}.

6.11 Element smallness

6.11.1 First formula to measure element smallness

Firstly, the mean element width and height of each group g of a screen s is calculated from the $\text{width}_{u,g,s}$ and $\text{heigth}_{u,g,s}$ of each element u in this group by calculating the arithmetic mean over all elements $U_{g,s}$. In order to calculate the mean element $\overline{width_s}$ and $\overline{height_s}$ for a screen s, each element width and height of a group g with the area $A_{g,s}$ are weighted as follows:

$$\overline{width_s} = \frac{1}{A_s} \sum_{g=1}^{G_s} A_{g,s} \frac{1}{U_{g,s}} \sum_{u=1}^{U_{g,s}} \text{width}_{u,g,s}$$

$$\overline{height_s} = \frac{1}{A_s} \sum_{g=1}^{G_s} A_{g,s} \frac{1}{U_{g,s}} \sum_{u=1}^{U_{g,s}} \text{height}_{u,g,s}$$

Where G_s is the weighted arithmetic mean of the width and height elements over all groups and $U_{g,s}$ is the number of GUI elements of group g on screen s. Then, $\overline{width_s}$ and $\overline{height_s}$ are combined to a single metric : E_s, the screen smallness (equation 6.30) (Riegler and Holzmann, 2018b):

$$E_s = \frac{\left(1 - \frac{\overline{\text{width}_s}}{\text{width}_s}\right) + \left(1 - \frac{\overline{\text{height}_s}}{\text{height}_s}\right)}{2} \tag{6.30}$$

Smaller objects tend to be harder to identify and interact with, so this metric needs a higher score for smaller elements and smaller score for higher elements. Therefore, $\overline{width_s}$ and $\overline{height_s}$ are normalized between 0 and 1 using the $width_s$ and $height_s$ of the screen s.

6.12 Order and complexity

6.12.1 Second formula to measure order and complexity

The second formula describing M, the measure of order and complexity, is defined as (equation 6.31) (Ngo, Samsudin, et al., 2000):

$$M = \frac{O_{BM} + O_{EM} + O_{SM} + O_{SQM}}{C} \in [0, 3] \qquad (C > 0) \tag{6.31}$$

Where O_{BM} is a measure of balance, O_{EM} is a measure of equilibrium, O_{SM} is a measure of symmetry, and O_{SQM} is a measure of sequence. The complexity C of a layout is defined as the number of its components.

With order, the objects are seen as forming one piece. The way to maximize screen order is to achieve balance, equilibrium, symmetry, and sequence while the way to minimize screen complexity is to reduce the number of screen objects (Ngo, Samsudin, et al., 2000). It is noteworthy that the second formula takes less element in consideration and is therefore easier but less precise than the first formula.

with the following definitions:

$$O_{BM} = \begin{cases} 0.5 & \text{if } w_L = w_R \text{ and } w_T = w_B \\ 0.25 & \text{if } w_L = w_R \text{ or } w_T = w_B \\ 0 & \text{Otherwise} \end{cases} \qquad \text{(from equation 6.2)}$$

$$O_{EM} = \begin{cases} 0.5 & \text{if } x_c = x_o \text{ and } x_c = x_o \\ 0.25 & \text{if } x_c = x_o \text{ or } x_c = x_o \\ 0 & \text{Otherwise} \end{cases} \qquad \text{(from equation 6.10)}$$

$$O_{SM} = \begin{cases} 1.5 & \text{if } (g_{UL} = g_{UR} \text{ and } g_{LL} = g_{LR}) \text{ and} \\ & (g_{UL} = g_{LL} \text{ and } g_{UR} = g_{LR}) \text{ and} \\ & (g_{UL} = g_{LR} \text{ and } g_{UR} = g_{LL}) \\ 1 & \text{if } (g_{UL} = g_{UR} \text{ and } g_{LL} = g_{LR}) \text{ and} \\ & (g_{UL} = g_{LL} \text{ and } g_{UR} = g_{LR}) \\ 0.5 & \text{if } (g_{UL} = g_{UR} \text{ and } g_{LL} = g_{LR}) \text{ or} \\ & (g_{UL} = g_{LL} \text{ and } g_{UR} = g_{LR}) \text{ or} \\ & (g_{UL} = g_{LR} \text{ and } g_{UR} = g_{LL}) \\ 0 & \text{Otherwise} \end{cases} \qquad \text{(from equation 6.12)}$$

$$O_{SQM} = \begin{cases} 0.5 & \text{if SQM} \geq 5 \\ 0 & \text{Otherwise} \end{cases} \qquad \text{(from equation 6.14)}$$

Notice that different order factors have different ranges, depending on their degree of importance. Balance and equilibrium form a collective factor with a range between 0 and 1. According to experts, the order of significance from most to least is symmetry, balance (or equilibrium) and sequence. Therefore, the weighting criteria for symmetry and sequence are 1.5 and 0.5, respectively. The range of the measure is from 0 to 3 (Ngo, Samsudin, et al., 2000).

6.12.2 Third formula to measure complexity

Guangfeng Song has developed an algorithm to measure complexity and visual segmentation of web pages. A piece of computer software was developed in Python to visually analyze web pages. The software start analyzing a web page by splitting it into two images: one image (TI) only contains English text at its original location on the page and the other image (NTI) only contains non-text information, also at its original location on the page. Thanks to these two images, the software is able to divide a web page into pieces and each piece into even smaller pieces until further division is not possible or not desirable. The result of this hierarchical segmentation process is a list of blocks for each web page. In addition to this process, a weighted average formula has been developed to derive a complexity estimate of each page. The formula is given below (equation 6.32) (Song, 2007):

$$C = \frac{\sum C_k S_k}{S_k} \tag{6.32}$$

Where C_k and S_k are the complexity and size of sub-block k respectively.

Bibliography

Alemerien, K., & Magel, K. (2014). Guievaluator: A metric-tool for evaluating the complexity of graphical user interfaces., In *Seke*.

Alemerien, K., & Magel, K. (2015). Slc: A visual cohesion metric to predict the usability of graphical user interfaces, In *Proceedings of the 30th annual acm symposium on applied computing*.

Altaboli, A., & Lin, Y. (2011). Investigating effects of screen layout elements on interface and screen design aesthetics. *Advances in Human-Computer Interaction*, *2011*.

Amerland, D. (2014). *Qu'est-ce que la recherche sémantique ?* Retrieved June 26, 2020, from https://la-cascade.io/quest-ce-que-la-recherche-semantique/

Bast, H., Buchhold, B., & Haussmann, E. (2016). Semantic search on text and knowledge bases. *Foundations and Trends® in Information Retrieval*, *10*, 119–271. https://doi.org/10.1561/1500000032

Bauerly, M., & Liu, Y. (2005). Development and validation of a symmetry metric for interface aesthetics, In *Proceedings of the human factors and ergonomics society annual meeting*. SAGE Publications Sage CA: Los Angeles, CA.

Bauerly, M., & Liu, Y. (2006). Computational modeling and experimental investigation of effects of compositional elements on interface and design aesthetics. *International Journal of Human-Computer Studies*, *64*(8), 670–682.

Bauerly, M., & Liu, Y. (2008). Effects of symmetry and number of compositional elements on interface and design aesthetics. *Intl. Journal of Human–Computer Interaction*, *24*(3), 275–287.

Birkhoff, G. D. (1933). *Aesthetic measure by george d. birkhoff*. Harvard University Press.

Brooks, R. (1965). Search time and color coding. *Psychonomic Science*, *2*(1-12), 281–282.

Casteleyn, S., Garrig'os, I., & Maz'on, J.-N. (2014). Ten years of rich internet applications: A systematic mapping study, and beyond. *ACM Trans. Web*, *8*(3). https://doi.org/10.1145/2626369

Chettaoui, N., & Bouhlel, M. S. (2017). I2evaluator: An aesthetic metric-tool for evaluating the usability of adaptive user interfaces, In *International conference on advanced intelligent systems and informatics*. Springer.

Colin, N., Landier, A., Mohnen, P., & Perrot, A. (2015). Économie numérique. *Notes du conseil danalyse economique*, (7), 1–12.

Constantine, L. L. (1996). Visual coherence and usability: A cohesion metric for assessing the quality of dialogue and screen designs, In *Proceedings sixth australian conference on computer-human interaction*. IEEE.

Dondis, D. (1973). *A primer of visual literacy*. MA, USA, The MIT Press.

Fessenden, T. (2019). *Visual design terms cheat sheet*. Retrieved July 5, 2020, from https://www.nngroup.com/articles/visual-design-cheat-sheet/

Fu, F., Chiu, S.-Y., & Su, C. H. (2007). Measuring the screen complexity of web pages, In *Symposium on human interface and the management of information*. Springer.

Gordon, K. (2020). *5 principles of visual design in ux*. Retrieved July 4, 2020, from https://www.nngroup.com/articles/principles-visual-design/

Harrington, S. J., Naveda, J. F., Jones, R. P., Roetling, P., & Thakkar, N. (2004). Aesthetic measures for automated document layout, In *Proceedings of the 2004 acm symposium on document engineering*.

Hartono, E., & Holsapple, C. W. (2019). Website visual design qualities: A threefold framework. *ACM Trans. Manage. Inf. Syst.*, *10*(1). https://doi.org/10.1145/3309708

ISO 25000. (2019). *The iso/iec 25000 series of standards*. Retrieved June 22, 2020, from https://iso25000.com/index.php/en/iso-25000-standards

Kar, T. K., & Zain, J. M. (2007). Measuring aesthetics of mandarin learning web pages : Are users ' perceptions congruent with the measured values of aesthetics-measurement application (ama) ?

Kitchenham, B., Pretorius, R., Budgen, D., Brereton, O. P., Turner, M., Niazi, M., & Linkman, S. (2010). Systematic literature reviews in software engineering–a tertiary study. *Information and software technology*, *52*(8), 792–805.

Kokol, P., Rozman, I., & Venuti, V. (1995). User interface metrics. *ACM SIGPLAN Notices*, *30*(4), 36–38.

Lai, C.-Y., Chen, P.-H., Shih, S.-W., Liu, Y., & Hong, J.-S. (2010). Computational models and experimental investigations of effects of balance and symmetry on the aesthetics of text-overlaid images. *International journal of human-computer studies*, *68*(1-2), 41–56.

Lau, J., & Wilson, D. N. (1998). Experience with software quality/usability metrics, In *Software quality management vi*. Springer.

Laublet, P., Reynaud, C., & Charlet, J. (2002). Sur quelques aspects du web sémantique. *Assises du GDR I*, *3*, 59–78.

Lavie, T., & Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of web sites. *International journal of human-computer studies*, *60*(3), 269–298.

Liberati, A., Altman, D., Tetzlaff, J., & Mulrow, C. (2009). Gotzsch e pc, ioannidis jp, etal. the prisma statement for reportin g systematic reviews and meta-analyses of studies that evaluat e health care interventions: Explanation and elaboration. *PLoS Med*, *6*, e1000100.

Liu, S. (2020). *Market share held by the leading computer (desktop/tablet/console) operating systems worldwide from january 2012 to february 2020*. Retrieved July 22, 2020, from https://www.statista.com/statistics/268237/global-market-share-held-by-operating-systems-since-2009/

López, S. G., Simarro, F. M., & López, P. G. (2013). Balores: A framework for quantitative user interface evaluation, In *New trends in interaction, virtual reality and modeling*. Springer.

Luria, S., Neri, D. F., & Jacobsen, A. R. (1986). The effects of set size on color matching using crt displays. *Human factors*, *28*(1), 49–61.

Lutteroth, C., & Weber, G. (2006). User interface layout with ordinal and linear constraints, In *Proceedings of the 7th australasian user interface conference-volume 50*.

Marcus, A. (1992). *Graphic design for electronic documents and user interfaces* (tech. rep.).

Mengist, W., Soromessa, T., & Legese, G. (2020). Method for conducting systematic literature review and meta-analysis for environmental science research. *MethodsX*, *7*, 100777.

Miniukovich, A., & De Angeli, A. (2014). Quantification of interface visual complexity, In *Proceedings of the 2014 international working conference on advanced visual interfaces*.

Morris, S. (2020). *Tech 101: Graphic design, ui design, visual design, and web design—what's the difference?* Retrieved July 4, 2020, from https://skillcrush.com/blog/what-is-visual-design/

Mullet, K., & Sano, D. (1995). *Designing visual interfaces: Communication oriented techniques*. USA, Pearson Education.

Ngo, D. C. L., & Byrne, J. G. (1998). Aesthetic measures for screen design, In *Proceedings 1998 australasian computer human interaction conference. ozchi'98 (cat. no. 98ex234)*. IEEE.

Ngo, D. C. L., & Byrne, J. G. (2001a). Another look at a model for evaluating interface aesthetics. *International Journal of Applied Mathematics and Computer Science*, *11*, 515–535.

Ngo, D. C. L., & Byrne, J. G. (2001b). Application of an aesthetic evaluation model to data entry screens. *Computers in Human Behavior*, *17*(2), 149–185.

Ngo, D. C. L., Samsudin, A., & Abdullah, R. (2000). Aesthetic measures for assessing graphic screens. *J. Inf. Sci. Eng*, *16*(1), 97–116.

Ngo, D. C. L., Teo, L. S., & Byrne, J. G. (2000a). Formalising guidelines for the design of screen layouts. *Displays*, *21*(1), 3–15.

Ngo, D. C. L., Teo, L. S., & Byrne, J. G. (2000b). A mathematical theory of interface aesthetics, In *Visual mathematics*. Mathematical Institute SASA.

Ngo, D. C. L., Teo, L. S., & Byrne, J. G. (2003). Modelling interface aesthetics. *Information Sciences*, *152*, 25–46.

Nielsen, J. (1989). Coordinating user interfaces for consistency. *ACM Sigchi Bulletin*, *20*(3), 63–65.

Noble, J., & Constantine, L. L. (1996). Interactive design metric visualization: Visual metric support for user interface design, In *Proceedings sixth australian conference on computer-human interaction*. IEEE.

Norman, D. A. (1988). *The psychology of everyday things*. Basic books.

Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research.

Ordoñez, K., Hilera, J., & Cueva, S. (n.d.). Model-driven development of accessible software: A systematic literature review.

Purchase, H. C., Hamer, J., Jamieson, A., & Ryan, O. (2011). Investigating objective measures of web page aesthetics and usability, In *Proceedings of the twelfth australasian user interface conference-volume 117*.

Riegler, A., & Holzmann, C. (2018a). Measuring Visual User Interface Complexity of Mobile Applications With Metrics. *Interacting with Computers*, *30*(3), 207–223. https://doi.org/10.1093/iwc/iwy008

Riegler, A., & Holzmann, C. (2018b). Measuring visual user interface complexity of mobile applications with metrics. *Interacting with Computers*, *30*(3), 207–223.

Salimun, C., Purchase, H. C., Simmons, D. R., & Brewster, S. (2010). The effect of aesthetically pleasing composition on visual search performance, In *Proceedings of the 6th nordic conference on human-computer interaction: Extending boundaries*.

Sears, A. (1995). Aide: A step toward metric-based interface development tools, In *Proceedings of the 8th annual acm symposium on user interface and software technology*.

Siang, T. Y. (2020). *Visual design*. Retrieved July 4, 2020, from https://www.interaction-design.org/literature/article/the-building-blocks-of-visual-design

Song, G. (2007). Analysis of web page complexity through visual segmentation, In *International conference on human-computer interaction*. Springer.

Soui, M., Chouchane, M., Mkaouer, M. W., Kessentini, M., & Ghedira, K. (2019). Assessing the quality of mobile graphical user interfaces using multi-objective optimization. *Soft Computing*, 1–30.

Stickel, C., Ebner, M., & Holzinger, A. (2010). The xaos metric–understanding visual complexity as measure of usability, In *Symposium of the austrian hci and usability engineering group*. Springer.

Streveler, D. J., & Wasserman, A. I. (1984). Quantitative measures of the spatial properties of screen designs, In *Proceedings of the ifip tc13 first international conference on human-computer interaction*.

U.S. Dept. of Health and Human Services. (2006). *The research based web design and usability guidelines, visual design basics*. Retrieved July 4, 2020, from https://www.usability.gov/what-and-why/visual-design.html

Vanderdonckt, J., & Gillo, X. (1999). Visual techniques for traditional and multimedia layouts. https://doi.org/10.1145/192309.192334

Villarreal-Narvaez, S., Vanderdonckt, J., Vatavu, R.-D., & Wobbrock, J. A. (2020). A systematic review of gesture elicitation studies: What can we learn from 216 studies, In *Proceedings of acm int. conf. on designing interactive systems (dis'20)*.

Walleriusz. (2015). *Principles of design*. Retrieved July 14, 2020, from https://fr.slideshare.net/walleriusz/principles-of-design-44496102/4

Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering, In *Proceedings of the 18^{th} international conference on evaluation and assessment in software engineering*.

Zain, J. M., Tey, M., & Goh, Y. (2011). Probing a self-developed aesthetics measurement application (sda) in measuring aesthetics of mandarin learning web page interfaces. *arXiv preprint arXiv:1101.1606*.

Zen, M. (2013). Metric-based evaluation of graphical user interfaces: Model, method, and software support, In *Proceedings of the 5th acm sigchi symposium on engineering interactive computing systems*.

Zen, M., & Vanderdonckt, J. (2014). Towards an evaluation of graphical user interfaces aesthetics based on metrics, In *2014 ieee eighth international conference on research challenges in information science (rcis)*. IEEE.

Zen, M., & Vanderdonckt, J. (2016). Assessing user interface aesthetics based on the inter-subjectivity of judgment (S. Faily, N. Jiang, H. Dogan, & J. Taylor, Eds.). In S. Faily, N. Jiang, H. Dogan, & J. Taylor (Eds.), *HCI 2016 - fusion! proceedings of the 30th international BCS human computer interaction conference, BCS HCI 2016, bournemouth university, poole, uk, 11-15 july 2016*, BCS. http://ewic.bcs.org/content/ConWebDoc/56903

Zheng, X. S., Chakraborty, I., Lin, J. J.-W., & Rauschenberger, R. (2009). Correlating low-level image statistics with users-rapid aesthetic and affective judgments of web pages, In *Proceedings of the sigchi conference on human factors in computing systems*.

A Systematic Literature Review of Visual Design Metrics for Graphical User Interfaces

Resume

Visual design of Graphical User Interfaces (GUI) aims to contribute to their usability by manipulating their visual components, such as their widgets, menus, contents and their layout by relying on a variety of techniques borrowed from general visual design and aesthetic properties. Therefore, a significant portion of GUI evaluation, but not the only one, concerns the visual aspects of GUIs. Many metrics attempts to characterize, express, and measure quality properties of the visual design of graphical user interfaces, but with different names, definitions, formulas, interpretations and empirical evidences. This thesis performs a Systematic Literature Review (SLR) for identifying and characterizing visual design metrics for graphical user interfaces on six scientific digital libraries, augmented by a snowballing procedure. We will identify a corpus of metrics that are systematically analyzed. Each identified metric will be described and accompanied by one or more formulas, the different formulas and metrics will be compared and their analysis will suggest some further consideration on how to apply and research further metrics in this area.

Adrien Habay

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Louvain School of Management
Place des Doyens, 1 bte L2.01.01, 1348 Louvain-la-Neuve
Boulevard Emile Devreux 6, 6000 Charleroi, Belgique
Chaussée de Binche 151, 7000 Mons, Belgique
www.uclouvain.be/lsm